News Recommendation using Advanced LLM Embedding Models

By

Student Name: Ahmed Fahim
ID: 0122320001

Submitted in partial fulfilment of the requirements
of the degree of Master of Science in Computer Science and Engineering

August 10, 2026

Department of Computer Science and Engineering
United International University

Approval Certificate

This thesis titled **'News Recommendation using Advanced LLM Embedding Models'** submitted by **'Ahmed Fahim'**, Student ID: **'0122320001'**, has been accepted as **Satisfactory** in fulfillment of the requirement for the degree of Master of Science in Computer Science and Engineering on 'August 8, 2026'.

Board of Examiners

1.

Supervisor

Dr. Mohammad Nurul Huda
Professor of the Dept. of CSE
United International University
Dhaka, Bangladesh

2.

Head Examiner

Dr. Jannatun Noor Mukta
Associate Professor of the Dept. of CSE
United International University
Dhaka, Bangladesh

3.

Examiner-I

Prof. Dr. Khondaker Abdullah Al Mamun
Professor of the Dept. of CSE
United International University
Dhaka, Bangladesh

4.

Examiner-II

Mr. Nahid Hossain
Assistant Professor of the Dept. of CSE
United International University
Dhaka, Bangladesh

5.

Ex-Officio

Dr. Dewan Md. Farid
Professor of the Dept. of CSE
United International University
Dhaka, Bangladesh

Declaration

This is to certify that the work entitled **'News Recommendation using Advanced LLM Embedding Models'** is the outcome of the research carried out by me under the supervision of **'Dr. Mohammad Nurul Huda, Professor of the Dept. of CSE'**

Student Name: Ahmed Fahim
MSCSE Program
Student ID: 0122320001
Dept. of Computer Science and Engineering
United International University
Dhaka, Bangladesh

In my capacity as supervisor of the candidate's thesis, I certify that the above statements are true to the best of my knowledge.

Dr. Mohammad Nurul Huda
Professor of the Dept. of CSE
United International University
Dhaka, Bangladesh

Abstract

Online news recommendation is a subset of recommendation systems used by online news portals to filter news articles according to user interest. News recommendation methods generally have three main components: news encoder, user encoder, and click predictor. After the success and popularity of BERT, many methods were developed that incorporate fine-tuned BERT for feature extraction. However, there have been further advancements in the field of LLM based embedding models, generalized embedding models, and decoder-based embedding models, that have not been studied yet for news recommendation systems. Most news recommendation studies consider only ranking a candidate set of news items for each user. In this study we compared this with other news recommendation tasks as well: user-news classification; all user classification for given news; candidate news classification for user; and candidate news ranking for user. We also studied multilingual news recommendation and the inclusion of a fake news classification component. In this study we investigated the effect of using five BERT-based models, five large decoder-based models, and one proprietary model as news encoders for news recommendation. We also studied the effect of different model inputs (title, category, and abstract) on news and user representation, as well as the effect of four user encoders:average pooling, attention pooling, dense layer + attention pooling, multi-head self-attention + attention pooling. Inclusion of the fake news classification component improved results for news recommendation tasks 1, 2 and 3, but reduced performance for task 4. For the first task, multilingual-e5-large-instruct with title, category, subcategory, and abstract input and dense layer + attention pooling for user encoder achieved the highest accuracy of 0.9149; for the second task, text-embedding-3-large with title input and multi-head self-attention + attention pooling for user encoder achieved the highest accuracy of 0.8902; for the third task, multilingual-e5-large-instruct with title as input and multi-head self-attention + attention pooling for user encoder achieved the highest accuracy of 0.8641; and for the fourth task, mxbai-embed-large-v1 with title, category and abstract as input and multi-head self-attention + attention pooling for user encoder achieved the highest AUC of 0.6648. Our findings indicate that frozen generalized embeddings can provide competitive news representations without embedding-model fine-tuning, although performance depends strongly on the recommendation task and user encoder.

Acknowledgements

All the praises and thanks are to Allah, the most gracious and the ever merciful. HE has given me the strength, courage, and patience to carry out this work.

Then, this thesis represents a great deal of time and effort not only on my part but also on the part of my supervisor, Prof. Dr. Mohammad Nurul Huda, who has been a constant guide and mentor. I am truly thankful for his continuous support and valuable feedback that helped me complete this research in the promising fields of Artificial Intelligence, Natural Language Processing and Machine Learning.

I am also thankful to my thesis committee members for providing insightful and constructive comments to improve the quality of this thesis.

Then thanks to family and friends for their unwavering love and emotional support that kept me going on this path.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

Recommender systems are a type of information filtering system designed to help users find items of interest from a large number of choices by suggesting items that are most pertinent to the user [1, 2, 3]. These are commonly employed in sites like e-commerce, online news portals and, movie and music streaming sites to help users make decisions quickly and to their liking. The term 'item' denotes whatever the system is suggesting to the user, which can be, news articles, movies, products etc. These systems are particularly beneficial for users who lack the knowledge or personal experience to search through and find items of interest from an overwhelming number of items a website may have to offer [3]. One of the most common example of recommender system is the e-commerce site Amazon.com which uses these systems to suggest items to users [4, 5].

Personalized recommendations help different users and groups get custom and diverse suggestions according to their preference. Non-personalized recommendations are also sometimes used, for example, the list of top 50 songs on Spotify. These recommendations are usually much easier to generate but not as valuable to the user because they lack personalization.

Recommender systems try to predict the most suitable set of items for the user based on their preference and taking into account other limitations (such as items being available in stock); these personalized recommendations are then shown to the user as a ranked list of items. The user preference is learnt from interactions of the user with the system. This can be in the form of explicit feedbacks such as star ratings or text reviews, or implicit feedbacks such as purchase or click history.

The idea of recommender system is from the observation that people often take advice from others in making everyday decisions [2, 3]. For instance, people take advice from friends and peers on which restaurants to dine at or which movies to watch; even in hiring processes of companies recommendation letters from past employers have significant impact.

Early recommender systems tried to replicate this behavior by making use of the recommendations from groups of users to find the best item to recommend to the current user. The recommendations were for items liked by users with similar tastes. This algorithm

is called collaborative filtering [2]. This is based on the assumption that if the current user has agreed with the preference or recommendation of some other group of users in the past then other items preferred or recommended by this group of users should also be preferred and suitable for the current user.

As e-commerce sites grew, they required the users to filter and find their desired item from an ever increasing sea of choices. This bolstered the need to build effective systems that could help users navigate this maze by providing them with tailored items.

Users are usually overwhelmed by the exponential growth in the number of range of information, products and services available online, which leads them to making sub optimal choices. The sea of choices instead of being helpful to the user by giving them the potential to make better choices, ends up causing decision fatigue and overall negative experience for the user. This phenomenon was studied in [6] and it was found that although up to a point more choice did lead to increased empowerment, autonomy and freedom to meet one's preferences, if we are faced with more choices than that, then it leads to decision fatigue, regret and an overall less satisfaction.

1.1 News Recommender Systems

We live in an internet era where nearly everyone has access to smartphone, computer, tablet or some other means of connecting to the internet. And as such, 86 % of adults in the US get news from digital sources, according to the report by Pew Research Center in 2024. Even though online news media might be more advanced than printed news, both have the same criteria for newsworthiness [7]. The causes for this might be the absence of fixed methodologies for providing the user with a diverse set of news within a suitable time-frame and deficiencies in the system's ability to make an accurate model of the user behavior. This makes it important to investigate and develop methods and algorithms to meet the information requirements of the readers by supplying them with timely and personalized news updates [8].

Many news companies like BBC, The Washington Post, The Guardian etc. allow their news readers to access their latest news articles from any place and at any time by providing online news portals. These online portals aim to engage large volume of traffic to their websites, by making use of recommender systems in order to provide personalized news recommendations to their users which enhances the user experience. In the domain of recommender systems "user experience" covers multiple aspects like how useful, usable and how good it is in satisfying users' personalization needs [9, 10].

However, it is tough recommending the right news stories to readers. The news domain faces unique challenges that other uses where recommenders are used do not face.

One of the biggest challenges is timeliness. Most news stories are relevant only for a very small amount of time. The popularity, recency, and current trends matter significantly, especially with the large volume of new stories appearing every second. Another major issue is how dynamic the user behaviors are. People's news preferences can be long-term

or short-term and can change gradually or suddenly.

Recently, there's also been a rise in the manipulation of news content, like the spread of false news and propaganda [11]. This creates a rising challenge in ensuring the quality control of news content.

Mobile devices and applications have transformed the way people come across news. Now, news aggregators such as Google and Microsoft News, and social media platforms, including Facebook and X, formerly Twitter, are the main sources for news consumption. Once incorporated into a news portal, personalized recommendation systems can deliver curated news feeds tailored to individual user preferences. This makes personalization an important feature for news readers.

Before exploring the specific challenges of news recommender systems, it's important to understand how the news domain differs from other recommendation areas, like movies, music, or books.

Average Consumption Time: News stories are typically consumed very quickly. The time taken for users to consume (read) a news story is estimated by the article length. A news article averages under 200 words, and a Pew research center report indicates that stories under 250 words take readers about 43 seconds to engage with. In contrast, articles over 5,000 words engage the reader for at least 270 seconds (4.5 minutes). This is significantly shorter than a typical movie (90-120 minutes), a song (3-5 minutes), or a book (much longer).

Lifespan of News Items: News articles are relevant only for a very small amount of time, often becoming irrelevant in minutes, hours, or a few days. This contrasts sharply with other items commonly suggested using recommender systems, such as, music, books, or movies, which can remain relevant for days, weeks, months, or even years. Additionally, comments or reviews for news items appear almost immediately after release, unlike the longer delay seen with other products, which is usually in the order of days, months or even years.

Catalog Size of News Items: News systems are constantly updated with new news content, sometimes receiving thousands of articles per hour. While music or movie services might have catalogs ranging from hundreds to thousands of items, these items, as discussed earlier, are usually relevant for much longer. Therefore the combination of onslaught of new news articles, and their short lifespan poses a significant challenge in news recommendation.

Expected Request-Response Rate: Providing the news in a timely manner is very important. This is considered as one of the distinguishing characteristics of the news domain. News aggregator sites, such as Google or Microsoft News can experience over 100 requests per second, and responses need to be sent within 100 milliseconds in order to provide real-time updates [12].

Sequential Consumption: Users often read news sequentially, looking for updates on various stories of interest. Listening to music also has a sequential consumption pattern, however, a major difference is that in music consumption, users may want to listen to the same song multiple times or repeated again after a few songs [13]. In news domain however,

readers prioritize staying informed of the latest updates about different or ongoing events they are interested in, rather than re-reading old stories [14].

Diversity: In movie or music domains, people generally stick to their preferred genre(s) and only at times check out other genres when in different contexts. In the news domain, however, diversity is not only useful but very important for a number of reasons. Firstly, a diverse list of news articles plays a significant role in maintaining readers' engagement throughout their online reading experience. Secondly, and perhaps more importantly, diversity in news recommendations plays a major role in exposing the readers to counter-attitudinal view points [15]. Therefore, diversity in news recommendations is of utmost importance for a democratic society [11].

Consumption Behavior: Unlike other common digital activities like watching movies or shows, news items are frequently consumed anonymously and often without explicit user profiles [16, 17]. Prior interactions are the vital component in many recommender system algorithms, making it difficult to provide useful and tailored suggestions. This challenge may be partially solved by using session based recommender system algorithms and implicit feedbacks like clicks, reading time and skipped/ ignored items [18, 19]. It should be noted that these implicit feedbacks are often noisy and not always indicative of user preference. For instance, long read times does not always imply good engagement with the contents of the news article; the long-read time might also be due to user-fatigue or that the user has simply moved to other tasks while leaving the browser tab open rather than actually engaging the content [20].

Privacy Concern: Online news portals gather and investigate the users' reading patterns to provide better personalized recommendations. However, this itself has resulted in threats to the users' privacy [21].

Reading Context: The context in which a news is read is very important, and this is very dynamic component and evolves with time and social factors [15]. Among the various contextual components, time [22] and location [23] are the most widely utilized in news recommendation systems. Research shows that more people visit news portals on weekdays than weekends [24]. Beyond these major components, a reader's context can also be closely linked to a latest event or trending news topic, current weather conditions, or even the user's mood or specific interests. For instance, during major sporting events like the FIFA, or the Olympic Games, even users who are generally not interested in sports news might suddenly develop a strong desire to stay up to date on games and the outcomes of completed ones, demonstrating how a major external event can temporarily alter a user's reading context and preferences.

Impact of Social Media: The advent and extensive impact of social media have greatly changed the way people look for and collect news stories [25]. Readers no longer simply read news articles, they also closely monitor social media for updates and to follow its effects. The public dialogue, duration, and reactions of the news can help journalists identify issues needing further coverage

Emotions: Emotions captures attention of the readers and evokes feelings that connect

them to the news story. While music and movies naturally draw out emotions that impact their taste and choices, emotions are becoming more and more significant in driving news consumption behavior. This acts as both a challenge to quality of news content and an opportunity for news recommender system to innovate [26].

Biases: News is primarily made to inform the reader, however, as Helberger [11] has shown, through varying style and tone of the text, news articles are able to bring forth bias in the reader. Quality news should present the reader with all the relevant details of a news story devoid of any bias, so the end user can form their own judgment and emotional attachment with the news story.

Multimodal News Information: In our current digital age, the Internet is a vital piece in spreading news and information. Social media in particular, plays a key role in this process, as it has grown into becoming a major source of major news events. Unlike traditional online news portals that primarily present the news in mainly text format, social media often include videos, podcasts and other forms that help explain the news, and its context. Furthermore, important global news is often translated into other languages as well, thereby reaching a global audience.

1.2 Recommender System Algorithms

Recommender systems typically fall under one of three main types of algorithms: collaborative filtering, content-based filtering, and hybrid approaches [8]. Regardless of which algorithm is used, two key elements are vital for making any effective recommender system: a thorough and complete knowledge of both user and item content, and deep understanding of the fundamental mechanics behind the interactions. Content-based filtering works by first mapping items and user profiles to a latent attribute space based on the content/ features of the items, and the history of interactions of the user, and then a direct comparison of the user's and item's profile used to make the recommendation. In contrast, the collaborative filtering approach does not make use of features of the items in making recommendations. Collaborative filtering makes use of intricate patterns of user behaviors, such as historical ratings given to items, and interaction history to make suggestions to the user.

Although common traditional recommendation algorithms can surely be applied to the news domain, several factors may limit their effectiveness. This is due to a combination of unique challenges intrinsic to the news environment. We need to consider the high dynamism of news content, where news articles become outdated within a short time; the short-lived relevance of news items for individual users; and the significant dependence of users' interests on the context of news reading, which can alter rapidly based on time, location, or even mood of the reader. Although collaborative filtering has the potential to address the challenge of dynamic content generation in news items by analyzing patterns in the user interactions, this requires having a large number of user interactions, generally stored in the form of interaction histories, to make useful recommendations. And since news items have a short life span, by the time we have enough interactions on a news

article to make accurate recommendations, the news article usually becomes irrelevant. On the other hand, content-based filtering can better handle users' evolving interests by updating the user profiles with information derived from the latest news articles they have read [27]. These regular updates allows content based recommendation systems to have faster responses to changes in the user preferences. However, these algorithms face major challenge in providing useful recommendations to temporary and anonymous users which are quite common in news recommendation systems. To effectively deal with these intrinsic limitations and challenges present in both collaborative filtering and content based filtering systems in news recommendation, researchers and system engineers turn to hybrid systems to provide an optimum solution for their users.

In [28] the authors studied the prevalence of each of these recommendation algorithms for news recommendation, and found that for news recommendation content based filtering was the most common, followed by hybrid systems and lastly by collaborative filtering. One potential reason for popularity of content based filtering methods is that they only require content metadata to generate recommendations, making it easy to develop and use.

1.3 Popular models for building news recommender systems

Over the years numerous approaches have been used to make news recommender systems. Among these models latent factors models, and in particular those based on factorization methods, have proven to be the most common and effective. More recently, deep learning-based solutions have become the next class of models that are rising in effectiveness and popularity. We'll briefly explore several news recommendation algorithms below.

1. Factorization models: Factorization methods are a very important class of algorithms in recommender systems. These models work by decomposing the user-item interaction matrix into a product of lower dimensional user and item matrices.

2. Matrix factorization: Matrix Factorization is a widely recognized recommendation algorithm based on collaborative filtering. Matrix factorization is very effective at discovering and making use of the latent features in the interactions between different entities, such as users and items. Matrix factorization methods can incorporate context based features like time, within the recommendation algorithm. In a recent study on news recommender systems [29], matrix factorization was enhanced to incorporate news-specific data and to model the dynamic nature of reader behavior.

3. Non-negative matrix factorization: Similar to matrix factorization, non-negative matrix factorization is a decomposition technique where the user-item interaction matrix R is broken down into the product of two matrices, U and V, representing the user and item features in the latent space. The main difference between matrix factorization and non-negative matrix factorization is that in the latter all of the elements in all the three matrices need to be non-negative. In news recommender

systems, each user has only interacted with a very small fraction of all the news items available, this leads to the user-item interaction matrices to be very often sparse in this field. In such cases, non-negative matrix factorization models generally outperform standard matrix factorization due to their fundamental ability to handle missing values [30]. However, if the ratings matrix isn't very sparse, regular singular value decomposition-based matrix factorization might give better results.

4. Tensor factorization: Tensor factorization models extend the matrix factorization model by incorporating latent vectors with extra dimensions. Tensor factorization-based recommender systems overcome the challenges of matrix factorization by incorporating extra contextual information about users and items, leading to more accurate and useful recommendations [31]. Therefore, tensor factorization methods are valuable in news recommendation cases where more contextual recommendations is needed, such as those involving time, location, and social interactions. However, including too many dimensions can lead to many lengthy calculations without corresponding rise in recommendation quality. In [32] the authors made a news recommendation system using tensor factorization was to integrate contextual information about news items and readers into the recommendation model.

5. Probabilistic matrix factorization: In 2007 Mnih and Salakhutdinov [33] made a variation of the matrix factorization model, called probabilistic matrix factorization model, that incorporates Gaussian observation noise into the model. This model is inspired from Bayesian learning method for parameter estimation and scales linearly as we add more samples/ observations. Probabilistic matrix factorization has good performance on large, sparse, and imbalanced datasets. This type of dataset is common in the news domain.

6. Bayesian personalized ranking: Traditional item prediction methods, like matrix factorization, often fall short in ranking the news items. Bayesian personalized ranking addresses this limitation by using pairs of items to make optimized ranked lists personalized for each user. In [34] Rendle showed that matrix factorization models can be used along with Bayesian personalized ranking model to provide ranked lists of items personalized for the user.

7. Neural extensions: Current research in recommender systems is heavily focused on incorporating neural networks and deep learning into the aforementioned successful latent factor methods. For instance, Neural Network Matrix Factorization [35] uses a neural network in probabilistic matrix instead of an inner product, enabling the learning of suitable non-linear functions of user and item latent factors. Deep matrix factorization [36] extends the traditional matrix factorization model by using non-linear projections to encode users and items into a shared low-dimensional vector space, and Neural collaborative filtering [37] expands the extends the collaborative filtering model. These models continue to influence news recommender system

researchers, leading to the development of several effective news recommendation models.

Most modern news recommendation systems use deep neural networks like Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and Graph Neural Networks (GNNs) because these models are very effective at capturing and representing information in texts. CNNs are very good at extracting specific textual elements from news [38, 39, 40], such as fine-grained and knowledge-based features. RNNs are effective at understanding the range of user preferences by analyzing the user's interaction sequences [41, 42] , including both short-lived and longer lasting preferences. GNNs are commonly used to create structural representations that link news and users [43, 44], like intent-aware user profiles. While these neural networks improve the recommender system performance over traditional latent factor based models, many studies have shown that they struggle with more complex data [45, 46, 47], for example, multi-modal information and deep news-user relationships.

Another important factor that affects performance of news recommendation is the ability to identify fake news articles. Fake news articles are harmful for the public for they spread misinformation, and they also lead to bad user experience. Therefore, news recommendation systems should help users avoid fake news articles by filtering out fake news articles from recommended lists. There have been many studies into effective models for fake news classification [48, 49] but the studies do not consider the effect on performance of news recommendation systems.

Multilingual news recommendation is also very important, particularly in countries where people tend to read news in multiple languages, typically in English and the native language. [50] has shown that since different languages still represent the same content, patterns and relations in the embedding space for one language are often similar to other languages as well. This typically allows models trained for one language to be easily adapted for other languages as well.

News recommender systems are typically used for ranking candidate news articles based on user's history. However, in our study we are incorporating other tasks as well. The news recommendation tasks we include in this study are:

1. Binary classification of whether user will read a candidate news article given user.
2. Classifying all users on who would read a candidate news article.
3. Binary classification of candidate set of news articles given user.
4. Ranking of candidate set of news articles given user.

1.4 Large Language Models in News Recommendation

The rise of Transformers [51, 52, 53] in Natural Language Processing (NLP) has led researchers to adopt discriminative Large Language Models (LLMs), including BERT [52],

RoBERTa [54], and BART [55], as news encoders. The discriminative LLMs are used to extract hidden semantic information within the news content [46, 56, 58] or to model the semantic relationship between news and user by using the specialized training strategy for discriminative LLMs [59, 60, 61]. All these approaches fine-tune BERT-based model for use as news encoders. Discriminative LLM-based approaches generally perform significantly better than simpler deep neural networks, such as CNNs, RNNs, and GNNs. Nevertheless, news recommender systems based on discriminative LLMs are limited by their knowledge gained during pre-training and require a fine-tuning step before the model can be effectively used. This makes it difficult for them to effectively deal with the cold-start problem and accurately represent news and user interests [52, 54, 60].

Recent generative Large Language Models (LLMs) such as OpenAI o3, LLaMA 4, and Deepseek R1 have hundreds of billions of parameters (as opposed to the discriminative LLMs having only hundreds of millions of parameters) and has been trained on large and extensive training data, so they offer superior semantic understanding and generation capabilities. There's been a rapid increase in the number of news recommender systems based on generative LLMs, with some achieving state-of-the-art performance [62, 63]. This is mainly because generative LLMs can generate relevant information to solve the cold-start problem and use their powerful reasoning to extract accurate news features and model user interests. However, news recommender systems based on generative LLMs often need a lot of training time and resources.

There has been many recent advances in text embedding models, and now we have many general purpose text embedding models that can be used directly for news representation. Moreover, recent advancements have also led to the use of large decoder-only language models to obtain even higher quality text embeddings. There hasn't been any study that makes use of general purpose text embeddings or large decoder only embedding models for news recommendation tasks. In this paper we address this deficiency by studying the use of general purpose text embeddings or large decoder only embedding models for news recommendation.

1.5 Research Questions

The purpose of this study is to answer the following research questions.

1. What is the performance of BERT-based generalized embedding models in news recommendation tasks?

2. What is the performance of large decoder-only generalized embedding models in news recommendation tasks?

3. What is the performance of closed-source embedding models (OpenAI embedding models) in news recommendation?

4. What is the performance of the news recommendation models across the different news recommendation tasks investigated in this study?
5. What is the effect of incorporating fake news classification on performance of news recommender systems?
6. What is the performance of news recommender systems for multilingual news recommendation?

1.6 Objectives of the study

The objectives of the study are:

1. Assess the difference in performance between fine-tuned and generalized BERT-based embedding models on news recommendation.
2. Assess the difference in performance between fine-tuned BERT-based embedding models and large decoder-only generalized embedding models.
3. Assess the performance of closed-source embedding models (OpenAI embedding models) in news recommendation.
4. Assess the importance of different components of news article (title, category, subcategory etc.) on the quality of news embeddings.
5. Assess the impact of dense layers, attention layers on the final aggregation of news articles for user representation.
6. Assess the impact of fake news classification model on the performance of news recommendation models.
7. Assess the impact of training in the target language for multilingual news recommendation.
8. Assess the performance of the news recommendation models across the different news recommendation tasks.

1.7 Thesis Organization

In this chapter we have provided an overview of news recommendation systems, the challenges involved, and our approach. In Chapter 2 we review the background work done in news recommender systems, in particular the two major components: news encoder and user encoder. We also review LLM-based embedding models; those that have been applied in news recommender systems and more recent advancements. In Chapter 3 we identify the limitations of existing research and discuss the reseach gaps that wiil be addressed in this thesis. In Chapter 4 we discuss the methodology, datasets used, and model variations

in our experiments. In Chapter 5 we present the experimental results, and the insights and observations from the results. In Chapter 6 we provide the conclusion, summarizing research findings, discussing the potential for application in real world news recommender systems, comparing our work to prior state-of-the-art models. We also acknowledge the limitations of the research and provide directions for future research.

Chapter 2

Literature Review

2.1 News Recommendation

Personalized news recommendation is now a very common feature found in many online news platforms [64, 65]. Unlike non-personalized methods of news recommendation [66], which recommend articles based on basic factors like news popularity [67], editor's picks [68], or geographic relevance [23, 69], personalized methods of news recommendation provide recommendations tailored to individual users' preferences that is better able to satisfy the user's information needs.

Personalized recommender systems generally fall under three categories:

- Collaborative filtering [70] suggests items to a user by identifying other users with similar preferences. Items that are liked by similar users should be liked by the active user as well.
- Content-based methods [71] suggest items by analyzing the content and characteristics of the items themselves.
- Hybrid systems [72] combine both collaborative and content-based approaches.

While collaborative filtering is not commonly used in the news recommendation domain, it is one of the early methods to be used, with its early application appearing in the Grouplens project for newsgroups [2, 73]. Although these methods are not very popular, they are still used in news aggregators like Google News [74] where, the authors have integrated both collaborative filtering and content-based techniques into a hybrid system.

Content-based recommendation methods are more common in news recommendation systems. Early examples include NewsWeeder [75] which was mainly for newsgroups, with NewsDude [76] and YourNews [77] being some fairly recent examples of content-based recommendation system implementations.

Hybrid systems for news recommendation include approaches like Liu et al.'s work [78], which extends Google News study [74] by analyzing user's click behavior to construct accurate user profiles and representations. Likewise Li et al. [79] made a contextual bandit

algorithm, which suggests news articles based on context about the user and and the news article.

A common framework of personalized news recommendation model development consists of the following steps:

1. News Modeling: The first step is analyzing the content of the news articles to form a suitable news representation.

2. User Modeling: The users have a history of news articles they have read; this is used along with the news representation obtained from the previous step to form a suitable user representation. Other contextual features may also be incorporated into the user model.

3. Click Prediction: Finally, the user and news representations are used to rank the relevant top k news articles to the user.

News modeling is the first and most vital component of news recommendation, and it involves analyzing the content and features of the news articles and extracting a suitable representation that is required for the subsequent steps. The next step user modeling is also very important, and it is essential for understanding individual user's news preferences from user profile like interaction history. Finally, after having the news and user representations from the news and user models respectively, the last step is to rank the relevant news articles using dot product, cosine similarity or even neural networks.

2.1.1 News Modeling

The purpose of news modeling is to analyze the characteristics and content of the news to create a suitable representation of the news. Accurate news modeling is a vital component of personalized news recommendation systems in extracting suitable news representation by analyzing news content. News content modeling techniques generally fall into two main types: feature-based and deep learning-based. Feature-based methods mainly make use of manually crafted features to represent news articles, whereas deep learning methods are formulated to learn suitable representations directly from raw news data.

Feature-based News Modeling

The main challenge in feature-based news modeling techniques is in crafting effective features to represent the news articles. There are four main types of features that are commonly used in feature based news modeling: Collaborative filtering features, Context features, Property features, and Content features.

Many collaborative filtering methods represent the news articles using collaborative filtering features like news IDs [2, 74]. However, since news websites keep publishing news articles rapidly and older ones become obsolete, ID-based representations face major cold-start problems, leading to relatively poor performance. Therefore, representing news articles with just their IDs is not a suitable approach [80].

Given the limitations of ID-based methods as described above, most methods now include content features to represent the news articles. For example, Gershman et al [81]. extracted Term Frequency-Inverse Document Frequency (TF-IDF) features from the news text for use in news modeling. Important entities (like known places, people etc.) and concepts are often more crucial for representing the content of news article than other individual words. Therefore, there has been a number of methods that utilize these for representing the content of the news article. For instance, Goossen et al. [82] proposed using Concept Frequency-Inverse Document Frequency (CF-IDF), which is a TF-IDF variant using concept frequencies from WordNet instead of term frequency, to represent news texts. Capelle et al. [83] introduced Synset Frequency-Inverse Document Frequency (SF-IDF), which modifies TF-IDF by using WordNet synonym sets instead of term frequencies. Moerland et al. [84] extended the SF-IDF model to to SF-IDF+ by incorporating relationships between the concepts. They expanded the synonym sets of concepts in the news articles by incorporating other concepts within WordNet that are related with the included concepts.

Beyond semantic features, some researches investigated extracting useful content features to improve the news content modeling [85, 86, 87]. For instance, Garcin et al. [88] joined the news title, summary and main content together and used Latent Dirichlet Allocation (LDA) on it to extract relevant topics from the news article. Parizi et al. [87] extracted emotion features from the news sentences used the Ekman model, which represents emotion as six emotion categories, for use as complimentary features to use alongside TF-IDF. Parizi et al. [89] also made a variant of this method that makes use of sentiment orientation (positive, neutral, or negative). Furthermore, studies have investigated utilizing image/video-related information, such as news videos, to act as additional information providing extra insights into the news content.

Apart from content features, other feature genres are also used for news modeling. They can be divided into two main types: property features and context features. Property features, like categories, publishers and locations often reflect inherent characteristics of the news. News category is the most commonly used property feature because it is very helpful in understanding the news content and preferences of the user. Liu et al. [78] for example, used the news topic categories to represent the news. News categories are usually not produced with the news article itself, rather they need to be manually provided by human editors, therefore the news categories might not be explicitly available in many news recommendation scenarios. Therefore some methods cluster the news into categories based on their content, such as in the SCENE [86] recommender system, which uses hierarchically clustering on the news topic features that have been extracted using LDA. Integration of news categories or clusters within the news model helps recommenders be aware of the news topics and categories and provide better recommendations. News location is another important and popular property feature for filtering news to only the locations the user is interested in. For instance, Tavakolifard et al. [90] integrated a location filter in the news recommendation system that uses geographic information of the news to filter news

articles based on the user's locations. Additionally, same news by different publishers can have differences in content and topics, and so a number of news recommendation methods use the news publisher information as an extra feature to improve the news model [79, 91].

In contrast to property features which are normally static once the news have been published, context features are dynamic in nature. Popularity, which represent the overall attractiveness of news, and recency, which indicates the freshness of the news, are two important context features used in some news recommendation methods. For example, MONERS [85] was a news recommender system that represented the news articles by their categories, estimated importance of the news given by the news provider, and recency of the news story. Gershman et al. [81] used popularity of the news article, age (recency) of the news, TF-IDF of words in the news texts, and the named entities to represent the news items for news recommender system. Jonnalagedda et al. [92] made use of timelines on Twitter for news modeling. They obtained the categories of the news articles as well as their popularity from Twitter and used them in the news representation. While recency feature only considers the time between publishing of the news article and displaying it to the user, the timestamp of when the news was displayed to user can provide finer-grained information like the season, month, day of week, and even the time of day when news was displayed. Therefore, several approaches incorporate impression timestamps within the news model [19, 93], with Ilievski et al. [19] including the weekday and hour when news displayed to user. Beyond these, a number of methods also investigate using features like weather [94], the click-through rate (CTR) [93], and fact/opinion bias [95] within the news content to improve the news representations.

A number of hybrid methods combine news IDs with some additional features for modeling the news articles [96]. NewsWeeder [75], for instance, uses news IDs and bag of word based features. Claypool et al. [97] used the IDs of the news articles and keywords of the news for news modeling. Liu et al. [78] represented the news articles with news IDs and categories of the topics. Saranya et al.[98] used news IDs, news topics, frequency of clicks, and the relative importance of a news belonging to different news categories. Combining content-based and ID-based techniques helps reduce the news cold start problem to some degree and has been extensively studied by incorporating other information such as property features of the news [99, 100], ontologies [101], news sessions [16], and knowledge graphs [102].

Deep learning-based News Modeling

Deep learning has revolutionized the news modeling landscape, with many recent papers employing neural networks to have the model learn useful news representations themselves. The idea is to move away from traditional manually generated features like TF-IDF towards using neural network based NLP models to automatically learn useful news representations. For instance, Okura et al. [64] used variation of the denoising autoencoder to learn useful representations for the news in the form of embeddings which was used in embedding based news recommendation EBNR. RA-DSSM [103] is a neural network based news recommendation method that used a similar architecture to the DSSM [104]. First it

creates the news representation using doc2vec tool [105], and then makes use of a two layer feed forward network for model the news representation. Kumar et al. [106] used word2vec embeddings and then used 3D CNN for news modeling. [107] later used same approach. However, these early neural network based NLP models often unable to fully extract the complex semantic information within news texts.

Many of the methods tried afterwards integrated more modern neural network based NLP models like CNNs [39, 108] and self-attention [109]. For example, WE3CN [110] used 2D CNNs for learning useful representations of the news. NPA [111] used CNNs on the news titles for creating context based word representations, and also using a personalized attention network to select important words based on the user's preferences. NRMS [109], and some similar works later on [112, 113, 114, 115, 116], used a multi-head self-attention network to learn the word representations and then used an additive attention network to represent the news. NRNF [117] employed self-attention to model the contexts of the words in both the news titles and bodies, alongside an interactive attention network for modeling the connection between news title and body. Mao et al. [44] used a similar co-attention mechanism for modeling the interaction between news titles and and the abstracts. FedRec [118] combined multi-head self-attention networks and CNNs for learning useful news representations from the news titles. Despite the advancements, many of these methods made use of non-contextualized word embeddings like GloVe [119] and shallow text models, potentially limiting their ability to extract complex semantic representation from the news. WG4Rec [120] introduced a word-graph based method to model the news text. It made a word graph based on their co-occurrence, semantic similarity, and news co-click data, and used a GNN to learn the word embeddings. While these neural network based methods significantly improved text understanding, the models remained relatively shallow and so they were unable to extract deep semantic representations from the news text.

Apart from modeling the semantic information from the news articles, some methods investigate on using keywords or entities in the news texts to supplement the news modeling with additional information complementary knowledge. A direct approach is to treat the entities in the news as text and incorporate them with news text modeling [121]. For example, Gao et al. [122] developed an approach incorporating knowledge graph in the news recommendation method and using hierarchical attention networks. The method involved learning a word based representation of the news by using a word attention network that used the embeddings of the keywords in the news text as the attention queries. These representations of the news are combined with the average of the news entities within their context, and the embeddings of the news entities separately. Finally these three types of representations are combined with the help of a item attention network based on their effectiveness in modeling the news. DAN [123] used two CNN networks in parallel and max pooling to learn useful news representations from the news titles and entities. Saskr [124] used the average of the word embeddings of the entities to learn news representations from the titles and bodies of news articles. DNA [125] used the body of the news, the news ID, keywords and entities in the news text to learn useful news representations.

It used doc2vec [105] to convert the sentences in the news body into their embeddings, which are then combined into a single embedding to represent the news article by using a candidate-aware sentence-level attention network. Element representations were formed by averaging word embeddings and synthesized by an element-level attention network. Finally, the embeddings of the text, ID, and elements of all the pieces of the news are combined to a single news representation. HieRec [126] used text self-attention for modeling the context in news titles and entity self-attention for modeling the relations in the entities in the news texts. While these methods are effective in unifying knowledge entities and texts, they are still unable to suitably model the relation between entities, they often fell short in fully exploiting the relatedness between entities.

Knowledge graph embeddings are another way of incorporating entity information in news recommendation [40, 127, 128, 129, 130]. For instance, DKN [40] used a knowledge-aware CNN to learn useful news representations from the news titles and entities present in the news titles. A knowledge graph using the TransD [91] knowledge graph embedding method are used to learn the representations of the embeddings themselves. Liu et al. [127] used Microsoft Satori knowledge graph as the basis to construct a new news-relevant knowledge graph. The knowledge graph was made by first extracting topic entities and knowledge entities from the news and then connecting entities clicked by same user, entities appearing in the same news article, and entities appearing in the same browsing session so that in the end the knowledge graph will have improved relations between the entities. They joined the entity embeddings learned using TransE [131] method with the news text embeddings learned using DSSM and LDA. CAGE [128] supplements the text embeddings learned using CNN with TransE embeddings of entities and uses one hop neighbors of entities to make subgraphs of the knowledge graph. However, mainly the low level relations between the entities are summarized in the knowledge graphs. To model deeper relationships between the entities, TEKGR [132] supplemented the knowledge graphs using topical relations between the entities. It uses the news text and its contents to predict the topic of the news and then improves the knowledge graph using the predicted topic and finally uses graph neural networks to learn knowledge representations of the news enhanced by topic information. KRED [133] first uses graph attention networks to learn entity embeddings from the knowledge graph and then integrates it with other entity features like position, category and frequency and finally chooses entities as per the texts representation of the news. KIM [134] integrates an interactive news modeling approach that incorporates knowledge graph. The approach uses graph co-attention network to help model the relations between the entities in the news and the entities in candidate news and the entities of neighbors of clicked news. KOPRA [135] used the TransE embeddings of news knowledge entities to represent the news articles. In this method hidden entity representations are learnt using a recurrent graph convolution network. These methods improve news recommendation by allowing the models to encode richer knowledge based information than models that only use text based information.

A number of methods explore ways of incorporating other types of information in the

news beyond just the news texts for news modeling in order to have better representations of the news articles. Among the other types of information, many methods have used tags and topic categories [14, 38, 136, 137, 138]. DeepJoNN [138] uses a character level CNN to learn news representations news categories, entities, keywords and news IDs. Park et al. [14] trained a doc2vec model [105] using proprietary corpus model for transforming the news articles into vector representations and then these news representations were used to create user representations using an LSTM network. Additionally, the news categories predicted by a CNN model [139] is also used in the news representation. TANR [38] used both an attention network and a CNN on the news titles to learn representations of the news. This was also used in [140, 141]. TANR also learns topic based news representations by including an additional task of predicting the news topics. NAML [137] used attentive multi-view learning for news recommendation, where different views of the news are the different kinds of information in the news. This method incorporates news categories, subcategories, bodies and news titles as different views, and then in the end a view level attention network is used to aggregate the embeddings obtained from the views into a single embedding. [142, 143] also used similar approach to model candidate news. LSTUR [42] used both CNN and attention network and CNN for transforming the news titles and uses non-linear transformations to category and subcategory embeddings for making embeddings and integrates them into the model. CHAMELEON [144, 145] uses CNN with different sizes of the kernel on the news body to learn representations of the news. These text based representations are combined with features from the news metadata such as categories, tags and topics by using a fully connected layer. This method also includes an auxiliary task of predicting news metadata features.

Several approaches make use of other content based news information apart from the just the topical features. For instance, SentiRec [112] learns sentiment based news representations by taking into account the sentiment orientation of the news articles. In this method the sentiment scores of the news were calculated using the VADER [146] algorithm. MM-Rec [147] used a model that incorporated both vision and language based information, ViLBERT [148], to learn multi-modal representations of the news from news texts and images. IMRec [149] modeled visual impression information of the news article including image regions, text regions, spatial position of different words and arrangement of different fields for use in news recommendation. These methods characterize the news content in multiple aspects, and thereby learns more accurate and useful news representations.

Another important category of extra information for use in news recommendation is context based features, like position within the recommended list and popularity. For instance, PP-Rec [150] incorporated features extracted from the news titles and entities, along with the popularity information of the news for modeling the news. This generates news popularity, recency and CTR from the news titles using a gating mechanism for use in a combined popularity score for the news. TSHGNN [151] integrated the active time spent by user on the news articles into news modeling. DCAN [152] used CTR in close to real time and the time since publication of the news predict the position of the news articles in

the recommended list. CTX [153] studied use of context based like freshness, popularity and CTR of the news on click prediction and found that these can have an stronger impact than signals based on personalized interest. DebiasedRec [141] used attention networks and CNN on the news titles to learn representations of the news content. It also learns a bias model to extract news bias representations based on the positions and size of the news displayed to the user on the websites. By using the additional context information these methods are usually better able to understand the interaction patterns of the user. generally enhance understanding of user interaction patterns, However, some of the news context features are unlikely to be available in many practical news recommendation scenarios, making it difficult to exploit these features.

A few methods makes use of graphs to learn news representations. For instance, IGNN [154] used KCNN [40] for learning text-based representations of the news from news titles, and for learning graph-based representations of the news from the news-user graph. GERL [155] combined an additive attention and multi-head self-attention networks for learning representations of the news title, then combined these news title representations with the news category embeddings for news modeling. MVL [156] makes use of two views for news recommendation. It uses the news title, category and body for the content view and and a graph view to improve the representation using the neighbors of the news article on the news-user graph. Additionally, the method also uses information from first and second order neighbors in the news user graph using a graph attention network for better news representation. GNUD [157] learns representations of the news text by using news encoders similar to those used in DAN, and learns disentangled representations of the news from the news-user graph by using a graph convolution network with a regularization term for the preference disentanglement. Apart from the news-user graph used the methods discussed above, some methods also use heterogeneous graphs to incorporate rich collaborative information [43, 158]. For instance, GNewsRec [43] is a hybrid method that makes use of heterogeneous graph information by incorporating graph information of news, users and categories of news topics. It uses a graph neural network with two layers on the heterogeneous news-user-topic graph for learning the graph based news representation and uses same architecture as DAN for learning representations on the news text. These methods can effectively use high-order graph information for improved news modeling, However, these models face challenges in cold-start scenarios because the new news article generated doesn't have many connections to other users or news articles in the news-user graph, particularly the ones used in training.

LLMs like BERT was very successful in the field of NLP, and as such, many recent works investigate improving news recommendation by fine-tuning pre-trained LLMs to enhance news modeling [46, 59, 60, 61, 159]. For instance, PLM-NR [46] improved English and multilingual recommendation of news articles in offline results and also significantly improved the recommendations in Microsoft News in online results by using a number of pre-trained language models for news modeling. UNBERT [61] joined the texts of news articles and used them as input to the BERT model for news modeling. MANNeR [161]

uses a regular BERT-based module and a separate aspect specific model to incorporate news aspects like category, sentiment etc. into the news model. Some studies have also incorporated generative LLMs in news modeling, primarily for making additional information for news representation. For example, Yada et al. [162] used generative LLMs like ChatGPT for generating category descriptions that are for news modeling. Chen et al. [62] used generative LLMs to make representations of the news with rich semantic information and then used Wikidata knowledge graph for handling long-tail distribution. These studies show that LLMs are very effective in understanding the news articles and making useful representations of the news articles for news recommendation.

2.1.2 User Modeling

User modeling is a very important step in for personalized news recommendation for understanding each user's personal interests in the news. This is usually done by user modeling methods applied on the user's behavior on the news websites and articles [38]. User modeling is done primarily based on the news articles that the user has interacted with, and additionally incorporating context features like age and location of user to improve the user model. Similar to news modeling as was discussed above, we can also classify user modeling methods into feature based user modeling and deep learning based user modeling. As before, feature based user modeling methods use manually crafted or heuristic based features for modeling the user interest. Deep learning based methods rely on techniques to automatically find useful features or patterns from the user interactions to model the user's interest. These two approaches to user modeling is described in detail below.

Feature-based User Modeling

Feature-based user modeling uses manually generated features to represent users. For collaborative filtering based methods, users are represented by their IDs [2, 74], similar to news modeling. However, methods based on the user ID often face challenges due to data sparsity, therefore, most approaches of user modeling make use of user interactions instead of the user ID. A common approach is to use the features of the news articles user has interacted with to build the user representation. For instance, Goossen et al. [82] represented the user by using the Concept Frequency-Inverse Document Frequency (CF-IDF) based features of the news articles user has interacted with to form the user representation. Capelle et al. [83] used the Synset Frequency-Inverse Document Frequency (SF-IDF) based features of the news articles user has interacted with for user modeling. Garcin et al. [88] proposed to model users by combining the LDA-based features of the news articles user has interacted with to model user interests. However, these methods tend to under-perform in modeling user interests in cases where the news interactions are sparse.

Apart from news features, other supplementary user information is often incorporated in user modeling. For example, the MONERS [85] news recommender system groups the users into segments, and represents the users using the preferences of the user segments on

the categories of the news and the news articles. Demographic features of the users like age, gender, and occupation are also valuable for user modeling, as different demographic groups often have different news interests. Therefore, features based on demographic information of users are integrated into a number of news recommendation methods [19, 85, 94]. For example, Yeung et al. [94] used gender, age, occupation status and also the socio-economic grade of the user to predict the categories of news the user is likely to be interested in. Methods incorporate these features to identify varied preferences across categories. Chu et al. [93] used modeled the characteristics of the user using the user's gender and age categories. Location information of the user is also very important for accurately modeling the user preferences and has been used in many location-aware news recommenders [163], which also contributes to accurate user modeling in location-aware systems. However, some of these demographic features have privacy concerns, and so users may be reluctant to provide accurate demographic information for news recommendation.

Since an user's click on a news article does not always mean that the user is interested in that particular news, some approaches also make use of other forms of user feedback and behavior for news modeling. For instance, Gershman et al. [81] developed a system that modeled user preferences based on the news they read thoroughly, which are taken as positive samples of user interest, as well as the stories they dismiss or scroll past, both of which are taken as negative samples of user interest. The amount of time a user spends on a clicked article is another very important indicator of the user interest, which Yi et al. [164] utilized by weighting clicked news articles based on the time users spent on them for modeling user's interests. Beyond these actions, other behavioral patterns, such as how and when users access news articles, have been used by a number of systems to understand the user's news reading behavior and making useful user models [86, 98].

A number of methods also use graph-based information, such as news-user interaction graphs, for user modeling [165]. For instance, Li et al. [166] made a news personalization technique that uses hypergraphs to capture complex, high-order relationships between various pieces of news information, and represents the users as subgraphs within the hypergraph. Garcin et al. [167] proposed using context trees, which are made from sequences of articles, and topics, and topic distributions of the news articles read by user, to model users' interests. Another approach by Trevisiol et al. [18] involved making a browsing graph from the users' news reading histories on Yahoo News. Joseph et al. [168] represented the users by taking their clicked news articles as subgraphs of a larger knowledge graph, which is made by linking the entities. These graph-based methods are useful for taking into account the higher-level connections to better understand the interests of users, and hence enhance user modeling.

Some approaches integrate user IDs with other types of features to create better, more useful user representations [96]. NewsWeeder [75], for example, used a combination of bag-of-words based features of the news articles clicked by user and the user IDs for user representation. Claypool et al. [97] took a similar approach, integrating both keywords based features of the news articles clicked by user with the user IDs for user modeling. Liu

et al. [78] represented users through a combination of user IDs and interest features of users predicted using a Bayesian model. These methods help overcome the limitations of using only on ID-based modeling while still incorporating the valuable personal information that user IDs can represent.

Given that interests of users can evolve over time, some approaches are designed to model both long-term and short-term preferences [101, 169]. An early example of this is NewsDude [170], which employed a hybrid model. This system captured short-term interests from news articles that have recently viewed been read by the user and modeled long-term interests by identifying the most important words within each news category based on TF-IDF scores. More recently, Li et al. [171] made LOGO, a news recommendation technique that also models both short term and long term interests of the user. LOGO represents long-term interest of the user as a weighted sum of the topic distributions of the news articles clicked by a user and short-term interest as the topic distribution of the most recent news article clicked by user. Viana et al. [172] proposed another method for incorporating short term and long term user interests in the user model. In the method long-term interest is determined by the frequency with which a user engages with news articles with a specific tag, while short-term interest is defined by a number of recently clicked news articles. By capturing both long-term and short-term interests, these methods are able to make a more dynamic and accurate representation of user interests unlike methods that only make use of either short term or long term user interests.

Deep Learning-based User Modeling

In recent years, deep learning techniques have become very common in personalized news recommendation systems for user modeling for being able to replace the manual feature engineering approaches with automatic methods. The majority of current methods deduce the interests of the user by analyzing their history of clicked news. Several approaches aim to combine the representations of previously clicked articles for the user model [136]. For instance, Khattar et al. [103] proposed weighting the sum of clicked news representations with an exponential decay function, which gives greater importance to more recent clicks. The methods NAML [137] and KRED [133] employ a news-level attention network to learn user representations from news articles clicked by user. AMM [59] uses a similar attention mechanism to consolidate the news features, but it uses both the historical news articles as well as the candidate news article for modeling user preferences. DKN [40] used a candidate-aware attention network to learn user representations, calculating the importance of each clicked news article in user's reading history based on its relevance to the candidate news article. TEKGR [132] also made use of candidate aware attention over the news articles in user's history for user modeling. Liu et al. [127] made the user embeddings by using a straightforward time-decayed average of the embeddings of news articles clicked by user. MM-Rec [147] used a multi-modal candidate-aware attention network to select news articles from the user's click history for user modeling based on their multi-modal relationship with candidate news article. HieRec [126] made the user representation using a hierarchical approach, first creating a user interest model at the

subtopic level by combining news features within the same subtopic, then merging these user interest representations from the subtopic level into broader topic-level representations, and finally combining the user interest representations from topic level into an overall representation. DebiasedRec [141] integrates the effect of presentation bias information on the user's click behavior using an attention model that combines the news representations from user's click history into a final debiased user representation. While these techniques are useful at capturing important user behaviors for modeling user preferences, they fall short in modeling the relationships between different news articles in user's click history, which could provide important insights into user interests.

To address this limitation, many methods incorporate the context based features available for news click behaviors. RNNs are a widely used tool for modeling the sequential relation between items, like the history of news clicks of user [64, 123, 160, 173]. For example, EBNR [64] uses a GRU network to learn user representations from the news representations of their browsed news articles. RA-DSSM [173] processes the news click sequence from user's history using a bi-directional LSTM network and then applies a attention network at the news-level to create a user representation. DAN [123] generates representations of the user interest from the history of clicked news articles incorporating both attentive LSTM which creates sequential embeddings for the user's history and a candidate aware attention which creates user interest embeddings. The primary strength of using user models based on RNNs is that they are very effective at capturing the dynamic nature of user interests. However, they tend to be less effective at capturing a user's overall interests. Furthermore, as noted by some studies [174], treating news recommendation as a standard sequential problem may not be appropriate, given that users often prefer diversity between the news they have read in the past and what they will click on next.

A variety of other common deep learning architectures, including CNNs [108], self-attention networks [109], and co-attention networks [152], have been used for user modeling. For instance, WE3CN [110] uses a 3D CNN model to learn user representations from 3D tensor representations of their clicked news items. To improve efficiency, SFI [152] later introduced a hard selection module to make the 3D CNN approach less computationally intensive. The NRMS model [109] uses a multi-headed self attention network that works at the news level to create contextual news representations and then forms the final user representation with an additive attention network. This particular technique has been used by many other systems, such as IMRec [149], FairRec [116], and SentiRec [112], with one variation using the representation of "CLS" token of Transformers [61]. FIM [39] introduces a technique for fine-grained modeling of user interests that uses a 3D CNN model to capture word-level relations between different news articles read by user. UniRec [115] first learns a user embedding for news ranking with the NRMS [109] model and then uses this embedding as a query in an attention network to select and combine a set of basis interest embeddings for news recall. KIM [134] employs a co-encoder for the user news relation to model interactions between candidate news article and news article in user's click history, allowing it to collaboratively learn both a news representation that

takes the user into account and a user representation that takes the candidate news into account. PP-Rec [150] implements a popularity-aware method for user modeling, first using self-attention to model the context of user news clicks and then a joint content-popularity attention network to select clicked news based on both relevance and popularity to model the user's interests. RMBERT [159] captures the complex interactions between user news interaction behaviors and candidate news articles using a reasoning memory network [175] to model user interests. Li et al. made MINER [58] in which the user was modeled using multiple embeddings for the same user to capture different views of the user interest. These advanced methods are effective at capturing the relations between different user news click behaviors to create more useful user models.

More recently, graph neural networks have been used to model the context of user news click behaviors by extracting their high-order relationships [44]. For instance, CAGE [128] first refines representations of news click behavior by using a Graph Convolutional Network to model relations between different news click behaviors within a news session, and then builds representation of the user using GRU network to construct the user representation. User-as-Graph [142] was one of the early works in news recommendation that represents each user using a personalized heterogeneous graph built from the user's news click history, framing the task of modeling user interest as a graph pooling problem. The study solves the problem by using HG-Pool, which is a heterogeneous graph pooling method, to iteratively aggregate the personalized heterogeneous graph for extracting representations of the user interest. EEG [143] models uses an entity graph to model each user. First it uses a graph neural network to model hidden entity representations and then combines them into a entity based representation of the user using an attention network. which are then aggregated by an attention network. KOPRA [135] also represents users with entity graphs but processes them with recurrent graph convolution networks, modeling long-term interest with the entire graph and short-term interest with subgraphs deduced from the from recently clicked news articles. The approach also includes an additional step of entity neighbor pruning for selecting entity neighbors that align with user interests. CNE-SUE [44] applies a graph convolutional network to different subgraphs of the user's news click behavior graph for modeling different representations of the user interest for various clusters of user behavior. The method also uses an attention network to nodes within a cluster for pooling node representations, and then an attention mechanism within different clusters to combine cluster representations. These graph neural network-based approaches are very effective at extracting latent user interests by capturing the rich, high-order relations within a user's click history.

Apart from the user's news click behaviors, some approaches also make use of the user's ID information [125, 138]. For instance, the NPA model [111], uses the embeddings of user IDs to make attention queries for a personalized, news-level attention network that selects important news articles from user's history based on user's news reading characteristics. LSTUR [42] models long-term interests of the user using embeddings learned for the user ID and models short-term interests of the user with a GRU network on the news

representations. In the paper, the authors explored two approaches of combining the long term and short term user representations: either concatenating the long term and short term representation vectors or using the long-term representation vector to initialize the hidden state of the GRU module used for learning short term representation. This approach was adopted by some other models as well like CUPMAR [130] and KG-LSTUP [129]. While these methods can provide better personalized news recommendations for active users who have sufficient news click history to train their user ID embeddings, they still struggle with the cold-start problem, facing difficulties in modeling new users who lack a well-tuned user embedding.

The previously mentioned techniques all primarily model user interests using news click data of the user. However, relying only on news click data is problematic because news click behavior is often noisy and doesn't reliably indicate genuine user interest. It is also difficult to build a complete and effective user model from news click data alone. Therefore, some techniques have explored integrating other types of user information to improve user interest modeling [176]. One important approach is incorporating contextual features to enrich the user models. For instance, the CHAMELEON [144, 145] system incorporates a number of user context features such as the location, time of day, device type, and referring source. It uses an Update Gate RNN network to learn representations of the user within a single session, with the final recommendation scores determined by the cosine similarity between the representations of the user and and candidate news article. The context features provides useful information about user's preference and by incorporating them in the user model the resulting user representations better capture user interests and thus lead to improved news recommendation.

Another major strategy in user modeling is to integrate different types of user behaviors [158]. For instance, NRHUB [140], for example, considers heterogeneous actions including news clicks, browsed webpages, and search queries for user modeling. It treats these different behaviors as different representations of the user, learning a separate embedding for each behavior. This process uses both a CNN and an attention network to learn representations for each user behavior and a behavior-specific attention network to select important actions for the user embedding. These behavior-specific embeddings are then combined into a single user interest embedding by a final attention network. WG4Rec [120] also studied the effectiveness of using the user's webpage browsing behavior for modeling user interest. CPRS [113], makes use of the user's news click and news reading behaviors to model the user interests. It uses the titles of news articles clicked by user to model the user's news click behavior, and uses the body of clicked news along with a reading speed metric, which is estimated from the time taken by user to read a news article and the article length, to model the user's reading satisfaction. NRNF [117] separates clicks on news articles into positive and negative samples for a particular user based on a threshold on the time spent by the user on the news article. The model then learns positive and negative representations of the user interest using separate transformer and attention networks. FeedRec [114] models user interest using a diverse set of feedback signals such

as clicks, non-clicks, quick closes, completed reads, shares, and dislikes. It uses different homogeneous transformer modules to model the relations within the same type of feedback and a heterogeneous transformer module to capture the relations between the different types of feedback. Additionally it also has a strong-to-weak attention module that utilizes the strong feedback representations to combine positive and negative information about the interest of the user from weak feedback. By extracting complementary information about the user interest from these multiple behavior types, these methods can typically create more effective representations of of user interests.

Additionally, there are several methods that create user representations on graphs that capture the collaborative information between users and news articles. For instance, IGNN [154] learns a graph-based representation of the user interest using a graph neural network and a content-based representation by taking the average of the embeddings of the news articles clicked by user. The graph-based and content-based representations of the user interest are then concatenated to from the final user interest representation. GNewsRec [43] learns short term representations of the user interest using a component similar to the DAN model [123], and learns long term interests using a two layer graph neural network on a heterogeneous graph of the user-news-topic relations. Finally the short term and long term user representations of the user interest are concatenated to form the final user interest representation. GERL [155] creates a content-based representation of the user interest from click history using additive attention and multi-head self-attention networks. The method also creates a graph-based representation of the user interest by extracting high order information from the user news graph by using graph attention network, and finally combines the content based and graph based representations to form the final user representation. Similarly, MVL [156] uses a multi-view approach, learning representations of user interest in the content view using attention networks and learning a representation of user interest in the graph view by using graph attention network on the user-news graph. GNUD [157] learns representations of the user interest from the user-news graph by using a disentangled graph convolution network. The methods can improve user modeling by making use of the high order information present in graphs. GBAN [158] learns user embeddings using an LSTM network and combines it with heterogeneous graph embeddings. In order to improve user representation in the graph embedding, the method also includes subgraph core and coritivity scores for estimating the importance of a target user-news pair in the subgraph. These approaches can effectively utilize the high-order relations between users and news articles and also incorporate meta features for improved user modeling. However, these models suffer from cold start problems as they cannot effectively represent new users who are not present in the training data.

2.1.3 Fake News Classification

Many methods used TF-IDF based features for fake news classification. Aphiwongsophon and Chongstitvatana [177] extracted TF-IDF based features from Twitter based fake news

dataset and used Naïve bayes, SVM and neural network models for fake news classification. Singh et al. [178] used TF-IDF and word embedding based features and used random forest, XGBoost, and AdaForest for fake news classification. Ni et al. [179] used document frequency based features with logistic regression, and SVM and random forest for the classification. Albahr and Albahr [180] used unigram, bigram, and trigram based features and decision trees, random forests, decision trees and neural networks for classification. Mouratidis et al. [181] showed that BERT based embedding models significantly improve accuracy of classification compared to other methods.

2.1.4 Multilingual News Recommendation

Most news recommendation studies focus on monolingual news recommendation. For example, MIND [65] and EB-NeRD [182] are two very popular datasets for news recommendation, but they are both monolingual, in English and Danish respectively. These datasets are very helpful for the progress of news recommendation models, however, these datasets leave behind news in low resource languages and polyglot users who read news in multiple languages. To overcome Iana et. al. [50] released the xMIND dataset for help in multilingual news recommendation. The paper also shows the performance of cross-lingual transfer in zero-shot and few-shot evaluation in different target languages.

2.2 Embedding Models

Text embedding models are key component of modern news recommender systems. Modern transformer based embedding models can make effective representations of the news articles which leads to improved performance of news recommender systems. Moreover, since user representations are often made from news representations of clicked news articles, better news representations lead to improvements in user representations as well. Since the introduction of BERT model [52] for text embedding in 2018, models following a similar architecture have been prevalent in text embedding tasks. However, recently it was shown that with some modifications and fine-tuning, decoder only models like LLAMA [183], Mistral [184] and Qwen [185] can be used to make effective text embeddings. Both BERT-based embedding models and text embedding models made from decoder only LLMs are described below.

BERT-based models

BERT [52] models use a Bidirectional Encoder architecture which lets each token attend to context from both its left and right sides. The embeddings for each token can be used for any token level tasks and the embedding of a special [CLS] token attached to the start of the sequence can be used for any sequence level task. For news representation we take the embedding of the [CLS] token as the news embedding. The BERT model was pre-trained on the BookCorpus English Wikipedia datasets using Masked Language Modeling (MLM) and Next Sentence Prediction (NSP) tasks. There are two sizes of BERT

models available: BERTbase with 110M parameters and BERTLarge with 340M parameters. Several other embedding models have been made following the similar architecture. For example, RoBERTa [54] updated the pre-training process of BERT, replacing the static masking in the MLM task with dynamic masking, and removing the NSP task altogether. The model was also trained using larger batch size and on more data. XLM-RoBERTa [186] is a similar model trained on 100 languages so that the model has multilingual capabilities. A major limitation of BERT and these variations is that these models is that for sequence comparison tasks, both sequences need to be fed to the model, i.e. the embeddings from these models cannot be directly used in embedding-specific tasks like retrieval or clustering.

In recent years, methods were developed that overcome this limitation. SBERT [187] was one of the first models to address this challenge. They used Siamese and triplet network structures to teach the BERT model to generate meaningful embeddings that can directly be used for pairwise tasks like retrieval or clustering. SimCSE [188] made a simpler training method using unsupervised and supervised contrastive learning. E5 [189] models are pre-trained on a large text pairs dataset using contrastive learning and further fine-tuned on MS-MARCO [190] and Natural Questions (NQ) [191] datasets. Multilingual E5 [192] extended the E5 models with multilingual capabilities by training on data from 93 languages. BGE-M3 [193] models also pre-trained on a large number of multilingual data, and used task-specific instructions for better training in the fine-tuning stage. The models described above extend the capabilities of the BERT models, allowing them to achieve good performance on embedding related tasks like news recommendation.

Decoder-based models

Large decoder-based language models are trained on large quantities of web and high quality data, and achieve state-of-the-art performance on most NLP tasks. However, for many years it was believed that these models would perform worse than bidirectional models for general embedding tasks [194]. The main reasons were that the causal attention mask limits the model's ability to learn useful representations, and that large decoder only models often lead to very high dimensional embeddings, which may face limitations due to the curse of dimensionality [195]. However, soon methods were developed to make use of the large decoder only models into embedding models. LLM2Vec [196] outlines a method that can be used to convert any decoder-only model into embedding model. The first step is removing the causal mask in the decoder model, so that model can perform bidirectional attention, and afterwards training using masked next token prediction task, and finally followed by unsupervised contrastive learning. The paper also compared different pooling methods: last token pooling, EOS token pooling, mean pooling, and weighted mean pooling [197]. E5 Mistral [198] used the Mistral 7B model [184] as base and converted it into an embedding model by fine-tuning the model on synthetic training data without removing the causal attention mask. This approach had similar performance across the different pooling methods and therefore chose the simple EOS token pooling for the final model. NV-Embed [195] also uses the Mistral 7B [184] model as base, however, along with enabling bidirectional attention, they also introduce a latent attention layer for the final pooling.

The Qwen3 embedding model [199] used the Qwen3 foundation model [200] as base and fine-tunes it on high quality synthetic data. The model doesn't remove the causal attention mask, and uses the embedding of the EOS token as the text embedding. These decoder-only embedding models are currently state-of-the-art embedding models, however, currently there is no study assessing their effectiveness in news recommendation tasks.

Chapter 3

Gap Analysis

This chapter identifies the gaps in prior research that will be addresed in this thesis. Chapter 2 already reviewed the relevant methods and studies, and as such, in this chapter we will primarily discuss the unresolved issues addressed by the research questions.

3.1 Synthesis of Existing Research

There has already been lot of research in the field of news recommendation system for representing news content and user interests using CNNs, RNNs, attention networks, knowledge-graphs, and fine-tuned BERT models [40, 42, 46, 109, 137]. There were also major advancements in the field of transformer-based embedding models that led to development of BERT-based and decoder-based embedding models with strong performance across many embedding benchmarks [187, 189, 192, 195, 196, 198, 199, 205]. However, no prior reseach assessed the performance of these new generalized embedding models on news recommendation tasks.

3.2 Limitations of Existing Approaches

The first gap is the use of frozen generalized embeddings for news recommendation. Existing language-model-based news recommenders generally adapt fine-tune BERT models for use as news encoder [46, 57, 60, 61]. In the literature review, we have seen lack of research on the use of frozen generalized embedding models directly for news recommendation without any fine-tuning. There is also no study on the controlled comparison among encoder-based embeddings, large decoder-based embeddings, and proprietary embedding models. Although the models have strong performance on a general benchmark such as MTEB, it has not been shown yet whether they also have good performance in a news recommendation setting [195, 198, 199, 205].

The second gap is related to news recommendation tasks. Most datasets and models for news recommendation mainly focus on candidate-news ranking [46, 58, 61, 65]. There has not been studies that use other news recommendation tasks, and there is there is

therefore insufficient evidence on whether the relative performance in candidate news ranking translates directly to different news recommendation tasks.

The third gap is related to fake news classification for news recommendation. Fake-news research commonly focuses on article-level classification using machine-learning or contextual text representations [177, 178, 179, 180, 181]. The accuracy of fake news classification alone does not indicate whether filtering predicted fake news improves or harms personalized recommendation. The effect of fake news classification component on news recommendation performance has not been studied in prior reseach.

The fourth gap is multilingual news recommendation across different news recommendation tasks. Large news recommendation datasets are commonly developed for a single language, such as MIND [65] and EB-NeRD [182]. Although xMIND [50] was introduced to support multilingual and cross-lingual news recommendation, the evaluation was focused only on candidate news ranking, and no similar evaluation was carried out for the other news recommendation tasks in other reseach.

3.3 Areas Requiring Further Investigation

These limitations indicate the need for a evaluation of the use of generalized embedding models for news recommendation with ablation studies on the parts of news article to use for embedding generation and the complexity of user encoder. The evaluation should compare encoder-based, decoder-based, and proprietary embeddings; vary the parts of news article used as input for embedding models, and the user encoder; assess the effect of including fake-news classification component; and evaluate the models across different news recommendation tasks; and compare cross-lingual transfer with target-language training.

3.4 Rationale for the Proposed Research

This thesis addresses these gaps by using generalized embedding models as frozen news encoders and evaluating their performance. The generalized embedding models used in the study includes BERT-based embedding models, large decoder-based embedding models, and a proprietary embedding model. Three combinations of parts of news articles and four user-encoders are evaluated so that the effects of article content, dense transformations, and attention mechanisms can be examined separately from the selected embedding model.

The same framework is evaluated on four different news recommendation tasks: user–news classification, user classification for a news article, candidate-news classification, and candidate-news ranking. The study also measures the effect of incorporating a fake-news classification component and compares English-trained multilingual transfer with training in selected target languages across the different news recommendation tasks using xMIND [50]. These experiments directly correspond to the research questions and objectives stated in Chapter 1 and provide evidence that cannot be obtained from comparisons across independently designed systems.

3.5 Summary

The main research gap is the lack of a controlled assessment of frozen generalized text embeddings for news recommendation across different news recommendation tasks. Related research gaps include comparison among BERT-based, decoder-based and proprietary embeddings, the effects of news and user-encoding choices, the effect of fake-news classification component, and multilingual transfer in the different news recommendation tasks. The following chapter describes the methodology used to investigate these issues.

Chapter 4

Methodology

In this chapter we describe our approach and give overview of baseline methods we are comparing it against.

4.1 Problem Formulation

The goal is to rank a set of candidate news articles, N^c, for each user based on his/her news reading interest. For this we need to calculate the preference score, s, of each candidate news article, n^c, for a given user, u. The preference score is a measure of the likelihood that the user will click the candidate news article. The preference score is then used to rank the set of candidate news articles, N^c, for each user. In our approach we represent each user by the ordered set of news articles, N^u, from the user's news click history, $N^u = [n_1^u, n_2^u, \cdots, n_M^u]$, where M is the number of news articles in user's history. Similar to other studies we use only the 50 most recent items in user's history, i.e. the maximum value of M is 50. Each news article is represented only by its embedding, h, obtained from the generalized embedding model. An user encoder is used to get the user embedding, e, from the news embeddings in user's history, and finally preference score, s, is obtained as the cosine similarity between user embedding, e^u and embedding of candidate news article, h^c.

In this study we will be using this model for 4 news recommendation tasks. The tasks are described below:

- **Task 1:** User-news classification. The task is described in the Figure 4.2 given below.

- **Task 2:** All user classification for given news. The task is described in the Figure 4.3 given below.

- **Task 3:** Candidate news classification for user. The task is described in the Figure 4.4 given below.

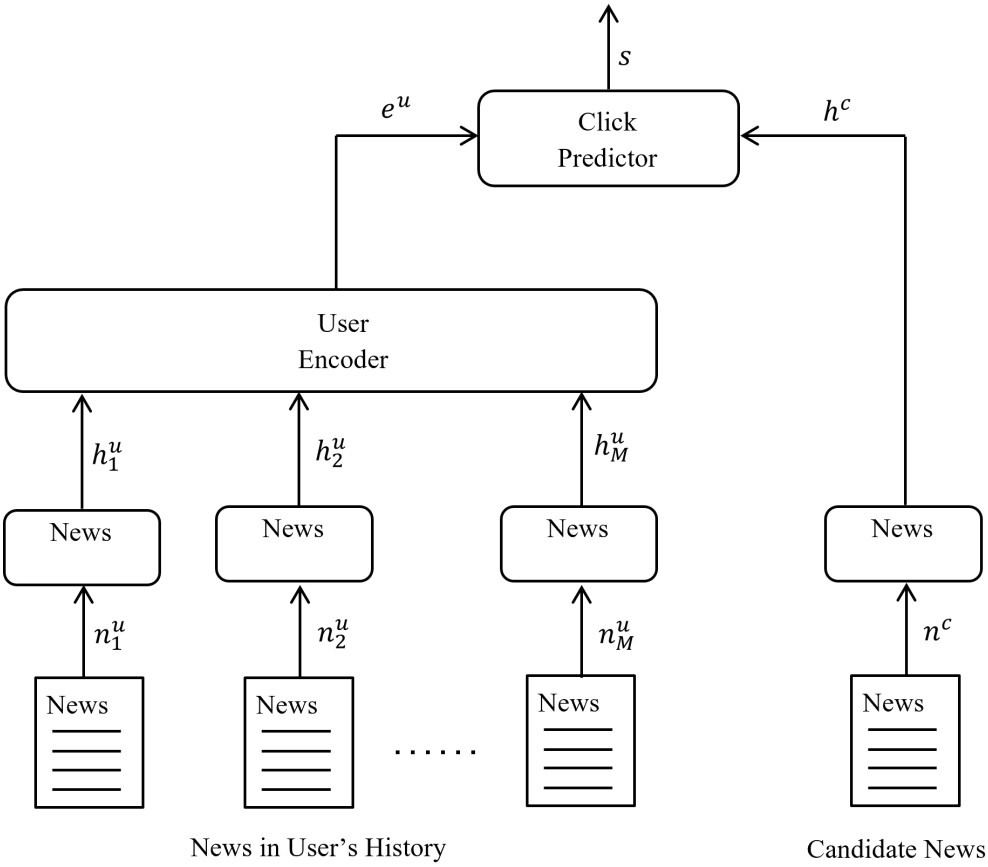

Figure 4.1: Framework for News Recommendation Model

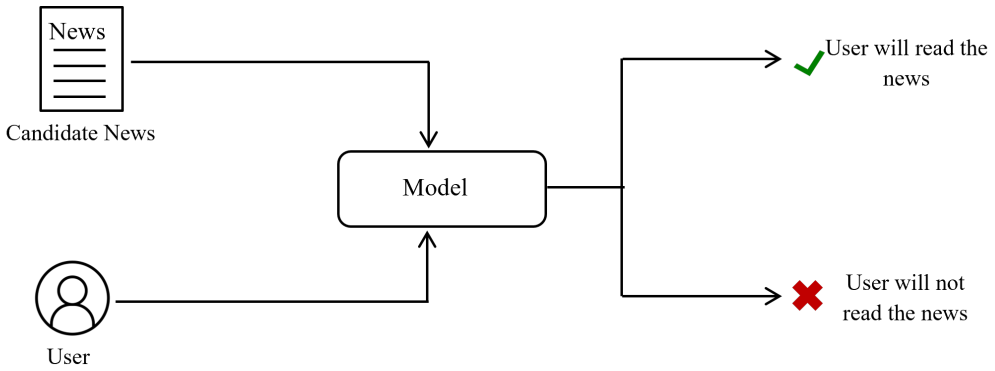

Figure 4.2: Illustration of Task 1

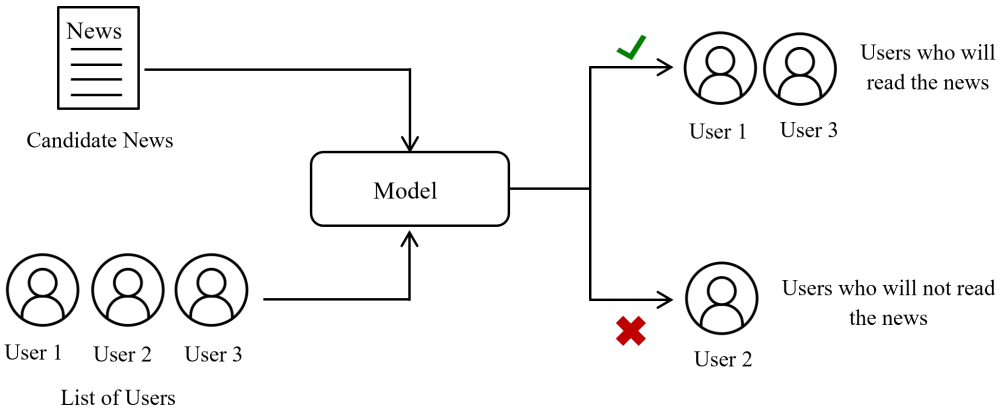

Figure 4.3: Illustration of Task 2

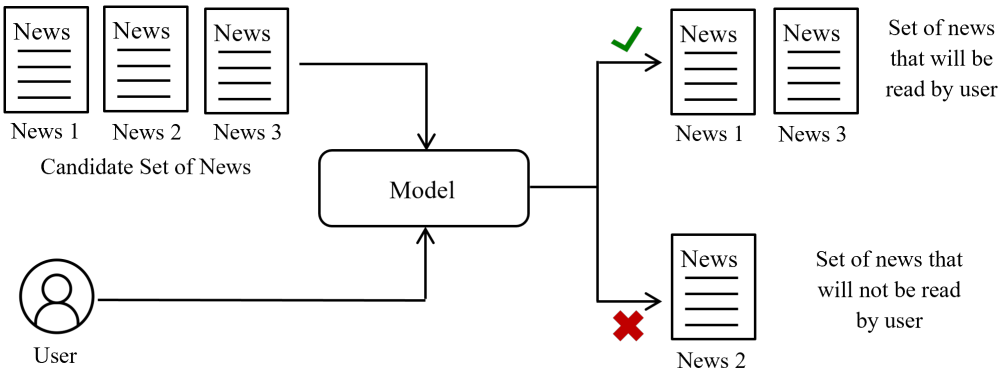

Figure 4.4: Illustration of Task 3

- **Task 4:** Candidate news ranking for user. The task is described in the Figure 4.5 given below. This is the most common news recommendation task and is widely used for news recommendation studies. Therefore, for this task we will also be comparing performance against baseline methods.

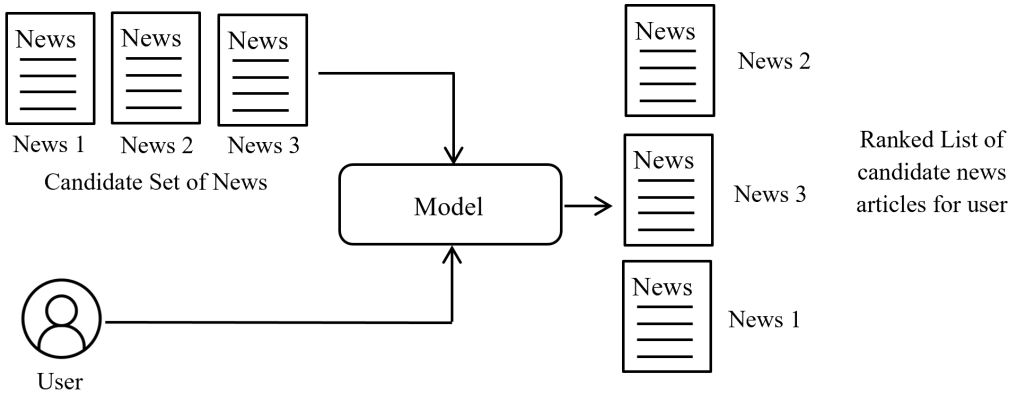

Figure 4.5: Illustration of Task 4

4.2 News Encoder

Most common news recommendation methods use only the news title and the news category for representing news articles [42, 58]. The main reasons for this is because the news title and category contain most of the useful information for news representation, and since they need to learn the news representation making use of the relatively larger news abstract or text is difficult. However, in our approach we are directly using pre-trained generalized embedding models to obtain the news representation. Therefore, we are using combinations of the news title with category and the abstract in obtaining the embeddings using the pre-trained models.

The model input variations we considered are:

1. Title

   Example prompt:
   News Title: The Brands Queen Elizabeth, Prince Charles, and Prince Philip Swear By

2. Title + Category and Subcategory

   Example prompt:
   News Title: The Brands Queen Elizabeth, Prince Charles, and Prince Philip Swear By
   Category: lifestyle
   SubCategory: lifestyleroyals

3. Title + Category and Subcategory + Abstract

Example prompt:
News Title: The Brands Queen Elizabeth, Prince Charles, and Prince Philip Swear By
Category: lifestyle
SubCategory: lifestyleroyals
Abstract: Shop the notebooks, jackets, and more that the royals can't live without.
The embedding models we are using for this study is given below.

Multilingual E5 large instruct [192]

The model was first initialized using xlm-roberta-large [186] (560M parameter) and then trained in two stages: first a weakly supervised contrastive pre-training followed by supervised fine-tuning. The details of the training process is given below:

- Training Data:
  - Pre-training: The pre-training data consists of around 1 billion multilingual text pairs from: Wikipedia, mC4, Multilingual CC News, NLLB, Reddit, S2ORC, Stackexchange, xP3 and Misc. SBERT Data (SimpleWiki, WikiAnswers, AG-News, AltLex, AmazonQA, AmazonReview, CNN/DailyMail, CodeSearchNet, Flickr30k, GooAQ, NPR, SearchQA, SentenceCompression, Specter, WikiHow, XSum, and YahooAnswers).
  - Fine-tuning: The fine-tuning data consists of around 1.6 million labeled samples from ELI5 (sample at 20%), HotpotQA, FEVER, MIRACL, MSMARCO passage ranking and document ranking (sample at 20%) , NQ, NLLB (sample at 100k), NLI [188], SQuAD, TriviaQA, Quora Duplicate Questions (sample at 10%), MrTyDi, and DuReader datasets. It also includes additional synthetic data generated by GPT 3.5 / 4. This extra data comprises of 150k instructions of over 93 languages.
- Training Method:
  - Pre-training: The pre-training was done using InfoNCE contrastive loss with in-batch negative samples.
  - Fine-tuning: The fine-tuning was done using InfoNCE contrastive loss for in-batch negative samples, as well as mined hard negative samples and knowledge distillation from a cross-encoder model.

The model has a maximum input sequence length of 512 tokens, and has an embedding dimension of 1024. Although this is BERT-based model, it uses average pooling instead of [CLS] token pooling to get the final embedding.

BGE-large-en-v1.5 [201]

This is a BERT-based model with 335 million parameters. The model is first pre-trained using retromae followed by fine-tuning using contrastive learning on large scale pairs data. Additional details on the model training is not provided. The model has a maximum input

token length of 512, and an embedding dimension of 1024. It uses [CLS] token pooling to get the final embedding.

UAE-Large-V1 [202]

This is a BERT-based model with 335 million parameters. The model is trained on SNLI and MultiNLI dataset using a combination of in-batch negative objective, cosine objective and angle objective. The model has a maximum input token length of 512, and an embedding dimension of 1024. It uses [CLS] token pooling to get the final embedding.

Mxbai-embed-large-v1 from Mixedbread [203]

This is also BERT-based model with 335 million parameters. The model is first pre-trained on more than 700 million pairs using contrastive learning, and then fine-tuned more than 30 million high quality triplets using Angle loss [202]. Further details of the training procedure wasn't shared by the model provider.

This model also has a maximum sequence input token length of 512, and an embedding dimension of 1024. This uses [CLS] token pooling to get the final embedding.

GIST-large-Embedding-v0 [204]

This is a BERT-based model (335 million parameters), obtained by fine-tuning the bge-large-en-v1.5 [201] model on triplets obtained from the MEDI dataset and guided in sample triplets mined from the MTEB classification datasets [205] (except the Amazon Polarity dataset). The UAE-Large-V1 [202] model was used to obtain the triplets from the MTEB classification dataset. The model was trained using InfoNCE loss.

The model has a maximum input token length of 512, and an embedding dimension of 1024. It uses [CLS] token pooling to get the final embedding.

E5-mistral-7b-instruct [198]

This is a decoder-based embedding model. It was first initialized using mistral 7b and then trained using about 1.8 million examples of mixture of synthetic data and published datasets. The details of the training process is given below:

- Training Data: GPT 3.5 Turbo and GPT 4 are used to generate a diverse set of text embedding data across 93 languages. Along with the synthetic data the model is also trained using the training set sampled from ELI5, HotpotQA, FEVER, MIRACL, MSMARCO passage ranking and document ranking, NQ, NLI [188], SQuAD, TriviaQA, Quora Duplicate Questions, MrTyDi, DuReader, and T2Ranking. For the datasets without hard negatives mE5 base was used to mine top 100 negative samples.

- Training Method: The model was trained using InfoNCE loss using in-batch negative samples and hard negative samples.

The model has a maximum input sequence length of 32768 tokens, and has an embedding dimension of 4096. The causal attention mask of the decoder architecture wasn't removed, and the final embedding was obtained from the embedding of the [EOS] token added at the end of the text to embed.

Qwen3 embedding models (0.6B, 4B, 8B) [199]

The Qwen3 embedding models are also decoder-based embedding models built on the Qwen3 foundation models [200]. The model is trained in two stages, first using large number of synthetic pairs for weakly supervised pre-training, followed by supervised fine-tuning using high quality synthetic pairs. Furthermore, model merging is used with sampled checkpoints from the fine-tuning stage to obtain a robust and generalizable model. The details of the training process is given below:

- Training Data: Qwen3-32B model was used to generate the synthetic data for training. For the weakly supervised pre-training 150 million pairs of synthetic data was generated. This was filtered to 12 million high quality samples for fine-tuning.
- Training Method: For both the pre-training and fine-tuning stage InfoNCE loss was used. And then model merging based on spherical linear interpolation (slerp) using sampled checkpoints across the fine-tuning stage was used to obtain a robust and generalizable model.

The model has a maximum input sequence length of 32768 tokens, and has an embedding dimension of 1024 for the 0.6B model, 2560 for the 4B model, and 4096 for the 8B model. The causal attention mask of the decoder architecture wasn't removed, and the final embedding was obtained from the embedding of the [EOS] token added at the end of the text to embed.

NV-Embed-v2 [195]

This is a decoder-based embedding model that comprises of mistral 7B model [184] followed by a Latent Attention layer for final pooling. The causal attention of mistral 7B is replaced with bidirectional attention. The model is first trained using contrastive learning with instructions on retrieval datasets, followed by instruction tuning on a blend of retrieval and non-retrieval datasets. The details of the training process is given below:

- Training Data: The model is trained on the following retrieval datasets: MS-MARCO, HotpotQA, Natural Question, PAQ, Stack Exchange, Natural Language Inference, SQuAD, ArguAna, BioASQ, FiQA, FEVER, HoVer, SciFact, NFCorpus, MIRACL and Mr.TyDi. The classification datasets used are: AmazonReviews, Amazon-Counterfactual, Banking77, Emotion, IMDB, MTOPDomain/ MTOPIntent, ToxicConversations, TweetSentiment-Extraction, AmazonPolarity, MassiveScenario/MassiveIntent. The clustering datasets used are: raw_arxiv, raw_biorxiv and raw_medrxiv datasets from MTEB Huggingface datasets, TwentyNewsgroups [75], Reddit, StackExchange, RedditP2P, StackExchangeP2P. The semantic textual similarity datasets used are: STS12, STS22 and STS-Benchmark. They also create a synthetic dataset of 120000 samples across 60000 synthetic tasks using the Mixtral-8x22B-Instruct model. For all the public datasets, only the training subset was used filtering for samples that don't match with the MTEB evaluation dataset.

- Training Method: The model is trained in two stages. In the first stage contrastive training with instructions was used, utilizing in-batch and curated hard negatives samples from the retrieval datasets. In the second stage retrieval and non-retrieval datasets was used with curated hard negative samples for training. The curated hard negative samples were obtained using the E5 mistral 7B instruct model.

The model has a maximum input sequence length of 32768 tokens, and has an embedding dimension of 4096.

Text-embedding-3-large (from OpenAI)

This is the top proprietary embedding model from OpenAI. It has a max input length of 8192 tokens and an embedding dimension of 3072. Other details about the model is not provided.

Table 4.1: MTEB (Eng V2) scores from MTEB online leaderboard on August 10, 2025

Model Name	**Model Type**	**Number of Parameters**	**Embedding Dimensions**	**Max Tokens**	**MTEB**
multilingual-e5-large-instruct	BERT-based	560.0M	1024	512	65.53
bge-large-en-v1.5	BERT-based	335.0M	1024	512	65.89
UAE-Large-V1	BERT-based	335.0M	1024	512	66.4
mxbai-embed-large-v1	BERT-based	335.0M	1024	512	66.26
GIST-large-Embedding-v0	BERT-based	335.0M	1024	512	66.25
e5-mistral-7b-instruct	Decoder-based	7.0B	4096	32768	67.97
Qwen3-Embedding-0.6B	Decoder-based	595.0M	1024	32768	70.7
Qwen3-Embedding-4B	Decoder-based	4.0B	2560	32768	74.6
Qwen3-Embedding-8B	Decoder-based	7.0B	4096	32768	75.22
NV-Embed-v2	Decoder-based	7.0B	4096	32768	69.81
text-embedding-3-large (OpenAI)	Unknown	Unknown	3072	8192	66.43

Table 4.1 provides a summary of the generalized embedding models and their performance on the MTEB dataset. MTEB [205] is a benchmark dataset for evaluating performance of generalized embedding models across a range of different tasks and domains.

4.3 User Encoder

In our approach we represent each user by the ordered set of news articles $N^u = [n_1^u, n_2^u, \cdots, n_M^u]$, where M is the number of news articles in user's history. Each news article, n, is represented by its embedding, h, from the generalized embedding model. The purpose of the user encoder is to process and aggregate the embeddings of the news articles in user's history $[h_1^u, h_2^u, \cdots, h_M^u]$, into a final user embedding that captures user's interests.

We use the following models for aggregation of user's news history into user embedding, e^u.

Model 1: Average Pooling

In this approach we simply take the mean of embeddings of the news articles in user's history to form the user embedding. For this model there is no training required.

$$e = \frac{1}{M} \sum_{i=1}^{M} h_i \tag{4.1}$$

Model 2: Attention Pooling

In this method we use weighted combination of news embeddings in user history by attention mechanism. The weight of each news embedding in user history, α_i is given by the following equation.

$$a_i = qtanh(V \times h_i + v_b) \tag{4.2}$$

$$\alpha_i = \frac{exp(a_i)}{\sum_{j=1}^{M} a_j} \tag{4.3}$$

Here V and v_b are projection parameters and q is the attention query vector. We use a hidden dimension of 200 in our experiments. The final user representation is obtained by summing the news embeddings weighted by their corresponding weights.

$$e = \sum_{i=1}^{M} \alpha_i h_i \tag{4.4}$$

Model 3: Dense Layer + Attention Pooling

In this approach, news embeddings from user's history are first passed through a Dense layer in a residual connection before going through Attention Pooling described in Model 2.

$$r = W \times h + b \tag{4.5}$$

$$h^d = h + tanh(r) \tag{4.6}$$

Here, W and b are trainable parameters of the dense layer, and the output r is passed through a $tanh$ activation and is joined with the original embedding, h , in a residual

connection. The final output h^d is passed to an Attention Pooling layer as described in Model 2 to make the user embedding e.

Model 4: Multi-Head Self-Attention + Attention Pooling

In this approach, we use a Multi-Head Self Attention mechanism to learn better representation.

$$Attention(Q,\ K,\ V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{4.7}$$

Here Q, K and V represent the query, key and value matrices respectively, d_k is the dimension of the key vector. Multi-Head Self-Attention applies n attention functions in parallel to produce the output representations which are concatenated and passed into a dense layer.

$$head_i = Attention\left(hW_i^Q, hW_i^K, hW_i^V\right) \tag{4.8}$$

$$r = concat\left([head_1, head_2, \cdots, head_n]\right)W^O + b^O \tag{4.9}$$

$$h^a = h + tanh(r) \tag{4.10}$$

Here W_i^Q, W_i^K, W_i^V, W^O and b^O are learnable parameters and i represents the i-th head. In our experiments we use 8 heads in the Multi Head Self Attention. The output of the self attention, r, is passed through a $tanh$ activation and is joined with the original embedding, h , using a residual connection, and the final output, h^a is passed to an Attention Pooling layer as described in Model 2 to make the user embedding e.

4.4 Click Prediction

The final preference score between the user and the candidate news article is obtained by taking the cosine similarity between the user embedding and the news embedding for the candidate news article.

$$s = \frac{e^u \bullet h^c}{|e^u|\,|h^c|} \tag{4.11}$$

4.5 Model Training

We train the user encoder models using InfoNCE loss, using a batch size of 512, for 5 epochs using AdamW optimizer [206]. For Model 2 (Attention Pooling) we used a learning rate of 1e-5 and for models 3, and 4 (Dense Layer + Attention Pooling and Multi-Head Self-Attention + Attention Pooling) we used a learning rate of 5e-5. For each news article clicked in the training dataset, D, we randomly sample L negative samples from the same news session. The preference score is calculated for the positive sample and the negative

news samples. The scores are used to calculate the InfoNCE loss using the equation given below. In our experiments we use negative sampling ratio, L of 4.

$$L_{NCE} = -\sum_{i=1}^{|D|} \log \frac{exp(s_i^+)}{exp(s_i^+) \ + \sum_{j=1}^{L} exp(s_i^j)} \tag{4.12}$$

4.6 Experiments

4.6.1 Dataset

Table 4.2: Statistics of the datasets

	MIND-small Train	MIND-small Dev
# User	50,000	50,000
# News	51,282	42,416
# Impressions	156,965	73,152
# Positive samples	236,344	111,383
# Negative samples	5,607,100	2,629,615

We conducted our experiments using a real-world dataset MIND-small [65]. The MIND-large dataset consists of click logs of a sample of 1 million users over 6 weeks on the Microsoft News website. MIND-small is a smaller version of the dataset obtained by sampling 50000 users and their behavior logs. The behavior logs contain the user id, time of impression, candidate news items shown to the user, and history of news articles clicked by user. The training and validation datasets also contain the outcome of the candidate news items shown to the user, i.e. whether they were clicked by user or not. For each news article the dataset includes, the title, category, sub-category, abstract, as well as entities extracted from the title and abstract.

We evaluate the model performance using the metrics: Accuracy, Precision, Recall, F1 Score, and AUC for news recommendation tasks 1, 2, and 3; and AUC, MRR, nDCG@5, and nDCG@10 for news recommendation task 4. The metrics for task 4 were as recommended by the MIND dataset. We evaluate these metrics for all the combinations of generalized embedding models, model input variations, and user encoder models across all the 4 news recommendation tasks described above. Finally we compare the best version of each generalized embedding model for each task. For task 4 we also compare the performance of our studied models against the baseline methods described below.

4.6.2 Fake News Classification Model

For fake news classification we trained our model using the ISOT fake news dataset [207, 207]. First we extracted the representation of the news articles in the ISOT dataset using the generalized embedding model using news title only as input. We trained a

two-layer dense neural network on the extracted embeddings, giving a accuracy of over 0.98 in the validation dataset. This trained fake news classification model was used to classify the news articles in the MIND-small dataset as either real or fake.

For classification tasks in this study the fake news articles were automatically assigned negative class, and for the ranking task the fake news articles were moved to the bottom of the list. We performed experiments with and without the inclusion of the fake news classification component to assess its performance.

4.6.3 Multilingual News Recommendation

For Multilingual News Recommendation we used the xMIND dataset [50]. The dataset consists of the MIND dataset in 14 geographically and linguistically diverse set of languages. In this analysis we used three of the languages: Chinese (language code: zh), Turkish (language code: tr) and Vietnamese (language code: vi).

We followed the process outlined in the xMIND paper, and did two variations for multilingual news recommendation. We compared the performance of training on English only and evaluation on the target language, versus both training and evaluation in the target language. This was done across all the model variations across all the 4 news recommendation tasks. We did not include fake news classification component for multilingual news recommendation, due to lack of suitable data for multilingual fake news classification.

We used five generalized embedding models for multilingual news recommendation analysis. These are:

1. multilingual-e5-large-instruct
2. Qwen3-Embedding-0.6B
3. Qwen3-Embedding-4B
4. Qwen3-Embedding-8B
5. Text-embedding-3-large (from OpenAI)

4.6.4 Baseline methods for Task 4 (Candidate News Ranking for User)

Feature based methods

These method uses hand-crafted features and ID based features for news recommendation.

LibFM [208]: This is a simple feature based method in which the user ID, news ID and TF-IDF features extracted from news articles were used with a factorization machine model to get the preference of users for the candidate news items.

DeepFM [209]: This is another feature based method similar to LibFM [208] above and using the same features as LibFM. It uses a combination of factorization machine model with a deep neural network for preference prediction.

Neural recommendation methods

These methods use neural networks for automatic feature extraction from the news contents.

DKN [40]: It used a knowledge-aware CNN on the news title and title entities for news representation and a candidate aware attention network on user's news reading history for user representation. The candidate and user representations are concatenated together and then uses a deep neural network for final score prediction.

NPA [111]: The model first uses a CNN on the news articles for the word representation, and then uses a personalized attention network on the word representations to form the final news representation. Another personalized attention network is used on the news representations in user's history to form the user representation. The final preference score is obtained by a dot product between news representation of the candidate news article and the user representation.

NAML [137]: In this model, the news title, body, category, and subcategory is processed into different views of the news article. A multi-view attention network is used on these views to form the news representation. A news level attention network is used on the user's news history to form the user representation. A dot product between the candidate news representation and the user representation is used to get the final score.

LSTUR [42]: In this model a CNN and attention network is used on the word embedding of the news title is used to obtain the title embedding. This is concatenated with category and subcategory embeddings to form the news representation. The long term interest of the user is learned using a user embedding, which is the used to initialize the hidden state of the GRU network used to learn the short-term user interest from user's history. The final output of the GRU network is used as the user representation. The final score is obtained from the dot product between the user representation and the candidate news representation.

NRMS [109]: It uses a multi-head self-attention network on the news title followed by an additive attention network for the news representation. A news level multi-head attention network is used on the news articles in user's history followed by an additive attention network to form the user representation. The final click probability score is obtained by using a dot product between the user representation and the candidate news representation.

Bert Enhanced Models

Most of the neural recommendation methods above use pre-trained word embeddings or learns the embeddings using simple neural networks. The BERT enhanced models fine-tunes last few layers of a pre-trained BERT model for the news embeddings.

LSTUR + BERT [46]: This is similar to the LSTUR [42] model describe above, but replaces the CNN on word embeddings with a BERT model for getting the news title embeddings.

NRMS + BERT [46]: It is similar to the NRMS [109] model described above, but uses a BERT model for obtaining the news embedding.

UNBERT [61]: In this method the user is represented by concatenating the titles of news articles in user's history with special tokens in between. This user sentence is concatenated with title of candidate news article along with some special tokens, and is

fed as input to a encoder model which then provides the final click prediction score. The model itself incorporates both word level and news level matching for better accuracy.

MINER [58]: In this model, BERT is used for embedding news titles and the news category is embedded separately. The user embedding is represented as a multi interest user embedding obtained from the news embeddings in user's history weighted by the category embedding of the candidate news article. The final model score is obtained using the dot product of user and news embedding.

Chapter 5

Results and Discussion

5.1 Main Results

5.1.1 Task 1: User-News Classification

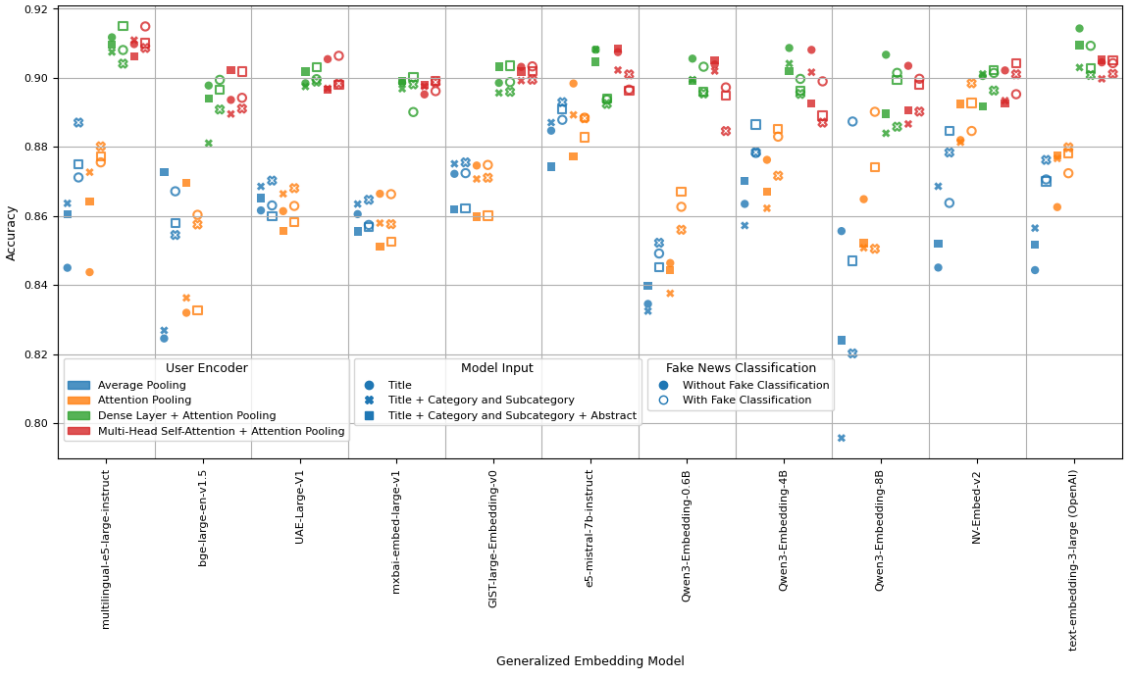

Figure 5.1: User-News Classification with vs without Fake News Classification Model

Figure 5.1 given above shows the performance of all the model variations for user-news classification. The values can be found in Tables A.1 and A.2. From the figure we can see that the user encoders "Dense Layer + Attention Pooling" and "Multi-Head Self-Attention + Attention Pooling" tend to give significantly higher accuracy than either "Average Pooling" or "Attention Pooling". This shows that the embeddings produced from some generalized embedding models is not properly aligned for user-news classification task; however, they still have the required information and by using "Dense Layer + Attention Pooling" and "Multi-Head Self-Attention + Attention Pooling" as user encoders, the raw embeddings

can be properly aligned to give very good performance in user-news classification.

The generalized embedding model also has a significant effect on the performance, especially with the user encoders "Average Pooling" and "Attention Pooling". This shows that the embeddings produced by some generalized embedding models is properly aligned for user-news classification task even without any transformation, while others require some transformation to align the embeddings. The inclusion of the fake news classification model also improved the performance of many of the models. The effect of the model input is less clear.

The best model for the task is using the "multilingual-e5-large-instruct" generalized embedding model using "Title + Category and Subcategory + Abstract" as the embedding model input with "Dense Layer + Attention Pooling" user encoder and including the fake news classification model.

The effect of each of the factors is analyzed in more detail below.

Effect of Generalized Embedding Model

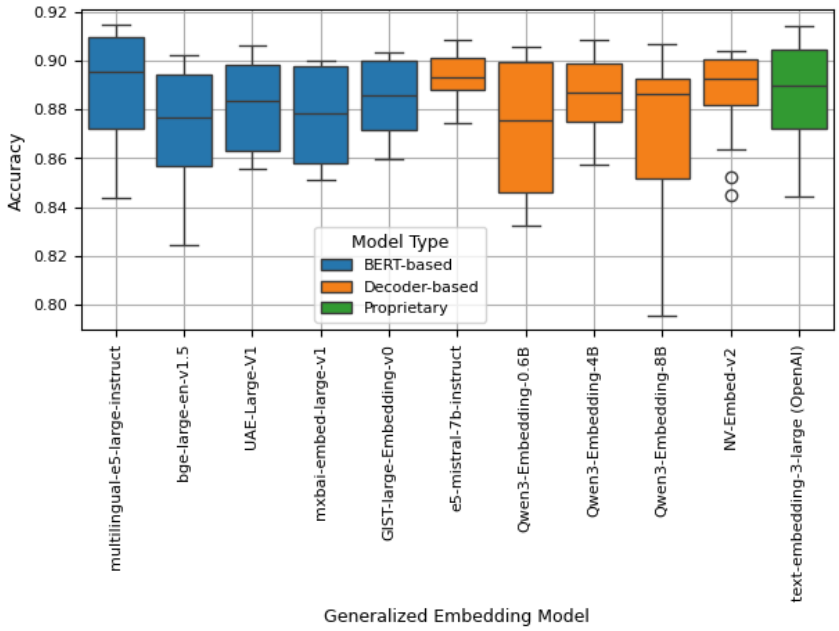

Figure 5.2: User-News Classification by Generalized Embedding Model

The effect of the generalized embedding model used for obtaining the embeddings are given in figure 5.2. We can see that the best performances are from "multilingual-e5-large-instruct" and "text-embedding-3-large (OpenAI)". The performance of large decoder-based models is not significantly better than smaller BERT-based models. Some of the decoder-based models "e5-mistral-7b-instruct" and "NV-Embed-v2" have good performance across all the variations in user encoder, model input and fake news classification model.

Effect of User Encoder

Figure 5.3 shows the effect of different user encoders on model performance. We can see in the figure that "Dense Layer + Attention Pooling" and "Multi-Head Self-Attention + Attention Pooling" have significantly better performance than "Average Pooling" or

"Attention Pooling". "Attention Pooling" is able to weigh the news embeddings from user history which gives slightly improved performance over "Average Pooling". The very close performance between "Dense Layer + Attention Pooling" and "Multi-Head Self-Attention + Attention Pooling" shows that even a simple trained linear transform through a Dense Layer is enough to properly align the embeddings for user-news classification.

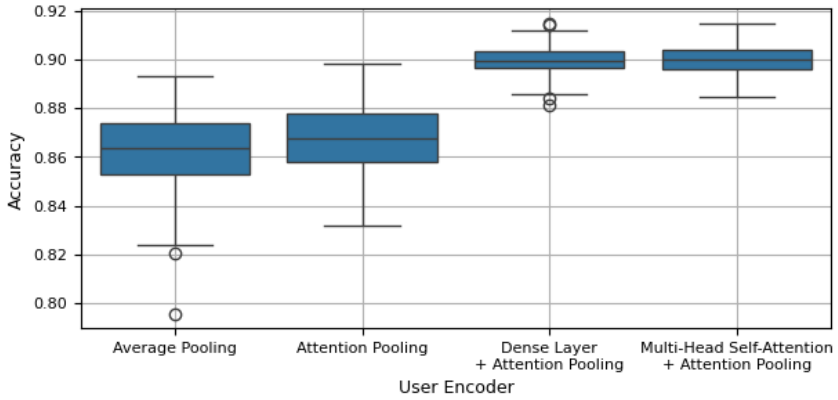

Figure 5.3: User-News Classification by User Encoder

Effect of Model Input

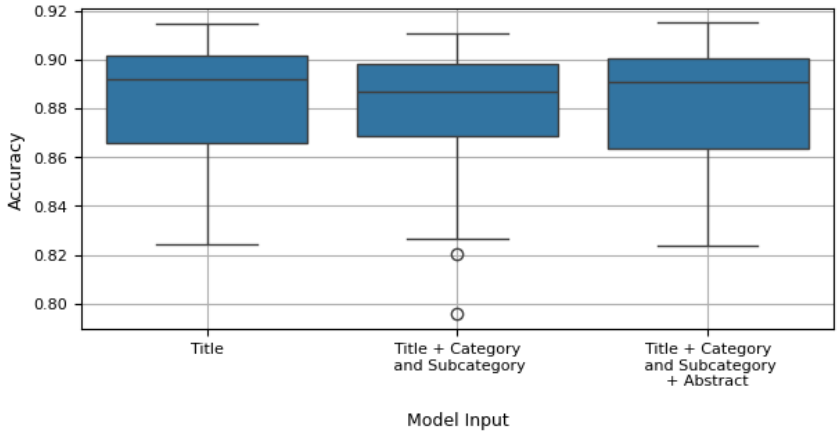

Figure 5.4: User-News Classification by Model Input

Figure 5.4 shows the effect of model input on user-news classification performance. We can see from the figure that the effect is not very significant. Using "Title + Category and Subcategory" gives slightly lower performance than using "Title" only or "Title + Category and Subcategory + Abstract".

Effect of Fake News Classification Component

Figure 5.5 shows the effect of inclusion of fake news classification component on user-news classification performance. The figure shows that including the fake news classification component leads to a modest increase in performance.

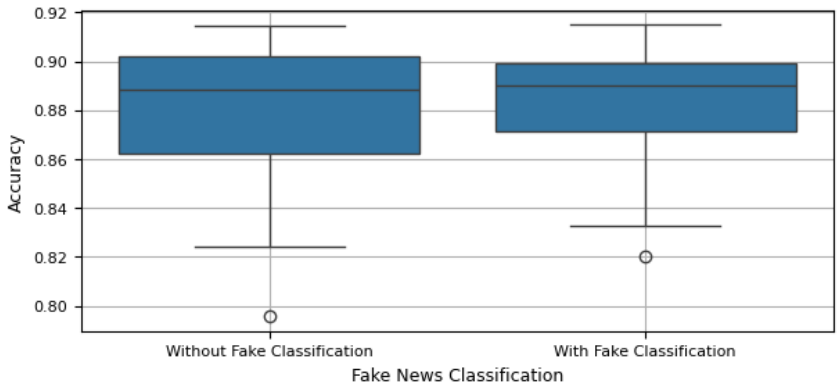

Figure 5.5: User-News Classification by Inclusion of Fake News Classification Component

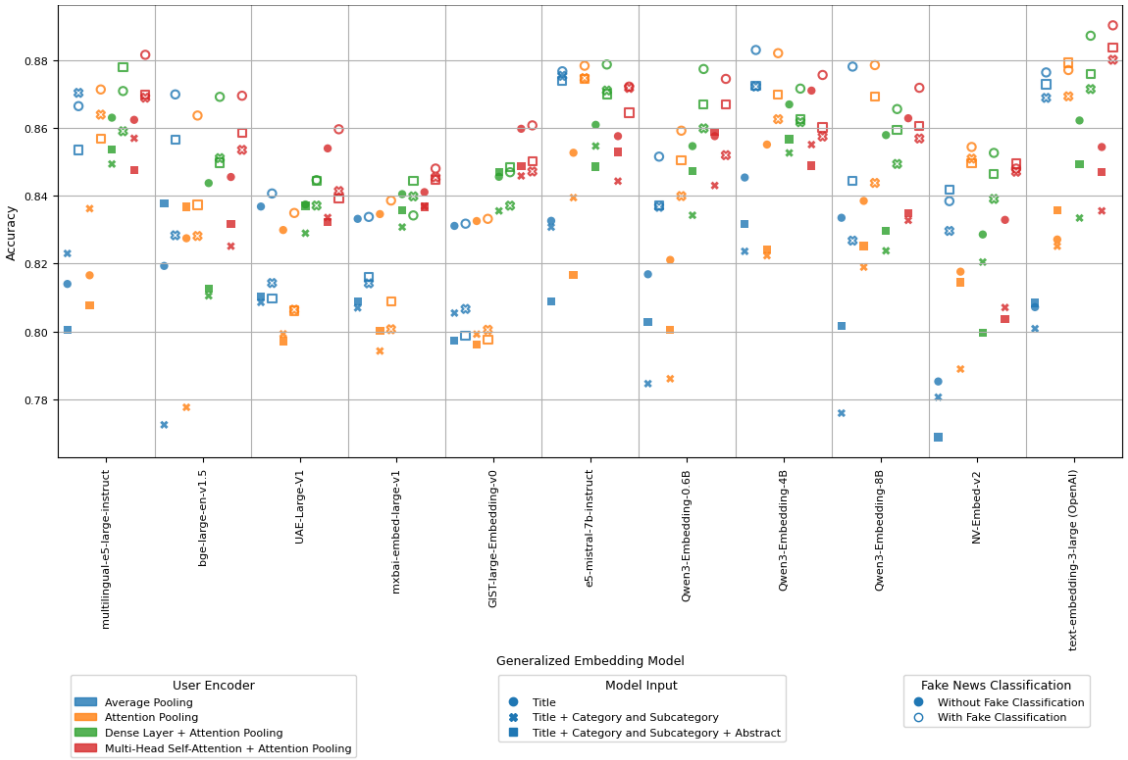

Figure 5.6: All User Classification for given News with vs without Fake News Classification Model

5.1.2 Task 2: All User Classification for given News

The figure 5.6 given above shows the performance of all the model variations for classifying all users for given news. The values can be found in Tables A.3 and A.4. From the figure we can see that including the fake news classification component significantly improves performance of the models. Identification of fake news articles helps identify news articles that users are unlikely to read. In this task, not recommending them to any users at all provides significant boost in performance.

Similar to Task 1, the user encoders "Dense Layer + Attention Pooling" and "Multi-Head Self-Attention + Attention Pooling" have significantly better performance than "Average Pooling" and "Attention Pooling". Using "Title" as the model input gives some better performance than the other model input variations.

The best model for the task is using the "text-embedding-3-large (OpenAI)" generalized embedding model using "Title" as the embedding model input with "Multi-Head Self-Attention + Attention Pooling" user encoder and including the fake news classification model.

The effect of each of the factors is analyzed in more detail below.

Effect of Generalized Embedding Model

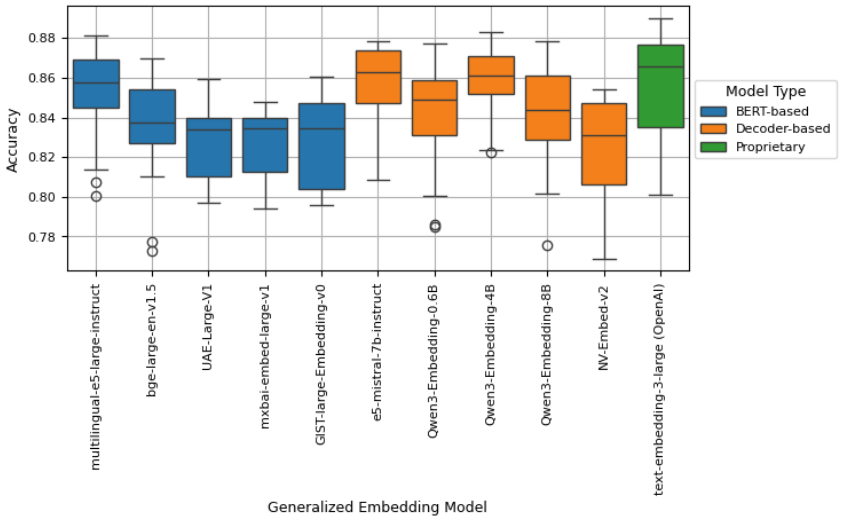

Figure 5.7: All User Classification for given News by Generalized Embedding Model

Figure 5.7 shows the effect of the generalized embedding model used for obtaining the embeddings. The best performance is from the "text-embedding-3-large (OpenAI)" embedding model. Unlike Task 1, here the performance of large decoder-based models is slightly better than smaller BERT-based models. One interesting observation is that the performance of the "Qwen3-Embedding-4B" model is significantly better than the "Qwen3-Embedding-0.6B" and "Qwen3-Embedding-8B" models. This shows that increasing number of parameters doesn't always lead to corresponding increase in performance.

Effect of User Encoder

Figure 5.8 shows the effect of different user encoders on classifying all users for given

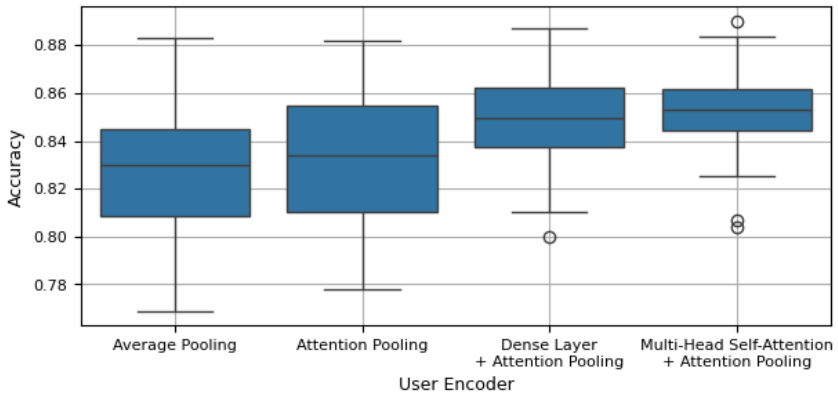

Figure 5.8: All User Classification for given News by User Encoder

news. The effects are very similar to those observed in Task 1. One difference to note is that "Multi-Head Self-Attention + Attention Pooling" has slightly better performance than "Dense Layer + Attention Pooling".

Effect of Model Input

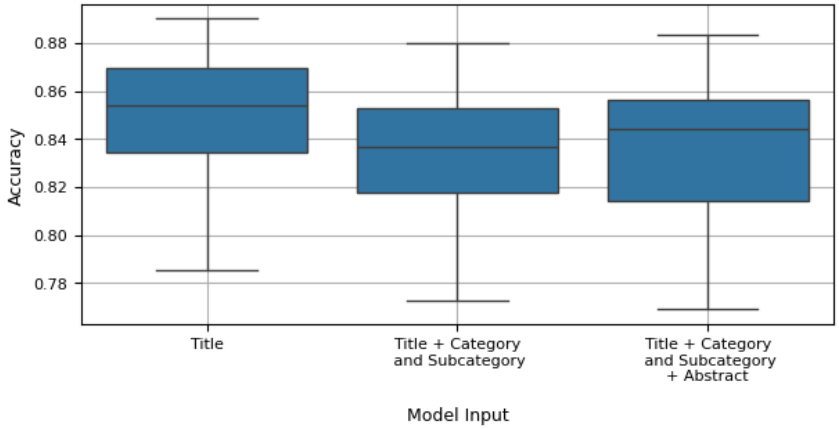

Figure 5.9: All User Classification for given News by Model Input

Figure 5.9 shows the effect of model input on the performance of classifying all users for given news. We can see from the figure that the effect is not very significant. Unlike Task 1, using "Title" gives better performance than using "Title + Category and Subcategory" or "Title + Category and Subcategory + Abstract" as the embedding model input.

Effect of Fake News Classification Component

Figure 5.10 shows the effect of inclusion of fake news classification component on the performance of classification all users for given news. The figure shows that including the fake news classification component provides a significant boost in performance. Fake news articles have a much lower chance of being read, and so not recommending them to users, helps increase the accuracy.

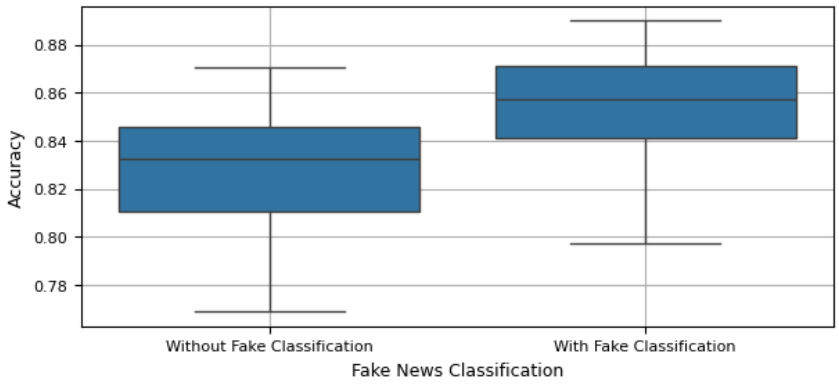

Figure 5.10: All User Classification for given News by Inclusion of Fake News Classification Component

5.1.3 Task 3: Candidate News Classification for User

Figure 5.11 shows the performance of all the model variations for candidate news classification for user. The values can be found in Tables A.5 and A.6. The main observations from this figure is very similar to those from figure 5.1 for Task 1, however, the scores are slightly lower.

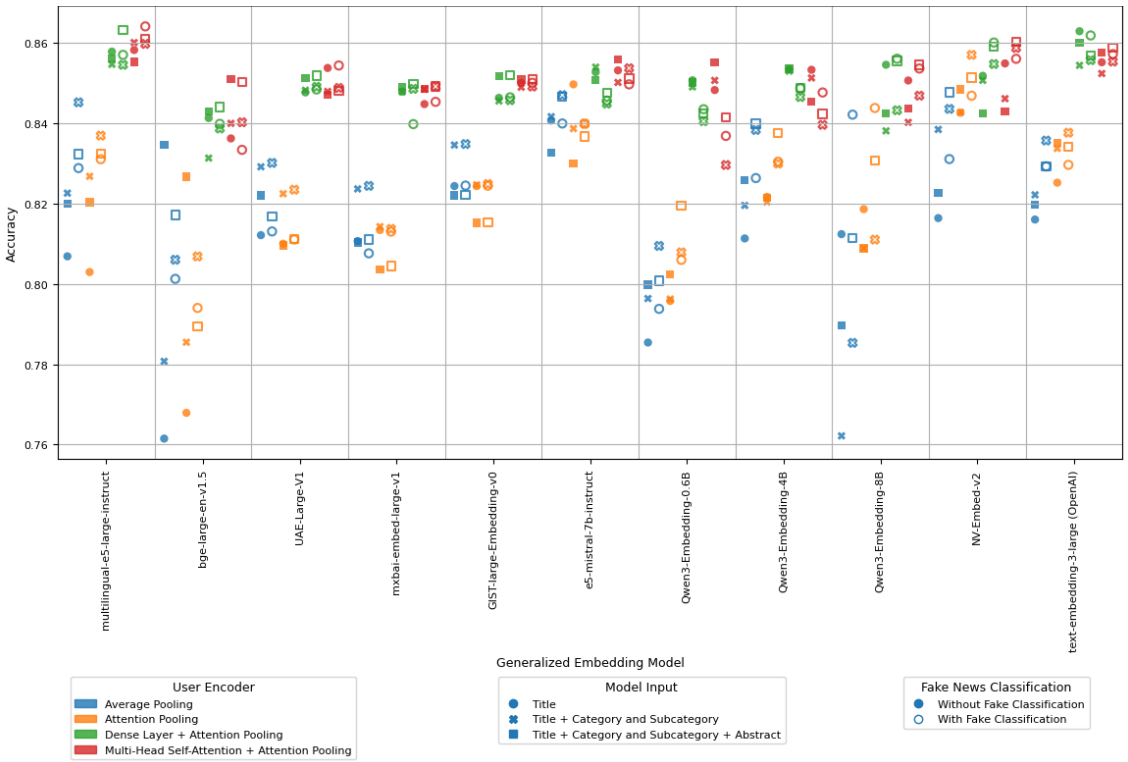

Figure 5.11: Candidate News Classification for User with vs without Fake News Classification Model

The best model for the task is using the “multilingual-e5-large-instruct” generalized embedding model using “Title” as the embedding model input with “Multi-Head Self-

Attention + Attention Pooling" user encoder and including the fake news classification model.

The effect of each of the factors is analyzed in more detail below

Effect of Generalized Embedding Model

The effect of the generalized embedding model used for obtaining the embeddings are given in figure 5.12. The trends observed are very similar to those observed for figure 5.2 in Task 1.

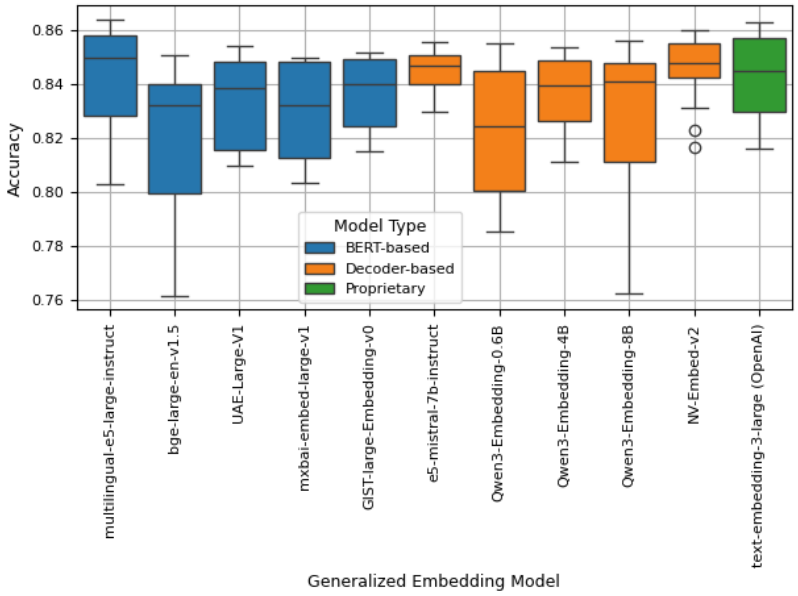

Figure 5.12: Candidate News Classification for User by Generalized Embedding Model

Effect of User Encoder

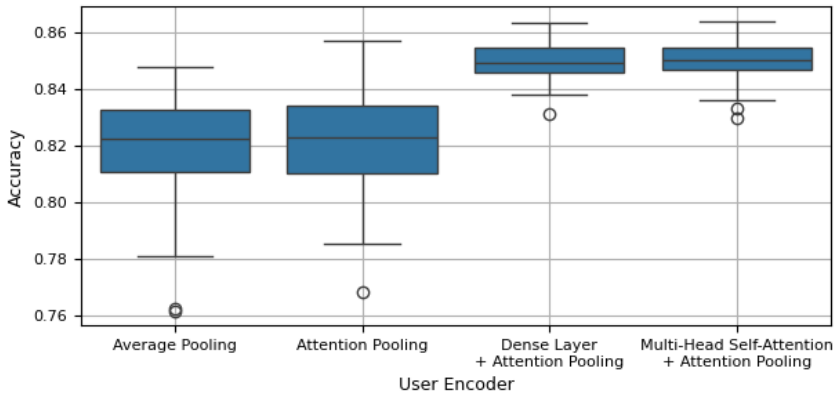

Figure 5.13: Candidate News Classification for User by User Encoder

Figure 5.13 shows the effect of different user encoders on candidate news classification. Here also the trends are similar to those observed in figure 5.3 for Task 1.

Effect of Model Input

Figure 5.14 shows the effect of model input on the performance of candidate news

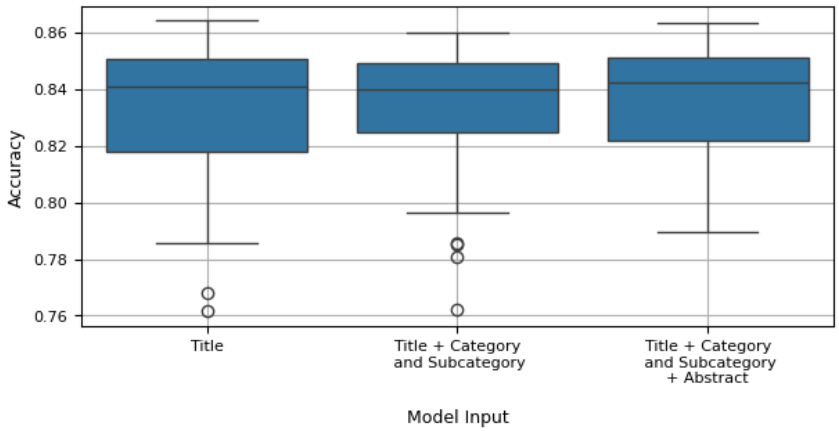

Figure 5.14: Candidate News Classification for User by Model Input

classification. We can see from the figure that all the model inputs have very similar performance.

Effect of Fake News Classification Component

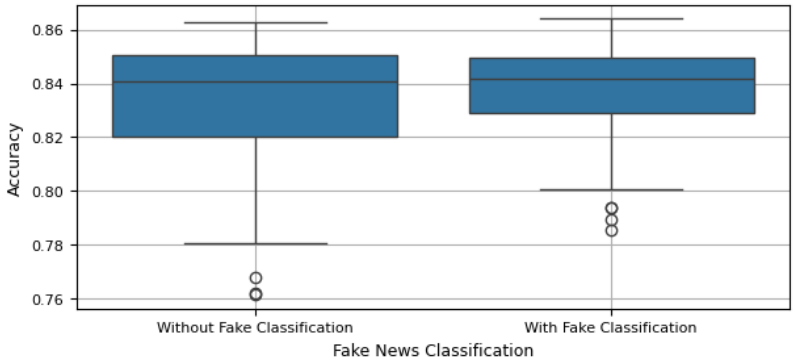

Figure 5.15: Candidate News Classification for User by Inclusion of Fake News Classification Component

Figure 5.15 shows the effect of inclusion of fake news classification component on candidate news classification performance. The figure shows that including the fake news classification component doesn't have significant effect on performance.

5.1.4 Task 4: Candidate News Ranking for User

The figure 5.16 given below shows the performance of all the model variations as well as the baseline methods for candidate news ranking. The values can be found in Tables A.7, A.8, and A.9. From the figure we can see that, unlike all the previous tasks studied, including the fake news classification component leads to significantly lower performance of the models. In ranking tasks, fake news articles are ranked at the bottom of the list, and in the cases where user still reads the fake news articles it negatively impacts the ranking metrics. The effect of the fake news classification component varies significantly across the different generalized embedding models used. Some generalized embedding

models like “UAE-Large-V1”, “mxbai-embed-large-v1”, and “GIST-large-Embedding-v0” are not significantly affected by the fake news classification component, however, models like “multilingual-e5-large-instruct” and “text-embedding-3-large (OpenAI)” are greatly affected by it.

Unlike the previous tasks, here the effect of user encoder is considerably less. The user encoders “Dense Layer + Attention Pooling” and “Multi-Head Self-Attention + Attention Pooling” has slightly better performance than “Average Pooling” and “Attention Pooling” but the difference is not as big as in the previous news recommendation tasks.

The best model for the task is using the “mxbai-embed-large-v1” generalized embedding model using “Title + Category or Subcategory + Abstract” as the embedding model input with “Multi-Head Self-Attention + Attention Pooling” user encoder and without the fake news classification component. In comparison to the baseline models, most of the models in our experiments that don’t include the fake news classification component have better performance than the feature based baseline models. Some of our experimental models are also able to perform better than the neural recommendation baseline methods. However, the BERT-enhanced baseline models have better performance than the models in our experiments.

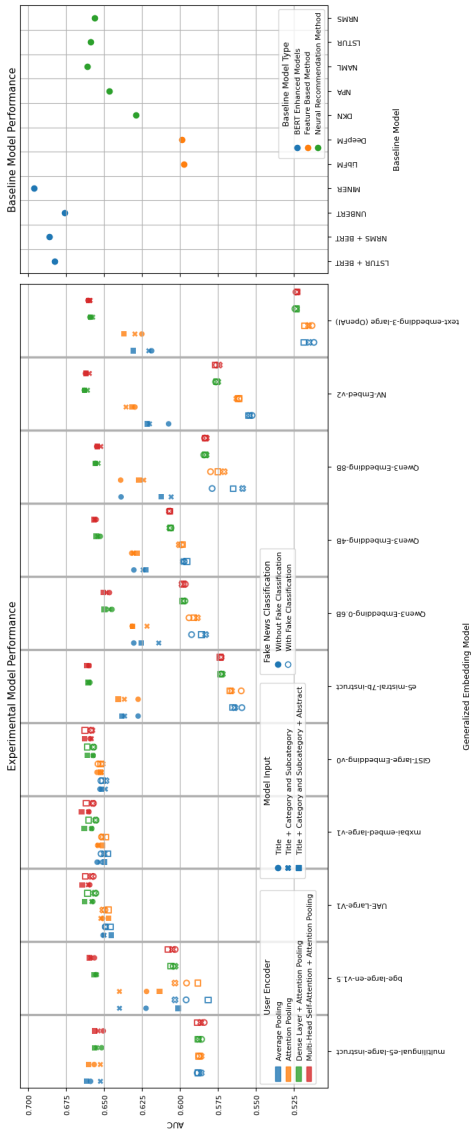

Figure 5.16: Candidate News Ranking for User with vs without Fake News Classification Model

The effect of each of the factors is analyzed in more detail below.

Effect of Generalized Embedding Model

Figure 5.17 shows the effect of the generalized embedding model used for obtaining the embeddings on candidate news ranking. The best performance is from the embedding models "UAE-Large-V1", "mxbai-embed-large-v1", and "GIST-large-Embedding-v0". Unlike the previous tasks, the large range in performance of some the models is due to the fake news classification component, as discussed in the observations from figure 5.16. Also for this task the large decoder-based models have slightly lower performance than the smaller BERT-based models.

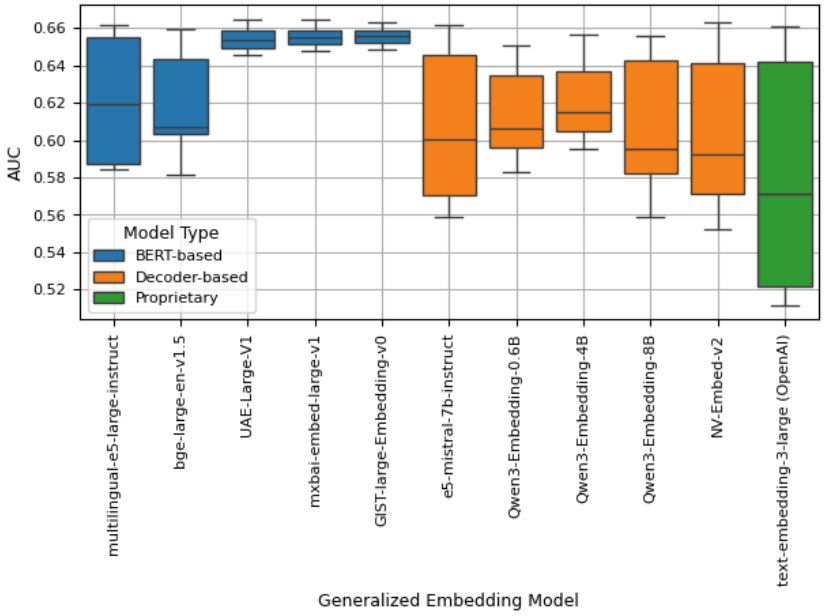

Figure 5.17: Candidate News Ranking for User by Generalized Embedding Model

Effect of User Encoder

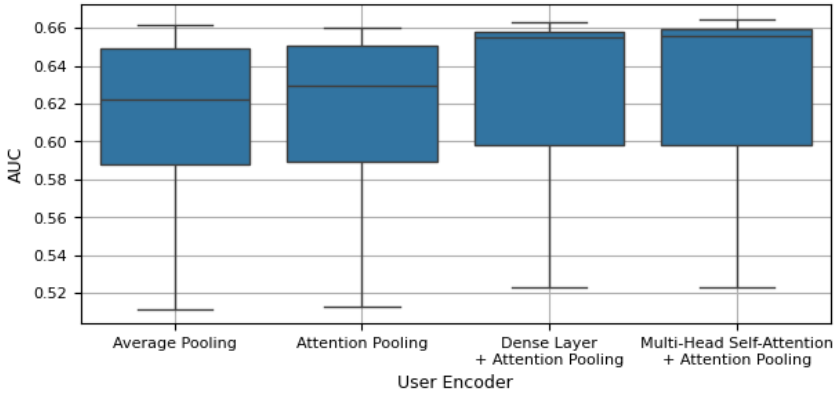

Figure 5.18: Candidate News Ranking for User by User Encoder

Figure 5.18 shows the effect of different user encoders on candidate news ranking. Here the user encoder has less effect than the Tasks 1, 2 and 3. Although "Multi-Head

Self-Attention + Attention Pooling" and "Dense Layer + Attention Pooling" has slightly better performance than "Average Pooling" and "Attention Pooling", the difference is much less compared to other news recommendation tasks.

Effect of Model Input

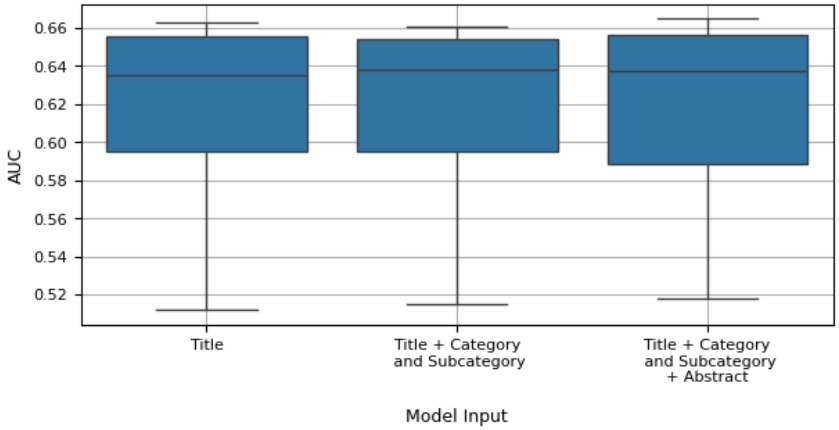

Figure 5.19: Candidate News Ranking for User by Model Input

Figure 5.19 shows the effect of model input on the performance of candidate news ranking. We can see from the figure that all the model inputs have very similar performance.

Effect of Fake News Classification Component

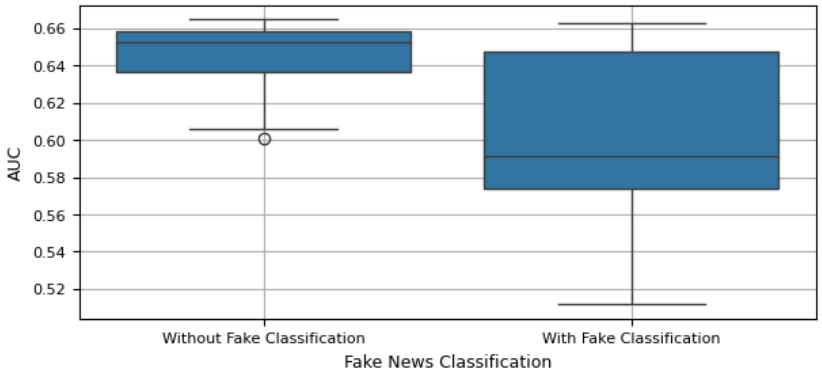

Figure 5.20: Candidate News Ranking for User by Inclusion of Fake News Classification Component

Figure 5.20 shows the effect of inclusion of fake news classification component on candidate news ranking performance. The figure shows that including the fake news classification component leads to significantly lower performance.

5.2 Multilingual News Recommendation

5.2.1 Task 1: User-News Classification

Figure 5.21 shows the performance of all the model variations for multilingual user-news classification. The figure 5.21 (a) shows the results for training in English and evaluation

in the target language, and 26 (b) shows the results for both training and evaluation in the target language. Figure 5.22 shows the performance for different languages grouped by whether training was done in English only or in the target language. The values can be found in Table A.10, and A.11. From the figure we can see that Vietnamese has good performance even when trained in English only in this task. The performance improves across all the languages when trained on the target language, but the improvement is not very significant.

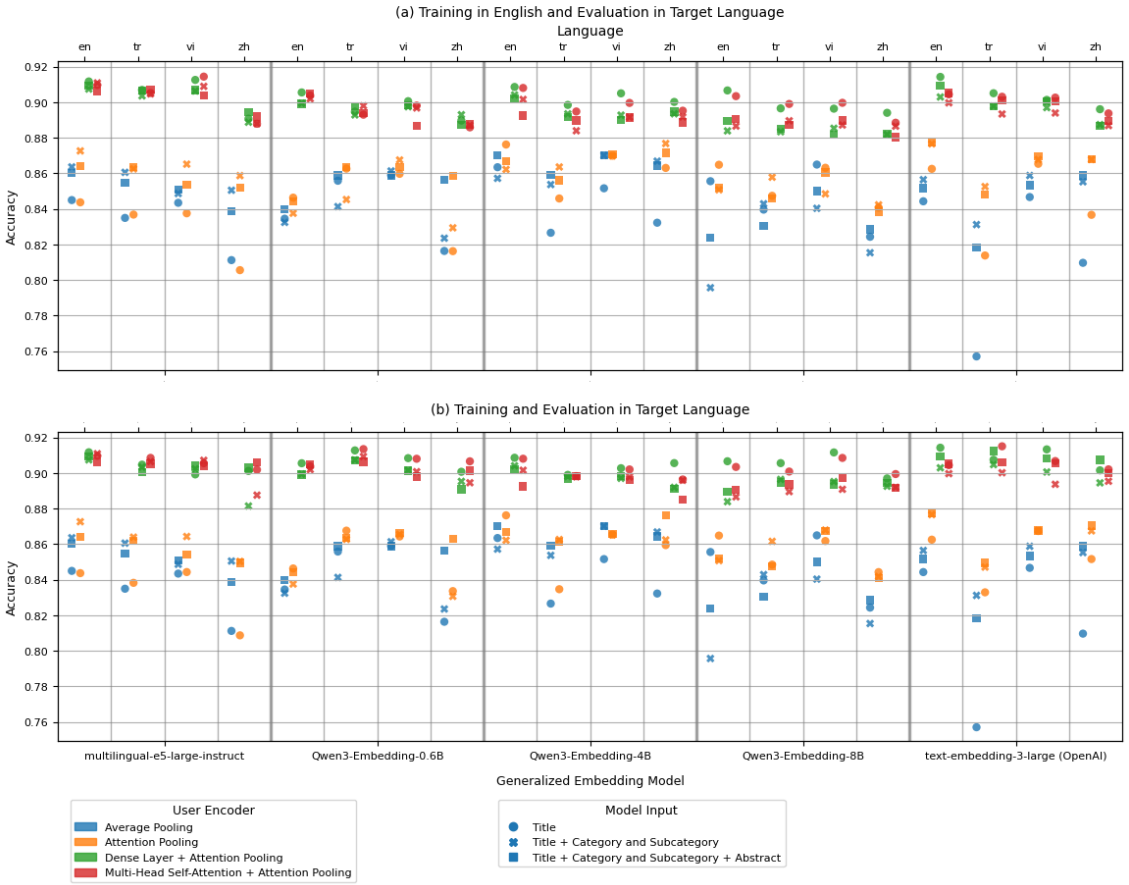

Figure 5.21: Multilingual User-News Classification

5.2.2 Task 2: All User Classification for given News

Figure 5.23 above shows the performance of all the model variations for multilingual classification of all users for given news. The sub-figure (a) shows the results for training in English and evaluation in the target language, and sub-figure (b) shows the results for both training and evaluation in the target language. Figure 5.24 shows the performance for different languages grouped by whether training was done in English only or in the target language. The values can be found in Table A.12, and A.13. From the figure we can see that Chinese has poor performance even when trained in target language in this task. There is slight increase in performance across all the languages when trained on the target

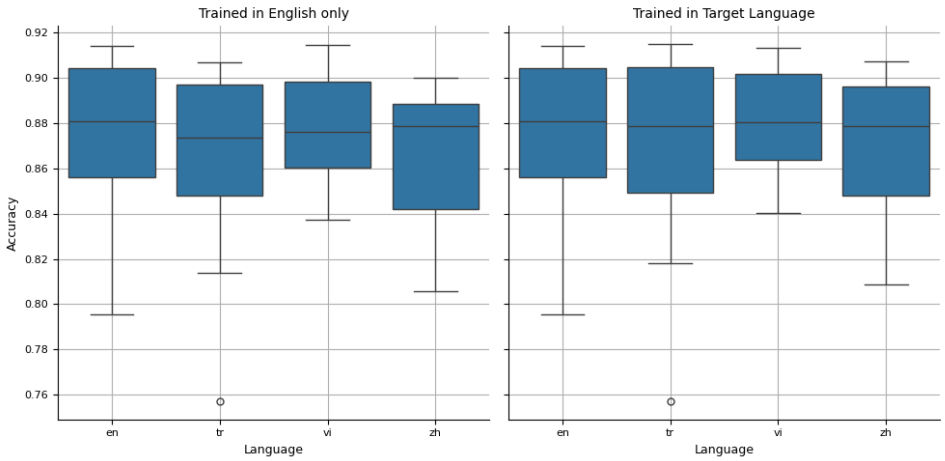

Figure 5.22: Multilingual User-News Classification by training in English only vs in Target Language

language.

5.2.3 Task 3: Candidate News Classification for User

Figure 5.25 above shows the performance of all the model variations for multilingual candidate news classification. The figure 5.25 (a) shows the results for training in English and evaluation in the target language, and figure 5.25 (b) shows the results for both training and evaluation in the target language. Figure 5.26 shows the performance for different languages grouped by whether training was done in English only or in the target language. The values can be found in Table A.14, and A.15. From the figure we can see that for this task Turkish has poor performance even when trained in target language. There is slight increase in performance across all the languages when trained on the target language.

5.2.4 Task 4: Candidate News Ranking for user

Figure 5.27 shows the performance of all the model variations for multilingual candidate news ranking. The figure 5.27 (a) shows the results for training in English and evaluation in the target language, and figure 5.27 (b) shows the results for both training and evaluation in the target language. Figure 5.28 shows the performance for different languages grouped by whether training was done in English only or in the target language. The values can be found in Table A.16, and A.17. From the figure we can see that for this task Chinese has poor performance when trained in English only, but improves significantly when trained on target language. The other languages also has considerable increase in performance when trained in target language compared to training in English only.

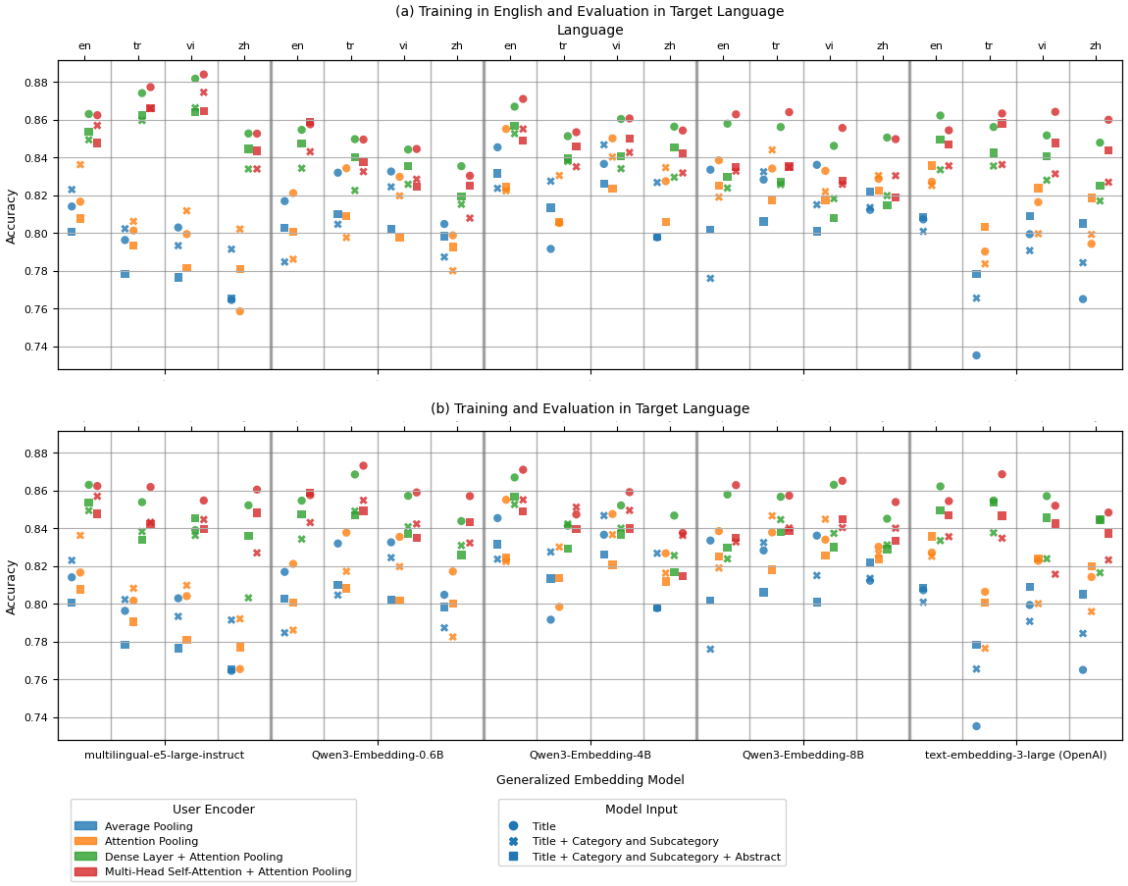

Figure 5.23: Multilingual Classification of all Users for Given News

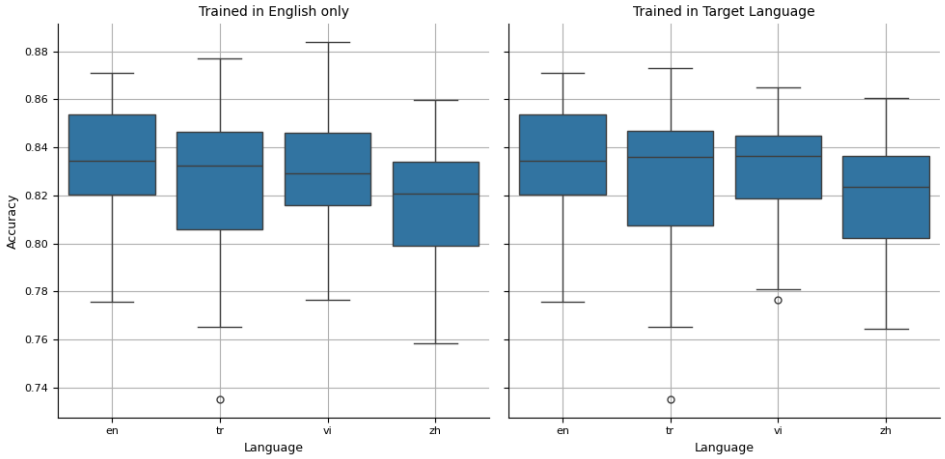

Figure 5.24: Multilingual Classification of all Users for given News by training in English only vs in Target Language

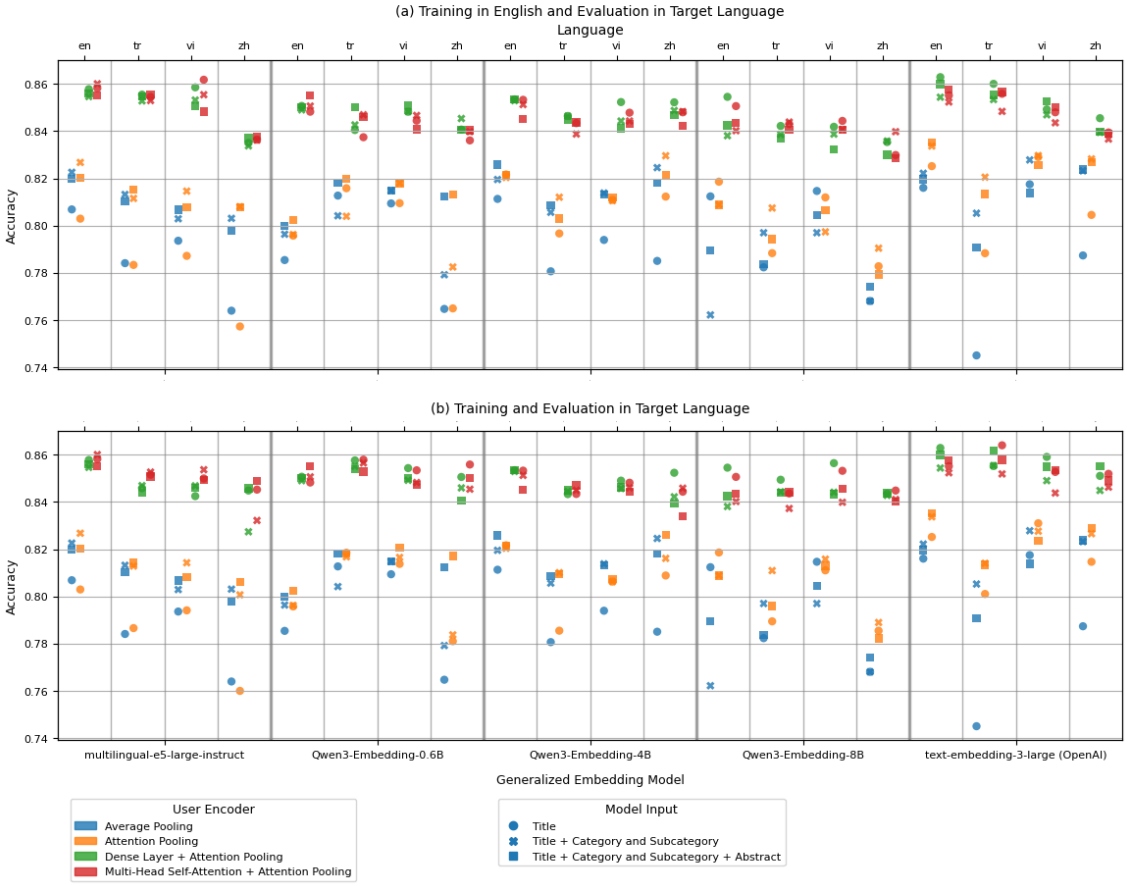

Figure 5.25: Multilingual Candidate News Classification for User

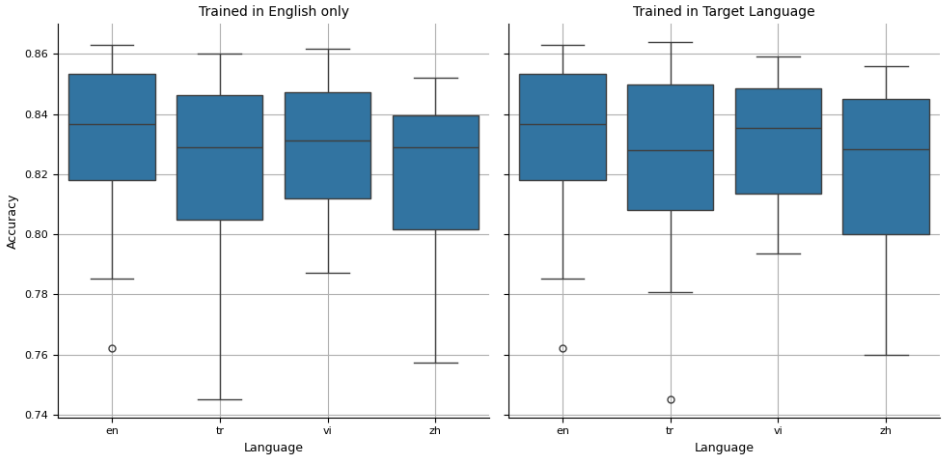

Figure 5.26: Multilingual Candidate News Classification for User by training in English only vs in Target Language

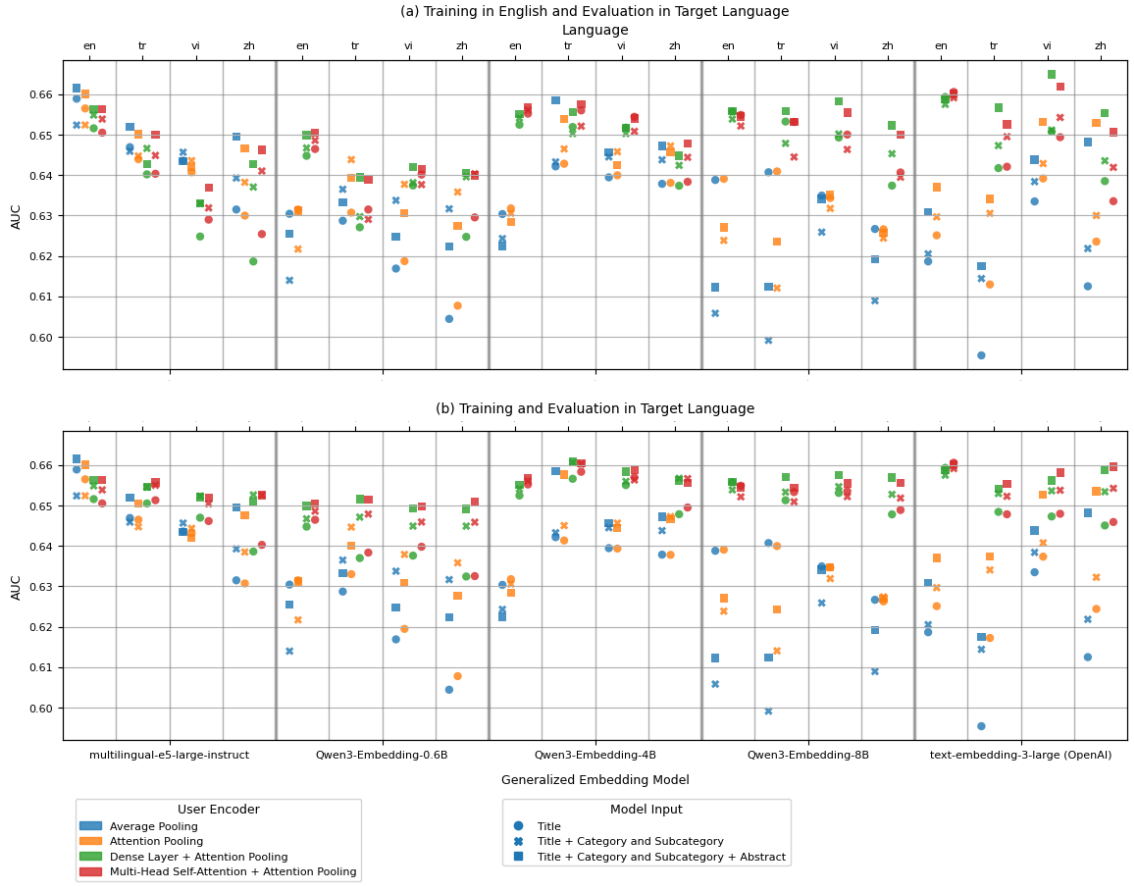

Figure 5.27: Multilingual Candidate News Ranking for User

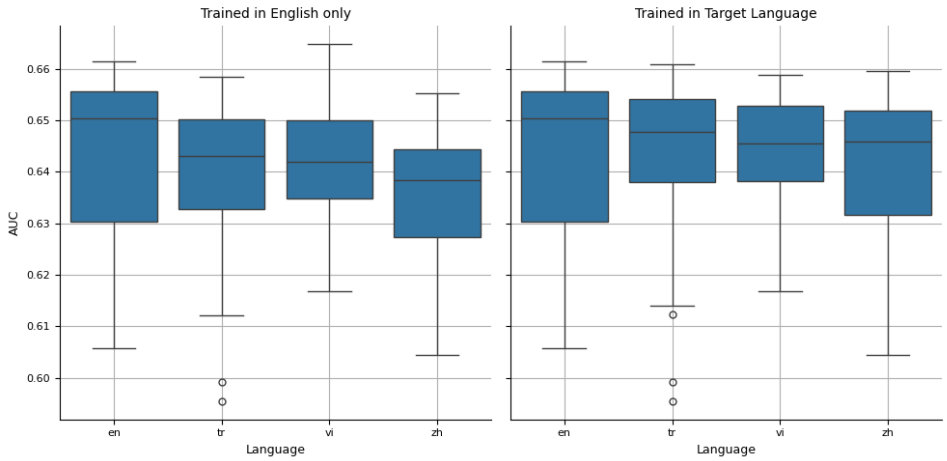

Figure 5.28: Multilingual Candidate News Ranking for User by training in English only vs in Target Language

5.3 Time taken for components

The news articles used in the dataset as well as in any production news recommendation is available beforehand. Therefore, the features that depend on the news only can be pre-computed to save time during inference. All the times are measured in 40 GB A100 GPU.

Time for pre-computed parts

Table 5.1: Time Taken for Embedding for different Model Input (pre-computed)

Generalized Embedding Model	**Model Input**		
	Title	**Title + Category and Subcategory**	**Title + Category and Subcategory + Abstract**
multilingual-e5-large-instruct	20.3 min	23.9 min	28 min
bge-large-en-v1.5	14.2 min	17.4 min	23.4 min
UAE-Large-V1	14.1 min	15.4 min	23.9 min
mxbai-embed-large-v1	14.1 min	15.5 min	22.9 min
GIST-large-Embedding-v0	14.1 min	15.6 min	23.1 min
e5-mistral-7b-instruct	188 min	187.8 min	280.9 min
Qwen3-Embedding-0.6B	18.8 min	20.3 min	33.4 min
Qwen3-Embedding-4B	73.6 min	84.8 min	151.6 min
Qwen3-Embedding-8B	147.1 min	170.4 min	247.1 min
NV-Embed-v2	17 min	18.4 min	34.4 min
text-embedding-3-large (OpenAI)	-	-	-

Table 5.1 gives the time for obtaining the embedding using different generalized embedding model for different model inputs for all the 65,238 news articles. We can see that the large decoder-based models take significantly longer than smaller BERT-based models. Also the time taken for getting embedding for model inputs "Title + Category and Subcategory + Abstract" is significantly longer than for "Title + Category and Subcategory" and "Title". Classifying whether a news article is fake or not can also be pre-computed. Table 5.2 gives the time taken to classify all the 65,238 news articles from the embedding obtained in previous step. The time taken mainly depends on the embedding dimensions. All the BERT-based models and the "Qwen3-Embedding-0.6B" model have a embedding dimension of 1024 and take about 0.24 s. The large decoder-based models have a embedding dimension of 4096 and take about 0.75 s.

Table 5.2: Time Taken for Fake News Classification (pre-computed)

Generalized Embedding Model	**Time Taken**
multilingual-e5-large-instruct	0.25 s

Generalized Embedding Model	Time Taken
bge-large-en-v1.5	0.28 s
UAE-Large-V1	0.25 s
mxbai-embed-large-v1	0.21 s
GIST-large-Embedding-v0	0.22 s
e5-mistral-7b-instruct	0.77 s
Qwen3-Embedding-0.6B	0.24 s
Qwen3-Embedding-4B	0.41 s
Qwen3-Embedding-8B	0.71 s
NV-Embed-v2	0.76 s
text-embedding-3-large (OpenAI)	0.50 s

Time for Online inference

Table 5.3: Time Taken for the User Encoder

Generalized Embedding Model	Average Pooling	Attention Pooling	User Encoder Dense Layer + Attention Pooling	Multi-Head Self-Attention + Attention Pooling
multilingual-e5-large-instruct	17.8 s	18.4 s	18.6 s	20.4 s
bge-large-en-v1.5	17.8 s	18.5 s	18.7 s	20.7 s
UAE-Large-V1	17.9 s	17.8 s	19 s	20.3 s
mxbai-embed-large-v1	17.2 s	17.7 s	18.8 s	20.2 s
GIST-large-Embedding-v0	17.3 s	18 s	18.5 s	20.6 s
e5-mistral-7b-instruct	47 s	49 s	64.8 s	127.6 s
Qwen3-Embedding-0.6B	17.3 s	17.7 s	18.9 s	20.1 s
Qwen3-Embedding-4B	39.6 s	44.1 s	46.8 s	60.9 s
Qwen3-Embedding-8B	47.5 s	48.5 s	64.7 s	127.7 s
NV-Embed-v2	46.8 s	48.4 s	64.5 s	127.4 s
text-embedding-3-large (OpenAI)	39.6 s	46.3 s	52.3 s	81.3 s

The User Encoder component is the main part of the model that cannot be pre-computed and needs online inference. Table 5.3 shows the time taken for different user encoders using embeddings from different generalized embedding models for 73,152 samples in the test set. The time taken increases as we increase the complexity of the user encoder from "Average Pooling" to "Attention Pooling" to "Dense Layer + Attention Pooling" and finally "Multi-Head Self-Attention + Attention Pooling". Similar to the fake news classification component, here also the time increases as the embedding dimension increases.

Chapter 6

Conclusions, Limitations, and Future Work

This chapter summarizes the key findings of the study, and its contributions. It also discusses its limitations and provides direction for future work. Through our experiments we found that using generalized embedding models we can get really good performance, without having to fine-tune any embedding model. The models we have trained are simple, and since the user encoder and news encoder are independent, we can store the user embedding and news embeddings separately, so that at inference time, we only need to find the cosine similarity to find the best news article for the user.

6.1 Summary of Findings

- Our results indicate that generalized embedding models are able to extract useful representation from news articles and our experimental models only slightly underperform compared to state-of-the-art models that fine-tune BERT.
- The results also show that larger decoder-based models, with 7B to 8B parameters do not perform significantly better than smaller BERT-based models with 300M to 600M parameters for news recommendation tasks. Even though the large decoder-based model has significantly higher score on the MTEB benchmark, it doesn't translate to similar gain in news recommendation task.
- For most tasks and embeddings, some transform is needed from the user encoder in order for good performance. Hence, for most cases, the user encoders "Dense Layer + Attention Pooling" and "Multi-Head Self-Attention + Attention Pooling" performs significantly better than "Average Pooling" or "Attention Pooling". Even though "Multi-Head Self-Attention + Attention Pooling" is a bit more complex than "Dense Layer + Attention Pooling", the difference in their performance is not significant.
- We found that the news title is usually enough for getting good performance. For some embedding models adding the news category and abstract as part of the model

input provided slight boost in performance, but in most cases, the gains of adding the extra context is very little.

- For most of the news recommendation tasks the inclusion of fake news classification component gives a modest boost in performance. The component produced a modest performance improvement for news recommendation task 1, a larger improvement for task 2, little or no meaningful improvement for task 3, but a significant drop in performance for task 4.

- News recommendation tasks 1 and 3 show very similar trends in performance across the variations. Task 3 was also similar, with the major difference that using title only as model input give better performance, and that the impact of incorporating fake news classification model was less significant. News recommendation task 4 showed some different trend, with major difference being that incorporating fake news classification model led to drop in performance.

- We also showed that for all the news recommendation tasks, models trained on English only also perform well on other target languages, moreover, when the models are trained and evaluated on the target language, the performance approaches that of training and evaluation in the English language.

- Overall the best performing model across all the news recommendation tasks is using the "text-embedding-3-large (OpenAI)" generalized embedding model using "Title" as the embedding model input with "Dense Layer + Attention Pooling" user encoder and without the fake news classification component.

6.2 Contributions of the Thesis

- Systematic comparison of 11 generalized embedding models (5 BERT-based, 5 decoder based, and 1 proprietary) for use as frozen news encoders.

- Evaluation of performance across four different news recommendation tasks, and identification of best model for each task.

- Analysis of effects of three news input configurations and four different user encoders.

- Investigation into multilingual news recommendation with and without training in target language.

- Analysis of the effect of fake news classification on performance for each of the news recommendation tasks.

6.3 Limitations

- Most news recommendation datasets are primarily made for task 4 of our approach. There is lack of high quality datasets for the other news recommendation tasks.
- There are few high quality fake news classification datasets, and the dataset used four our analysis wasn't high quality. Therefore the fake news classification model may not generalize properly to the MIND dataset.
- There was no multilingual fake news classification dataset available. Hence, we could not incorporate the fake news classification component in our multilingual news recommendation analysis.
- The evaluation was carried out on offline datasets; performance of the approaches in online production systems was not investigated in our study.

6.4 Future Work

- Researchers should develop high quality fake news classification datasets both for English and multilingual use case.
- Further studies should use the rich representation of generalized embedding models and focus on developing better user encoders and making better use of the temporal component of user click history.
- Investigate diversity, fairness, novelty, and user trust in addition to click prediction.
- Further work should also explore different embedding techniques and architectural changes in the embedding models.
- Evaluate the use of generalized embedding models in online production news recommendation.
- Develop datasets for other news recommendation tasks instead of just candidate news ranking.

References

[1] R. Burke, "Hybrid web recommender systems," in The adaptive web: Methods and strategies of web personalization, P. Brusilovsky, A. Kobsa, and W. Nejdl, Eds., Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, pp. 377–408. doi: 10.1007/978-3-540-72079-9_12.

[2] P. Resnick, N. Iacovou, M. Suchak, P. Bergstrom, and J. Riedl, "GroupLens: an open architecture for collaborative filtering of netnews," in Proceedings of the 1994 ACM conference on computer supported cooperative work, in Cscw '94. Chapel Hill, North Carolina, USA and New York, NY, USA: Association for Computing Machinery, 1994, pp. 175–186. doi: 10.1145/192844.192905.

[3] P. Resnick and H. R. Varian, "Recommender systems," Commun. ACM, vol. 40, no. 3, pp. 56–58, Mar. 1997, doi: 10.1145/245108.245121.

[4] G. Linden, B. Smith, and J. York, "Amazon.com recommendations: item-to-item collaborative filtering," IEEE Internet Computing, vol. 7, no. 1, pp. 76–80, Jan. 2003, doi: 10.1109/MIC.2003.1167344.

[5] B. Smith and G. Linden, "Two decades of recommender systems at amazon.com," IEEE Internet Computing, vol. 21, no. 3, pp. 12–18, May 2017, doi: 10.1109/MIC.2017.72.

[6] B. Schwartz, The paradox of choice: Why more is less. in Harper perennial. HarperCollins, 2003. [Online]. Available: `https://books.google.com.bd/books?id=zutxr7rGc_QC`

[7] P. J. Shoemaker, "News and newsworthiness: A commentary," Communications, vol. 31, no. 1, pp. 105–111, 2006, doi: doi:10.1515/COMMUN.2006.007.

[8] G. Adomavicius and A. Tuzhilin, "Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions," IEEE Transactions on Knowledge and Data Engineering, vol. 17, no. 6, pp. 734–749, 2005, doi: 10.1109/TKDE.2005.99.

[9] B. P. Knijnenburg, M. C. Willemsen, Z. Gantner, H. Soncu, and C. Newell, "Explaining the user experience of recommender systems," User Modeling and User-Adapted Interaction, vol. 22, no. 4, pp. 441–504, Oct. 2012, doi: 10.1007/s11257-011-9118-4.

[10] J. A. Konstan and J. Riedl, “Recommender systems: from algorithms to user experience,” User Modeling and User-Adapted Interaction, vol. 22, no. 1, pp. 101–123, Apr. 2012, doi: 10.1007/s11257-011-9112-x.

[11] N. Helberger, “On the democratic role of news recommenders,” Digital Journalism, vol. 7, no. 8, pp. 993–1012, 2019, doi: 10.1080/21670811.2019.1623700.

[12] B. Kille, A. Lommatzsch, F. Hopfgartner, M. Larson, and T. Brodt, “CLEF 2017 NewsREEL Overview: Offline and Online Evaluation of Stream-based News Recommender Systems”.

[13] M. Schedl, H. Zamani, C.-W. Chen, Y. Deldjoo, and M. Elahi, “Current challenges and visions in music recommender systems research,” International Journal of Multimedia Information Retrieval, vol. 7, no. 2, pp. 95–116, Jun. 2018, doi: 10.1007/s13735-018-0154-2.

[14] K. Park, J. Lee, and J. Choi, “Deep neural networks for news recommendations,” in Proceedings of the 2017 ACM on conference on information and knowledge management, in Cikm ’17. Singapore, Singapore and New York, NY, USA: Association for Computing Machinery, 2017, pp. 2255–2258. doi: 10.1145/3132847.3133154.

[15] S. Raza and C. Ding, “A regularized model to trade-off between accuracy and diversity in a news recommender system,” in 2020 IEEE international conference on big data (big data), 2020, pp. 551–560. doi: 10.1109/BigData50022.2020.9378340.

[16] G. Sottocornola, P. Symeonidis, and M. Zanker, “Session-based news recommendations,” in Companion proceedings of the the web conference 2018, in Www ’18. Lyon, France and Republic and Canton of Geneva, CHE: International World Wide Web Conferences Steering Committee, 2018, pp. 1395–1399. doi: 10.1145/3184558.3191582.

[17] D. Doychev, R. Rafter, A. Lawlor, and B. Smyth, “News recommenders: Real-time, real-life experiences,” in User modeling, adaptation and personalization, F. Ricci, K. Bontcheva, O. Conlan, and S. Lawless, Eds., Cham: Springer International Publishing, 2015, pp. 337–342.

[18] M. Trevisiol, L. M. Aiello, R. Schifanella, and A. Jaimes, “Cold-start news recommendation with domain-dependent browse graph,” in Proceedings of the 8th ACM conference on recommender systems, in RecSys ’14. Foster City, Silicon Valley, California, USA and New York, NY, USA: Association for Computing Machinery, 2014, pp. 81–88. doi: 10.1145/2645710.2645726.

[19] I. Ilievski and S. Roy, “Personalized news recommendation based on implicit feedback,” in Proceedings of the 2013 international news recommender systems workshop and challenge, in Nrs ’13. Kowloon, Hong Kong and New York, NY, USA: Association for Computing Machinery, 2013, pp. 10–15. doi: 10.1145/2516641.2516644.

[20] H. Ma, X. Liu, and Z. Shen, “User fatigue in online news recommendation,” in

Proceedings of the 25th international conference on world wide web, in Www '16. Montréal, Québec, Canada and Republic and Canton of Geneva, CHE: International World Wide Web Conferences Steering Committee, 2016, pp. 1363–1372. doi: 10.1145/2872427.2874813.

[21] M. S. Desarkar and N. Shinde, "Diversification in news recommendation for privacy concerned users," in 2014 international conference on data science and advanced analytics (DSAA), 2014, pp. 135–141. doi: 10.1109/DSAA.2014.7058064.

[22] Y. Park, J. Oh, and H. Yu, "RecTime: Real-Time recommender system for online broadcasting," Information Sciences, vol. 409–410, pp. 1–16, 2017, doi: `https://doi.org/10.1016/j.ins.2017.04.038`.

[23] Y. A. Asikin and W. Worndl, "Stories around You: Location-based Serendipitous Recommendation of News Articles".

[24] A. Lommatzsch, B. Kille, and S. Albayrak, "Incorporating context and trends in news recommender systems," in Proceedings of the international conference on web intelligence, in Wi '17. Leipzig, Germany and New York, NY, USA: Association for Computing Machinery, 2017, pp. 1062–1068. doi: 10.1145/3106426.3109433.

[25] A. Cucchiarelli, C. Morbidoni, G. Stilo, and P. Velardi, "What to write and why: a recommender for news media," in Proceedings of the 33rd annual ACM symposium on applied computing, in Sac '18. Pau, France and New York, NY, USA: Association for Computing Machinery, 2018, pp. 1321–1330. doi: 10.1145/3167132.3167274.

[26] C. Beckett and M. Deuze, "On the role of emotion in the future of journalism," Social Media + Society, vol. 2, no. 3, p. 2056305116662395, 2016, doi: 10.1177/2056305116662395.

[27] Z. Wang, K. Hahn, Y. Kim, S. Song, and J.-M. Seo, "A news-topic recommender system based on keywords extraction," Multimedia Tools and Applications, vol. 77, no. 4, pp. 4339–4353, Feb. 2018, doi: 10.1007/s11042-017-5513-0.

[28] S. Raza and C. Ding, "News recommender system: a review of recent progress, challenges, and opportunities," Artif Intell Rev, vol. 55, no. 1, pp. 749–800, 2022, doi: 10.1007/s10462-021-10043-x.

[29] S. Raza and C. Ding, "News recommender system considering temporal dynamics and news taxonomy," in 2019 IEEE international conference on big data (big data), 2019, pp. 920–929. doi: 10.1109/BigData47090.2019.9005459.

[30] N. Gillis, Nonnegative matrix factorization. Philadelphia, PA: Society for Industrial and Applied Mathematics, 2020. doi: 10.1137/1.9781611976410.

[31] E. Frolov and I. Oseledets, "Tensor methods and recommender systems," WIREs Data Mining and Knowledge Discovery, vol. 7, no. 3, p. e1201, 2017, doi: `https://doi.org/10.1002/widm.1201`.

[32] S. Wang, B. Zou, C. Li, K. Zhao, Q. Liu, and H. Chen, "CROWN: a context-aware RecOmmender for web news," in 2015 IEEE 31st international conference on data engineering, 2015, pp. 1420–1423. doi: 10.1109/ICDE.2015.7113391.

[33] R. Salakhutdinov and A. Mnih, "Probabilistic matrix factorization," in Proceedings of the 21st international conference on neural information processing systems, in NIPS'07. Vancouver, British Columbia, Canada and Red Hook, NY, USA: Curran Associates Inc., 2007, pp. 1257–1264.

[34] S. Rendle, C. Freudenthaler, Z. Gantner, and L. Schmidt-Thieme, "BPR: Bayesian personalized ranking from implicit feedback," in Proceedings of the twenty-fifth conference on uncertainty in artificial intelligence, in Uai '09. Montreal, Quebec, Canada and Arlington, Virginia, USA: AUAI Press, 2009, pp. 452–461.

[35] G. K. Dziugaite and D. M. Roy, "Neural network matrix factorization," CoRR, vol. abs/1511.06443, 2015, [Online]. Available: http://arxiv.org/abs/1511.06443

[36] H.-J. Xue, X.-Y. Dai, J. Zhang, S. Huang, and J. Chen, "Deep matrix factorization models for recommender systems," in Proceedings of the 26th international joint conference on artificial intelligence, in IJCAI'17. Melbourne, Australia: AAAI Press, 2017, pp. 3203–3209.

[37] X. He, L. Liao, H. Zhang, L. Nie, X. Hu, and T.-S. Chua, "Neural Collaborative Filtering," in Proceedings of the 26th International Conference on World Wide Web, Perth Australia: International World Wide Web Conferences Steering Committee, Apr. 2017, pp. 173–182. doi: 10.1145/3038912.3052569.

[38] C. Wu, F. Wu, M. An, Y. Huang, and X. Xie, "Neural news recommendation with topic-aware news representation," in Proceedings of the 57th annual meeting of the association for computational linguistics, A. Korhonen, D. Traum, and L. Màrquez, Eds., Florence, Italy: Association for Computational Linguistics, Jul. 2019, pp. 1154–1159. doi: 10.18653/v1/P19-1110.

[39] H. Wang, F. Wu, Z. Liu, and X. Xie, "Fine-grained interest matching for neural news recommendation," in Proceedings of the 58th annual meeting of the association for computational linguistics, D. Jurafsky, J. Chai, N. Schluter, and J. Tetreault, Eds., Online: Association for Computational Linguistics, Jul. 2020, pp. 836–845. doi: 10.18653/v1/2020.acl-main.77.

[40] H. Wang, F. Zhang, X. Xie, and M. Guo, "DKN: Deep knowledge-aware network for news recommendation," in Proceedings of the 2018 world wide web conference, in Www '18. Lyon, France and Republic and Canton of Geneva, CHE: International World Wide Web Conferences Steering Committee, 2018, pp. 1835–1844. doi: 10.1145/3178876.3186175.

[41] R. Wang, S. Wang, W. Lu, and X. Peng, "News recommendation via multi-interest news sequence modelling," in ICASSP 2022 - 2022 IEEE international conference

on acoustics, speech and signal processing (ICASSP), 2022, pp. 7942–7946. doi: 10.1109/ICASSP43922.2022.9747149.

[42] M. An, F. Wu, C. Wu, K. Zhang, Z. Liu, and X. Xie, "Neural news recommendation with long- and short-term user representations," in Proceedings of the 57th annual meeting of the association for computational linguistics, A. Korhonen, D. Traum, and L. Màrquez, Eds., Florence, Italy: Association for Computational Linguistics, Jul. 2019, pp. 336–345. doi: 10.18653/v1/P19-1033.

[43] L. Hu, C. Li, C. Shi, C. Yang, and C. Shao, "Graph neural news recommendation with long-term and short-term interest modeling," Information Processing & Management, vol. 57, no. 2, p. 102142, 2020, doi: https://doi.org/10.1016/j.ipm.2019.102142.

[44] Z. Mao, X. Zeng, and K.-F. Wong, "Neural news recommendation with collaborative news encoding and structural user encoding," in Findings of the association for computational linguistics: EMNLP 2021, M.-F. Moens, X. Huang, L. Specia, and S. W. Yih, Eds., Punta Cana, Dominican Republic: Association for Computational Linguistics, Nov. 2021, pp. 46–55. doi: 10.18653/v1/2021.findings-emnlp.5.

[45] Q. Liu et al., "Multimodal recommender systems: a survey," ACM Comput. Surv., vol. 57, no. 2, Oct. 2024, doi: 10.1145/3695461.

[46] C. Wu, F. Wu, T. Qi, and Y. Huang, "Empowering news recommendation with pre-trained language models," in Proceedings of the 44th international ACM SIGIR conference on research and development in information retrieval, in Sigir '21. Virtual Event, Canada and New York, NY, USA: Association for Computing Machinery, 2021, pp. 1652–1656. doi: 10.1145/3404835.3463069.

[47] A. Vaswani et al., "Attention is all you need," in Advances in neural information processing systems, I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, Eds., Curran Associates, Inc., 2017. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf

[48] I. Saha, S. Puthran, I. Saha, and S. Puthran, "Fake News Detection: A Comprehensive Review and a Novel Framework," in 2024 OITS International Conference on Information Technology (OCIT), 2024, pp. 463–469. doi: 10.1109/OCIT65031.2024.00087.

[49] F. A. Alshuwaier, F. A. Alsulaiman, F. A. Alshuwaier, and F. A. Alsulaiman, "Fake News Detection Using Machine Learning and Deep Learning Algorithms: A Comprehensive Review and Future Perspectives," Computers, vol. 14, no. 9, 2025, doi: 10.3390/computers14090394.

[50] A. Iana et al., "MIND Your Language: A Multilingual Dataset for Cross-Lingual News Recommendation (Extended Abstract)," in KI 2024: Advances in Artificial Intelligence, A. Hotho and S. Rudolph, Eds., Cham: Springer Nature Switzerland,

2024, pp. 335–340.

[51] L. Wu et al., "A survey on large language models for recommendation," World Wide Web, vol. 27, no. 5, p. 60, Aug. 2024, doi: 10.1007/s11280-024-01291-2.

[52] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," in Proceedings of the 2019 conference of the north American chapter of the association for computational linguistics: Human language technologies, volume 1 (long and short papers), J. Burstein, C. Doran, and T. Solorio, Eds., Minneapolis, Minnesota: Association for Computational Linguistics, Jun. 2019, pp. 4171–4186. doi: 10.18653/v1/N19-1423.

[53] T. Wolf et al., "Transformers: State-of-the-art natural language processing," in Proceedings of the 2020 conference on empirical methods in natural language processing: System demonstrations, Q. Liu and D. Schlangen, Eds., Online: Association for Computational Linguistics, Oct. 2020, pp. 38–45. doi: 10.18653/v1/2020.emnlp-demos.6.

[54] Y. Liu et al., "RoBERTa: a robustly optimized BERT pretraining approach," ArXiv, vol. abs/1907.11692, 2019, [Online]. Available: https://api.semanticscholar.org/CorpusID:198953378

[55] M. Lewis et al., "BART: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension," in Annual meeting of the association for computational linguistics, 2019. [Online]. Available: https://api.semanticscholar.org/CorpusID:204960716

[56] Z. Mao, H. Wang, Y. Du, and K.-F. Wong, "UniTRec: a unified text-to-text transformer and joint contrastive learning framework for text-based recommendation," in Annual meeting of the association for computational linguistics, 2023. [Online]. Available: https://api.semanticscholar.org/CorpusID:258888030

[57] Q. Bi, J. Li, L. Shang, X. Jiang, Q. Liu, and H. Yang, "MTRec: Multi-task learning over BERT for news recommendation," in Findings, 2022. [Online]. Available: https://api.semanticscholar.org/CorpusID:248779874

[58] J. Li et al., "MINER: Multi-interest matching network for news recommendation," in Findings, 2022. [Online]. Available: https://api.semanticscholar.org/CorpusID:248780372

[59] Q. Zhang, Q. Jia, C. Wang, J. Li, Z. Wang, and X. He, "AMM: Attentive multi-field matching for news recommendation," Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval, 2021, [Online]. Available: https://api.semanticscholar.org/CorpusID:235792523

[60] C. Wu, F. Wu, Y. Yu, T. Qi, Y. Huang, and Q. Liu, "NewsBERT: Distilling pre-trained language model for intelligent news application," in Conference on empirical methods in natural language processing, 2021. [Online]. Available: https://api.se

manticscholar.org/CorpusID:231855304

[61] Q. Zhang et al., "UNBERT: User-news matching BERT for news recommendation," in International joint conference on artificial intelligence, 2021. [Online]. Available: https://api.semanticscholar.org/CorpusID:237101122

[62] H. Chen et al., "LKPNR: Large language models and knowledge graph for personalized news recommendation framework," Computers, Materials & Continua, 2024, [Online]. Available: https://api.semanticscholar.org/CorpusID:269940497

[63] S. Gao et al., "Generative news recommendation," Proceedings of the ACM Web Conference 2024, 2024, [Online]. Available: https://api.semanticscholar.org/CorpusID:268253020

[64] S. Okura, Y. Tagami, S. Ono, and A. Tajima, "Embedding-based news recommendation for millions of users," in Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining, in Kdd '17. Halifax, NS, Canada and New York, NY, USA: Association for Computing Machinery, 2017, pp. 1933–1942. doi: 10.1145/3097983.3098108.

[65] F. Wu et al., "MIND: a large-scale dataset for news recommendation," in Proceedings of the 58th annual meeting of the association for computational linguistics, D. Jurafsky, J. Chai, N. Schluter, and J. Tetreault, Eds., Online: Association for Computational Linguistics, Jul. 2020, pp. 3597–3606. doi: 10.18653/v1/2020.acl-main.331.

[66] V. Lavrenko, M. Schmill, D. Lawrie, P. Ogilvie, D. Jensen, and J. Allan, "Language models for financial news recommendation," in Proceedings of the ninth international conference on information and knowledge management, in Cikm '00. McLean, Virginia, USA and New York, NY, USA: Association for Computing Machinery, 2000, pp. 389–396. doi: 10.1145/354756.354845.

[67] J. Yang, "Effects of popularity-based news recommendations ('most-viewed') on users' exposure to online news," Media Psychology, vol. 19, no. 2, pp. 243–271, 2016, doi: 10.1080/15213269.2015.1006333.

[68] X. Wang et al., "Dynamic attention deep model for article recommendation by learning human editors' demonstration," in Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining, in Kdd '17. Halifax, NS, Canada and New York, NY, USA: Association for Computing Machinery, 2017, pp. 2051–2059. doi: 10.1145/3097983.3098096.

[69] C. Chen, X. Meng, Z. Xu, and T. Lukasiewicz, "Location-aware personalized news recommendation with deep semantic analysis," IEEE Access, vol. 5, pp. 1624–1638, 2017, doi: 10.1109/ACCESS.2017.2655150.

[70] X. Su and T. M. Khoshgoftaar, "A survey of collaborative filtering techniques," Adv. in Artif. Intell., vol. 2009, Jan. 2009, doi: 10.1155/2009/421425.

[71] P. Lops, M. de Gemmis, and G. Semeraro, "Content-based recommender systems: State of the art and trends," in Recommender systems handbook, F. Ricci, L. Rokach, B. Shapira, and P. B. Kantor, Eds., Boston, MA: Springer US, 2011, pp. 73–105. doi: 10.1007/978-0-387-85820-3_3.

[72] R. Burke, "Hybrid Recommender Systems: Survey and Experiments," User Modeling and User-Adapted Interaction, vol. 12, no. 4, pp. 331–370, Nov. 2002, doi: 10.1023/A:1021240730564.

[73] J. A. Konstan, B. N. Miller, D. Maltz, J. L. Herlocker, L. R. Gordon, and J. Riedl, "GroupLens: applying collaborative filtering to Usenet news," Commun. ACM, vol. 40, no. 3, pp. 77–87, Mar. 1997, doi: 10.1145/245108.245126.

[74] A. S. Das, M. Datar, A. Garg, and S. Rajaram, "Google news personalization: scalable online collaborative filtering," in Proceedings of the 16th international conference on world wide web, in Www '07. Banff, Alberta, Canada and New York, NY, USA: Association for Computing Machinery, 2007, pp. 271–280. doi: 10.1145/1242572.1242610.

[75] K. Lang, "NewsWeeder: Learning to filter netnews," in Machine learning proceedings 1995, A. Prieditis and S. Russell, Eds., San Francisco (CA): Morgan Kaufmann, 1995, pp. 331–339. doi: https://doi.org/10.1016/B978-1-55860-377-6.50048-7.

[76] D. Billsus and M. J. Pazzani, "A hybrid user model for news story classification," in UM99 user modeling, J. Kay, Ed., Vienna: Springer Vienna, 1999, pp. 99–108.

[77] J. Ahn, P. Brusilovsky, J. Grady, D. He, and S. Y. Syn, "Open user profiles for adaptive news systems: help or harm?," in Proceedings of the 16th international conference on world wide web, in Www '07. Banff, Alberta, Canada and New York, NY, USA: Association for Computing Machinery, 2007, pp. 11–20. doi: 10.1145/1242572.1242575.

[78] J. Liu, P. Dolan, and E. R. Pedersen, "Personalized news recommendation based on click behavior," in Proceedings of the 15th international conference on intelligent user interfaces, in Iui '10. Hong Kong, China and New York, NY, USA: Association for Computing Machinery, 2010, pp. 31–40. doi: 10.1145/1719970.1719976.

[79] L. Li, W. Chu, J. Langford, and R. E. Schapire, "A contextual-bandit approach to personalized news article recommendation," in Proceedings of the 19th international conference on world wide web, in Www '10. Raleigh, North Carolina, USA and New York, NY, USA: Association for Computing Machinery, 2010, pp. 661–670. doi: 10.1145/1772690.1772758.

[80] J. Domann and A. Lommatzsch, "A highly available real-time news recommender based on apache spark," in Experimental IR meets multilinguality, multimodality, and interaction, G. J. F. Jones, S. Lawless, J. Gonzalo, L. Kelly, L. Goeuriot, T. Mandl, L. Cappellato, and N. Ferro, Eds., Cham: Springer International Publishing, 2017, pp. 161–172.

[81] A. Gershman, T. Wolfe, E. Fink, and J. G. Carbonell, "News personalization using support vector machines," 2011. [Online]. Available: https://api.semanticscholar.org/CorpusID:11836012

[82] F. Goossen, W. IJntema, F. Frasincar, F. Hogenboom, and U. Kaymak, "News personalization using the CF-IDF semantic recommender," in Proceedings of the international conference on web intelligence, mining and semantics, in Wims '11. Sogndal, Norway and New York, NY, USA: Association for Computing Machinery, 2011. doi: 10.1145/1988688.1988701.

[83] M. Capelle, F. Frasincar, M. Moerland, and F. Hogenboom, "Semantics-based news recommendation," in Proceedings of the 2nd international conference on web intelligence, mining and semantics, in Wims '12. Craiova, Romania and New York, NY, USA: Association for Computing Machinery, 2012. doi: 10.1145/2254129.2254163.

[84] M. Moerland, F. Hogenboom, M. Capelle, and F. Frasincar, "Semantics-based news recommendation with SF-IDF+," in Proceedings of the 3rd international conference on web intelligence, mining and semantics, in Wims '13. Madrid, Spain and New York, NY, USA: Association for Computing Machinery, 2013. doi: 10.1145/2479787.2479795.

[85] H. J. Lee and S. J. Park, "MONERS: A news recommender for the mobile web," Expert Systems with Applications, vol. 32, no. 1, pp. 143–150, 2007, doi: https://doi.org/10.1016/j.eswa.2005.11.010.

[86] L. Li, D. Wang, T. Li, D. Knox, and B. Padmanabhan, "SCENE: a scalable two-stage personalized news recommendation system," in Proceedings of the 34th international ACM SIGIR conference on research and development in information retrieval, in Sigir '11. Beijing, China and New York, NY, USA: Association for Computing Machinery, 2011, pp. 125–134. doi: 10.1145/2009916.2009937.

[87] A. H. Parizi and M. Kazemifard, "Emotional news recommender system," in 2015 sixth international conference of cognitive science (ICCS), Apr. 2015, pp. 37–41. doi: 10.1109/COGSCI.2015.7426666.

[88] F. Garcin, K. Zhou, B. Faltings, and V. Schickel, "Personalized news recommendation based on collaborative filtering," in 2012 IEEE/WIC/ACM international conferences on web intelligence and intelligent agent technology, 2012, pp. 437–441. doi: 10.1109/WI-IAT.2012.95.

[89] A. H. Parizi, M. Kazemifard, and M. Asghari, "Emonews: An emotional news recommender system," vol. 14, pp. 392–402, Dec. 2016.

[90] M. Tavakolifard, J. A. Gulla, K. C. Almeroth, J. E. Ingvaldesn, G. Nygreen, and E. Berg, "Tailored news in the palm of your hand: a multi-perspective transparent approach to news recommendation," in Proceedings of the 22nd international conference on world wide web, in WWW '13 companion. Rio de Janeiro, Brazil and

New York, NY, USA: Association for Computing Machinery, 2013, pp. 305–308. doi: 10.1145/2487788.2487930.

[91] G. Ji, S. He, L. Xu, K. Liu, and J. Zhao, "Knowledge graph embedding via dynamic mapping matrix," in Proceedings of the 53rd annual meeting of the association for computational linguistics and the 7th international joint conference on natural language processing (volume 1: Long papers), C. Zong and M. Strube, Eds., Beijing, China: Association for Computational Linguistics, Jul. 2015, pp. 687–696. doi: 10.3115/v1/P15-1067.

[92] N. Jonnalagedda and S. Gauch, "Personalized news recommendation using twitter," in 2013 IEEE/WIC/ACM international joint conferences on web intelligence (WI) and intelligent agent technologies (IAT), 2013, pp. 21–25. doi: 10.1109/WI-IAT.2013.144.

[93] W. Chu and S.-T. Park, "Personalized recommendation on dynamic content using predictive bilinear models," in Proceedings of the 18th international conference on world wide web, in Www '09. Madrid, Spain and New York, NY, USA: Association for Computing Machinery, 2009, pp. 691–700. doi: 10.1145/1526709.1526802.

[94] K. F. Yeung and Y. Yang, "A proactive personalized mobile news recommendation system," in 2010 developments in e-systems engineering, 2010, pp. 207–212. doi: 10.1109/DeSE.2010.40.

[95] A. Patankar, J. Bose, and H. Khanna, "A bias aware news recommendation system," in 2019 IEEE 13th international conference on semantic computing (ICSC), 2019, pp. 232–238. doi: 10.1109/ICOSC.2019.8665610.

[96] A. Lommatzsch, "Real-time news recommendation using context-aware ensembles," in Advances in information retrieval, M. de Rijke, T. Kenter, A. P. de Vries, C. Zhai, F. de Jong, K. Radinsky, and K. Hofmann, Eds., Cham: Springer International Publishing, 2014, pp. 51–62.

[97] M. Claypool, A. Gokhale, T. Miranda, P. Murnikov, D. Netes, and M. M. Sartin, "Combining content-based and collaborative filters in an online newspaper," in Annual international ACM SIGIR conference on research and development in information retrieval, 1999. [Online]. Available: `https://api.semanticscholar.org/CorpusID:14001779`

[98] G. S. S. Saranya.k.g, "A personalized online news recommendation system," International Journal of Computer Applications, vol. 57, no. 18, pp. 6–14, Nov. 2012, doi: 10.5120/9212-3758.

[99] A. Darvishy, H. Ibrahim, F. Sidi, and A. Mustapha, "HYPNER: a hybrid approach for personalized news recommendation," IEEE Access, vol. 8, pp. 46877–46894, 2020, doi: 10.1109/ACCESS.2020.2978505.

[100] H. Wen, L. Fang, and L. Guan, "A hybrid approach for personalized recommendation

of news on the Web," Expert Systems with Applications, vol. 39, no. 5, pp. 5806–5814, Apr. 2012, doi: 10.1016/j.eswa.2011.11.087.

[101] I. Cantador, P. Castells, and A. Bellogín, "An enhanced semantic layer for hybrid recommender systems: Application to news recommendation," Int. J. Semant. Web Inf. Syst., vol. 7, no. 1, pp. 44–78, Jan. 2011, doi: 10.4018/jswis.2011010103.

[102] K. Zhang, X. Xin, P. Luo, and P. Guot, "Fine-grained news recommendation by fusing matrix factorization, topic analysis and knowledge graph representation," in 2017 IEEE international conference on systems, man, and cybernetics (SMC), 2017, pp. 918–923. doi: 10.1109/SMC.2017.8122727.

[103] V. Kumar, D. Khattar, S. Gupta, M. Gupta, and V. Varma, "User profiling based deep neural network for temporal news recommendation," in 2017 IEEE international conference on data mining workshops (ICDMW), 2017, pp. 765–772. doi: 10.1109/ICDMW.2017.106.

[104] P.-S. Huang, X. He, J. Gao, L. Deng, A. Acero, and L. Heck, "Learning deep structured semantic models for web search using clickthrough data," in Proceedings of the 22nd ACM international conference on information & knowledge management, in Cikm '13. San Francisco, California, USA and New York, NY, USA: Association for Computing Machinery, 2013, pp. 2333–2338. doi: 10.1145/2505515.2505665.

[105] Q. Le and T. Mikolov, "Distributed representations of sentences and documents," in Proceedings of the 31st international conference on international conference on machine learning - volume 32, in ICML'14. Beijing, China: JMLR.org, 2014, pp. II–1188–II–1196.

[106] V. Kumar, D. Khattar, S. Gupta, and V. Varma, "Word semantics based 3-D convolutional neural networks for news recommendation," in 2017 IEEE international conference on data mining workshops (ICDMW), 2017, pp. 761–764. doi: 10.1109/ICDMW.2017.105.

[107] Y. Xiao, "News recommendation system based on user interest and deep network," Procedia Computer Science, vol. 243, pp. 1105–1114, 2024, doi: `https://doi.org/10.1016/j.procs.2024.09.131`.

[108] P. Zhang, Z. Dou, and J. Yao, "Learning to select historical news articles for interaction based neural news recommendation," ArXiv, vol. abs/2110.06459, 2021, [Online]. Available: `https://api.semanticscholar.org/CorpusID:238744243`

[109] C. Wu, F. Wu, S. Ge, T. Qi, Y. Huang, and X. Xie, "Neural news recommendation with multi-head self-attention," in Proceedings of the 2019 conference on empirical methods in natural language processing and the 9th international joint conference on natural language processing (EMNLP-IJCNLP), K. Inui, J. Jiang, V. Ng, and X. Wan, Eds., Hong Kong, China: Association for Computational Linguistics, Nov. 2019, pp. 6389–6394. doi: 10.18653/v1/D19-1671.

[110] D. Khattar, V. Kumar, V. Varma, and M. Gupta, "Weave&rec: a word embedding based 3-D convolutional network for news recommendation," in Proceedings of the 27th ACM international conference on information and knowledge management, in Cikm '18. Torino, Italy and New York, NY, USA: Association for Computing Machinery, 2018, pp. 1855–1858. doi: 10.1145/3269206.3269307.

[111] C. Wu, F. Wu, M. An, J. Huang, Y. Huang, and X. Xie, "NPA: Neural News Recommendation with Personalized Attention," in Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, Anchorage AK USA: ACM, Jul. 2019, pp. 2576–2584. doi: 10.1145/3292500.3330665.

[112] C. Wu, F. Wu, T. Qi, and Y. Huang, "SentiRec: Sentiment diversity-aware neural news recommendation," in Proceedings of the 1st conference of the asia-pacific chapter of the association for computational linguistics and the 10th international joint conference on natural language processing, K.-F. Wong, K. Knight, and H. Wu, Eds., Suzhou, China: Association for Computational Linguistics, Dec. 2020, pp. 44–53. doi: 10.18653/v1/2020.aacl-main.6.

[113] C. Wu, F. Wu, T. Qi, and Y. Huang, "User modeling with click preference and reading satisfaction for news recommendation," in Proceedings of the twenty-ninth international joint conference on artificial intelligence, in IJCAI'20. Yokohama, Yokohama, Japan, 2021.

[114] C. Wu et al., "FeedRec: News Feed Recommendation with Various User Feedbacks," in Proceedings of the ACM Web Conference 2022, Virtual Event, Lyon France: ACM, Apr. 2022, pp. 2088–2097. doi: 10.1145/3485447.3512082.

[115] C. Wu, F. Wu, T. Qi, and Y. Huang, "Two birds with one stone: Unified model learning for both recall and ranking in news recommendation," in Findings of the association for computational linguistics: ACL 2022, S. Muresan, P. Nakov, and A. Villavicencio, Eds., Dublin, Ireland: Association for Computational Linguistics, May 2022, pp. 3474–3480. doi: 10.18653/v1/2022.findings-acl.274.

[116] C. Wu, F. Wu, X. Wang, Y. Huang, and X. Xie, "FairRec: Fairness-aware News Recommendation with Decomposed Adversarial Learning," Apr. 15, 2021, arXiv: arXiv:2006.16742. doi: 10.48550/arXiv.2006.16742.

[117] C. Wu, F. Wu, Y. Huang, and X. Xie, "Neural news recommendation with negative feedback," CCF Transactions on Pervasive Computing and Interaction, vol. 2, no. 3, pp. 178–188, Oct. 2020, doi: 10.1007/s42486-020-00044-0.

[118] T. Qi, F. Wu, C. Wu, Y. Huang, and X. Xie, "Privacy-preserving news recommendation model learning," in Findings of the association for computational linguistics: EMNLP 2020, T. Cohn, Y. He, and Y. Liu, Eds., Online: Association for Computational Linguistics, Nov. 2020, pp. 1423–1432. doi: 10.18653/v1/2020.findings-emnlp.128.

[119] J. Pennington, R. Socher, and C. Manning, "GloVe: Global vectors for word representation," in Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP), A. Moschitti, B. Pang, and W. Daelemans, Eds., Doha, Qatar: Association for Computational Linguistics, Oct. 2014, pp. 1532–1543. doi: 10.3115/v1/D14-1162.

[120] S. Shi et al., "WG4Rec: Modeling textual content with word graph for news recommendation," in Proceedings of the 30th ACM international conference on information & knowledge management, in Cikm '21. Virtual Event, Queensland, Australia and New York, NY, USA: Association for Computing Machinery, 2021, pp. 1651–1660. doi: 10.1145/3459637.3482401.

[121] Z. Ji, M. Wu, J. Liu, and J. E. Armendáriz Íñigo, "Attention-based graph neural network for news recommendation," in 2021 international joint conference on neural networks (IJCNN), 2021, pp. 1–8. doi: 10.1109/IJCNN52387.2021.9534339.

[122] J. Gao et al., "Fine-grained deep knowledge-aware network for news recommendation with self-attention," in 2018 IEEE/WIC/ACM international conference on web intelligence (WI), 2018, pp. 81–88. doi: 10.1109/WI.2018.0-104.

[123] Q. Zhu, X. Zhou, Z. Song, J. Tan, and L. Guo, "DAN: deep attention neural network for news recommendation," in Proceedings of the thirty-third AAAI conference on artificial intelligence and thirty-first innovative applications of artificial intelligence conference and ninth AAAI symposium on educational advances in artificial intelligence, in AAAI'19/IAAI'19/EAAI'19. Honolulu, Hawaii, USA: AAAI Press, 2019. doi: 10.1609/aaai.v33i01.33015973.

[124] Q. Chu, G. Liu, H. Sun, and C. Zhou, "Next news recommendation via knowledge-aware sequential model," in Chinese computational linguistics: 18th china national conference, CCL 2019, kunming, china, october 18–20, 2019, proceedings, Kunming, China and Berlin, Heidelberg: Springer-Verlag, 2019, pp. 221–232. doi: 10.1007/978-3-030-32381-3_18.

[125] H. Zhang, X. Chen, and S. Ma, "Dynamic news recommendation with hierarchical attention network," in 2019 IEEE international conference on data mining (ICDM), 2019, pp. 1456–1461. doi: 10.1109/ICDM.2019.00190.

[126] T. Qi et al., "HieRec: Hierarchical user interest modeling for personalized news recommendation," in Proceedings of the 59th annual meeting of the association for computational linguistics and the 11th international joint conference on natural language processing (volume 1: Long papers), C. Zong, F. Xia, W. Li, and R. Navigli, Eds., Online: Association for Computational Linguistics, Aug. 2021, pp. 5446–5456. doi: 10.18653/v1/2021.acl-long.423.

[127] D. Liu et al., "News graph: An enhanced knowledge graph for news recommendation," in KaRS@CIKM, 2019. [Online]. Available: `https://api.semanticscholar.org/`

CorpusID:204777685

[128] H.-S. Sheu and S. Li, "Context-aware graph embedding for session-based news recommendation," in Proceedings of the 14th ACM conference on recommender systems, in RecSys '20. Virtual Event, Brazil and New York, NY, USA: Association for Computing Machinery, 2020, pp. 657–662. doi: 10.1145/3383313.3418477.

[129] Y. Sun et al., "A hybrid approach to news recommendation based on knowledge graph and long short-term user preferences," in 2021 IEEE international conference on services computing (SCC), 2021, pp. 165–173. doi: 10.1109/SCC53864.2021.00029.

[130] D. H. Tran et al., "Deep news recommendation with contextual user profiling and multifaceted article representation," in Web information systems engineering – WISE 2021, W. Zhang, L. Zou, Z. Maamar, and L. Chen, Eds., Cham: Springer International Publishing, 2021, pp. 237–251.

[131] A. Bordes, N. Usunier, A. Garcia-Durán, J. Weston, and O. Yakhnenko, "Translating embeddings for modeling multi-relational data," in Proceedings of the 27th international conference on neural information processing systems - volume 2, in NIPS'13. Lake Tahoe, Nevada and Red Hook, NY, USA: Curran Associates Inc., 2013, pp. 2787–2795.

[132] D. Lee, B. Oh, S. Seo, and K.-H. Lee, "News recommendation with topic-enriched knowledge graphs," in Proceedings of the 29th ACM international conference on information & knowledge management, in Cikm '20. Virtual Event, Ireland and New York, NY, USA: Association for Computing Machinery, 2020, pp. 695–704. doi: 10.1145/3340531.3411932.

[133] D. Liu et al., "KRED: Knowledge-aware document representation for news recommendations," in Proceedings of the 14th ACM conference on recommender systems, in RecSys '20. Virtual Event, Brazil and New York, NY, USA: Association for Computing Machinery, 2020, pp. 200–209. doi: 10.1145/3383313.3412237.

[134] T. Qi, F. Wu, C. Wu, and Y. Huang, "Personalized news recommendation with knowledge-aware interactive matching," in Proceedings of the 44th international ACM SIGIR conference on research and development in information retrieval, in Sigir '21. Virtual Event, Canada and New York, NY, USA: Association for Computing Machinery, 2021, pp. 61–70. doi: 10.1145/3404835.3462861.

[135] Y. Tian et al., "Joint knowledge pruning and recurrent graph convolution for news recommendation," in Proceedings of the 44th international ACM SIGIR conference on research and development in information retrieval, in Sigir '21. Virtual Event, Canada and New York, NY, USA: Association for Computing Machinery, 2021, pp. 51–60. doi: 10.1145/3404835.3462912.

[136] S. Han, H. Huang, and J. Liu, "Neural news recommendation with event extraction," 2021. [Online]. Available: https://arxiv.org/abs/2111.05068

[137] C. Wu, F. Wu, M. An, J. Huang, Y. Huang, and X. Xie, "Neural news recommendation with attentive multi-view learning," in Proceedings of the 28th international joint conference on artificial intelligence, in IJCAI'19. Macao, China: AAAI Press, 2019, pp. 3863–3869.

[138] L. Zhang, P. Liu, and J. A. Gulla, "A deep joint network for session-based news recommendations with contextual augmentation," in Proceedings of the 29th on hypertext and social media, in Ht '18. Baltimore, MD, USA and New York, NY, USA: Association for Computing Machinery, 2018, pp. 201–209. doi: 10.1145/3209542.3209557.

[139] Y. Kim, "Convolutional neural networks for sentence classification," in Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP), A. Moschitti, B. Pang, and W. Daelemans, Eds., Doha, Qatar: Association for Computational Linguistics, Oct. 2014, pp. 1746–1751. doi: 10.3115/v1/D14-1181.

[140] C. Wu et al., "Neural news recommendation with heterogeneous user behavior," in Proceedings of the 2019 conference on empirical methods in natural language processing and the 9th international joint conference on natural language processing (EMNLP-IJCNLP), K. Inui, J. Jiang, V. Ng, and X. Wan, Eds., Hong Kong, China: Association for Computational Linguistics, Nov. 2019, pp. 4874–4883. doi: 10.18653/v1/D19-1493.

[141] J. Yi, F. Wu, C. Wu, Q. Li, G. Sun, and X. Xie, "DebiasedRec: Bias-aware user modeling and click prediction for personalized news recommendation," 2021. [Online]. Available: `https://arxiv.org/abs/2104.07360`

[142] C. Wu, F. Wu, Y. Huang, and X. Xie, "User-as-graph: User modeling with heterogeneous graph pooling for news recommendation," in Proceedings of the thirtieth international joint conference on artificial intelligence, IJCAI-21, Z.-H. Zhou, Ed., International Joint Conferences on Artificial Intelligence Organization, Aug. 2021, pp. 1624–1630. doi: 10.24963/ijcai.2021/224.

[143] X. Zhang, Q. Yang, and D. Xu, "Combining explicit entity graph with implicit text information for news recommendation," in Companion proceedings of the web conference 2021, in Www '21. Ljubljana, Slovenia and New York, NY, USA: Association for Computing Machinery, 2021, pp. 412–416. doi: 10.1145/3442442.3452329.

[144] G. de Souza Pereira Moreira, F. Ferreira, and A. M. da Cunha, "News session-based recommendations using deep neural networks," in Proceedings of the 3rd workshop on deep learning for recommender systems, in Dlrs 2018. Vancouver, BC, Canada and New York, NY, USA: Association for Computing Machinery, 2018, pp. 15–23. doi: 10.1145/3270323.3270328.

[145] G. D. S. P. Moreira, D. Jannach, and A. M. D. Cunha, "Contextual hybrid session-based news recommendation with recurrent neural networks," IEEE Access, vol. 7, pp. 169185–169203, 2019, doi: 10.1109/ACCESS.2019.2954957.

[146] C. Hutto and E. Gilbert, "VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text," ICWSM, vol. 8, no. 1, pp. 216–225, May 2014, doi: 10.1609/icwsm.v8i1.14550.

[147] C. Wu, F. Wu, T. Qi, C. Zhang, Y. Huang, and T. Xu, "MM-rec: Visiolinguistic model empowered multimodal news recommendation," in Proceedings of the 45th international ACM SIGIR conference on research and development in information retrieval, in Sigir '22. Madrid, Spain and New York, NY, USA: Association for Computing Machinery, 2022, pp. 2560–2564. doi: 10.1145/3477495.3531896.

[148] J. Lu, D. Batra, D. Parikh, and S. Lee, "ViLBERT: pretraining task-agnostic visiolinguistic representations for vision-and-language tasks," in Proceedings of the 33rd international conference on neural information processing systems, Red Hook, NY, USA: Curran Associates Inc., 2019.

[149] J. Xun et al., "Why do we click: Visual impression-aware news recommendation," in Proceedings of the 29th ACM international conference on multimedia, in Mm '21. Virtual Event, China and New York, NY, USA: Association for Computing Machinery, 2021, pp. 3881–3890. doi: 10.1145/3474085.3475514.

[150] T. Qi, F. Wu, C. Wu, and Y. Huang, "PP-rec: News recommendation with personalized user interest and time-aware news popularity," in Proceedings of the 59th annual meeting of the association for computational linguistics and the 11th international joint conference on natural language processing (volume 1: Long papers), C. Zong, F. Xia, W. Li, and R. Navigli, Eds., Online: Association for Computational Linguistics, Aug. 2021, pp. 5457–5467. doi: 10.18653/v1/2021.acl-long.424.

[151] Z. Ji, M. Wu, H. Yang, and J. E. A. Íñigo, "Temporal sensitive heterogeneous graph neural network for news recommendation," Future Generation Computer Systems, vol. 125, pp. 324–333, 2021, doi: https://doi.org/10.1016/j.future.2021.06.007.

[152] L. Meng, C. Shi, S. Hao, and X. Su, "DCAN: Deep co-attention network by modeling user preference and news lifecycle for news recommendation," in Database systems for advanced applications: 26th international conference, DASFAA 2021, taipei, taiwan, april 11–14, 2021, proceedings, part III, Taipei, Taiwan and Berlin, Heidelberg: Springer-Verlag, 2021, pp. 100–114. doi: 10.1007/978-3-030-73200-4_7.

[153] S. Cho, H. Lim, K. Park, S. Yoo, and E. Park, "On the overlooked significance of underutilized contextual features in recent news recommendation models," 2021. [Online]. Available: https://arxiv.org/abs/2112.14370

[154] Y. Qian et al., "Interaction graph neural network for news recommendation," in Web information systems engineering – WISE 2019: 20th international conference, hong kong, china, january 19–22, 2020, proceedings, Hong Kong, China and Berlin, Heidelberg: Springer-Verlag, 2020, pp. 599–614. doi: 10.1007/978-3-030-34223-4_38.

[155] S. Ge, C. Wu, F. Wu, T. Qi, and Y. Huang, "Graph enhanced representation learning

for news recommendation," in Proceedings of the web conference 2020, in Www '20. Taipei, Taiwan and New York, NY, USA: Association for Computing Machinery, 2020, pp. 2863–2869. doi: 10.1145/3366423.3380050.

[156] T. Y. S. S. Santosh, A. Saha, and N. Ganguly, "MVL: Multi-view learning for news recommendation," in Proceedings of the 43rd international ACM SIGIR conference on research and development in information retrieval, in Sigir '20. Virtual Event, China and New York, NY, USA: Association for Computing Machinery, 2020, pp. 1873–1876. doi: 10.1145/3397271.3401294.

[157] L. Hu et al., "Graph neural news recommendation with unsupervised preference disentanglement," in Proceedings of the 58th annual meeting of the association for computational linguistics, D. Jurafsky, J. Chai, N. Schluter, and J. Tetreault, Eds., Online: Association for Computational Linguistics, Jul. 2020, pp. 4255–4264. doi: 10.18653/v1/2020.acl-main.392.

[158] M. Ma, S. Na, H. Wang, C. Chen, and J. Xu, "The graph-based behavior-aware recommendation for interactive news," Applied Intelligence, vol. 52, no. 2, pp. 1913–1929, Jan. 2022, doi: 10.1007/s10489-021-02497-x.

[159] Q. Jia, J. Li, Q. Zhang, X. He, and J. Zhu, "RMBERT: News recommendation via recurrent reasoning memory network over BERT," in Proceedings of the 44th international ACM SIGIR conference on research and development in information retrieval, in Sigir '21. Virtual Event, Canada and New York, NY, USA: Association for Computing Machinery, 2021, pp. 1773–1777. doi: 10.1145/3404835.3463234.

[160] S. Raza and C. Ding, "Deep neural network to tradeoff between accuracy and diversity in a news recommender system," in 2021 IEEE international conference on big data (big data), 2021, pp. 5246–5256. doi: 10.1109/BigData52589.2021.9671467.

[161] A. Iana, G. Glavaš, and H. Paulheim, "Train once, use flexibly: a modular framework for multi-aspect neural news recommendation," in Findings of the association for computational linguistics: EMNLP 2024, Y. Al-Onaizan, M. Bansal, and Y.-N. Chen, Eds., Miami, Florida, USA: Association for Computational Linguistics, Nov. 2024, pp. 9555–9571. doi: 10.18653/v1/2024.findings-emnlp.558.

[162] Y. Yada and H. Yamana, "News recommendation with category description by a large language model," 2024. [Online]. Available: `https://arxiv.org/abs/2405.13007`

[163] Y. Noh, Y.-H. Oh, and S.-B. Park, "A location-based personalized news recommendation," in 2014 international conference on big data and smart computing (BIGCOMP), 2014, pp. 99–104. doi: 10.1109/BIGCOMP.2014.6741416.

[164] X. Yi, L. Hong, E. Zhong, N. N. Liu, and S. Rajan, "Beyond clicks: dwell time for personalization," in Proceedings of the 8th ACM conference on recommender systems, in RecSys '14. Foster City, Silicon Valley, California, USA and New York, NY, USA: Association for Computing Machinery, 2014, pp. 113–120. doi: 10.1145/2645710.2645724.

[165] A. Gharahighehi, C. Vens, and K. Pliakos, "Multi-stakeholder news recommendation using hypergraph learning," in ECML PKDD 2020 workshops, I. Koprinska, M. Kamp, A. Appice, C. Loglisci, L. Antonie, A. Zimmermann, R. Guidotti, Ö. Özgöbek, R. P. Ribeiro, R. Gavaldà, J. Gama, L. Adilova, Y. Krishnamurthy, P. M. Ferreira, D. Malerba, I. Medeiros, M. Ceci, G. Manco, E. Masciari, Z. W. Ras, P. Christen, E. Ntoutsi, E. Schubert, A. Zimek, A. Monreale, P. Biecek, S. Rinzivillo, B. Kille, A. Lommatzsch, and J. A. Gulla, Eds., Cham: Springer International Publishing, 2020, pp. 531–535.

[166] L. Li and T. Li, "News recommendation via hypergraph learning: encapsulation of user behavior and news content," in Proceedings of the sixth ACM international conference on web search and data mining, in Wsdm '13. Rome, Italy and New York, NY, USA: Association for Computing Machinery, 2013, pp. 305–314. doi: 10.1145/2433396.2433436.

[167] F. Garcin, C. Dimitrakakis, and B. Faltings, "Personalized news recommendation with context trees," in Proceedings of the 7th ACM conference on recommender systems, in RecSys '13. Hong Kong, China and New York, NY, USA: Association for Computing Machinery, 2013, pp. 105–112. doi: 10.1145/2507157.2507166.

[168] K. Joseph and H. Jiang, "Content based news recommendation via shortest entity distance over knowledge graphs," in Companion proceedings of the 2019 world wide web conference, in Www '19. San Francisco, USA and New York, NY, USA: Association for Computing Machinery, 2019, pp. 690–699. doi: 10.1145/3308560.3317703.

[169] L. Li, L. Zheng, F. Yang, and T. Li, "Modeling and broadening temporal user interest in personalized news recommendation," Expert Syst. Appl., vol. 41, no. 7, pp. 3168–3177, Jun. 2014, doi: 10.1016/j.eswa.2013.11.020.

[170] D. Billsus and M. J. Pazzani, "User modeling for adaptive news access," User Modeling and User-Adapted Interaction, vol. 10, pp. 147–180, 2000, [Online]. Available: https://api.semanticscholar.org/CorpusID:3363124

[171] L. Li, L. Zheng, and T. Li, "LOGO: a long-short user interest integration in personalized news recommendation," in Proceedings of the fifth ACM conference on recommender systems, in RecSys '11. Chicago, Illinois, USA and New York, NY, USA: Association for Computing Machinery, 2011, pp. 317–320. doi: 10.1145/2043932.2043992.

[172] P. Viana and M. Soares, "A hybrid approach for personalized news recommendation in a mobility scenario using long-short user interest," Int. J. Artif. Intell. Tools, vol. 26, pp. 1760012:1-1760012:29, 2017, [Online]. Available: https://api.semanticscholar.org/CorpusID:31019129

[173] V. Kumar, D. Khattar, S. Gupta, M. Gupta, and V. Varma, "Deep neural architecture for news recommendation," in Conference and labs of the evaluation forum, 2017. [Online]. Available: https://api.semanticscholar.org/CorpusID:27220979

[174] C. Wu, F. Wu, T. Qi, C. Li, and Y. Huang, "Is news recommendation a sequential recommendation task?," in Proceedings of the 45th international ACM SIGIR conference on research and development in information retrieval, in Sigir '22. Madrid, Spain and New York, NY, USA: Association for Computing Machinery, 2022, pp. 2382–2386. doi: 10.1145/3477495.3531862.

[175] J. Fu et al., "Recurrent memory reasoning network for expert finding in community question answering," in Proceedings of the 13th international conference on web search and data mining, in Wsdm '20. Houston, TX, USA and New York, NY, USA: Association for Computing Machinery, 2020, pp. 187–195. doi: 10.1145/3336191.3371817.

[176] L. Zhang, P. Liu, and J. A. Gulla, "Dynamic attention-integrated neural network for session-based news recommendation," Machine Learning, vol. 108, no. 10, pp. 1851–1875, Oct. 2019, doi: 10.1007/s10994-018-05777-9.

[177] S. Aphiwongsophon, P. Chongstitvatana, S. Aphiwongsophon, and P. Chongstitvatana, "Detecting Fake News with Machine Learning Method," in 2018 15th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), 2018, pp. 528–531. doi: 10.1109/ECTICon.2018.8620051.

[178] D. Singh et al., "Fake News Detection Using Ensemble Learning Models," in Proceedings of Data Analytics and Management, A. Swaroop, Z. Polkowski, S. D. Correia, and B. Virdee, Eds., Singapore: Springer Nature Singapore, 2023, pp. 53–66.

[179] B. Ni et al., "Improving Generalizability of Fake News Detection Methods using Propensity Score Matching." 2020. [Online]. Available: `https://arxiv.org/abs/2002.00838`

[180] A. Albahr, M. Albahar, A. Albahr, and M. Albahar, "An Empirical Comparison of Fake News Detection using different Machine Learning Algorithms," International Journal of Advanced Computer Science and Applications, vol. 11, no. 9, 2020, doi: 10.14569/IJACSA.2020.0110917.

[181] D. Mouratidis, A. Kanavos, K. Kermanidis, D. Mouratidis, A. Kanavos, and K. Kermanidis, "From Misinformation to Insight: Machine Learning Strategies for Fake News Detection," Information, vol. 16, no. 3, 2025, doi: 10.3390/info16030189.

[182] J. Kruse et al., "EB-NeRD a large-scale dataset for news recommendation," in Proceedings of the Recommender Systems Challenge 2024, in RecSysChallenge '24. New York, NY, USA: Association for Computing Machinery, 2024, pp. 1–11. doi: 10.1145/3687151.3687152.

[183] H. Touvron et al., "LLaMA: Open and efficient foundation language models," 2023. [Online]. Available: `https://arxiv.org/abs/2302.13971`

[184] A. Q. Jiang et al., "Mistral 7B," 2023. [Online]. Available: `https://arxiv.org/ab`

s/2310.06825

[185] J. Bai et al., "Qwen technical report," 2023. [Online]. Available: https://arxiv.org/abs/2309.16609

[186] A. Conneau et al., "Unsupervised cross-lingual representation learning at scale," 2020. [Online]. Available: https://arxiv.org/abs/1911.02116

[187] N. Reimers and I. Gurevych, "Sentence-BERT: Sentence embeddings using siamese BERT-networks," 2019. [Online]. Available: https://arxiv.org/abs/1908.10084

[188] T. Gao, X. Yao, and D. Chen, "SimCSE: Simple contrastive learning of sentence embeddings," in Proceedings of the 2021 conference on empirical methods in natural language processing, M.-F. Moens, X. Huang, L. Specia, and S. W. Yih, Eds., Online and Punta Cana, Dominican Republic: Association for Computational Linguistics, Nov. 2021, pp. 6894–6910. doi: 10.18653/v1/2021.emnlp-main.552.

[189] L. Wang et al., "Text embeddings by weakly-supervised contrastive pre-training," arXiv, Dec. 2022. [Online]. Available: https://www.microsoft.com/en-us/research/publication/text-embeddings-by-weakly-supervised-contrastive-pre-training/

[190] T. Nguyen et al., "MS MARCO: a human generated machine reading comprehension dataset," Nov. 2016, [Online]. Available: https://www.microsoft.com/en-us/research/publication/ms-marco-human-generated-machine-reading-comprehension-dataset/

[191] T. Kwiatkowski et al., "Natural questions: a benchmark for question answering research," Transactions of the Association of Computational Linguistics, 2019.

[192] L. Wang et al., "Multilingual E5 text embeddings: a technical report," Microsoft, MSR-TR-2024-45, Feb. 2024. [Online]. Available: https://www.microsoft.com/en-us/research/publication/multilingual-e5-text-embeddings-a-technical-report/

[193] J. Chen, S. Xiao, P. Zhang, K. Luo, D. Lian, and Z. Liu, "M3-embedding: Multi-linguality, multi-functionality, multi-granularity text embeddings through self-knowledge distillation," in Findings of the association for computational linguistics: ACL 2024, L.-W. Ku, A. Martins, and V. Srikumar, Eds., Bangkok, Thailand: Association for Computational Linguistics, Aug. 2024, pp. 2318–2335. doi: 10.18653/v1/2024.findings-acl.137.

[194] T. B. Brown et al., "Language models are few-shot learners," in Proceedings of the 34th international conference on neural information processing systems, in Nips '20. Vancouver, BC, Canada and Red Hook, NY, USA: Curran Associates Inc., 2020.

[195] C. Lee et al., "NV-embed: Improved techniques for training llms as generalist embedding models," 2025. [Online]. Available: https://arxiv.org/abs/2405.174

28

[196] P. BehnamGhader, V. Adlakha, M. Mosbach, D. Bahdanau, N. Chapados, and S. Reddy, "LLM2Vec: Large language models are secretly powerful text encoders," 2024. [Online]. Available: https://arxiv.org/abs/2404.05961

[197] N. Muennighoff, "SGPT: GPT sentence embeddings for semantic search," 2022. [Online]. Available: https://arxiv.org/abs/2202.08904

[198] L. Wang et al., "Improving text embeddings with large language models," in Proceedings of the 62nd annual meeting of the association for computational linguistics (volume 1: Long papers), L.-W. Ku, A. Martins, and V. Srikumar, Eds., Bangkok, Thailand: Association for Computational Linguistics, Aug. 2024, pp. 11897–11916. doi: 10.18653/v1/2024.acl-long.642.

[199] Y. Zhang et al., "Qwen3 embedding: Advancing text embedding and reranking through foundation models," 2025. [Online]. Available: https://arxiv.org/abs/2506.05176

[200] A. Yang et al., "Qwen3 technical report," 2025. [Online]. Available: https://arxiv.org/abs/2505.09388

[201] S. Xiao, Z. Liu, P. Zhang, and N. Muennighoff, "C-pack: Packaged resources to advance general chinese embedding," 2023.

[202] X. Li and J. Li, "AoE: Angle-optimized embeddings for semantic textual similarity," in Proceedings of the 62nd annual meeting of the association for computational linguistics (volume 1: Long papers), L.-W. Ku, A. Martins, and V. Srikumar, Eds., Bangkok, Thailand: Association for Computational Linguistics, Aug. 2024, pp. 1825–1839. doi: 10.18653/v1/2024.acl-long.101.

[203] S. Lee, A. Shakir, D. Koenig, and J. Lipp, "Open source strikes bread - new fluffy embeddings model." [Online]. Available: https://www.mixedbread.ai/blog/mxbai-embed-large-v1

[204] A. V. Solatorio, "GISTEmbed: Guided in-sample selection of training negatives for text embedding fine-tuning," 2024. [Online]. Available: https://arxiv.org/abs/2402.16829

[205] N. Muennighoff, N. Tazi, L. Magne, and N. Reimers, "MTEB: Massive text embedding benchmark," in Proceedings of the 17th conference of the european chapter of the association for computational linguistics, A. Vlachos and I. Augenstein, Eds., Dubrovnik, Croatia: Association for Computational Linguistics, May 2023, pp. 2014–2037. doi: 10.18653/v1/2023.eacl-main.148.

[206] I. Loshchilov and F. Hutter, "Decoupled weight decay regularization," 2019. [Online]. Available: https://arxiv.org/abs/1711.05101

[207] H. Ahmed, I. Traore, S. Saad, H. Ahmed, I. Traore, and S. Saad, "Detecting opinion

spams and fake news using text classification," SECURITY AND PRIVACY, vol. 1, no. 1, p. e9, 2018, doi: https://doi.org/10.1002/spy2.9.

[208] S. Rendle, "Factorization machines with libFM," ACM Trans. Intell. Syst. Technol., vol. 3, no. 3, May 2012, doi: 10.1145/2168752.2168771.

[209] H. Guo, R. TANG, Y. Ye, Z. Li, and X. He, "DeepFM: a factorization-machine based neural network for CTR prediction," in Proceedings of the twenty-sixth international joint conference on artificial intelligence, IJCAI-17, 2017, pp. 1725–1731. doi: 10.24963/ijcai.2017/239.

Appendix

A.1 Main Result

A.1.1 Task 1: User-News Classification

The following table gives the experimental results for Task 1 without the fake news classification component.

Table A.1: Performance of Models on User-News Classification without Fake News Classification Component

Generalized Embedding Model	User Encoder	Model Input	F1 Score	AUC	Accuracy	Precision	Recall
multilingual-e5-large-instruct	Average Pooling	Title	0.1409	0.6461	0.8450	0.0909	0.3128
		Title + Category and Subcategory	0.1425	0.6477	0.8636	0.0957	0.2788
		Title + Category and Subcategory + Abstract	0.1482	0.6552	0.8604	0.0986	0.2989
	Attention Pooling	Title	0.1451	0.6484	0.8437	0.0933	0.3264
		Title + Category and Subcategory	0.1530	0.6526	0.8727	0.1049	0.2831
		Title + Category and Subcategory + Abstract	0.1554	0.6581	0.8641	0.1039	0.3075
	Dense Layer + Attention Pooling	Title	0.1530	0.6684	0.9117	0.1254	0.1961
		Title + Category and Subcategory	0.1558	0.6697	0.9075	0.1238	0.2102
		Title + Category and Subcategory + Abstract	0.1577	0.6731	0.9095	0.1268	0.2083
	Multi-Head Self-Attention + Attention Pooling	Title	0.1514	0.6601	0.9098	0.1225	0.1980
		Title + Category and Subcategory	0.1512	0.6619	0.9109	0.1234	0.1953
		Title + Category and Subcategory + Abstract	0.1532	0.6651	0.9063	0.1210	0.2086
bge-large-en-v1.5	Average Pooling	Title	0.1356	0.6438	0.8245	0.0848	0.3387
		Title + Category and Subcategory	0.1395	0.6478	0.8268	0.0874	0.3454
		Title + Category and Subcategory + Abstract	0.1119	0.5845	0.8727	0.0781	0.1974
	Attention Pooling	Title	0.1367	0.6441	0.8319	0.0864	0.3274
		Title + Category and Subcategory	0.1419	0.6499	0.8363	0.0902	0.3333
		Title + Category and Subcategory + Abstract	0.1275	0.5981	0.8697	0.0876	0.2344
	Dense Layer + Attention Pooling	Title	0.1577	0.6699	0.8978	0.1186	0.2356
		Title + Category and Subcategory	0.1554	0.6708	0.8811	0.1092	0.2693
		Title + Category and Subcategory + Abstract	0.1527	0.6691	0.8940	0.1131	0.2349
	Multi-Head Self-Attention + Attention Pooling	Title	0.1562	0.6667	0.8936	0.1152	0.2423
		Title + Category and Subcategory	0.1534	0.6706	0.8895	0.1114	0.2463
		Title + Category and Subcategory + Abstract	0.1490	0.6653	0.9022	0.1153	0.2107
UAE-Large-V1	Average Pooling	Title	0.1474	0.6531	0.8616	0.0983	0.2945
		Title + Category and Subcategory	0.1401	0.6482	0.8685	0.0954	0.2636
		Title + Category and Subcategory + Abstract	0.1415	0.6393	0.8652	0.0954	0.2734
	Attention Pooling	Title	0.1524	0.6562	0.8614	0.1014	0.3066
		Title + Category and Subcategory	0.1500	0.6532	0.8664	0.1011	0.2900

Generalized Embedding Model	User Encoder	Model Input	F1 Score	AUC	Accuracy	Precision	Recall
UAE-Large-V1	Attention Pooling	Title + Category and Subcategory + Abstract	0.1500	0.6456	0.8556	0.0985	0.3134
	Dense Layer + Attention Pooling	Title	0.1566	0.6753	0.8984	0.1182	0.2321
		Title + Category and Subcategory	0.1552	0.6745	0.8975	0.1167	0.2317
		Title + Category and Subcategory + Abstract	0.1560	0.6781	0.9018	0.1199	0.2232
	Multi-Head Self-Attention + Attention Pooling	Title	0.1555	0.6704	0.9054	0.1220	0.2143
		Title + Category and Subcategory	0.1542	0.6727	0.8969	0.1156	0.2311
		Title + Category and Subcategory + Abstract	0.1552	0.6757	0.8966	0.1162	0.2336
mxbai-embed-large-v1	Average Pooling	Title	0.1492	0.6590	0.8605	0.0992	0.3011
		Title + Category and Subcategory	0.1420	0.6534	0.8634	0.0953	0.2781
		Title + Category and Subcategory + Abstract	0.1445	0.6486	0.8555	0.0951	0.3002
	Attention Pooling	Title	0.1552	0.6611	0.8664	0.1044	0.3019
		Title + Category and Subcategory	0.1492	0.6574	0.8579	0.0986	0.3066
		Title + Category and Subcategory + Abstract	0.1509	0.6536	0.8511	0.0982	0.3257
	Dense Layer + Attention Pooling	Title	0.1562	0.6741	0.8986	0.1180	0.2309
		Title + Category and Subcategory	0.1545	0.6733	0.8968	0.1158	0.2320
		Title + Category and Subcategory + Abstract	0.1558	0.6787	0.8991	0.1180	0.2292
	Multi-Head Self-Attention + Attention Pooling	Title	0.1557	0.6704	0.8952	0.1157	0.2377
		Title + Category and Subcategory	0.1532	0.6727	0.8975	0.1153	0.2282
		Title + Category and Subcategory + Abstract	0.1542	0.6759	0.8979	0.1162	0.2289
GIST-large-Embedding-v0	Average Pooling	Title	0.1467	0.6526	0.8722	0.1007	0.2704
		Title + Category and Subcategory	0.1409	0.6483	0.8751	0.0978	0.2522
		Title + Category and Subcategory + Abstract	0.1408	0.6444	0.8618	0.0942	0.2786
	Attention Pooling	Title	0.1520	0.6570	0.8746	0.1048	0.2766
		Title + Category and Subcategory	0.1518	0.6551	0.8707	0.1035	0.2848
		Title + Category and Subcategory + Abstract	0.1503	0.6510	0.8598	0.0997	0.3052
	Dense Layer + Attention Pooling	Title	0.1571	0.6738	0.8985	0.1186	0.2327
		Title + Category and Subcategory	0.1549	0.6722	0.8956	0.1154	0.2353
		Title + Category and Subcategory + Abstract	0.1563	0.6784	0.9032	0.1210	0.2205
	Multi-Head Self-Attention + Attention Pooling	Title	0.1565	0.6721	0.9032	0.1211	0.2210
		Title + Category and Subcategory	0.1537	0.6731	0.8992	0.1166	0.2252
		Title + Category and Subcategory + Abstract	0.1557	0.6757	0.9018	0.1196	0.2229
e5-mistral-7b-instruct	Average Pooling	Title	0.1247	0.6228	0.8847	0.0902	0.2022

Generalized Embedding Model	User Encoder	Model Input	F1 Score	AUC	Accuracy	Precision	Recall
e5-mistral-7b-instruct	Average Pooling	Title	0.1342	0.6348	0.8870	0.0974	0.2154
		Title + Category and Subcategory	0.1318	0.6268	0.8743	0.0916	0.2348
		Title + Category and Subcategory + Abstract					
	Attention Pooling	Title	0.1302	0.6261	0.8984	0.0998	0.1872
		Title + Category and Subcategory	0.1505	0.6410	0.8893	0.1093	0.2413
		Title + Category and Subcategory + Abstract	0.1442	0.6345	0.8773	0.1006	0.2543
	Dense Layer + Attention Pooling	Title	0.1599	0.6815	0.9082	0.1273	0.2152
		Title + Category and Subcategory	0.1594	0.6798	0.9082	0.1269	0.2141
		Title + Category and Subcategory + Abstract	0.1588	0.6820	0.9046	0.1237	0.2215
	Multi-Head Self-Attention + Attention Pooling	Title	0.1584	0.6754	0.9074	0.1256	0.2144
		Title + Category and Subcategory	0.1589	0.6745	0.9022	0.1221	0.2272
		Title + Category and Subcategory + Abstract	0.1561	0.6766	0.9084	0.1248	0.2084
Qwen3-Embedding-0.6B	Average Pooling	Title	0.1340	0.6364	0.8345	0.0851	0.3152
		Title + Category and Subcategory	0.1197	0.6145	0.8324	0.0761	0.2804
		Title + Category and Subcategory + Abstract	0.1256	0.6261	0.8398	0.0807	0.2832
	Attention Pooling	Title	0.1391	0.6394	0.8464	0.0900	0.3054
		Title + Category and Subcategory	0.1267	0.6260	0.8376	0.0811	0.2900
		Title + Category and Subcategory + Abstract	0.1307	0.6324	0.8442	0.0845	0.2882
	Dense Layer + Attention Pooling	Title	0.1369	0.6641	0.9056	0.1089	0.1843
		Title + Category and Subcategory	0.1429	0.6665	0.8995	0.1093	0.2062
		Title + Category and Subcategory + Abstract	0.1430	0.6637	0.8991	0.1092	0.2070
	Multi-Head Self-Attention + Attention Pooling	Title	0.1379	0.6633	0.9038	0.1085	0.1893
		Title + Category and Subcategory	0.1406	0.6673	0.9020	0.1092	0.1973
		Title + Category and Subcategory + Abstract	0.1397	0.6634	0.9050	0.1105	0.1899
Qwen3-Embedding-4B	Average Pooling	Title	0.1400	0.6425	0.8635	0.0941	0.2736
		Title + Category and Subcategory	0.1200	0.6202	0.8572	0.0801	0.2396
		Title + Category and Subcategory + Abstract	0.1265	0.6258	0.8701	0.0870	0.2314
	Attention Pooling	Title	0.1462	0.6442	0.8762	0.1016	0.2607
		Title + Category and Subcategory	0.1318	0.6311	0.8622	0.0886	0.2574
		Title + Category and Subcategory + Abstract	0.1345	0.6331	0.8670	0.0914	0.2545
	Dense Layer + Attention Pooling	Title	0.1475	0.6676	0.9086	0.1188	0.1945
		Title + Category and Subcategory	0.1505	0.6698	0.9041	0.1176	0.2090
		Title + Category and Subcategory + Abstract	0.1475	0.6664	0.9019	0.1140	0.2088

Generalized Embedding Model	User Encoder	Model Input	F1 Score	AUC	Accuracy	Precision	Recall
Qwen3-Embedding-4B	Multi-Head Self-Attention + Attention Pooling	Title	0.1480	0.6695	0.9081	0.1187	0.1965
		Title + Category and Subcategory	0.1508	0.6690	0.9016	0.1162	0.2150
		Title + Category and Subcategory + Abstract	0.1488	0.6656	0.8926	0.1097	0.2310
Qwen3-Embedding-8B	Average Pooling	Title	0.1368	0.6423	0.8556	0.0903	0.2815
		Title + Category and Subcategory	0.1184	0.6115	0.7957	0.0718	0.3376
		Title + Category and Subcategory + Abstract	0.1196	0.6160	0.8240	0.0751	0.2943
	Attention Pooling	Title	0.1411	0.6448	0.8648	0.0951	0.2733
		Title + Category and Subcategory	0.1320	0.6319	0.8508	0.0864	0.2792
		Title + Category and Subcategory + Abstract	0.1340	0.6346	0.8521	0.0879	0.2816
	Dense Layer + Attention Pooling	Title	0.1466	0.6721	0.9067	0.1166	0.1972
		Title + Category and Subcategory	0.1476	0.6617	0.8839	0.1052	0.2472
		Title + Category and Subcategory + Abstract	0.1522	0.6678	0.8896	0.1106	0.2439
	Multi-Head Self-Attention + Attention Pooling	Title	0.1489	0.6717	0.9035	0.1160	0.2077
		Title + Category and Subcategory	0.1465	0.6609	0.8866	0.1056	0.2395
		Title + Category and Subcategory + Abstract	0.1516	0.6656	0.8906	0.1107	0.2406
NV-Embed-v2	Average Pooling	Title	0.1095	0.5943	0.8450	0.0714	0.2344
		Title + Category and Subcategory	0.1117	0.5995	0.8686	0.0770	0.2034
		Title + Category and Subcategory + Abstract	0.1193	0.6060	0.8520	0.0787	0.2466
	Attention Pooling	Title	0.1271	0.6120	0.8820	0.0908	0.2114
		Title + Category and Subcategory	0.1320	0.6175	0.8813	0.0939	0.2220
		Title + Category and Subcategory + Abstract	0.1361	0.6201	0.8924	0.1010	0.2086
	Dense Layer + Attention Pooling	Title	0.1535	0.6681	0.9005	0.1173	0.2219
		Title + Category and Subcategory	0.1524	0.6701	0.9011	0.1169	0.2187
		Title + Category and Subcategory + Abstract	0.1539	0.6725	0.8917	0.1127	0.2424
	Multi-Head Self-Attention + Attention Pooling	Title	0.1500	0.6616	0.9021	0.1159	0.2124
		Title + Category and Subcategory	0.1504	0.6616	0.8934	0.1112	0.2323
		Title + Category and Subcategory + Abstract	0.1543	0.6688	0.8925	0.1134	0.2413
text-embedding-3-large (OpenAI)	Average Pooling	Title	0.1117	0.6029	0.8443	0.0727	0.2409
		Title + Category and Subcategory	0.1220	0.6162	0.8565	0.0812	0.2454
		Title + Category and Subcategory + Abstract	0.1252	0.6148	0.8516	0.0823	0.2613
	Attention Pooling	Title	0.1218	0.6128	0.8625	0.0822	0.2345
		Title + Category and Subcategory	0.1350	0.6290	0.8766	0.0944	0.2369

Generalized Embedding Model	User Encoder	Model Input	F1 Score	AUC	Accuracy	Precision	Recall
text-embedding-3-large (OpenAI)	Dense Layer + Attention Pooling	Title + Category and Subcategory + Abstract	0.1368	0.6257	0.8775	0.0959	0.2389
		Title	0.1520	0.6721	0.9143	0.1271	0.1890
		Title + Category and Subcategory	0.1547	0.6696	0.9030	0.1197	0.2184
		Title + Category and Subcategory + Abstract	0.1532	0.6712	0.9095	0.1236	0.2015
	Multi-Head Self-Attention + Attention Pooling	Title	0.1537	0.6699	0.9045	0.1201	0.2135
		Title + Category and Subcategory	0.1528	0.6659	0.8997	0.1163	0.2225
		Title + Category and Subcategory + Abstract	0.1527	0.6681	0.9052	0.1199	0.2103

The following table gives the experimental results for Task 1 with the fake news classification component.

Table A.2: Performance of Models on User-News Classification with Fake News Classification Component

Generalized Embedding Model	User Encoder	Model Input	F1 Score	AUC	Accuracy	Precision	Recall
multilingual-e5-large-instruct	Average Pooling	Title	0.1266	0.5842	0.8711	0.0873	0.2298
		Title + Category and Subcategory	0.1245	0.5852	0.8870	0.0909	0.1976
		Title + Category and Subcategory + Abstract	0.1321	0.5875	0.8750	0.0920	0.2340
	Attention Pooling	Title	0.1320	0.5858	0.8755	0.0921	0.2329
		Title + Category and Subcategory	0.1340	0.5880	0.8801	0.0949	0.2283
		Title + Category and Subcategory + Abstract	0.1377	0.5894	0.8773	0.0964	0.2412
	Dense Layer + Attention Pooling	Title	0.1407	0.5925	0.9081	0.1135	0.1853
		Title + Category and Subcategory	0.1421	0.5936	0.9041	0.1116	0.1955
		Title + Category and Subcategory + Abstract	0.1418	0.5949	0.9149	0.1202	0.1730
	Multi-Head Self-Attention + Attention Pooling	Title	0.1343	0.5883	0.9149	0.1144	0.1624
		Title + Category and Subcategory	0.1365	0.5900	0.9086	0.1108	0.1777
		Title + Category and Subcategory + Abstract	0.1378	0.5914	0.9101	0.1129	0.1768
bge-large-en-v1.5	Average Pooling	Title	0.1412	0.6164	0.8671	0.0957	0.2688
		Title + Category and Subcategory	0.1360	0.6118	0.8545	0.0896	0.2819
		Title + Category and Subcategory + Abstract	0.1084	0.5819	0.8580	0.0728	0.2125
	Attention Pooling	Title	0.1414	0.6166	0.8603	0.0942	0.2831
		Title + Category and Subcategory	0.1395	0.6139	0.8575	0.0924	0.2843
		Title + Category and Subcategory + Abstract	0.1163	0.5887	0.8327	0.0740	0.2709

Generalized Embedding Model	User Encoder	Model Input	F1 Score	AUC	Accuracy	Precision	Recall
bge-large-en-v1.5	Dense Layer + Attention Pooling	Title	0.1529	0.6213	0.8994	0.1162	0.2233
		Title + Category and Subcategory	0.1507	0.6212	0.8908	0.1102	0.2383
		Title + Category and Subcategory + Abstract	0.1494	0.6209	0.8966	0.1122	0.2234
	Multi-Head Self-Attention + Attention Pooling	Title	0.1497	0.6190	0.8942	0.1111	0.2291
		Title + Category and Subcategory	0.1487	0.6204	0.8911	0.1090	0.2341
		Title + Category and Subcategory + Abstract	0.1472	0.6187	0.9018	0.1137	0.2087
UAE-Large-V1	Average Pooling	Title	0.1478	0.6558	0.8630	0.0989	0.2923
		Title + Category and Subcategory	0.1406	0.6517	0.8702	0.0962	0.2612
		Title + Category and Subcategory + Abstract	0.1426	0.6448	0.8599	0.0949	0.2868
	Attention Pooling	Title	0.1528	0.6589	0.8629	0.1020	0.3044
		Title + Category and Subcategory	0.1503	0.6564	0.8680	0.1018	0.2873
		Title + Category and Subcategory + Abstract	0.1510	0.6503	0.8581	0.0997	0.3104
	Dense Layer + Attention Pooling	Title	0.1570	0.6761	0.8996	0.1191	0.2300
		Title + Category and Subcategory	0.1554	0.6756	0.8989	0.1176	0.2289
		Title + Category and Subcategory + Abstract	0.1563	0.6790	0.9030	0.1209	0.2211
	Multi-Head Self-Attention + Attention Pooling	Title	0.1557	0.6714	0.9064	0.1229	0.2124
		Title + Category and Subcategory	0.1544	0.6740	0.8984	0.1166	0.2284
		Title + Category and Subcategory + Abstract	0.1556	0.6769	0.8980	0.1172	0.2312
mxbai-embed-large-v1	Average Pooling	Title	0.1490	0.6592	0.8574	0.0983	0.3073
		Title + Category and Subcategory	0.1421	0.6537	0.8647	0.0957	0.2758
		Title + Category and Subcategory + Abstract	0.1448	0.6498	0.8568	0.0956	0.2984
	Attention Pooling	Title	0.1553	0.6615	0.8663	0.1045	0.3024
		Title + Category and Subcategory	0.1493	0.6578	0.8576	0.0986	0.3074
		Title + Category and Subcategory + Abstract	0.1513	0.6547	0.8525	0.0987	0.3235
	Dense Layer + Attention Pooling	Title	0.1567	0.6740	0.8901	0.1138	0.2513
		Title + Category and Subcategory	0.1547	0.6734	0.8981	0.1167	0.2295
		Title + Category and Subcategory + Abstract	0.1561	0.6784	0.9003	0.1190	0.2270
	Multi-Head Self-Attention + Attention Pooling	Title	0.1559	0.6705	0.8962	0.1164	0.2361
		Title + Category and Subcategory	0.1534	0.6728	0.8987	0.1162	0.2259
		Title + Category and Subcategory + Abstract	0.1545	0.6759	0.8991	0.1171	0.2269
GIST-large-Embedding-v0	Average Pooling	Title	0.1467	0.6532	0.8724	0.1007	0.2700
		Title + Category and Subcategory	0.1410	0.6493	0.8755	0.0980	0.2514

Generalized Embedding Model	User Encoder	Model Input	F1 Score	AUC	Accuracy	Precision	Recall
GIST-large-Embedding-v0	Average Pooling	Title + Category and Subcategory + Abstract	0.1409	0.6455	0.8622	0.0943	0.2779
	Attention Pooling	Title	0.1520	0.6577	0.8748	0.1049	0.2763
		Title + Category and Subcategory	0.1519	0.6561	0.8711	0.1037	0.2842
		Title + Category and Subcategory + Abstract	0.1504	0.6521	0.8601	0.0998	0.3047
	Dense Layer + Attention Pooling	Title	0.1571	0.6745	0.8988	0.1187	0.2322
		Title + Category and Subcategory	0.1549	0.6730	0.8959	0.1156	0.2347
		Title + Category and Subcategory + Abstract	0.1564	0.6791	0.9035	0.1213	0.2200
	Multi-Head Self-Attention + Attention Pooling	Title	0.1565	0.6726	0.9033	0.1212	0.2208
		Title + Category and Subcategory	0.1537	0.6739	0.8994	0.1168	0.2247
		Title + Category and Subcategory + Abstract	0.1558	0.6765	0.9020	0.1199	0.2224
e5-mistral-7b-instruct	Average Pooling	Title	0.1059	0.5564	0.8878	0.0783	0.1635
		Title + Category and Subcategory	0.1106	0.5608	0.8929	0.0835	0.1638
		Title + Category and Subcategory + Abstract	0.1089	0.5591	0.8908	0.0814	0.1642
	Attention Pooling	Title	0.1101	0.5563	0.8884	0.0815	0.1698
		Title + Category and Subcategory	0.1234	0.5635	0.8883	0.0906	0.1936
		Title + Category and Subcategory + Abstract	0.1192	0.5616	0.8827	0.0858	0.1952
	Dense Layer + Attention Pooling	Title	0.1282	0.5725	0.8937	0.0961	0.1923
		Title + Category and Subcategory	0.1281	0.5727	0.8924	0.0955	0.1944
		Title + Category and Subcategory + Abstract	0.1278	0.5735	0.8939	0.0959	0.1912
	Multi-Head Self-Attention + Attention Pooling	Title	0.1284	0.5711	0.8966	0.0976	0.1875
		Title + Category and Subcategory	0.1274	0.5719	0.9010	0.0992	0.1778
		Title + Category and Subcategory + Abstract	0.1273	0.5725	0.8963	0.0967	0.1861
Qwen3-Embedding-0.6B	Average Pooling	Title	0.1330	0.6027	0.8491	0.0868	0.2849
		Title + Category and Subcategory	0.1175	0.5877	0.8522	0.0776	0.2421
		Title + Category and Subcategory + Abstract	0.1214	0.5924	0.8452	0.0789	0.2632
	Attention Pooling	Title	0.1379	0.6049	0.8626	0.0926	0.2704
		Title + Category and Subcategory	0.1248	0.5957	0.8560	0.0828	0.2526
		Title + Category and Subcategory + Abstract	0.1270	0.5974	0.8670	0.0866	0.2379
	Dense Layer + Attention Pooling	Title	0.1381	0.6119	0.9032	0.1082	0.1908
		Title + Category and Subcategory	0.1430	0.6134	0.8955	0.1072	0.2145
		Title + Category and Subcategory + Abstract	0.1419	0.6112	0.8960	0.1068	0.2117
		Title	0.1395	0.6111	0.8972	0.1057	0.2050

Generalized Embedding Model	User Encoder	Model Input	F1 Score	AUC	Accuracy	Precision	Recall
Qwen3-Embedding-0.6B	Multi-Head Self-Attention + Attention Pooling	Title + Category and Subcategory	0.1431	0.6134	0.8846	0.1024	0.2371
		Title + Category and Subcategory + Abstract	0.1409	0.6110	0.8949	0.1055	0.2121
Qwen3-Embedding-4B	Average Pooling	Title	0.1360	0.6070	0.8782	0.0956	0.2359
		Title + Category and Subcategory	0.1165	0.5990	0.8784	0.0827	0.1973
		Title + Category and Subcategory + Abstract	0.1266	0.6017	0.8864	0.0921	0.2026
	Attention Pooling	Title	0.1414	0.6085	0.8830	0.1007	0.2371
		Title + Category and Subcategory	0.1310	0.6045	0.8716	0.0904	0.2381
		Title + Category and Subcategory + Abstract	0.1363	0.6053	0.8853	0.0982	0.2228
	Dense Layer + Attention Pooling	Title	0.1476	0.6155	0.8997	0.1127	0.2137
		Title + Category and Subcategory	0.1494	0.6169	0.8953	0.1115	0.2261
		Title + Category and Subcategory + Abstract	0.1463	0.6147	0.8963	0.1099	0.2186
	Multi-Head Self-Attention + Attention Pooling	Title	0.1478	0.6157	0.8990	0.1124	0.2156
		Title + Category and Subcategory	0.1473	0.6158	0.8870	0.1062	0.2401
		Title + Category and Subcategory + Abstract	0.1472	0.6144	0.8890	0.1070	0.2359
Qwen3-Embedding-8B	Average Pooling	Title	0.1182	0.5725	0.8873	0.0867	0.1858
		Title + Category and Subcategory	0.1018	0.5564	0.8202	0.0639	0.2508
		Title + Category and Subcategory + Abstract	0.1046	0.5605	0.8470	0.0686	0.2199
	Attention Pooling	Title	0.1216	0.5736	0.8902	0.0901	0.1871
		Title + Category and Subcategory	0.1129	0.5669	0.8505	0.0744	0.2341
		Title + Category and Subcategory + Abstract	0.1166	0.5706	0.8742	0.0816	0.2044
	Dense Layer + Attention Pooling	Title	0.1239	0.5790	0.9014	0.0970	0.1716
		Title + Category and Subcategory	0.1211	0.5750	0.8858	0.0881	0.1936
		Title + Category and Subcategory + Abstract	0.1252	0.5761	0.8995	0.0968	0.1770
	Multi-Head Self-Attention + Attention Pooling	Title	0.1249	0.5780	0.8997	0.0968	0.1761
		Title + Category and Subcategory	0.1198	0.5745	0.8902	0.0889	0.1839
		Title + Category and Subcategory + Abstract	0.1245	0.5757	0.8980	0.0956	0.1783
NV-Embed-v2	Average Pooling	Title	0.0976	0.5420	0.8638	0.0668	0.1813
		Title + Category and Subcategory	0.1017	0.5435	0.8784	0.0727	0.1694
		Title + Category and Subcategory + Abstract	0.1033	0.5438	0.8845	0.0755	0.1636
	Attention Pooling	Title	0.1077	0.5471	0.8846	0.0785	0.1713
		Title + Category and Subcategory	0.1157	0.5506	0.8984	0.0895	0.1635
		Title + Category and Subcategory + Abstract	0.1154	0.5498	0.8927	0.0868	0.1722

Generalized Embedding Model	User Encoder	Model Input	F1 Score	AUC	Accuracy	Precision	Recall
NV-Embed-v2	Dense Layer + Attention Pooling	Title	0.1213	0.5611	0.9014	0.0951	0.1674
		Title + Category and Subcategory	0.1242	0.5634	0.8963	0.0946	0.1810
		Title + Category and Subcategory + Abstract	0.1247	0.5645	0.9023	0.0981	0.1714
	Multi-Head Self-Attention + Attention Pooling	Title	0.1198	0.5587	0.8953	0.0909	0.1753
		Title + Category and Subcategory	0.1225	0.5608	0.9010	0.0957	0.1700
		Title + Category and Subcategory + Abstract	0.1252	0.5636	0.9042	0.0996	0.1687
text-embedding-3-large (OpenAI)	Average Pooling	Title	0.0895	0.5202	0.8706	0.0627	0.1566
		Title + Category and Subcategory	0.0948	0.5224	0.8762	0.0675	0.1595
		Title + Category and Subcategory + Abstract	0.0988	0.5231	0.8700	0.0688	0.1754
	Attention Pooling	Title	0.0913	0.5202	0.8724	0.0642	0.1578
		Title + Category and Subcategory	0.0967	0.5227	0.8799	0.0696	0.1583
		Title + Category and Subcategory + Abstract	0.0995	0.5225	0.8780	0.0711	0.1658
	Dense Layer + Attention Pooling	Title	0.1108	0.5318	0.9093	0.0921	0.1391
		Title + Category and Subcategory	0.1092	0.5312	0.9008	0.0860	0.1496
		Title + Category and Subcategory + Abstract	0.1098	0.5311	0.9028	0.0875	0.1475
	Multi-Head Self-Attention + Attention Pooling	Title	0.1094	0.5305	0.9043	0.0879	0.1447
		Title + Category and Subcategory	0.1078	0.5298	0.9011	0.0851	0.1470
		Title + Category and Subcategory + Abstract	0.1082	0.5298	0.9050	0.0875	0.1418

A.1.2 Task 2: All User Classification for given News

The following table gives the experimental results for Task 2 without the fake news classification component.

Table A.3: Performance of Models on All User Classification for given News without Fake News Classification Component

Generalized Embedding Model	User Encoder	Model Input	F1 Score	AUC	Accuracy	Precision	Recall
multilingual-e5-large-instruct	Average Pooling	Title	0.0389	0.5008	0.8140	0.0404	0.1476
		Title + Category and Subcategory	0.0384	0.5019	0.8230	0.0399	0.1354
		Title + Category and Subcategory + Abstract	0.0393	0.5002	0.8004	0.0390	0.1526
	Attention Pooling	Title	0.0383	0.5019	0.8165	0.0403	0.1471
		Title + Category and Subcategory	0.0357	0.5017	0.8362	0.0397	0.1223

Generalized Embedding Model	User Encoder	Model Input	F1 Score	AUC	Accuracy	Precision	Recall
multilingual-e5-large-instruct	Attention Pooling	Title + Category and Subcategory + Abstract	0.0388	0.5012	0.8077	0.0399	0.1478
	Dense Layer + Attention Pooling	Title	0.0327	0.5013	0.8630	0.0388	0.0923
		Title + Category and Subcategory	0.0356	0.5014	0.8493	0.0386	0.1047
		Title + Category and Subcategory + Abstract	0.0354	0.5010	0.8535	0.0397	0.0997
	Multi-Head Self-Attention + Attention Pooling	Title	0.0349	0.5011	0.8623	0.0411	0.0955
		Title + Category and Subcategory	0.0342	0.5001	0.8569	0.0376	0.0979
		Title + Category and Subcategory + Abstract	0.0369	0.5023	0.8475	0.0400	0.1080
bge-large-en-v1.5	Average Pooling	Title	0.0347	0.5053	0.8193	0.0415	0.1503
		Title + Category and Subcategory	0.0424	0.5030	0.7725	0.0409	0.1895
		Title + Category and Subcategory + Abstract	0.0315	0.5050	0.8377	0.0419	0.1306
	Attention Pooling	Title	0.0331	0.5052	0.8274	0.0410	0.1401
		Title + Category and Subcategory	0.0407	0.5039	0.7776	0.0413	0.1821
		Title + Category and Subcategory + Abstract	0.0321	0.5061	0.8367	0.0399	0.1287
	Dense Layer + Attention Pooling	Title	0.0353	0.5051	0.8437	0.0404	0.1155
		Title + Category and Subcategory	0.0412	0.5068	0.8105	0.0416	0.1472
		Title + Category and Subcategory + Abstract	0.0404	0.5062	0.8126	0.0405	0.1379
	Multi-Head Self-Attention + Attention Pooling	Title	0.0349	0.5058	0.8455	0.0401	0.1146
		Title + Category and Subcategory	0.0378	0.5059	0.8251	0.0389	0.1306
		Title + Category and Subcategory + Abstract	0.0384	0.5063	0.8317	0.0405	0.1233
UAE-Large-V1	Average Pooling	Title	0.0390	0.5063	0.8368	0.0422	0.1313
		Title + Category and Subcategory	0.0444	0.5037	0.8086	0.0441	0.1570
		Title + Category and Subcategory + Abstract	0.0406	0.5058	0.8103	0.0400	0.1489
	Attention Pooling	Title	0.0373	0.5054	0.8299	0.0403	0.1320
		Title + Category and Subcategory	0.0422	0.5049	0.7993	0.0423	0.1583
		Title + Category and Subcategory + Abstract	0.0418	0.5048	0.7972	0.0412	0.1620
	Dense Layer + Attention Pooling	Title	0.0376	0.5078	0.8374	0.0414	0.1186
		Title + Category and Subcategory	0.0380	0.5059	0.8289	0.0409	0.1219
		Title + Category and Subcategory + Abstract	0.0387	0.5052	0.8370	0.0419	0.1180
	Multi-Head Self-Attention + Attention Pooling	Title	0.0340	0.5065	0.8539	0.0394	0.1019
		Title + Category and Subcategory	0.0380	0.5055	0.8336	0.0408	0.1205
		Title + Category and Subcategory + Abstract	0.0392	0.5047	0.8322	0.0414	0.1223
mxbai-embed-large-v1	Average Pooling	Title	0.0386	0.5041	0.8332	0.0409	0.1318

Generalized Embedding Model	User Encoder	Model Input	F1 Score	AUC	Accuracy	Precision	Recall
mxbai-embed-large-v1	Average Pooling	Title + Category and Subcategory	0.0423	0.5020	0.8069	0.0424	0.1534
		Title + Category and Subcategory + Abstract	0.0406	0.5041	0.8089	0.0403	0.1484
	Attention Pooling	Title	0.0363	0.5056	0.8346	0.0400	0.1267
		Title + Category and Subcategory	0.0429	0.5043	0.7942	0.0426	0.1649
		Title + Category and Subcategory + Abstract	0.0406	0.5033	0.8003	0.0402	0.1549
	Dense Layer + Attention Pooling	Title	0.0369	0.5076	0.8405	0.0400	0.1159
		Title + Category and Subcategory	0.0377	0.5056	0.8307	0.0408	0.1198
		Title + Category and Subcategory + Abstract	0.0388	0.5061	0.8357	0.0412	0.1196
	Multi-Head Self-Attention + Attention Pooling	Title	0.0366	0.5065	0.8410	0.0404	0.1154
		Title + Category and Subcategory	0.0372	0.5053	0.8369	0.0409	0.1178
		Title + Category and Subcategory + Abstract	0.0380	0.5060	0.8367	0.0403	0.1170
GIST-large-Embedding-v0	Average Pooling	Title	0.0397	0.5048	0.8311	0.0406	0.1317
		Title + Category and Subcategory	0.0434	0.5030	0.8054	0.0403	0.1537
		Title + Category and Subcategory + Abstract	0.0437	0.5052	0.7972	0.0397	0.1630
	Attention Pooling	Title	0.0374	0.5060	0.8325	0.0395	0.1276
		Title + Category and Subcategory	0.0430	0.5055	0.7992	0.0418	0.1560
		Title + Category and Subcategory + Abstract	0.0422	0.5034	0.7961	0.0400	0.1612
	Dense Layer + Attention Pooling	Title	0.0356	0.5052	0.8456	0.0393	0.1106
		Title + Category and Subcategory	0.0382	0.5035	0.8355	0.0418	0.1179
		Title + Category and Subcategory + Abstract	0.0380	0.5067	0.8470	0.0422	0.1113
	Multi-Head Self-Attention + Attention Pooling	Title	0.0329	0.5049	0.8597	0.0396	0.0963
		Title + Category and Subcategory	0.0369	0.5047	0.8458	0.0415	0.1113
		Title + Category and Subcategory + Abstract	0.0376	0.5058	0.8488	0.0412	0.1107
e5-mistral-7b-instruct	Average Pooling	Title	0.0338	0.4998	0.8326	0.0356	0.1264
		Title + Category and Subcategory	0.0367	0.5008	0.8307	0.0388	0.1235
		Title + Category and Subcategory + Abstract	0.0389	0.5025	0.8088	0.0375	0.1438
	Attention Pooling	Title	0.0314	0.5008	0.8526	0.0370	0.1061
		Title + Category and Subcategory	0.0357	0.5011	0.8394	0.0403	0.1206
		Title + Category and Subcategory + Abstract	0.0382	0.5038	0.8166	0.0384	0.1389
	Dense Layer + Attention Pooling	Title	0.0328	0.5007	0.8609	0.0403	0.0965
		Title + Category and Subcategory	0.0351	0.5017	0.8546	0.0409	0.1020
		Title + Category and Subcategory + Abstract	0.0355	0.5029	0.8486	0.0395	0.1076

Generalized Embedding Model	User Encoder	Model Input	F1 Score	AUC	Accuracy	Precision	Recall
e5-mistral-7b-instruct	Multi-Head Self-Attention + Attention Pooling	Title	0.0346	0.5021	0.8575	0.0417	0.1004
		Title + Category and Subcategory	0.0366	0.5022	0.8443	0.0400	0.1098
		Title + Category and Subcategory + Abstract	0.0368	0.5042	0.8528	0.0417	0.1034
Qwen3-Embedding-0.6B	Average Pooling	Title	0.0392	0.5097	0.8168	0.0409	0.1508
		Title + Category and Subcategory	0.0420	0.5064	0.7846	0.0401	0.1732
		Title + Category and Subcategory + Abstract	0.0414	0.5034	0.8027	0.0436	0.1630
	Attention Pooling	Title	0.0390	0.5101	0.8211	0.0416	0.1471
		Title + Category and Subcategory	0.0416	0.5060	0.7860	0.0401	0.1727
		Title + Category and Subcategory + Abstract	0.0411	0.5031	0.8005	0.0420	0.1637
	Dense Layer + Attention Pooling	Title	0.0326	0.5090	0.8546	0.0376	0.0980
		Title + Category and Subcategory	0.0383	0.5057	0.8342	0.0416	0.1224
		Title + Category and Subcategory + Abstract	0.0364	0.5070	0.8473	0.0407	0.1130
	Multi-Head Self-Attention + Attention Pooling	Title	0.0322	0.5073	0.8575	0.0398	0.0971
		Title + Category and Subcategory	0.0380	0.5054	0.8430	0.0424	0.1171
		Title + Category and Subcategory + Abstract	0.0344	0.5088	0.8587	0.0411	0.1023
Qwen3-Embedding-4B	Average Pooling	Title	0.0301	0.5026	0.8454	0.0362	0.1200
		Title + Category and Subcategory	0.0378	0.5015	0.8236	0.0427	0.1379
		Title + Category and Subcategory + Abstract	0.0350	0.5033	0.8316	0.0415	0.1274
	Attention Pooling	Title	0.0297	0.5017	0.8551	0.0364	0.1116
		Title + Category and Subcategory	0.0362	0.5014	0.8223	0.0400	0.1368
		Title + Category and Subcategory + Abstract	0.0348	0.5019	0.8241	0.0404	0.1337
	Dense Layer + Attention Pooling	Title	0.0281	0.4970	0.8669	0.0351	0.0858
		Title + Category and Subcategory	0.0323	0.4986	0.8526	0.0371	0.1007
		Title + Category and Subcategory + Abstract	0.0350	0.5020	0.8566	0.0412	0.1044
	Multi-Head Self-Attention + Attention Pooling	Title	0.0281	0.4974	0.8710	0.0367	0.0845
		Title + Category and Subcategory	0.0336	0.4987	0.8550	0.0396	0.1017
		Title + Category and Subcategory + Abstract	0.0371	0.5018	0.8489	0.0418	0.1141
Qwen3-Embedding-8B	Average Pooling	Title	0.0373	0.5017	0.8335	0.0411	0.1295
		Title + Category and Subcategory	0.0476	0.5045	0.7759	0.0449	0.1990
		Title + Category and Subcategory + Abstract	0.0408	0.5048	0.8015	0.0401	0.1635
	Attention Pooling	Title	0.0349	0.5014	0.8385	0.0388	0.1211
		Title + Category and Subcategory	0.0390	0.5048	0.8189	0.0414	0.1410

Generalized Embedding Model	User Encoder	Model Input	F1 Score	AUC	Accuracy	Precision	Recall
Qwen3-Embedding-8B	Attention Pooling	Title + Category and Subcategory + Abstract	0.0379	0.5055	0.8252	0.0422	0.1327
	Dense Layer + Attention Pooling	Title	0.0347	0.5027	0.8579	0.0401	0.0988
		Title + Category and Subcategory	0.0396	0.5053	0.8238	0.0394	0.1303
		Title + Category and Subcategory + Abstract	0.0395	0.5037	0.8296	0.0397	0.1289
	Multi-Head Self-Attention + Attention Pooling	Title	0.0354	0.5025	0.8628	0.0423	0.0977
		Title + Category and Subcategory	0.0404	0.5047	0.8327	0.0414	0.1259
		Title + Category and Subcategory + Abstract	0.0415	0.5045	0.8349	0.0418	0.1283
NV-Embed-v2	Average Pooling	Title	0.0444	0.5008	0.7852	0.0412	0.1878
		Title + Category and Subcategory	0.0464	0.5023	0.7806	0.0433	0.1870
		Title + Category and Subcategory + Abstract	0.0448	0.5015	0.7689	0.0405	0.1955
	Attention Pooling	Title	0.0397	0.5022	0.8176	0.0410	0.1423
		Title + Category and Subcategory	0.0421	0.5014	0.7889	0.0411	0.1659
		Title + Category and Subcategory + Abstract	0.0380	0.5016	0.8145	0.0405	0.1418
	Dense Layer + Attention Pooling	Title	0.0394	0.5018	0.8286	0.0403	0.1251
		Title + Category and Subcategory	0.0396	0.5050	0.8204	0.0397	0.1328
		Title + Category and Subcategory + Abstract	0.0410	0.5055	0.7996	0.0395	0.1496
	Multi-Head Self-Attention + Attention Pooling	Title	0.0390	0.5029	0.8329	0.0397	0.1209
		Title + Category and Subcategory	0.0418	0.5032	0.8070	0.0402	0.1432
		Title + Category and Subcategory + Abstract	0.0406	0.5052	0.8037	0.0391	0.1462
text-embedding-3-large (OpenAI)	Average Pooling	Title	0.0400	0.5000	0.8071	0.0383	0.1493
		Title + Category and Subcategory	0.0428	0.5018	0.8008	0.0400	0.1565
		Title + Category and Subcategory + Abstract	0.0413	0.5000	0.8084	0.0395	0.1513
	Attention Pooling	Title	0.0374	0.5011	0.8271	0.0380	0.1293
		Title + Category and Subcategory	0.0386	0.5032	0.8251	0.0389	0.1309
		Title + Category and Subcategory + Abstract	0.0377	0.5013	0.8357	0.0386	0.1223
	Dense Layer + Attention Pooling	Title	0.0370	0.4988	0.8622	0.0429	0.0979
		Title + Category and Subcategory	0.0415	0.5016	0.8334	0.0437	0.1240
		Title + Category and Subcategory + Abstract	0.0386	0.4994	0.8493	0.0420	0.1083
	Multi-Head Self-Attention + Attention Pooling	Title	0.0389	0.4994	0.8543	0.0437	0.1062
		Title + Category and Subcategory	0.0416	0.5000	0.8355	0.0444	0.1243
		Title + Category and Subcategory + Abstract	0.0390	0.5009	0.8470	0.0423	0.1115

The following table gives the experimental results for Task 2 with the fake news classification component.

Table A.4: Performance of Models on All User Classification for given News with Fake News Classification Component

Generalized Embedding Model	User Encoder	Model Input	F1 Score	AUC	Accuracy	Precision	Recall
multilingual-e5-large-instruct	Average Pooling	Title	0.0280	0.4970	0.8664	0.0430	0.0991
		Title + Category and Subcategory	0.0262	0.4976	0.8703	0.0407	0.0876
		Title + Category and Subcategory + Abstract	0.0269	0.4982	0.8535	0.0389	0.1036
	Attention Pooling	Title	0.0259	0.4971	0.8713	0.0414	0.0913
		Title + Category and Subcategory	0.0279	0.4974	0.8639	0.0427	0.0962
		Title + Category and Subcategory + Abstract	0.0264	0.4976	0.8567	0.0380	0.1020
	Dense Layer + Attention Pooling	Title	0.0275	0.4973	0.8708	0.0403	0.0836
		Title + Category and Subcategory	0.0289	0.4967	0.8590	0.0411	0.0920
		Title + Category and Subcategory + Abstract	0.0261	0.4989	0.8778	0.0402	0.0752
	Multi-Head Self-Attention + Attention Pooling	Title	0.0263	0.4968	0.8815	0.0417	0.0724
		Title + Category and Subcategory	0.0279	0.4967	0.8689	0.0409	0.0848
		Title + Category and Subcategory + Abstract	0.0278	0.4976	0.8697	0.0405	0.0844
bge-large-en-v1.5	Average Pooling	Title	0.0230	0.4997	0.8698	0.0401	0.0952
		Title + Category and Subcategory	0.0287	0.4988	0.8283	0.0382	0.1303
		Title + Category and Subcategory + Abstract	0.0237	0.4981	0.8565	0.0393	0.1136
	Attention Pooling	Title	0.0238	0.4994	0.8636	0.0399	0.1021
		Title + Category and Subcategory	0.0302	0.4991	0.8281	0.0410	0.1333
		Title + Category and Subcategory + Abstract	0.0285	0.4988	0.8373	0.0432	0.1362
	Dense Layer + Attention Pooling	Title	0.0265	0.5017	0.8691	0.0394	0.0868
		Title + Category and Subcategory	0.0299	0.5035	0.8511	0.0391	0.1050
		Title + Category and Subcategory + Abstract	0.0305	0.5032	0.8497	0.0407	0.1038
	Multi-Head Self-Attention + Attention Pooling	Title	0.0262	0.5022	0.8695	0.0388	0.0867
		Title + Category and Subcategory	0.0300	0.5022	0.8535	0.0401	0.1034
		Title + Category and Subcategory + Abstract	0.0302	0.5018	0.8585	0.0412	0.0987
UAE-Large-V1	Average Pooling	Title	0.0374	0.5051	0.8406	0.0413	0.1261
		Title + Category and Subcategory	0.0425	0.5025	0.8143	0.0438	0.1488
		Title + Category and Subcategory + Abstract	0.0403	0.5046	0.8096	0.0402	0.1522
	Attention Pooling	Title	0.0356	0.5041	0.8349	0.0393	0.1261

Generalized Embedding Model	User Encoder	Model Input	F1 Score	AUC	Accuracy	Precision	Recall
UAE-Large-V1	Attention Pooling	Title + Category and Subcategory	0.0405	0.5036	0.8064	0.0422	0.1503
		Title + Category and Subcategory + Abstract	0.0405	0.5039	0.8061	0.0416	0.1539
	Dense Layer + Attention Pooling	Title	0.0356	0.5056	0.8446	0.0412	0.1110
		Title + Category and Subcategory	0.0366	0.5042	0.8371	0.0417	0.1151
		Title + Category and Subcategory + Abstract	0.0366	0.5027	0.8443	0.0416	0.1112
	Multi-Head Self-Attention + Attention Pooling	Title	0.0321	0.5042	0.8596	0.0389	0.0956
		Title + Category and Subcategory	0.0365	0.5034	0.8414	0.0413	0.1138
		Title + Category and Subcategory + Abstract	0.0371	0.5024	0.8392	0.0409	0.1152
mxbai-embed-large-v1	Average Pooling	Title	0.0380	0.5020	0.8337	0.0404	0.1320
		Title + Category and Subcategory	0.0403	0.5002	0.8142	0.0419	0.1456
		Title + Category and Subcategory + Abstract	0.0380	0.5022	0.8161	0.0392	0.1396
	Attention Pooling	Title	0.0346	0.5031	0.8386	0.0390	0.1210
		Title + Category and Subcategory	0.0413	0.5025	0.8006	0.0423	0.1585
		Title + Category and Subcategory + Abstract	0.0384	0.5017	0.8088	0.0398	0.1455
	Dense Layer + Attention Pooling	Title	0.0380	0.5048	0.8342	0.0413	0.1223
		Title + Category and Subcategory	0.0358	0.5030	0.8398	0.0410	0.1113
		Title + Category and Subcategory + Abstract	0.0368	0.5027	0.8443	0.0413	0.1117
	Multi-Head Self-Attention + Attention Pooling	Title	0.0345	0.5037	0.8480	0.0399	0.1070
		Title + Category and Subcategory	0.0354	0.5031	0.8454	0.0411	0.1100
		Title + Category and Subcategory + Abstract	0.0358	0.5031	0.8446	0.0399	0.1092
GIST-large-Embedding-v0	Average Pooling	Title	0.0395	0.5045	0.8317	0.0406	0.1311
		Title + Category and Subcategory	0.0433	0.5031	0.8066	0.0406	0.1530
		Title + Category and Subcategory + Abstract	0.0435	0.5052	0.7988	0.0399	0.1619
	Attention Pooling	Title	0.0373	0.5057	0.8332	0.0392	0.1271
		Title + Category and Subcategory	0.0429	0.5057	0.8005	0.0420	0.1549
		Title + Category and Subcategory + Abstract	0.0420	0.5035	0.7975	0.0402	0.1599
	Dense Layer + Attention Pooling	Title	0.0353	0.5050	0.8469	0.0393	0.1093
		Title + Category and Subcategory	0.0378	0.5036	0.8371	0.0416	0.1163
		Title + Category and Subcategory + Abstract	0.0376	0.5065	0.8485	0.0420	0.1101
	Multi-Head Self-Attention + Attention Pooling	Title	0.0326	0.5047	0.8607	0.0393	0.0951
		Title + Category and Subcategory	0.0365	0.5048	0.8471	0.0414	0.1102
		Title + Category and Subcategory + Abstract	0.0374	0.5058	0.8502	0.0413	0.1100

Generalized Embedding Model	User Encoder	Model Input	F1 Score	AUC	Accuracy	Precision	Recall
e5-mistral-7b-instruct	Average Pooling	Title	0.0244	0.4939	0.8767	0.0427	0.0844
		Title + Category and Subcategory	0.0249	0.4943	0.8753	0.0415	0.0807
		Title + Category and Subcategory + Abstract	0.0246	0.4954	0.8738	0.0400	0.0812
	Attention Pooling	Title	0.0222	0.4946	0.8783	0.0397	0.0791
		Title + Category and Subcategory	0.0238	0.4936	0.8747	0.0418	0.0791
		Title + Category and Subcategory + Abstract	0.0243	0.4963	0.8745	0.0414	0.0829
	Dense Layer + Attention Pooling	Title	0.0249	0.4972	0.8787	0.0417	0.0759
		Title + Category and Subcategory	0.0280	0.4976	0.8710	0.0443	0.0858
		Title + Category and Subcategory + Abstract	0.0277	0.4991	0.8699	0.0445	0.0857
	Multi-Head Self-Attention + Attention Pooling	Title	0.0270	0.4976	0.8721	0.0418	0.0853
		Title + Category and Subcategory	0.0273	0.4974	0.8718	0.0433	0.0832
		Title + Category and Subcategory + Abstract	0.0289	0.4987	0.8644	0.0443	0.0906
Qwen3-Embedding-0.6B	Average Pooling	Title	0.0319	0.5075	0.8515	0.0442	0.1234
		Title + Category and Subcategory	0.0310	0.5047	0.8367	0.0400	0.1274
		Title + Category and Subcategory + Abstract	0.0327	0.5047	0.8371	0.0437	0.1352
	Attention Pooling	Title	0.0280	0.5077	0.8591	0.0410	0.1070
		Title + Category and Subcategory	0.0327	0.5042	0.8399	0.0440	0.1259
		Title + Category and Subcategory + Abstract	0.0309	0.5032	0.8504	0.0439	0.1171
	Dense Layer + Attention Pooling	Title	0.0265	0.5071	0.8773	0.0391	0.0804
		Title + Category and Subcategory	0.0300	0.5048	0.8598	0.0411	0.0985
		Title + Category and Subcategory + Abstract	0.0291	0.5060	0.8669	0.0407	0.0947
	Multi-Head Self-Attention + Attention Pooling	Title	0.0259	0.5059	0.8744	0.0382	0.0827
		Title + Category and Subcategory	0.0322	0.5040	0.8519	0.0425	0.1100
		Title + Category and Subcategory + Abstract	0.0297	0.5080	0.8668	0.0419	0.0962
Qwen3-Embedding-4B	Average Pooling	Title	0.0195	0.4981	0.8829	0.0372	0.0789
		Title + Category and Subcategory	0.0239	0.4966	0.8721	0.0384	0.0866
		Title + Category and Subcategory + Abstract	0.0228	0.4980	0.8723	0.0386	0.0854
	Attention Pooling	Title	0.0202	0.4979	0.8820	0.0377	0.0775
		Title + Category and Subcategory	0.0259	0.4960	0.8625	0.0403	0.0981
		Title + Category and Subcategory + Abstract	0.0262	0.4971	0.8697	0.0434	0.0924
	Dense Layer + Attention Pooling	Title	0.0246	0.4973	0.8716	0.0385	0.0822
		Title + Category and Subcategory	0.0286	0.4972	0.8616	0.0415	0.0920

Generalized Embedding Model	User Encoder	Model Input	F1 Score	AUC	Accuracy	Precision	Recall
Qwen3-Embedding-4B	Dense Layer + Attention Pooling	Title + Category and Subcategory + Abstract	0.0269	0.4980	0.8626	0.0380	0.0899
	Multi-Head Self-Attention + Attention Pooling	Title	0.0242	0.4973	0.8756	0.0385	0.0798
		Title + Category and Subcategory	0.0299	0.4968	0.8574	0.0413	0.0991
		Title + Category and Subcategory + Abstract	0.0276	0.4979	0.8601	0.0388	0.0956
Qwen3-Embedding-8B	Average Pooling	Title	0.0278	0.4972	0.8780	0.0452	0.0879
		Title + Category and Subcategory	0.0345	0.5002	0.8267	0.0446	0.1461
		Title + Category and Subcategory + Abstract	0.0324	0.5006	0.8444	0.0443	0.1260
	Attention Pooling	Title	0.0259	0.4978	0.8785	0.0418	0.0853
		Title + Category and Subcategory	0.0331	0.5006	0.8438	0.0453	0.1199
		Title + Category and Subcategory + Abstract	0.0292	0.5016	0.8692	0.0453	0.0972
	Dense Layer + Attention Pooling	Title	0.0294	0.4995	0.8655	0.0424	0.0913
		Title + Category and Subcategory	0.0336	0.5027	0.8493	0.0432	0.1117
		Title + Category and Subcategory + Abstract	0.0320	0.5006	0.8594	0.0429	0.1030
	Multi-Head Self-Attention + Attention Pooling	Title	0.0283	0.5008	0.8718	0.0424	0.0865
		Title + Category and Subcategory	0.0323	0.5013	0.8568	0.0424	0.1052
		Title + Category and Subcategory + Abstract	0.0327	0.5016	0.8606	0.0436	0.1039
NV-Embed-v2	Average Pooling	Title	0.0318	0.4988	0.8384	0.0403	0.1310
		Title + Category and Subcategory	0.0336	0.4993	0.8296	0.0421	0.1365
		Title + Category and Subcategory + Abstract	0.0305	0.4985	0.8418	0.0400	0.1182
	Attention Pooling	Title	0.0316	0.4974	0.8543	0.0433	0.1077
		Title + Category and Subcategory	0.0315	0.4979	0.8509	0.0442	0.1065
		Title + Category and Subcategory + Abstract	0.0292	0.4972	0.8496	0.0402	0.1040
	Dense Layer + Attention Pooling	Title	0.0309	0.4996	0.8526	0.0410	0.1001
		Title + Category and Subcategory	0.0345	0.5009	0.8390	0.0440	0.1144
		Title + Category and Subcategory + Abstract	0.0327	0.5014	0.8465	0.0431	0.1055
	Multi-Head Self-Attention + Attention Pooling	Title	0.0324	0.5011	0.8480	0.0408	0.1076
		Title + Category and Subcategory	0.0323	0.5018	0.8470	0.0419	0.1049
		Title + Category and Subcategory + Abstract	0.0323	0.5017	0.8494	0.0431	0.1030
text-embedding-3-large (OpenAI)	Average Pooling	Title	0.0230	0.4978	0.8763	0.0402	0.0862
		Title + Category and Subcategory	0.0250	0.4986	0.8688	0.0419	0.0923
		Title + Category and Subcategory + Abstract	0.0247	0.4968	0.8729	0.0427	0.0913
	Attention Pooling	Title	0.0244	0.4980	0.8770	0.0439	0.0838

Generalized Embedding Model	User Encoder	Model Input	F1 Score	AUC	Accuracy	Precision	Recall
text-embedding-3-large (OpenAI)	Attention Pooling	Title + Category and Subcategory	0.0252	0.4991	0.8693	0.0437	0.0869
		Title + Category and Subcategory + Abstract	0.0250	0.4981	0.8792	0.0447	0.0823
	Dense Layer + Attention Pooling	Title	0.0254	0.5005	0.8871	0.0451	0.0724
		Title + Category and Subcategory	0.0256	0.5001	0.8714	0.0426	0.0848
		Title + Category and Subcategory + Abstract	0.0263	0.5002	0.8758	0.0441	0.0833
	Multi-Head Self-Attention + Attention Pooling	Title	0.0253	0.5005	0.8902	0.0462	0.0709
		Title + Category and Subcategory	0.0252	0.4992	0.8800	0.0436	0.0792
		Title + Category and Subcategory + Abstract	0.0255	0.5008	0.8837	0.0446	0.0766

A.1.3 Task 3: Candidate News Classification for User

The following table gives the experimental results for Task 3 without the fake news classification component.

Table A.5: Performance of Models on Candidate News Classification for User without Fake News Classification Component

Generalized Embedding Model	User Encoder	Model Input	F1 Score	AUC	Accuracy	Precision	Recall
multilingual-e5-large-instruct	Average Pooling	Title	0.1450	0.6589	0.8069	0.1509	0.2973
		Title + Category and Subcategory	0.1410	0.6524	0.8226	0.1571	0.2659
		Title + Category and Subcategory + Abstract	0.1488	0.6615	0.8199	0.1621	0.2853
	Attention Pooling	Title	0.1516	0.6565	0.8029	0.1542	0.3123
		Title + Category and Subcategory	0.1498	0.6524	0.8268	0.1658	0.2731
		Title + Category and Subcategory + Abstract	0.1568	0.6601	0.8203	0.1680	0.2958
	Dense Layer + Attention Pooling	Title	0.1370	0.6516	0.8578	0.1805	0.1982
		Title + Category and Subcategory	0.1388	0.6548	0.8546	0.1795	0.2120
		Title + Category and Subcategory + Abstract	0.1369	0.6563	0.8561	0.1810	0.2106
	Multi-Head Self-Attention + Attention Pooling	Title	0.1356	0.6505	0.8582	0.1818	0.1994
		Title + Category and Subcategory	0.1299	0.6539	0.8601	0.1819	0.1966
		Title + Category and Subcategory + Abstract	0.1336	0.6563	0.8552	0.1773	0.2098
bge-large-en-v1.5	Average Pooling	Title	0.1521	0.6223	0.7615	0.1361	0.3394
		Title + Category and Subcategory	0.1553	0.6398	0.7807	0.1440	0.3371
		Title + Category and Subcategory + Abstract	0.1087	0.6011	0.8347	0.1361	0.1882

Generalized Embedding Model	User Encoder	Model Input	F1 Score	AUC	Accuracy	Precision	Recall
bge-large-en-v1.5	Attention Pooling	Title	0.1506	0.6220	0.7679	0.1365	0.3285
		Title + Category and Subcategory	0.1537	0.6399	0.7855	0.1428	0.3262
		Title + Category and Subcategory + Abstract	0.1233	0.6135	0.8267	0.1458	0.2256
	Dense Layer + Attention Pooling	Title	0.1430	0.6553	0.8413	0.1743	0.2410
		Title + Category and Subcategory	0.1550	0.6550	0.8313	0.1697	0.2725
		Title + Category and Subcategory + Abstract	0.1456	0.6566	0.8429	0.1718	0.2407
	Multi-Head Self-Attention + Attention Pooling	Title	0.1422	0.6563	0.8362	0.1724	0.2494
		Title + Category and Subcategory	0.1463	0.6588	0.8399	0.1723	0.2494
		Title + Category and Subcategory + Abstract	0.1351	0.6598	0.8510	0.1750	0.2164
UAE-Large-V1	Average Pooling	Title	0.1474	0.6506	0.8121	0.1546	0.2868
		Title + Category and Subcategory	0.1412	0.6499	0.8291	0.1584	0.2553
		Title + Category and Subcategory + Abstract	0.1402	0.6453	0.8220	0.1558	0.2644
	Attention Pooling	Title	0.1537	0.6509	0.8100	0.1572	0.2996
		Title + Category and Subcategory	0.1528	0.6521	0.8224	0.1615	0.2828
		Title + Category and Subcategory + Abstract	0.1552	0.6470	0.8096	0.1596	0.3045
	Dense Layer + Attention Pooling	Title	0.1480	0.6574	0.8476	0.1765	0.2358
		Title + Category and Subcategory	0.1439	0.6581	0.8482	0.1763	0.2343
		Title + Category and Subcategory + Abstract	0.1404	0.6629	0.8511	0.1777	0.2256
	Multi-Head Self-Attention + Attention Pooling	Title	0.1401	0.6594	0.8537	0.1789	0.2188
		Title + Category and Subcategory	0.1433	0.6604	0.8479	0.1759	0.2332
		Title + Category and Subcategory + Abstract	0.1440	0.6644	0.8471	0.1759	0.2355
mxbai-embed-large-v1	Average Pooling	Title	0.1494	0.6546	0.8107	0.1547	0.2930
		Title + Category and Subcategory	0.1447	0.6527	0.8236	0.1572	0.2696
		Title + Category and Subcategory + Abstract	0.1479	0.6496	0.8103	0.1544	0.2914
	Attention Pooling	Title	0.1536	0.6542	0.8134	0.1573	0.2953
		Title + Category and Subcategory	0.1552	0.6539	0.8143	0.1581	0.2987
		Title + Category and Subcategory + Abstract	0.1584	0.6511	0.8036	0.1578	0.3178
	Dense Layer + Attention Pooling	Title	0.1466	0.6581	0.8479	0.1769	0.2343
		Title + Category and Subcategory	0.1426	0.6583	0.8478	0.1760	0.2339
		Title + Category and Subcategory + Abstract	0.1420	0.6631	0.8491	0.1776	0.2311
	Multi-Head Self-Attention + Attention Pooling	Title	0.1484	0.6600	0.8447	0.1752	0.2418
		Title + Category and Subcategory	0.1405	0.6603	0.8484	0.1756	0.2301

Generalized Embedding Model	User Encoder	Model Input	F1 Score	AUC	Accuracy	Precision	Recall
mxbai-embed-large-v1	Multi-Head Self-Attention + Attention Pooling	Title + Category and Subcategory + Abstract	0.1410	0.6648	0.8485	0.1763	0.2301
GIST-large-Embedding-v0	Average Pooling	Title	0.1390	0.6527	0.8243	0.1551	0.2614
		Title + Category and Subcategory	0.1358	0.6491	0.8346	0.1592	0.2439
		Title + Category and Subcategory + Abstract	0.1410	0.6513	0.8220	0.1575	0.2684
	Attention Pooling	Title	0.1437	0.6546	0.8243	0.1579	0.2681
		Title + Category and Subcategory	0.1507	0.6516	0.8247	0.1621	0.2780
		Title + Category and Subcategory + Abstract	0.1537	0.6523	0.8152	0.1614	0.2965
	Dense Layer + Attention Pooling	Title	0.1483	0.6572	0.8463	0.1750	0.2366
		Title + Category and Subcategory	0.1446	0.6573	0.8455	0.1738	0.2375
		Title + Category and Subcategory + Abstract	0.1403	0.6612	0.8518	0.1781	0.2231
	Multi-Head Self-Attention + Attention Pooling	Title	0.1441	0.6591	0.8500	0.1758	0.2258
		Title + Category and Subcategory	0.1410	0.6584	0.8489	0.1751	0.2278
		Title + Category and Subcategory + Abstract	0.1416	0.6628	0.8509	0.1779	0.2256
e5-mistral-7b-instruct	Average Pooling	Title	0.1077	0.6276	0.8407	0.1378	0.1908
		Title + Category and Subcategory	0.1231	0.6365	0.8416	0.1545	0.2087
		Title + Category and Subcategory + Abstract	0.1245	0.6381	0.8326	0.1491	0.2230
	Attention Pooling	Title	0.1078	0.6276	0.8497	0.1458	0.1773
		Title + Category and Subcategory	0.1409	0.6366	0.8386	0.1668	0.2363
		Title + Category and Subcategory + Abstract	0.1366	0.6407	0.8300	0.1570	0.2444
	Dense Layer + Attention Pooling	Title	0.1440	0.6593	0.8527	0.1797	0.2183
		Title + Category and Subcategory	0.1399	0.6604	0.8539	0.1816	0.2171
		Title + Category and Subcategory + Abstract	0.1424	0.6604	0.8508	0.1772	0.2236
	Multi-Head Self-Attention + Attention Pooling	Title	0.1453	0.6599	0.8531	0.1824	0.2187
		Title + Category and Subcategory	0.1475	0.6599	0.8501	0.1821	0.2298
		Title + Category and Subcategory + Abstract	0.1386	0.6616	0.8558	0.1829	0.2110
Qwen3-Embedding-0.6B	Average Pooling	Title	0.1436	0.6304	0.7854	0.1399	0.3045
		Title + Category and Subcategory	0.1346	0.6140	0.7963	0.1384	0.2703
		Title + Category and Subcategory + Abstract	0.1348	0.6256	0.7998	0.1413	0.2728
	Attention Pooling	Title	0.1455	0.6315	0.7957	0.1449	0.2953
		Title + Category and Subcategory	0.1411	0.6217	0.7962	0.1430	0.2808
		Title + Category and Subcategory + Abstract	0.1390	0.6312	0.8024	0.1463	0.2777
	Dense Layer + Attention Pooling	Title	0.1188	0.6448	0.8507	0.1541	0.1842

Generalized Embedding Model	User Encoder	Model Input	F1 Score	AUC	Accuracy	Precision	Recall
Qwen3-Embedding-0.6B	Dense Layer + Attention Pooling	Title + Category and Subcategory	0.1315	0.6467	0.8490	0.1632	0.2079
		Title + Category and Subcategory + Abstract	0.1276	0.6498	0.8501	0.1671	0.2062
	Multi-Head Self-Attention + Attention Pooling	Title	0.1221	0.6465	0.8482	0.1557	0.1916
		Title + Category and Subcategory	0.1293	0.6486	0.8506	0.1617	0.2002
		Title + Category and Subcategory + Abstract	0.1219	0.6505	0.8551	0.1676	0.1897
Qwen3-Embedding-4B	Average Pooling	Title	0.1431	0.6304	0.8113	0.1505	0.2678
		Title + Category and Subcategory	0.1242	0.6243	0.8195	0.1394	0.2293
		Title + Category and Subcategory + Abstract	0.1231	0.6225	0.8258	0.1415	0.2219
	Attention Pooling	Title	0.1433	0.6318	0.8216	0.1568	0.2554
		Title + Category and Subcategory	0.1386	0.6306	0.8204	0.1517	0.2496
		Title + Category and Subcategory + Abstract	0.1348	0.6284	0.8214	0.1506	0.2461
	Dense Layer + Attention Pooling	Title	0.1266	0.6525	0.8535	0.1746	0.1948
		Title + Category and Subcategory	0.1364	0.6541	0.8529	0.1753	0.2109
		Title + Category and Subcategory + Abstract	0.1320	0.6551	0.8535	0.1770	0.2083
	Multi-Head Self-Attention + Attention Pooling	Title	0.1272	0.6552	0.8533	0.1752	0.1975
		Title + Category and Subcategory	0.1383	0.6561	0.8513	0.1753	0.2170
		Title + Category and Subcategory + Abstract	0.1391	0.6568	0.8453	0.1745	0.2309
Qwen3-Embedding-8B	Average Pooling	Title	0.1404	0.6388	0.8124	0.1501	0.2715
		Title + Category and Subcategory	0.1554	0.6058	0.7622	0.1417	0.3330
		Title + Category and Subcategory + Abstract	0.1458	0.6123	0.7896	0.1422	0.2874
	Attention Pooling	Title	0.1420	0.6391	0.8186	0.1543	0.2643
		Title + Category and Subcategory	0.1532	0.6238	0.8089	0.1550	0.2770
		Title + Category and Subcategory + Abstract	0.1520	0.6271	0.8088	0.1541	0.2779
	Dense Layer + Attention Pooling	Title	0.1312	0.6557	0.8545	0.1732	0.1983
		Title + Category and Subcategory	0.1489	0.6539	0.8381	0.1704	0.2488
		Title + Category and Subcategory + Abstract	0.1500	0.6557	0.8424	0.1736	0.2463
	Multi-Head Self-Attention + Attention Pooling	Title	0.1371	0.6549	0.8506	0.1716	0.2103
		Title + Category and Subcategory	0.1474	0.6521	0.8402	0.1682	0.2408
		Title + Category and Subcategory + Abstract	0.1497	0.6545	0.8436	0.1729	0.2423
NV-Embed-v2	Average Pooling	Title	0.1140	0.6075	0.8164	0.1331	0.2203
		Title + Category and Subcategory	0.1083	0.6200	0.8384	0.1415	0.1930
		Title + Category and Subcategory + Abstract	0.1260	0.6214	0.8227	0.1467	0.2371

Generalized Embedding Model	User Encoder	Model Input	F1 Score	AUC	Accuracy	Precision	Recall
NV-Embed-v2	Attention Pooling	Title	0.1155	0.6297	0.8426	0.1515	0.1987
		Title + Category and Subcategory	0.1266	0.6355	0.8427	0.1610	0.2141
		Title + Category and Subcategory + Abstract	0.1252	0.6316	0.8484	0.1628	0.2032
	Dense Layer + Attention Pooling	Title	0.1421	0.6630	0.8518	0.1804	0.2247
		Title + Category and Subcategory	0.1375	0.6609	0.8506	0.1789	0.2232
		Title + Category and Subcategory + Abstract	0.1468	0.6627	0.8425	0.1759	0.2468
	Multi-Head Self-Attention + Attention Pooling	Title	0.1370	0.6623	0.8549	0.1836	0.2137
		Title + Category and Subcategory	0.1425	0.6597	0.8461	0.1784	0.2350
		Title + Category and Subcategory + Abstract	0.1465	0.6621	0.8429	0.1772	0.2442
text-embedding-3-large (OpenAI)	Average Pooling	Title	0.1150	0.6186	0.8160	0.1357	0.2241
		Title + Category and Subcategory	0.1283	0.6205	0.8222	0.1455	0.2353
		Title + Category and Subcategory + Abstract	0.1272	0.6308	0.8196	0.1485	0.2444
	Attention Pooling	Title	0.1199	0.6251	0.8252	0.1440	0.2195
		Title + Category and Subcategory	0.1338	0.6297	0.8337	0.1558	0.2287
		Title + Category and Subcategory + Abstract	0.1283	0.6370	0.8351	0.1567	0.2261
	Dense Layer + Attention Pooling	Title	0.1314	0.6593	0.8629	0.1850	0.1900
		Title + Category and Subcategory	0.1419	0.6575	0.8544	0.1822	0.2207
		Title + Category and Subcategory + Abstract	0.1343	0.6588	0.8600	0.1855	0.2026
	Multi-Head Self-Attention + Attention Pooling	Title	0.1404	0.6606	0.8551	0.1811	0.2149
		Title + Category and Subcategory	0.1409	0.6591	0.8523	0.1810	0.2250
		Title + Category and Subcategory + Abstract	0.1369	0.6600	0.8575	0.1855	0.2115

The following table gives the experimental results for Task 3 with the fake news classification component.

Table A.6: Performance of Models on Candidate News Classification for User with Fake News Classification Component

Generalized Embedding Model	User Encoder	Model Input	F1 Score	AUC	Accuracy	Precision	Recall
multilingual-e5-large-instruct	Average Pooling	Title	0.1218	0.5886	0.8288	0.1386	0.2189
		Title + Category and Subcategory	0.1166	0.5859	0.8452	0.1482	0.1891
		Title + Category and Subcategory + Abstract	0.1272	0.5880	0.8323	0.1483	0.2238
	Attention Pooling	Title	0.1266	0.5875	0.8310	0.1443	0.2236

Generalized Embedding Model	User Encoder	Model Input	F1 Score	AUC	Accuracy	Precision	Recall
multilingual-e5-large-instruct	Attention Pooling	Title + Category and Subcategory	0.1317	0.5860	0.8369	0.1550	0.2210
		Title + Category and Subcategory + Abstract	0.1339	0.5876	0.8324	0.1540	0.2329
	Dense Layer + Attention Pooling	Title	0.1327	0.5857	0.8570	0.1707	0.1891
		Title + Category and Subcategory	0.1358	0.5876	0.8545	0.1704	0.1987
		Title + Category and Subcategory + Abstract	0.1250	0.5881	0.8632	0.1754	0.1765
	Multi-Head Self-Attention + Attention Pooling	Title	0.1200	0.5840	0.8641	0.1721	0.1655
		Title + Category and Subcategory	0.1248	0.5863	0.8597	0.1701	0.1799
		Title + Category and Subcategory + Abstract	0.1234	0.5884	0.8610	0.1723	0.1796
bge-large-en-v1.5	Average Pooling	Title	0.1482	0.5959	0.8012	0.1437	0.2735
		Title + Category and Subcategory	0.1501	0.6033	0.8060	0.1487	0.2798
		Title + Category and Subcategory + Abstract	0.1141	0.5813	0.8171	0.1276	0.2073
	Attention Pooling	Title	0.1515	0.5956	0.7940	0.1430	0.2882
		Title + Category and Subcategory	0.1527	0.6033	0.8068	0.1499	0.2826
		Title + Category and Subcategory + Abstract	0.1287	0.5880	0.7894	0.1306	0.2648
	Dense Layer + Attention Pooling	Title	0.1452	0.6045	0.8398	0.1694	0.2319
		Title + Category and Subcategory	0.1556	0.6031	0.8387	0.1720	0.2449
		Title + Category and Subcategory + Abstract	0.1517	0.6062	0.8440	0.1736	0.2325
	Multi-Head Self-Attention + Attention Pooling	Title	0.1435	0.6032	0.8334	0.1646	0.2384
		Title + Category and Subcategory	0.1527	0.6049	0.8402	0.1730	0.2403
		Title + Category and Subcategory + Abstract	0.1450	0.6079	0.8503	0.1773	0.2171
UAE-Large-V1	Average Pooling	Title	0.1472	0.6492	0.8130	0.1548	0.2848
		Title + Category and Subcategory	0.1410	0.6489	0.8301	0.1586	0.2532
		Title + Category and Subcategory + Abstract	0.1441	0.6457	0.8167	0.1561	0.2774
	Attention Pooling	Title	0.1535	0.6494	0.8110	0.1576	0.2974
		Title + Category and Subcategory	0.1525	0.6508	0.8234	0.1617	0.2803
		Title + Category and Subcategory + Abstract	0.1551	0.6469	0.8111	0.1601	0.3018
	Dense Layer + Attention Pooling	Title	0.1476	0.6555	0.8484	0.1769	0.2338
		Title + Category and Subcategory	0.1434	0.6562	0.8490	0.1767	0.2315
		Title + Category and Subcategory + Abstract	0.1400	0.6609	0.8519	0.1781	0.2236
	Multi-Head Self-Attention + Attention Pooling	Title	0.1397	0.6571	0.8543	0.1792	0.2169
		Title + Category and Subcategory	0.1429	0.6583	0.8488	0.1765	0.2306
		Title + Category and Subcategory + Abstract	0.1437	0.6624	0.8480	0.1764	0.2332

Generalized Embedding Model	User Encoder	Model Input	F1 Score	AUC	Accuracy	Precision	Recall
mxbai-embed-large-v1	Average Pooling	Title	0.1506	0.6520	0.8076	0.1541	0.2990
		Title + Category and Subcategory	0.1442	0.6500	0.8244	0.1571	0.2672
		Title + Category and Subcategory + Abstract	0.1478	0.6475	0.8110	0.1545	0.2898
	Attention Pooling	Title	0.1537	0.6517	0.8130	0.1572	0.2957
		Title + Category and Subcategory	0.1556	0.6512	0.8137	0.1581	0.2994
		Title + Category and Subcategory + Abstract	0.1582	0.6489	0.8045	0.1579	0.3157
	Dense Layer + Attention Pooling	Title	0.1535	0.6553	0.8398	0.1732	0.2543
		Title + Category and Subcategory	0.1421	0.6555	0.8485	0.1762	0.2314
		Title + Category and Subcategory + Abstract	0.1416	0.6601	0.8498	0.1779	0.2289
	Multi-Head Self-Attention + Attention Pooling	Title	0.1481	0.6571	0.8453	0.1754	0.2400
		Title + Category and Subcategory	0.1401	0.6574	0.8491	0.1759	0.2277
		Title + Category and Subcategory + Abstract	0.1406	0.6618	0.8492	0.1766	0.2280
GIST-large-Embedding-v0	Average Pooling	Title	0.1389	0.6521	0.8245	0.1551	0.2610
		Title + Category and Subcategory	0.1356	0.6487	0.8348	0.1593	0.2432
		Title + Category and Subcategory + Abstract	0.1409	0.6511	0.8222	0.1576	0.2678
	Attention Pooling	Title	0.1435	0.6540	0.8244	0.1579	0.2677
		Title + Category and Subcategory	0.1506	0.6512	0.8249	0.1622	0.2774
		Title + Category and Subcategory + Abstract	0.1537	0.6521	0.8153	0.1614	0.2962
	Dense Layer + Attention Pooling	Title	0.1482	0.6569	0.8464	0.1752	0.2361
		Title + Category and Subcategory	0.1444	0.6570	0.8457	0.1739	0.2369
		Title + Category and Subcategory + Abstract	0.1401	0.6610	0.8519	0.1783	0.2226
	Multi-Head Self-Attention + Attention Pooling	Title	0.1440	0.6586	0.8500	0.1759	0.2256
		Title + Category and Subcategory	0.1409	0.6580	0.8491	0.1753	0.2273
		Title + Category and Subcategory + Abstract	0.1414	0.6626	0.8510	0.1781	0.2251
e5-mistral-7b-instruct	Average Pooling	Title	0.0871	0.5592	0.8399	0.1140	0.1520
		Title + Category and Subcategory	0.0945	0.5638	0.8470	0.1254	0.1538
		Title + Category and Subcategory + Abstract	0.0926	0.5652	0.8466	0.1253	0.1528
	Attention Pooling	Title	0.0907	0.5596	0.8398	0.1194	0.1590
		Title + Category and Subcategory	0.1112	0.5660	0.8399	0.1368	0.1853
		Title + Category and Subcategory + Abstract	0.1071	0.5670	0.8367	0.1332	0.1846
	Dense Layer + Attention Pooling	Title	0.1150	0.5727	0.8456	0.1416	0.1901
		Title + Category and Subcategory	0.1167	0.5720	0.8449	0.1432	0.1924

Generalized Embedding Model	User Encoder	Model Input	F1 Score	AUC	Accuracy	Precision	Recall
e5-mistral-7b-instruct	Dense Layer + Attention Pooling	Title + Category and Subcategory + Abstract	0.1151	0.5731	0.8475	0.1445	0.1886
	Multi-Head Self-Attention + Attention Pooling	Title	0.1154	0.5730	0.8496	0.1461	0.1858
		Title + Category and Subcategory	0.1129	0.5728	0.8537	0.1486	0.1763
		Title + Category and Subcategory + Abstract	0.1149	0.5735	0.8511	0.1489	0.1839
Qwen3-Embedding-0.6B	Average Pooling	Title	0.1338	0.5923	0.7938	0.1310	0.2748
		Title + Category and Subcategory	0.1242	0.5831	0.8094	0.1312	0.2319
		Title + Category and Subcategory + Abstract	0.1264	0.5857	0.8008	0.1306	0.2525
	Attention Pooling	Title	0.1351	0.5940	0.8060	0.1373	0.2616
		Title + Category and Subcategory	0.1299	0.5885	0.8078	0.1351	0.2428
		Title + Category and Subcategory + Abstract	0.1248	0.5908	0.8194	0.1387	0.2283
	Dense Layer + Attention Pooling	Title	0.1217	0.5967	0.8435	0.1457	0.1915
		Title + Category and Subcategory	0.1349	0.5978	0.8405	0.1551	0.2165
		Title + Category and Subcategory + Abstract	0.1313	0.5981	0.8424	0.1581	0.2125
	Multi-Head Self-Attention + Attention Pooling	Title	0.1262	0.5965	0.8368	0.1451	0.2077
		Title + Category and Subcategory	0.1435	0.5986	0.8296	0.1511	0.2396
		Title + Category and Subcategory + Abstract	0.1316	0.5980	0.8414	0.1573	0.2133
Qwen3-Embedding-4B	Average Pooling	Title	0.1363	0.5972	0.8263	0.1512	0.2334
		Title + Category and Subcategory	0.1165	0.5979	0.8383	0.1429	0.1906
		Title + Category and Subcategory + Abstract	0.1200	0.5955	0.8399	0.1457	0.1963
	Attention Pooling	Title	0.1403	0.5990	0.8304	0.1582	0.2348
		Title + Category and Subcategory	0.1378	0.6004	0.8299	0.1555	0.2324
		Title + Category and Subcategory + Abstract	0.1324	0.5984	0.8375	0.1561	0.2181
	Dense Layer + Attention Pooling	Title	0.1402	0.6059	0.8487	0.1746	0.2158
		Title + Category and Subcategory	0.1473	0.6066	0.8465	0.1761	0.2290
		Title + Category and Subcategory + Abstract	0.1411	0.6063	0.8488	0.1775	0.2202
	Multi-Head Self-Attention + Attention Pooling	Title	0.1400	0.6068	0.8476	0.1724	0.2179
		Title + Category and Subcategory	0.1504	0.6070	0.8396	0.1706	0.2426
		Title + Category and Subcategory + Abstract	0.1464	0.6066	0.8423	0.1748	0.2377
Qwen3-Embedding-8B	Average Pooling	Title	0.1068	0.5788	0.8421	0.1373	0.1766
		Title + Category and Subcategory	0.1164	0.5587	0.7853	0.1173	0.2422
		Title + Category and Subcategory + Abstract	0.1110	0.5650	0.8114	0.1211	0.2111
	Attention Pooling	Title	0.1093	0.5798	0.8438	0.1405	0.1784

Generalized Embedding Model	User Encoder	Model Input	F1 Score	AUC	Accuracy	Precision	Recall
Qwen3-Embedding-8B	Attention Pooling	Title + Category and Subcategory	0.1209	0.5708	0.8110	0.1321	0.2269
		Title + Category and Subcategory + Abstract	0.1135	0.5750	0.8306	0.1323	0.1974
	Dense Layer + Attention Pooling	Title	0.1074	0.5842	0.8560	0.1520	0.1700
		Title + Category and Subcategory	0.1107	0.5827	0.8432	0.1485	0.1912
		Title + Category and Subcategory + Abstract	0.1090	0.5832	0.8555	0.1560	0.1756
	Multi-Head Self-Attention + Attention Pooling	Title	0.1089	0.5837	0.8536	0.1489	0.1753
		Title + Category and Subcategory	0.1065	0.5826	0.8468	0.1471	0.1815
		Title + Category and Subcategory + Abstract	0.1090	0.5833	0.8545	0.1549	0.1766
NV-Embed-v2	Average Pooling	Title	0.0953	0.5524	0.8311	0.1241	0.1699
		Title + Category and Subcategory	0.0960	0.5541	0.8435	0.1304	0.1600
		Title + Category and Subcategory + Abstract	0.0933	0.5546	0.8477	0.1310	0.1547
	Attention Pooling	Title	0.0967	0.5610	0.8468	0.1411	0.1595
		Title + Category and Subcategory	0.1056	0.5625	0.8570	0.1537	0.1556
		Title + Category and Subcategory + Abstract	0.1068	0.5611	0.8514	0.1492	0.1643
	Dense Layer + Attention Pooling	Title	0.1044	0.5765	0.8601	0.1611	0.1636
		Title + Category and Subcategory	0.1130	0.5750	0.8547	0.1583	0.1776
		Title + Category and Subcategory + Abstract	0.1103	0.5760	0.8590	0.1600	0.1688
	Multi-Head Self-Attention + Attention Pooling	Title	0.1061	0.5759	0.8560	0.1597	0.1703
		Title + Category and Subcategory	0.1090	0.5741	0.8587	0.1610	0.1666
		Title + Category and Subcategory + Abstract	0.1102	0.5768	0.8603	0.1622	0.1660
text-embedding-3-large (OpenAI)	Average Pooling	Title	0.0816	0.5117	0.8292	0.1052	0.1464
		Title + Category and Subcategory	0.0846	0.5148	0.8357	0.1104	0.1494
		Title + Category and Subcategory + Abstract	0.0897	0.5184	0.8293	0.1164	0.1621
	Attention Pooling	Title	0.0792	0.5129	0.8296	0.1053	0.1454
		Title + Category and Subcategory	0.0842	0.5152	0.8376	0.1133	0.1472
		Title + Category and Subcategory + Abstract	0.0853	0.5177	0.8341	0.1157	0.1524
	Dense Layer + Attention Pooling	Title	0.0888	0.5242	0.8618	0.1429	0.1344
		Title + Category and Subcategory	0.0914	0.5232	0.8557	0.1396	0.1443
		Title + Category and Subcategory + Abstract	0.0905	0.5232	0.8569	0.1403	0.1423
	Multi-Head Self-Attention + Attention Pooling	Title	0.0896	0.5235	0.8572	0.1381	0.1398
		Title + Category and Subcategory	0.0895	0.5228	0.8553	0.1385	0.1418

text-embedding-3-large (OpenAI)	Multi-Head Self-Attention + Attention Pooling	Title + Category and Subcategory + Abstract	0.0876	0.5228	0.8588	0.1411	0.1364

A.1.4 Task 4: Candidate News Ranking for User

The following table gives the experimental results for Task 4 without the fake news classification component.

Table A.7: Performance of Models on Candidate News Ranking for User without Fake News Classification Component

Generalized Embedding Model	User Encoder	Model Input	AUC	MRR	NDCG@5	NDCG@10
multilingual-e5-large-instruct	Average Pooling	Title	0.6589	0.3141	0.3483	0.4071
		Title + Category and Subcategory	0.6524	0.3168	0.3493	0.4088
		Title + Category and Subcategory + Abstract	0.6615	0.3231	0.3584	0.4165
	Attention Pooling	Title	0.6565	0.3169	0.3520	0.4104
		Title + Category and Subcategory	0.6524	0.3207	0.3549	0.4141
		Title + Category and Subcategory + Abstract	0.6601	0.3256	0.3622	0.4201
	Dense Layer + Attention Pooling	Title	0.6516	0.3197	0.3518	0.4130
		Title + Category and Subcategory	0.6548	0.3231	0.3562	0.4162
		Title + Category and Subcategory + Abstract	0.6563	0.3239	0.3570	0.4175
	Multi-Head Self-Attention + Attention Pooling	Title	0.6505	0.3217	0.3530	0.4138
		Title + Category and Subcategory	0.6539	0.3244	0.3564	0.4159
		Title + Category and Subcategory + Abstract	0.6563	0.3241	0.3568	0.4174
bge-large-en-v1.5	Average Pooling	Title	0.6223	0.2991	0.3261	0.3859
		Title + Category and Subcategory	0.6398	0.3133	0.3448	0.4014
		Title + Category and Subcategory + Abstract	0.6011	0.2945	0.3177	0.3753
	Attention Pooling	Title	0.6220	0.2988	0.3262	0.3858
		Title + Category and Subcategory	0.6399	0.3146	0.3466	0.4036
		Title + Category and Subcategory + Abstract	0.6135	0.3021	0.3295	0.3868
	Dense Layer + Attention Pooling	Title	0.6553	0.3213	0.3546	0.4142
		Title + Category and Subcategory	0.6550	0.3244	0.3567	0.4169
		Title + Category and Subcategory + Abstract	0.6566	0.3211	0.3544	0.4148
	Multi-Head Self-Attention + Attention Pooling	Title	0.6563	0.3216	0.3547	0.4144
		Title + Category and Subcategory	0.6588	0.3257	0.3585	0.4184

Generalized Embedding Model	User Encoder	Model Input	AUC	MRR	NDCG@5	NDCG@10
bge-large-en-v1.5	Multi-Head Self-Attention + Attention Pooling	Title + Category and Subcategory + Abstract	0.6598	0.3227	0.3562	0.4166
UAE-Large-V1	Average Pooling	Title	0.6506	0.3163	0.3464	0.4054
		Title + Category and Subcategory	0.6499	0.3143	0.3431	0.4033
		Title + Category and Subcategory + Abstract	0.6453	0.3154	0.3458	0.4039
	Attention Pooling	Title	0.6509	0.3209	0.3524	0.4104
		Title + Category and Subcategory	0.6521	0.3211	0.3524	0.4111
		Title + Category and Subcategory + Abstract	0.6470	0.3205	0.3533	0.4113
	Dense Layer + Attention Pooling	Title	0.6574	0.3260	0.3591	0.4184
		Title + Category and Subcategory	0.6581	0.3274	0.3605	0.4197
		Title + Category and Subcategory + Abstract	0.6629	0.3286	0.3633	0.4222
	Multi-Head Self-Attention + Attention Pooling	Title	0.6594	0.3263	0.3598	0.4188
		Title + Category and Subcategory	0.6604	0.3279	0.3612	0.4203
		Title + Category and Subcategory + Abstract	0.6644	0.3292	0.3641	0.4229
mxbai-embed-large-v1	Average Pooling	Title	0.6546	0.3182	0.3498	0.4087
		Title + Category and Subcategory	0.6527	0.3155	0.3451	0.4053
		Title + Category and Subcategory + Abstract	0.6496	0.3148	0.3461	0.4049
	Attention Pooling	Title	0.6542	0.3220	0.3552	0.4132
		Title + Category and Subcategory	0.6539	0.3214	0.3534	0.4120
		Title + Category and Subcategory + Abstract	0.6511	0.3195	0.3527	0.4111
	Dense Layer + Attention Pooling	Title	0.6581	0.3255	0.3590	0.4184
		Title + Category and Subcategory	0.6583	0.3263	0.3599	0.4195
		Title + Category and Subcategory + Abstract	0.6631	0.3280	0.3630	0.4219
	Multi-Head Self-Attention + Attention Pooling	Title	0.6600	0.3258	0.3596	0.4187
		Title + Category and Subcategory	0.6603	0.3271	0.3605	0.4200
		Title + Category and Subcategory + Abstract	0.6648	0.3292	0.3643	0.4232
GIST-large-Embedding-v0	Average Pooling	Title	0.6527	0.3159	0.3460	0.4056
		Title + Category and Subcategory	0.6491	0.3151	0.3430	0.4031
		Title + Category and Subcategory + Abstract	0.6513	0.3176	0.3478	0.4066
	Attention Pooling	Title	0.6546	0.3194	0.3515	0.4103
		Title + Category and Subcategory	0.6516	0.3217	0.3530	0.4117
		Title + Category and Subcategory + Abstract	0.6523	0.3218	0.3553	0.4131
	Dense Layer + Attention Pooling	Title	0.6572	0.3234	0.3571	0.4167

Generalized Embedding Model	User Encoder	Model Input	AUC	MRR	NDCG@5	NDCG@10
GIST-large-Embedding-v0	Dense Layer + Attention Pooling	Title	0.6573	0.3253	0.3587	0.4182
		Title + Category and Subcategory	0.6612	0.3269	0.3621	0.4214
		Title + Category and Subcategory + Abstract				
	Multi-Head Self-Attention + Attention Pooling	Title	0.6591	0.3232	0.3575	0.4167
		Title + Category and Subcategory	0.6584	0.3252	0.3584	0.4182
		Title + Category and Subcategory + Abstract	0.6628	0.3272	0.3624	0.4215
e5-mistral-7b-instruct	Average Pooling	Title	0.6276	0.2970	0.3235	0.3844
		Title + Category and Subcategory	0.6365	0.3080	0.3350	0.3956
		Title + Category and Subcategory + Abstract	0.6381	0.3071	0.3345	0.3957
	Attention Pooling	Title	0.6276	0.3003	0.3281	0.3891
		Title + Category and Subcategory	0.6366	0.3152	0.3462	0.4047
		Title + Category and Subcategory + Abstract	0.6407	0.3138	0.3445	0.4043
	Dense Layer + Attention Pooling	Title	0.6593	0.3231	0.3575	0.4176
		Title + Category and Subcategory	0.6604	0.3275	0.3611	0.4210
		Title + Category and Subcategory + Abstract	0.6604	0.3249	0.3602	0.4197
	Multi-Head Self-Attention + Attention Pooling	Title	0.6599	0.3250	0.3590	0.4190
		Title + Category and Subcategory	0.6599	0.3277	0.3613	0.4209
		Title + Category and Subcategory + Abstract	0.6616	0.3261	0.3609	0.4205
Qwen3-Embedding-0.6B	Average Pooling	Title	0.6304	0.3036	0.3340	0.3923
		Title + Category and Subcategory	0.6140	0.2925	0.3157	0.3757
		Title + Category and Subcategory + Abstract	0.6256	0.3030	0.3296	0.3892
	Attention Pooling	Title	0.6315	0.3042	0.3357	0.3940
		Title + Category and Subcategory	0.6217	0.2968	0.3220	0.3823
		Title + Category and Subcategory + Abstract	0.6312	0.3064	0.3346	0.3942
	Dense Layer + Attention Pooling	Title	0.6448	0.3011	0.3318	0.3943
		Title + Category and Subcategory	0.6467	0.3106	0.3416	0.4023
		Title + Category and Subcategory + Abstract	0.6498	0.3139	0.3453	0.4062
	Multi-Head Self-Attention + Attention Pooling	Title	0.6465	0.3020	0.3327	0.3949
		Title + Category and Subcategory	0.6486	0.3108	0.3421	0.4028
		Title + Category and Subcategory + Abstract	0.6505	0.3137	0.3450	0.4061
Qwen3-Embedding-4B	Average Pooling	Title	0.6304	0.3034	0.3333	0.3934
		Title + Category and Subcategory	0.6243	0.2973	0.3224	0.3834
		Title + Category and Subcategory + Abstract	0.6225	0.2993	0.3262	0.3867

Generalized Embedding Model	User Encoder	Model Input	AUC	MRR	NDCG@5	NDCG@10
Qwen3-Embedding-4B	Attention Pooling	Title	0.6318	0.3064	0.3371	0.3971
		Title + Category and Subcategory	0.6306	0.3046	0.3326	0.3924
		Title + Category and Subcategory + Abstract	0.6284	0.3048	0.3341	0.3935
	Dense Layer + Attention Pooling	Title	0.6525	0.3176	0.3500	0.4101
		Title + Category and Subcategory	0.6541	0.3206	0.3535	0.4129
		Title + Category and Subcategory + Abstract	0.6551	0.3230	0.3554	0.4154
	Multi-Head Self-Attention + Attention Pooling	Title	0.6552	0.3195	0.3510	0.4114
		Title + Category and Subcategory	0.6561	0.3228	0.3554	0.4145
		Title + Category and Subcategory + Abstract	0.6568	0.3240	0.3561	0.4161
Qwen3-Embedding-8B	Average Pooling	Title	0.6388	0.3105	0.3374	0.3974
		Title + Category and Subcategory	0.6058	0.2916	0.3098	0.3714
		Title + Category and Subcategory + Abstract	0.6123	0.2913	0.3088	0.3720
	Attention Pooling	Title	0.6391	0.3111	0.3393	0.3994
		Title + Category and Subcategory	0.6238	0.3030	0.3267	0.3868
		Title + Category and Subcategory + Abstract	0.6271	0.3025	0.3271	0.3874
	Dense Layer + Attention Pooling	Title	0.6557	0.3195	0.3517	0.4124
		Title + Category and Subcategory	0.6539	0.3249	0.3573	0.4160
		Title + Category and Subcategory + Abstract	0.6557	0.3230	0.3558	0.4163
	Multi-Head Self-Attention + Attention Pooling	Title	0.6549	0.3179	0.3494	0.4105
		Title + Category and Subcategory	0.6521	0.3231	0.3554	0.4144
		Title + Category and Subcategory + Abstract	0.6545	0.3224	0.3551	0.4155
NV-Embed-v2	Average Pooling	Title	0.6075	0.2852	0.3051	0.3691
		Title + Category and Subcategory	0.6200	0.2993	0.3214	0.3819
		Title + Category and Subcategory + Abstract	0.6214	0.3043	0.3260	0.3853
	Attention Pooling	Title	0.6297	0.3062	0.3320	0.3925
		Title + Category and Subcategory	0.6355	0.3140	0.3415	0.3997
		Title + Category and Subcategory + Abstract	0.6316	0.3151	0.3413	0.3996
	Dense Layer + Attention Pooling	Title	0.6630	0.3294	0.3642	0.4239
		Title + Category and Subcategory	0.6609	0.3280	0.3621	0.4218
		Title + Category and Subcategory + Abstract	0.6627	0.3282	0.3629	0.4226
	Multi-Head Self-Attention + Attention Pooling	Title	0.6623	0.3297	0.3640	0.4239
		Title + Category and Subcategory	0.6597	0.3287	0.3620	0.4210

Generalized Embedding Model	User Encoder	Model Input	AUC	MRR	NDCG@5	NDCG@10
NV-Embed-v2	Multi-Head Self-Attention + Attention Pooling	Title + Category and Subcategory + Abstract	0.6621	0.3280	0.3627	0.4227
text-embedding-3-large (OpenAI)	Average Pooling	Title	0.6186	0.2957	0.3161	0.3777
		Title + Category and Subcategory	0.6205	0.2964	0.3172	0.3796
		Title + Category and Subcategory + Abstract	0.6308	0.3084	0.3328	0.3929
	Attention Pooling	Title	0.6251	0.3019	0.3257	0.3863
		Title + Category and Subcategory	0.6297	0.3060	0.3319	0.3924
		Title + Category and Subcategory + Abstract	0.6370	0.3143	0.3418	0.4018
	Dense Layer + Attention Pooling	Title	0.6593	0.3266	0.3606	0.4210
		Title + Category and Subcategory	0.6575	0.3283	0.3620	0.4218
		Title + Category and Subcategory + Abstract	0.6588	0.3278	0.3622	0.4222
	Multi-Head Self-Attention + Attention Pooling	Title	0.6606	0.3275	0.3614	0.4218
		Title + Category and Subcategory	0.6591	0.3297	0.3634	0.4228
		Title + Category and Subcategory + Abstract	0.6600	0.3291	0.3636	0.4233

The following table gives the experimental results for Task 4 with the fake news classification component.

Table A.8: Performance of Models on Candidate News Ranking for User with Fake News Classification Component

Generalized Embedding Model	User Encoder	Model Input	AUC	MRR	NDCG@5	NDCG@10
multilingual-e5-large-instruct	Average Pooling	Title	0.5886	0.2935	0.3211	0.3755
		Title + Category and Subcategory	0.5859	0.2963	0.3224	0.3770
		Title + Category and Subcategory + Abstract	0.5880	0.3000	0.3279	0.3815
	Attention Pooling	Title	0.5875	0.2967	0.3251	0.3789
		Title + Category and Subcategory	0.5860	0.3003	0.3283	0.3818
		Title + Category and Subcategory + Abstract	0.5876	0.3028	0.3318	0.3850
	Dense Layer + Attention Pooling	Title	0.5857	0.3013	0.3271	0.3839
		Title + Category and Subcategory	0.5876	0.3040	0.3297	0.3863
		Title + Category and Subcategory + Abstract	0.5881	0.3043	0.3307	0.3867
	Multi-Head Self-Attention + Attention Pooling	Title	0.5840	0.2999	0.3240	0.3814
		Title + Category and Subcategory	0.5863	0.3023	0.3270	0.3834

Generalized Embedding Model	User Encoder	Model Input	AUC	MRR	NDCG@5	NDCG@10
multilingual-e5-large-instruct	Multi-Head Self-Attention + Attention Pooling	Title + Category and Subcategory + Abstract	0.5884	0.3038	0.3296	0.3854
bge-large-en-v1.5	Average Pooling	Title	0.5959	0.2987	0.3253	0.3818
		Title + Category and Subcategory	0.6033	0.3110	0.3382	0.3918
		Title + Category and Subcategory + Abstract	0.5813	0.2844	0.3053	0.3604
	Attention Pooling	Title	0.5956	0.2987	0.3255	0.3818
		Title + Category and Subcategory	0.6033	0.3121	0.3404	0.3939
		Title + Category and Subcategory + Abstract	0.5880	0.2931	0.3167	0.3711
	Dense Layer + Attention Pooling	Title	0.6045	0.3148	0.3431	0.3991
		Title + Category and Subcategory	0.6031	0.3133	0.3406	0.3973
		Title + Category and Subcategory + Abstract	0.6062	0.3154	0.3441	0.3994
	Multi-Head Self-Attention + Attention Pooling	Title	0.6032	0.3129	0.3410	0.3970
		Title + Category and Subcategory	0.6049	0.3149	0.3422	0.3990
		Title + Category and Subcategory + Abstract	0.6079	0.3171	0.3457	0.4009
UAE-Large-V1	Average Pooling	Title	0.6492	0.3162	0.3463	0.4051
		Title + Category and Subcategory	0.6489	0.3142	0.3429	0.4028
		Title + Category and Subcategory + Abstract	0.6457	0.3158	0.3464	0.4044
	Attention Pooling	Title	0.6494	0.3208	0.3524	0.4103
		Title + Category and Subcategory	0.6508	0.3210	0.3521	0.4107
		Title + Category and Subcategory + Abstract	0.6469	0.3207	0.3535	0.4114
	Dense Layer + Attention Pooling	Title	0.6555	0.3256	0.3587	0.4180
		Title + Category and Subcategory	0.6562	0.3270	0.3600	0.4191
		Title + Category and Subcategory + Abstract	0.6609	0.3283	0.3630	0.4218
	Multi-Head Self-Attention + Attention Pooling	Title	0.6571	0.3260	0.3594	0.4182
		Title + Category and Subcategory	0.6583	0.3275	0.3608	0.4196
		Title + Category and Subcategory + Abstract	0.6624	0.3288	0.3638	0.4224
mxbai-embed-large-v1	Average Pooling	Title	0.6520	0.3180	0.3497	0.4082
		Title + Category and Subcategory	0.6500	0.3152	0.3448	0.4048
		Title + Category and Subcategory + Abstract	0.6475	0.3145	0.3459	0.4044
	Attention Pooling	Title	0.6517	0.3218	0.3551	0.4128
		Title + Category and Subcategory	0.6512	0.3211	0.3531	0.4114
		Title + Category and Subcategory + Abstract	0.6489	0.3192	0.3526	0.4106
	Dense Layer + Attention Pooling	Title	0.6553	0.3251	0.3585	0.4176

Generalized Embedding Model	User Encoder	Model Input	AUC	MRR	NDCG@5	NDCG@10
mxbai-embed-large-v1	Dense Layer + Attention Pooling	Title	0.6555	0.3258	0.3593	0.4185
		Title + Category and Subcategory	0.6601	0.3274	0.3623	0.4210
		Title + Category and Subcategory + Abstract				
	Multi-Head Self-Attention + Attention Pooling	Title	0.6571	0.3254	0.3591	0.4180
		Title + Category and Subcategory	0.6574	0.3266	0.3599	0.4191
		Title + Category and Subcategory + Abstract	0.6618	0.3286	0.3637	0.4224
GIST-large-Embedding-v0	Average Pooling	Title	0.6521	0.3157	0.3458	0.4054
		Title + Category and Subcategory	0.6487	0.3150	0.3428	0.4028
		Title + Category and Subcategory + Abstract	0.6511	0.3175	0.3477	0.4064
	Attention Pooling	Title	0.6540	0.3193	0.3512	0.4100
		Title + Category and Subcategory	0.6512	0.3216	0.3529	0.4115
		Title + Category and Subcategory + Abstract	0.6521	0.3217	0.3552	0.4131
	Dense Layer + Attention Pooling	Title	0.6569	0.3233	0.3570	0.4166
		Title + Category and Subcategory	0.6570	0.3252	0.3586	0.4181
		Title + Category and Subcategory + Abstract	0.6610	0.3269	0.3621	0.4214
	Multi-Head Self-Attention + Attention Pooling	Title	0.6586	0.3231	0.3574	0.4166
		Title + Category and Subcategory	0.6580	0.3252	0.3583	0.4182
		Title + Category and Subcategory + Abstract	0.6626	0.3272	0.3623	0.4215
e5-mistral-7b-instruct	Average Pooling	Title	0.5592	0.2688	0.2865	0.3491
		Title + Category and Subcategory	0.5638	0.2775	0.2975	0.3571
		Title + Category and Subcategory + Abstract	0.5652	0.2780	0.2969	0.3573
	Attention Pooling	Title	0.5596	0.2727	0.2908	0.3526
		Title + Category and Subcategory	0.5660	0.2856	0.3069	0.3650
		Title + Category and Subcategory + Abstract	0.5670	0.2841	0.3050	0.3637
	Dense Layer + Attention Pooling	Title	0.5727	0.2913	0.3144	0.3731
		Title + Category and Subcategory	0.5720	0.2923	0.3147	0.3737
		Title + Category and Subcategory + Abstract	0.5731	0.2921	0.3154	0.3742
	Multi-Head Self-Attention + Attention Pooling	Title	0.5730	0.2920	0.3157	0.3739
		Title + Category and Subcategory	0.5728	0.2933	0.3160	0.3747
		Title + Category and Subcategory + Abstract	0.5735	0.2926	0.3162	0.3745
Qwen3-Embedding-0.6B	Average Pooling	Title	0.5923	0.2949	0.3260	0.3825
		Title + Category and Subcategory	0.5831	0.2837	0.3089	0.3670
		Title + Category and Subcategory + Abstract	0.5857	0.2913	0.3183	0.3749

Generalized Embedding Model	User Encoder	Model Input	AUC	MRR	NDCG@5	NDCG@10
Qwen3-Embedding-0.6B	Attention Pooling	Title	0.5940	0.2961	0.3284	0.3849
		Title + Category and Subcategory	0.5885	0.2886	0.3151	0.3739
		Title + Category and Subcategory + Abstract	0.5908	0.2946	0.3236	0.3799
	Dense Layer + Attention Pooling	Title	0.5967	0.2889	0.3209	0.3804
		Title + Category and Subcategory	0.5978	0.2978	0.3293	0.3874
		Title + Category and Subcategory + Abstract	0.5981	0.3005	0.3321	0.3897
	Multi-Head Self-Attention + Attention Pooling	Title	0.5965	0.2889	0.3214	0.3803
		Title + Category and Subcategory	0.5986	0.2981	0.3298	0.3879
		Title + Category and Subcategory + Abstract	0.5980	0.2999	0.3314	0.3893
Qwen3-Embedding-4B	Average Pooling	Title	0.5972	0.2992	0.3269	0.3849
		Title + Category and Subcategory	0.5979	0.2952	0.3197	0.3785
		Title + Category and Subcategory + Abstract	0.5955	0.2973	0.3250	0.3816
	Attention Pooling	Title	0.5990	0.3041	0.3325	0.3900
		Title + Category and Subcategory	0.6004	0.3027	0.3288	0.3872
		Title + Category and Subcategory + Abstract	0.5984	0.3021	0.3312	0.3875
	Dense Layer + Attention Pooling	Title	0.6059	0.3130	0.3408	0.3988
		Title + Category and Subcategory	0.6066	0.3137	0.3425	0.4000
		Title + Category and Subcategory + Abstract	0.6063	0.3154	0.3431	0.4008
	Multi-Head Self-Attention + Attention Pooling	Title	0.6068	0.3121	0.3398	0.3979
		Title + Category and Subcategory	0.6070	0.3135	0.3419	0.3995
		Title + Category and Subcategory + Abstract	0.6066	0.3143	0.3421	0.3998
Qwen3-Embedding-8B	Average Pooling	Title	0.5788	0.2882	0.3102	0.3688
		Title + Category and Subcategory	0.5587	0.2668	0.2820	0.3439
		Title + Category and Subcategory + Abstract	0.5650	0.2701	0.2863	0.3481
	Attention Pooling	Title	0.5798	0.2904	0.3132	0.3714
		Title + Category and Subcategory	0.5708	0.2824	0.3013	0.3608
		Title + Category and Subcategory + Abstract	0.5750	0.2824	0.3032	0.3628
	Dense Layer + Attention Pooling	Title	0.5842	0.2948	0.3197	0.3782
		Title + Category and Subcategory	0.5827	0.2964	0.3214	0.3797
		Title + Category and Subcategory + Abstract	0.5832	0.2951	0.3210	0.3791
	Multi-Head Self-Attention + Attention Pooling	Title	0.5837	0.2932	0.3174	0.3763
		Title + Category and Subcategory	0.5826	0.2959	0.3210	0.3790

Generalized Embedding Model	User Encoder	Model Input	AUC	MRR	NDCG@5	NDCG@10
Qwen3-Embedding-8B	Multi-Head Self-Attention + Attention Pooling	Title + Category and Subcategory + Abstract	0.5833	0.2954	0.3213	0.3794
NV-Embed-v2	Average Pooling	Title	0.5524	0.2645	0.2794	0.3429
		Title + Category and Subcategory	0.5541	0.2710	0.2874	0.3488
		Title + Category and Subcategory + Abstract	0.5546	0.2739	0.2899	0.3507
	Attention Pooling	Title	0.5610	0.2806	0.2983	0.3592
		Title + Category and Subcategory	0.5625	0.2845	0.3041	0.3631
		Title + Category and Subcategory + Abstract	0.5611	0.2849	0.3033	0.3623
	Dense Layer + Attention Pooling	Title	0.5765	0.2993	0.3239	0.3828
		Title + Category and Subcategory	0.5750	0.2985	0.3222	0.3814
		Title + Category and Subcategory + Abstract	0.5760	0.2986	0.3229	0.3817
	Multi-Head Self-Attention + Attention Pooling	Title	0.5759	0.2994	0.3240	0.3826
		Title + Category and Subcategory	0.5741	0.2974	0.3208	0.3801
		Title + Category and Subcategory + Abstract	0.5768	0.2994	0.3241	0.3829
text-embedding-3-large (OpenAI)	Average Pooling	Title	0.5117	0.2445	0.2514	0.3138
		Title + Category and Subcategory	0.5148	0.2470	0.2546	0.3167
		Title + Category and Subcategory + Abstract	0.5184	0.2574	0.2662	0.3270
	Attention Pooling	Title	0.5129	0.2492	0.2566	0.3189
		Title + Category and Subcategory	0.5152	0.2512	0.2592	0.3210
		Title + Category and Subcategory + Abstract	0.5177	0.2580	0.2675	0.3281
	Dense Layer + Attention Pooling	Title	0.5242	0.2701	0.2843	0.3436
		Title + Category and Subcategory	0.5232	0.2681	0.2817	0.3413
		Title + Category and Subcategory + Abstract	0.5232	0.2683	0.2821	0.3418
	Multi-Head Self-Attention + Attention Pooling	Title	0.5235	0.2684	0.2824	0.3419
		Title + Category and Subcategory	0.5228	0.2670	0.2803	0.3402
		Title + Category and Subcategory + Abstract	0.5228	0.2672	0.2806	0.3404

The following table gives the performance of baseline models for Task 4

Table A.9: Performance of Baseline Models on Candidate News Ranking for User

Baseline Model Type	Baseline Model Name	AUC	MRR	NDCG@5	NDCG@10
Feature Based Method	LibFM	0.5974	0.2633	0.2795	0.3429
Feature Based Method	DeepFM	0.5989	0.2621	0.2774	0.3406
Neural Recommendation Method	DKN	0.6290	0.2837	0.3099	0.3741
Neural Recommendation Method	NPA	0.6465	0.3001	0.3314	0.3947
Neural Recommendation Method	NAML	0.6612	0.3153	0.3488	0.4109
Neural Recommendation Method	LSTUR	0.6587	0.3078	0.3395	0.4015
Neural Recommendation Method	NRMS	0.6563	0.3096	0.3413	0.4052
BERT Enhanced Models	LSTUR + BERT	0.6828	0.3258	0.3599	0.4232
BERT Enhanced Models	NRMS + BERT	0.6860	0.3297	0.3655	0.4278
BERT Enhanced Models	UNBERT	0.6762	0.3172	0.3475	0.4102
BERT Enhanced Models	MINER	0.6961	0.3397	0.3762	0.4390

A.2 Multilingual News Recommendation

A.2.1 Task 1: User-News Classification

The results for training in English only and evaluation in the target language is given in the table below.

Table A.10: Performance of Models on Multilingual User-News Classification trained on English only

Generalized Embedding Model	User Encoder	Model Input	Language	F1 Score	AUC	Accuracy	Precision	Recall
multilingual-e5-large-instruct	Average Pooling	Title	en	0.1409	0.6461	0.8450	0.0909	0.3128
		Title + Category and Subcategory	en	0.1425	0.6477	0.8636	0.0957	0.2788
		Title + Category and Subcategory + Abstract	en	0.1482	0.6552	0.8604	0.0986	0.2989
	Attention Pooling	Title	en	0.1451	0.6484	0.8437	0.0933	0.3264
		Title + Category and Subcategory	en	0.1530	0.6526	0.8727	0.1049	0.2831
		Title + Category and Subcategory + Abstract	en	0.1554	0.6581	0.8641	0.1039	0.3075
	Dense Layer + Attention Pooling	Title	en	0.1530	0.6684	0.9117	0.1254	0.1961
		Title + Category and Subcategory	en	0.1558	0.6697	0.9075	0.1238	0.2102
		Title + Category and Subcategory + Abstract	en	0.1577	0.6731	0.9095	0.1268	0.2083
	Multi-Head Self-Attention + Attention Pooling	Title	en	0.1514	0.6601	0.9098	0.1225	0.1980
		Title + Category and Subcategory	en	0.1512	0.6619	0.9109	0.1234	0.1953

Generalized Embedding Model	User Encoder	Model Input	Language	F1 Score	AUC	Accuracy	Precision	Recall
multilingual-e5-large-instruct	Multi-Head Self-Attention + Attention Pooling	Title + Category and Subcategory + Abstract	en	0.1532	0.6651	0.9063	0.1210	0.2086
Qwen3-Embedding-0.6B	Average Pooling	Title	en	0.1340	0.6364	0.8345	0.0851	0.3152
		Title + Category and Subcategory	en	0.1197	0.6145	0.8324	0.0761	0.2804
		Title + Category and Subcategory + Abstract	en	0.1256	0.6261	0.8398	0.0807	0.2832
	Attention Pooling	Title	en	0.1391	0.6394	0.8464	0.0900	0.3054
		Title + Category and Subcategory	en	0.1267	0.6260	0.8376	0.0811	0.2900
		Title + Category and Subcategory + Abstract	en	0.1307	0.6324	0.8442	0.0845	0.2882
	Dense Layer + Attention Pooling	Title	en	0.1369	0.6641	0.9056	0.1089	0.1843
		Title + Category and Subcategory	en	0.1429	0.6665	0.8995	0.1093	0.2062
		Title + Category and Subcategory + Abstract	en	0.1430	0.6637	0.8991	0.1092	0.2070
	Multi-Head Self-Attention + Attention Pooling	Title	en	0.1379	0.6633	0.9038	0.1085	0.1893
		Title + Category and Subcategory	en	0.1406	0.6673	0.9020	0.1092	0.1973
		Title + Category and Subcategory + Abstract	en	0.1397	0.6634	0.9050	0.1105	0.1899
Qwen3-Embedding-4B	Average Pooling	Title	en	0.1400	0.6425	0.8635	0.0941	0.2736
		Title + Category and Subcategory	en	0.1200	0.6202	0.8572	0.0801	0.2396
		Title + Category and Subcategory + Abstract	en	0.1265	0.6258	0.8701	0.0870	0.2314
	Attention Pooling	Title	en	0.1462	0.6442	0.8762	0.1016	0.2607
		Title + Category and Subcategory	en	0.1318	0.6311	0.8622	0.0886	0.2574
		Title + Category and Subcategory + Abstract	en	0.1345	0.6331	0.8670	0.0914	0.2545
	Dense Layer + Attention Pooling	Title	en	0.1475	0.6676	0.9086	0.1188	0.1945
		Title + Category and Subcategory	en	0.1505	0.6698	0.9041	0.1176	0.2090
		Title + Category and Subcategory + Abstract	en	0.1475	0.6664	0.9019	0.1140	0.2088
	Multi-Head Self-Attention + Attention Pooling	Title	en	0.1480	0.6695	0.9081	0.1187	0.1965
		Title + Category and Subcategory	en	0.1508	0.6690	0.9016	0.1162	0.2150
		Title + Category and Subcategory + Abstract	en	0.1488	0.6656	0.8926	0.1097	0.2310
Qwen3-Embedding-8B	Average Pooling	Title	en	0.1368	0.6423	0.8556	0.0903	0.2815
		Title + Category and Subcategory	en	0.1184	0.6115	0.7957	0.0718	0.3376
		Title + Category and Subcategory + Abstract	en	0.1196	0.6160	0.8240	0.0751	0.2943
	Attention Pooling	Title	en	0.1411	0.6448	0.8648	0.0951	0.2733
		Title + Category and Subcategory	en	0.1320	0.6319	0.8508	0.0864	0.2792
		Title + Category and Subcategory + Abstract	en	0.1340	0.6346	0.8521	0.0879	0.2816
	Dense Layer + Attention Pooling	Title	en	0.1466	0.6721	0.9067	0.1166	0.1972

Generalized Embedding Model	User Encoder	Model Input	Language	F1 Score	AUC	Accuracy	Precision	Recall
Qwen3-Embedding-8B	Dense Layer + Attention Pooling	Title + Category and Subcategory	en	0.1476	0.6617	0.8839	0.1052	0.2472
		Title + Category and Subcategory + Abstract	en	0.1522	0.6678	0.8896	0.1106	0.2439
	Multi-Head Self-Attention + Attention Pooling	Title	en	0.1489	0.6717	0.9035	0.1160	0.2077
		Title + Category and Subcategory	en	0.1465	0.6609	0.8866	0.1056	0.2395
		Title + Category and Subcategory + Abstract	en	0.1516	0.6656	0.8906	0.1107	0.2406
text-embedding-3-large (OpenAI)	Average Pooling	Title	en	0.1117	0.6029	0.8443	0.0727	0.2409
		Title + Category and Subcategory	en	0.1220	0.6162	0.8565	0.0812	0.2454
		Title + Category and Subcategory + Abstract	en	0.1252	0.6148	0.8516	0.0823	0.2613
	Attention Pooling	Title	en	0.1218	0.6128	0.8625	0.0822	0.2345
		Title + Category and Subcategory	en	0.1350	0.6290	0.8766	0.0944	0.2369
		Title + Category and Subcategory + Abstract	en	0.1368	0.6257	0.8775	0.0959	0.2389
	Dense Layer + Attention Pooling	Title	en	0.1520	0.6721	0.9143	0.1271	0.1890
		Title + Category and Subcategory	en	0.1547	0.6696	0.9030	0.1197	0.2184
		Title + Category and Subcategory + Abstract	en	0.1532	0.6712	0.9095	0.1236	0.2015
	Multi-Head Self-Attention + Attention Pooling	Title	en	0.1537	0.6699	0.9045	0.1201	0.2135
		Title + Category and Subcategory	en	0.1528	0.6659	0.8997	0.1163	0.2225
		Title + Category and Subcategory + Abstract	en	0.1527	0.6681	0.9052	0.1199	0.2103
multilingual-e5-large-instruct	Average Pooling	Title	tr	0.1473	0.6505	0.8349	0.0932	0.3509
		Title + Category and Subcategory	tr	0.1484	0.6496	0.8606	0.0987	0.2990
		Title + Category and Subcategory + Abstract	tr	0.1463	0.6527	0.8548	0.0961	0.3062
	Attention Pooling	Title	tr	0.1505	0.6511	0.8368	0.0954	0.3557
		Title + Category and Subcategory	tr	0.1551	0.6521	0.8625	0.1033	0.3105
		Title + Category and Subcategory + Abstract	tr	0.1521	0.6541	0.8635	0.1018	0.3013
	Dense Layer + Attention Pooling	Title	tr	0.1329	0.6440	0.9070	0.1069	0.1754
		Title + Category and Subcategory	tr	0.1433	0.6520	0.9036	0.1121	0.1983
		Title + Category and Subcategory + Abstract	tr	0.1432	0.6534	0.9067	0.1142	0.1917
	Multi-Head Self-Attention + Attention Pooling	Title	tr	0.1338	0.6416	0.9052	0.1064	0.1801
		Title + Category and Subcategory	tr	0.1386	0.6493	0.9048	0.1096	0.1886
		Title + Category and Subcategory + Abstract	tr	0.1416	0.6540	0.9071	0.1133	0.1885
Qwen3-Embedding-0.6B	Average Pooling	Title	tr	0.1245	0.6335	0.8559	0.0827	0.2523
		Title + Category and Subcategory	tr	0.1287	0.6403	0.8413	0.0829	0.2885
		Title + Category and Subcategory + Abstract	tr	0.1301	0.6368	0.8588	0.0868	0.2598

Generalized Embedding Model	User Encoder	Model Input	Language	F1 Score	AUC	Accuracy	Precision	Recall
Qwen3-Embedding-0.6B	Attention Pooling	Title	tr	0.1286	0.6375	0.8627	0.0867	0.2494
		Title + Category and Subcategory	tr	0.1374	0.6488	0.8453	0.0888	0.3032
		Title + Category and Subcategory + Abstract	tr	0.1381	0.6442	0.8635	0.0929	0.2691
	Dense Layer + Attention Pooling	Title	tr	0.1361	0.6466	0.8947	0.1021	0.2040
		Title + Category and Subcategory	tr	0.1396	0.6470	0.8929	0.1036	0.2138
		Title + Category and Subcategory + Abstract	tr	0.1417	0.6514	0.8979	0.1076	0.2074
	Multi-Head Self-Attention + Attention Pooling	Title	tr	0.1389	0.6503	0.8930	0.1032	0.2123
		Title + Category and Subcategory	tr	0.1373	0.6451	0.8980	0.1046	0.1997
		Title + Category and Subcategory + Abstract	tr	0.1416	0.6515	0.8940	0.1055	0.2152
Qwen3-Embedding-4B	Average Pooling	Title	tr	0.1431	0.6550	0.8266	0.0895	0.3563
		Title + Category and Subcategory	tr	0.1450	0.6568	0.8537	0.0951	0.3052
		Title + Category and Subcategory + Abstract	tr	0.1578	0.6687	0.8592	0.1043	0.3247
	Attention Pooling	Title	tr	0.1489	0.6570	0.8459	0.0960	0.3318
		Title + Category and Subcategory	tr	0.1567	0.6622	0.8636	0.1046	0.3117
		Title + Category and Subcategory + Abstract	tr	0.1605	0.6677	0.8561	0.1052	0.3384
	Dense Layer + Attention Pooling	Title	tr	0.1550	0.6598	0.8985	0.1172	0.2291
		Title + Category and Subcategory	tr	0.1509	0.6561	0.8935	0.1116	0.2329
		Title + Category and Subcategory + Abstract	tr	0.1509	0.6616	0.8917	0.1108	0.2368
	Multi-Head Self-Attention + Attention Pooling	Title	tr	0.1554	0.6628	0.8948	0.1153	0.2381
		Title + Category and Subcategory	tr	0.1499	0.6562	0.8840	0.1067	0.2517
		Title + Category and Subcategory + Abstract	tr	0.1497	0.6603	0.8898	0.1090	0.2386
Qwen3-Embedding-8B	Average Pooling	Title	tr	0.1433	0.6506	0.8396	0.0915	0.3300
		Title + Category and Subcategory	tr	0.1200	0.6113	0.8429	0.0777	0.2636
		Title + Category and Subcategory + Abstract	tr	0.1288	0.6265	0.8303	0.0814	0.3086
	Attention Pooling	Title	tr	0.1461	0.6521	0.8474	0.0945	0.3213
		Title + Category and Subcategory	tr	0.1327	0.6275	0.8578	0.0882	0.2677
		Title + Category and Subcategory + Abstract	tr	0.1397	0.6402	0.8460	0.0903	0.3077
	Dense Layer + Attention Pooling	Title	tr	0.1529	0.6674	0.8966	0.1146	0.2297
		Title + Category and Subcategory	tr	0.1481	0.6568	0.8834	0.1053	0.2493
		Title + Category and Subcategory + Abstract	tr	0.1529	0.6685	0.8851	0.1092	0.2552
	Multi-Head Self-Attention + Attention Pooling	Title	tr	0.1545	0.6701	0.8991	0.1171	0.2270
		Title + Category and Subcategory	tr	0.1476	0.6535	0.8896	0.1075	0.2351

Generalized Embedding Model	User Encoder	Model Input	Language	F1 Score	AUC	Accuracy	Precision	Recall
Qwen3-Embedding-8B	Multi-Head Self-Attention + Attention Pooling	Title + Category and Subcategory + Abstract	tr	0.1487	0.6643	0.8873	0.1073	0.2424
text-embedding-3-large (OpenAI)	Average Pooling	Title	tr	0.0926	0.5706	0.7570	0.0546	0.3053
		Title + Category and Subcategory	tr	0.1040	0.5931	0.8312	0.0663	0.2412
		Title + Category and Subcategory + Abstract	tr	0.1132	0.5981	0.8183	0.0706	0.2855
	Attention Pooling	Title	tr	0.1021	0.5878	0.8138	0.0635	0.2605
		Title + Category and Subcategory	tr	0.1196	0.6124	0.8526	0.0789	0.2463
		Title + Category and Subcategory + Abstract	tr	0.1247	0.6125	0.8480	0.0814	0.2665
	Dense Layer + Attention Pooling	Title	tr	0.1353	0.6383	0.9051	0.1074	0.1826
		Title + Category and Subcategory	tr	0.1437	0.6470	0.8979	0.1090	0.2109
		Title + Category and Subcategory + Abstract	tr	0.1417	0.6528	0.8979	0.1076	0.2075
	Multi-Head Self-Attention + Attention Pooling	Title	tr	0.1366	0.6419	0.9031	0.1071	0.1886
		Title + Category and Subcategory	tr	0.1449	0.6514	0.8934	0.1075	0.2223
		Title + Category and Subcategory + Abstract	tr	0.1388	0.6543	0.9013	0.1075	0.1957
multilingual-e5-large-instruct	Average Pooling	Title	vi	0.1460	0.6435	0.8435	0.0938	0.3294
		Title + Category and Subcategory	vi	0.1457	0.6447	0.8487	0.0945	0.3177
		Title + Category and Subcategory + Abstract	vi	0.1434	0.6409	0.8509	0.0935	0.3070
	Attention Pooling	Title	vi	0.1476	0.6441	0.8375	0.0938	0.3462
		Title + Category and Subcategory	vi	0.1533	0.6466	0.8652	0.1029	0.3004
		Title + Category and Subcategory + Abstract	vi	0.1479	0.6432	0.8536	0.0968	0.3126
	Dense Layer + Attention Pooling	Title	vi	0.1267	0.6314	0.9126	0.1066	0.1560
		Title + Category and Subcategory	vi	0.1364	0.6400	0.9064	0.1091	0.1819
		Title + Category and Subcategory + Abstract	vi	0.1404	0.6471	0.9069	0.1123	0.1871
	Multi-Head Self-Attention + Attention Pooling	Title	vi	0.1273	0.6303	0.9145	0.1087	0.1535
		Title + Category and Subcategory	vi	0.1312	0.6359	0.9090	0.1072	0.1692
		Title + Category and Subcategory + Abstract	vi	0.1362	0.6443	0.9038	0.1072	0.1866
Qwen3-Embedding-0.6B	Average Pooling	Title	vi	0.1279	0.6261	0.8590	0.0854	0.2544
		Title + Category and Subcategory	vi	0.1389	0.6392	0.8614	0.0929	0.2752
		Title + Category and Subcategory + Abstract	vi	0.1324	0.6288	0.8588	0.0882	0.2650
	Attention Pooling	Title	vi	0.1317	0.6287	0.8598	0.0880	0.2618
		Title + Category and Subcategory	vi	0.1466	0.6456	0.8675	0.0993	0.2799
		Title + Category and Subcategory + Abstract	vi	0.1394	0.6363	0.8635	0.0937	0.2722
	Dense Layer + Attention Pooling	Title	vi	0.1420	0.6508	0.9006	0.1094	0.2024

Generalized Embedding Model	User Encoder	Model Input	Language	F1 Score	AUC	Accuracy	Precision	Recall
Qwen3-Embedding-0.6B	Dense Layer + Attention Pooling	Title + Category and Subcategory	vi	0.1453	0.6510	0.8975	0.1099	0.2144
		Title + Category and Subcategory + Abstract	vi	0.1430	0.6503	0.8983	0.1087	0.2088
	Multi-Head Self-Attention + Attention Pooling	Title	vi	0.1430	0.6520	0.8982	0.1086	0.2090
		Title + Category and Subcategory	vi	0.1447	0.6503	0.8968	0.1091	0.2148
		Title + Category and Subcategory + Abstract	vi	0.1431	0.6496	0.8867	0.1033	0.2327
Qwen3-Embedding-4B	Average Pooling	Title	vi	0.1459	0.6545	0.8516	0.0952	0.3118
		Title + Category and Subcategory	vi	0.1499	0.6621	0.8702	0.1021	0.2818
		Title + Category and Subcategory + Abstract	vi	0.1566	0.6633	0.8702	0.1064	0.2966
	Attention Pooling	Title	vi	0.1525	0.6565	0.8700	0.1037	0.2878
		Title + Category and Subcategory	vi	0.1593	0.6653	0.8699	0.1080	0.3035
		Title + Category and Subcategory + Abstract	vi	0.1613	0.6620	0.8706	0.1095	0.3062
	Dense Layer + Attention Pooling	Title	vi	0.1478	0.6640	0.9050	0.1164	0.2027
		Title + Category and Subcategory	vi	0.1499	0.6620	0.8928	0.1106	0.2326
		Title + Category and Subcategory + Abstract	vi	0.1487	0.6623	0.8900	0.1085	0.2365
	Multi-Head Self-Attention + Attention Pooling	Title	vi	0.1483	0.6658	0.8997	0.1132	0.2150
		Title + Category and Subcategory	vi	0.1480	0.6594	0.8916	0.1087	0.2317
		Title + Category and Subcategory + Abstract	vi	0.1481	0.6618	0.8915	0.1087	0.2322
Qwen3-Embedding-8B	Average Pooling	Title	vi	0.1404	0.6422	0.8650	0.0947	0.2714
		Title + Category and Subcategory	vi	0.1307	0.6396	0.8403	0.0839	0.2953
		Title + Category and Subcategory + Abstract	vi	0.1374	0.6437	0.8501	0.0897	0.2939
	Attention Pooling	Title	vi	0.1425	0.6435	0.8632	0.0956	0.2798
		Title + Category and Subcategory	vi	0.1436	0.6491	0.8485	0.0932	0.3127
		Title + Category and Subcategory + Abstract	vi	0.1490	0.6500	0.8605	0.0991	0.3004
	Dense Layer + Attention Pooling	Title	vi	0.1540	0.6692	0.8964	0.1153	0.2320
		Title + Category and Subcategory	vi	0.1511	0.6628	0.8854	0.1081	0.2510
		Title + Category and Subcategory + Abstract	vi	0.1553	0.6740	0.8817	0.1094	0.2677
	Multi-Head Self-Attention + Attention Pooling	Title	vi	0.1543	0.6703	0.8997	0.1174	0.2251
		Title + Category and Subcategory	vi	0.1509	0.6585	0.8872	0.1087	0.2465
		Title + Category and Subcategory + Abstract	vi	0.1520	0.6694	0.8898	0.1106	0.2430
text-embedding-3-large (OpenAI)	Average Pooling	Title	vi	0.1212	0.6091	0.8467	0.0790	0.2602
		Title + Category and Subcategory	vi	0.1224	0.6199	0.8588	0.0819	0.2424
		Title + Category and Subcategory + Abstract	vi	0.1370	0.6269	0.8533	0.0900	0.2864

Generalized Embedding Model	User Encoder	Model Input	Language	F1 Score	AUC	Accuracy	Precision	Recall
text-embedding-3-large (OpenAI)	Attention Pooling	Title	vi	0.1350	0.6189	0.8654	0.0913	0.2584
		Title + Category and Subcategory	vi	0.1395	0.6307	0.8678	0.0948	0.2637
		Title + Category and Subcategory + Abstract	vi	0.1470	0.6366	0.8698	0.1002	0.2761
	Dense Layer + Attention Pooling	Title	vi	0.1557	0.6545	0.9014	0.1194	0.2239
		Title + Category and Subcategory	vi	0.1560	0.6553	0.8970	0.1169	0.2341
		Title + Category and Subcategory + Abstract	vi	0.1582	0.6650	0.9005	0.1205	0.2301
	Multi-Head Self-Attention + Attention Pooling	Title	vi	0.1515	0.6545	0.9027	0.1173	0.2138
		Title + Category and Subcategory	vi	0.1549	0.6576	0.8940	0.1146	0.2391
		Title + Category and Subcategory + Abstract	vi	0.1555	0.6654	0.9004	0.1186	0.2256
multilingual-e5-large-instruct	Average Pooling	Title	zh	0.1346	0.6392	0.8112	0.0827	0.3613
		Title + Category and Subcategory	zh	0.1397	0.6467	0.8505	0.0912	0.2987
		Title + Category and Subcategory + Abstract	zh	0.1422	0.6491	0.8388	0.0907	0.3290
	Attention Pooling	Title	zh	0.1348	0.6394	0.8055	0.0823	0.3729
		Title + Category and Subcategory	zh	0.1459	0.6488	0.8587	0.0967	0.2970
		Title + Category and Subcategory + Abstract	zh	0.1468	0.6497	0.8519	0.0959	0.3136
	Dense Layer + Attention Pooling	Title	zh	0.1211	0.6261	0.8908	0.0900	0.1851
		Title + Category and Subcategory	zh	0.1431	0.6500	0.8888	0.1042	0.2285
		Title + Category and Subcategory + Abstract	zh	0.1499	0.6595	0.8945	0.1114	0.2288
	Multi-Head Self-Attention + Attention Pooling	Title	zh	0.1256	0.6292	0.8879	0.0919	0.1981
		Title + Category and Subcategory	zh	0.1421	0.6486	0.8880	0.1032	0.2283
		Title + Category and Subcategory + Abstract	zh	0.1471	0.6549	0.8919	0.1083	0.2294
Qwen3-Embedding-0.6B	Average Pooling	Title	zh	0.1175	0.6179	0.8163	0.0730	0.3007
		Title + Category and Subcategory	zh	0.1301	0.6379	0.8235	0.0813	0.3247
		Title + Category and Subcategory + Abstract	zh	0.1293	0.6258	0.8562	0.0857	0.2627
	Attention Pooling	Title	zh	0.1191	0.6205	0.8162	0.0740	0.3058
		Title + Category and Subcategory	zh	0.1365	0.6441	0.8294	0.0859	0.3320
		Title + Category and Subcategory + Abstract	zh	0.1339	0.6320	0.8585	0.0891	0.2690
	Dense Layer + Attention Pooling	Title	zh	0.1228	0.6337	0.8898	0.0907	0.1899
		Title + Category and Subcategory	zh	0.1382	0.6507	0.8930	0.1027	0.2111
		Title + Category and Subcategory + Abstract	zh	0.1394	0.6515	0.8874	0.1011	0.2245
	Multi-Head Self-Attention + Attention Pooling	Title	zh	0.1234	0.6377	0.8859	0.0897	0.1977
		Title + Category and Subcategory	zh	0.1392	0.6528	0.8871	0.1009	0.2246

Generalized Embedding Model	User Encoder	Model Input	Language	F1 Score	AUC	Accuracy	Precision	Recall
Qwen3-Embedding-0.6B	Multi-Head Self-Attention + Attention Pooling	Title + Category and Subcategory + Abstract	zh	0.1393	0.6516	0.8875	0.1011	0.2240
Qwen3-Embedding-4B	Average Pooling	Title	zh	0.1354	0.6410	0.8322	0.0856	0.3232
		Title + Category and Subcategory	zh	0.1374	0.6483	0.8669	0.0932	0.2608
		Title + Category and Subcategory + Abstract	zh	0.1441	0.6537	0.8645	0.0969	0.2807
	Attention Pooling	Title	zh	0.1385	0.6421	0.8631	0.0931	0.2709
		Title + Category and Subcategory	zh	0.1497	0.6548	0.8768	0.1040	0.2668
		Title + Category and Subcategory + Abstract	zh	0.1508	0.6559	0.8716	0.1031	0.2806
	Dense Layer + Attention Pooling	Title	zh	0.1308	0.6406	0.9002	0.1012	0.1848
		Title + Category and Subcategory	zh	0.1424	0.6469	0.8935	0.1058	0.2175
		Title + Category and Subcategory + Abstract	zh	0.1425	0.6539	0.8947	0.1065	0.2153
	Multi-Head Self-Attention + Attention Pooling	Title	zh	0.1294	0.6398	0.8953	0.0977	0.1915
		Title + Category and Subcategory	zh	0.1406	0.6469	0.8918	0.1038	0.2179
		Title + Category and Subcategory + Abstract	zh	0.1426	0.6540	0.8886	0.1037	0.2279
Qwen3-Embedding-8B	Average Pooling	Title	zh	0.1371	0.6467	0.8243	0.0856	0.3434
		Title + Category and Subcategory	zh	0.1194	0.6292	0.8154	0.0740	0.3080
		Title + Category and Subcategory + Abstract	zh	0.1320	0.6444	0.8284	0.0831	0.3211
	Attention Pooling	Title	zh	0.1406	0.6479	0.8408	0.0900	0.3205
		Title + Category and Subcategory	zh	0.1369	0.6461	0.8423	0.0880	0.3077
		Title + Category and Subcategory + Abstract	zh	0.1425	0.6529	0.8385	0.0909	0.3304
	Dense Layer + Attention Pooling	Title	zh	0.1480	0.6573	0.8941	0.1100	0.2265
		Title + Category and Subcategory	zh	0.1497	0.6565	0.8819	0.1058	0.2558
		Title + Category and Subcategory + Abstract	zh	0.1585	0.6693	0.8825	0.1118	0.2724
	Multi-Head Self-Attention + Attention Pooling	Title	zh	0.1494	0.6614	0.8885	0.1083	0.2409
		Title + Category and Subcategory	zh	0.1472	0.6511	0.8866	0.1060	0.2410
		Title + Category and Subcategory + Abstract	zh	0.1568	0.6675	0.8803	0.1099	0.2737
text-embedding-3-large (OpenAI)	Average Pooling	Title	zh	0.1078	0.5939	0.8097	0.0666	0.2829
		Title + Category and Subcategory	zh	0.1193	0.6150	0.8553	0.0792	0.2411
		Title + Category and Subcategory + Abstract	zh	0.1350	0.6332	0.8589	0.0899	0.2709
	Attention Pooling	Title	zh	0.1171	0.6069	0.8367	0.0750	0.2664
		Title + Category and Subcategory	zh	0.1333	0.6286	0.8679	0.0908	0.2500
		Title + Category and Subcategory + Abstract	zh	0.1437	0.6423	0.8679	0.0976	0.2729
	Dense Layer + Attention Pooling	Title	zh	0.1333	0.6464	0.8961	0.1008	0.1965

Generalized Embedding Model	User Encoder	Model Input	Language	F1 Score	AUC	Accuracy	Precision	Recall
text-embedding-3-large (OpenAI)	Dense Layer + Attention Pooling	Title + Category and Subcategory	zh	0.1485	0.6532	0.8875	0.1072	0.2415
		Title + Category and Subcategory + Abstract	zh	0.1519	0.6649	0.8867	0.1091	0.2497
	Multi-Head Self-Attention + Attention Pooling	Title	zh	0.1306	0.6473	0.8938	0.0978	0.1963
		Title + Category and Subcategory	zh	0.1470	0.6546	0.8869	0.1060	0.2399
		Title + Category and Subcategory + Abstract	zh	0.1503	0.6650	0.8897	0.1094	0.2401

The results for both training and evaluation in the target language is given in the table below.

Table A.11: Performance of Models on Multilingual User-News Classification trained and evaluated in target language

Generalized Embedding Model	User Encoder	Model Input	Language	F1 Score	AUC	Accuracy	Precision	Recall
multilingual-e5-large-instruct	Average Pooling	Title	en	0.1409	0.6461	0.8450	0.0909	0.3128
		Title + Category and Subcategory	en	0.1425	0.6477	0.8636	0.0957	0.2788
		Title + Category and Subcategory + Abstract	en	0.1482	0.6552	0.8604	0.0986	0.2989
	Attention Pooling	Title	en	0.1451	0.6484	0.8437	0.0933	0.3264
		Title + Category and Subcategory	en	0.1530	0.6526	0.8727	0.1049	0.2831
		Title + Category and Subcategory + Abstract	en	0.1554	0.6581	0.8641	0.1039	0.3075
	Dense Layer + Attention Pooling	Title	en	0.1530	0.6684	0.9117	0.1254	0.1961
		Title + Category and Subcategory	en	0.1558	0.6697	0.9075	0.1238	0.2102
		Title + Category and Subcategory + Abstract	en	0.1577	0.6731	0.9095	0.1268	0.2083
	Multi-Head Self-Attention + Attention Pooling	Title	en	0.1514	0.6601	0.9098	0.1225	0.1980
		Title + Category and Subcategory	en	0.1512	0.6619	0.9109	0.1234	0.1953
		Title + Category and Subcategory + Abstract	en	0.1532	0.6651	0.9063	0.1210	0.2086
Qwen3-Embedding-0.6B	Average Pooling	Title	en	0.1340	0.6364	0.8345	0.0851	0.3152
		Title + Category and Subcategory	en	0.1197	0.6145	0.8324	0.0761	0.2804
		Title + Category and Subcategory + Abstract	en	0.1256	0.6261	0.8398	0.0807	0.2832
	Attention Pooling	Title	en	0.1391	0.6394	0.8464	0.0900	0.3054
		Title + Category and Subcategory	en	0.1267	0.6260	0.8376	0.0811	0.2900
		Title + Category and Subcategory + Abstract	en	0.1307	0.6324	0.8442	0.0845	0.2882
	Dense Layer + Attention Pooling	Title	en	0.1369	0.6641	0.9056	0.1089	0.1843
		Title + Category and Subcategory	en	0.1429	0.6665	0.8995	0.1093	0.2062

Generalized Embedding Model	User Encoder	Model Input	Language	F1 Score	AUC	Accuracy	Precision	Recall
Qwen3-Embedding-0.6B	Dense Layer + Attention Pooling	Title + Category and Subcategory + Abstract	en	0.1430	0.6637	0.8991	0.1092	0.2070
	Multi-Head Self-Attention + Attention Pooling	Title	en	0.1379	0.6633	0.9038	0.1085	0.1893
		Title + Category and Subcategory	en	0.1406	0.6673	0.9020	0.1092	0.1973
		Title + Category and Subcategory + Abstract	en	0.1397	0.6634	0.9050	0.1105	0.1899
Qwen3-Embedding-4B	Average Pooling	Title	en	0.1400	0.6425	0.8635	0.0941	0.2736
		Title + Category and Subcategory	en	0.1200	0.6202	0.8572	0.0801	0.2396
		Title + Category and Subcategory + Abstract	en	0.1265	0.6258	0.8701	0.0870	0.2314
	Attention Pooling	Title	en	0.1462	0.6442	0.8762	0.1016	0.2607
		Title + Category and Subcategory	en	0.1318	0.6311	0.8622	0.0886	0.2574
		Title + Category and Subcategory + Abstract	en	0.1345	0.6331	0.8670	0.0914	0.2545
	Dense Layer + Attention Pooling	Title	en	0.1475	0.6676	0.9086	0.1188	0.1945
		Title + Category and Subcategory	en	0.1505	0.6698	0.9041	0.1176	0.2090
		Title + Category and Subcategory + Abstract	en	0.1475	0.6664	0.9019	0.1140	0.2088
	Multi-Head Self-Attention + Attention Pooling	Title	en	0.1480	0.6695	0.9081	0.1187	0.1965
		Title + Category and Subcategory	en	0.1508	0.6690	0.9016	0.1162	0.2150
		Title + Category and Subcategory + Abstract	en	0.1488	0.6656	0.8926	0.1097	0.2310
Qwen3-Embedding-8B	Average Pooling	Title	en	0.1368	0.6423	0.8556	0.0903	0.2815
		Title + Category and Subcategory	en	0.1184	0.6115	0.7957	0.0718	0.3376
		Title + Category and Subcategory + Abstract	en	0.1196	0.6160	0.8240	0.0751	0.2943
	Attention Pooling	Title	en	0.1411	0.6448	0.8648	0.0951	0.2733
		Title + Category and Subcategory	en	0.1320	0.6319	0.8508	0.0864	0.2792
		Title + Category and Subcategory + Abstract	en	0.1340	0.6346	0.8521	0.0879	0.2816
	Dense Layer + Attention Pooling	Title	en	0.1466	0.6721	0.9067	0.1166	0.1972
		Title + Category and Subcategory	en	0.1476	0.6617	0.8839	0.1052	0.2472
		Title + Category and Subcategory + Abstract	en	0.1522	0.6678	0.8896	0.1106	0.2439
	Multi-Head Self-Attention + Attention Pooling	Title	en	0.1489	0.6717	0.9035	0.1160	0.2077
		Title + Category and Subcategory	en	0.1465	0.6609	0.8866	0.1056	0.2395
		Title + Category and Subcategory + Abstract	en	0.1516	0.6656	0.8906	0.1107	0.2406
text-embedding-3-large (OpenAI)	Average Pooling	Title	en	0.1117	0.6029	0.8443	0.0727	0.2409
		Title + Category and Subcategory	en	0.1220	0.6162	0.8565	0.0812	0.2454
		Title + Category and Subcategory + Abstract	en	0.1252	0.6148	0.8516	0.0823	0.2613
	Attention Pooling	Title	en	0.1218	0.6128	0.8625	0.0822	0.2345

Generalized Embedding Model	User Encoder	Model Input	Language	F1 Score	AUC	Accuracy	Precision	Recall
text-embedding-3-large (OpenAI)	Attention Pooling	Title + Category and Subcategory	en	0.1350	0.6290	0.8766	0.0944	0.2369
		Title + Category and Subcategory + Abstract	en	0.1368	0.6257	0.8775	0.0959	0.2389
	Dense Layer + Attention Pooling	Title	en	0.1520	0.6721	0.9143	0.1271	0.1890
		Title + Category and Subcategory	en	0.1547	0.6696	0.9030	0.1197	0.2184
		Title + Category and Subcategory + Abstract	en	0.1532	0.6712	0.9095	0.1236	0.2015
	Multi-Head Self-Attention + Attention Pooling	Title	en	0.1537	0.6699	0.9045	0.1201	0.2135
		Title + Category and Subcategory	en	0.1528	0.6659	0.8997	0.1163	0.2225
		Title + Category and Subcategory + Abstract	en	0.1527	0.6681	0.9052	0.1199	0.2103
multilingual-e5-large-instruct	Average Pooling	Title	tr	0.1473	0.6505	0.8349	0.0932	0.3509
		Title + Category and Subcategory	tr	0.1484	0.6496	0.8606	0.0987	0.2990
		Title + Category and Subcategory + Abstract	tr	0.1463	0.6527	0.8548	0.0961	0.3062
	Attention Pooling	Title	tr	0.1485	0.6509	0.8383	0.0945	0.3472
		Title + Category and Subcategory	tr	0.1547	0.6520	0.8640	0.1035	0.3064
		Title + Category and Subcategory + Abstract	tr	0.1512	0.6540	0.8620	0.1008	0.3024
	Dense Layer + Attention Pooling	Title	tr	0.1547	0.6751	0.9048	0.1210	0.2145
		Title + Category and Subcategory	tr	0.1565	0.6740	0.9034	0.1213	0.2207
		Title + Category and Subcategory + Abstract	tr	0.1576	0.6773	0.9006	0.1202	0.2289
	Multi-Head Self-Attention + Attention Pooling	Title	tr	0.1500	0.6611	0.9086	0.1206	0.1985
		Title + Category and Subcategory	tr	0.1524	0.6611	0.9065	0.1206	0.2068
		Title + Category and Subcategory + Abstract	tr	0.1522	0.6633	0.9053	0.1196	0.2091
Qwen3-Embedding-0.6B	Average Pooling	Title	tr	0.1245	0.6335	0.8559	0.0827	0.2523
		Title + Category and Subcategory	tr	0.1287	0.6403	0.8413	0.0829	0.2885
		Title + Category and Subcategory + Abstract	tr	0.1301	0.6368	0.8588	0.0868	0.2598
	Attention Pooling	Title	tr	0.1329	0.6402	0.8676	0.0905	0.2496
		Title + Category and Subcategory	tr	0.1428	0.6511	0.8627	0.0957	0.2814
		Title + Category and Subcategory + Abstract	tr	0.1418	0.6462	0.8634	0.0952	0.2777
	Dense Layer + Attention Pooling	Title	tr	0.1339	0.6585	0.9128	0.1122	0.1659
		Title + Category and Subcategory	tr	0.1480	0.6654	0.9074	0.1182	0.1978
		Title + Category and Subcategory + Abstract	tr	0.1523	0.6713	0.9070	0.1209	0.2055
	Multi-Head Self-Attention + Attention Pooling	Title	tr	0.1317	0.6590	0.9135	0.1112	0.1614
		Title + Category and Subcategory	tr	0.1466	0.6664	0.9094	0.1187	0.1916

Generalized Embedding Model	User Encoder	Model Input	Language	F1 Score	AUC	Accuracy	Precision	Recall
Qwen3-Embedding-0.6B	Multi-Head Self-Attention + Attention Pooling	Title + Category and Subcategory + Abstract	tr	0.1509	0.6714	0.9064	0.1195	0.2047
Qwen3-Embedding-4B	Average Pooling	Title	tr	0.1431	0.6550	0.8266	0.0895	0.3563
		Title + Category and Subcategory	tr	0.1450	0.6568	0.8537	0.0951	0.3052
		Title + Category and Subcategory + Abstract	tr	0.1578	0.6687	0.8592	0.1043	0.3247
	Attention Pooling	Title	tr	0.1480	0.6574	0.8347	0.0936	0.3533
		Title + Category and Subcategory	tr	0.1582	0.6615	0.8626	0.1053	0.3178
		Title + Category and Subcategory + Abstract	tr	0.1602	0.6698	0.8615	0.1063	0.3251
	Dense Layer + Attention Pooling	Title	tr	0.1604	0.6778	0.8990	0.1212	0.2374
		Title + Category and Subcategory	tr	0.1627	0.6800	0.8981	0.1221	0.2438
		Title + Category and Subcategory + Abstract	tr	0.1593	0.6782	0.8965	0.1189	0.2413
	Multi-Head Self-Attention + Attention Pooling	Title	tr	0.1601	0.6766	0.8983	0.1205	0.2385
		Title + Category and Subcategory	tr	0.1615	0.6775	0.8979	0.1212	0.2419
		Title + Category and Subcategory + Abstract	tr	0.1581	0.6742	0.8984	0.1192	0.2348
Qwen3-Embedding-8B	Average Pooling	Title	tr	0.1433	0.6506	0.8396	0.0915	0.3300
		Title + Category and Subcategory	tr	0.1200	0.6113	0.8429	0.0777	0.2636
		Title + Category and Subcategory + Abstract	tr	0.1288	0.6265	0.8303	0.0814	0.3086
	Attention Pooling	Title	tr	0.1459	0.6515	0.8485	0.0946	0.3184
		Title + Category and Subcategory	tr	0.1344	0.6294	0.8616	0.0901	0.2645
		Title + Category and Subcategory + Abstract	tr	0.1403	0.6401	0.8474	0.0910	0.3064
	Dense Layer + Attention Pooling	Title	tr	0.1506	0.6693	0.9056	0.1187	0.2059
		Title + Category and Subcategory	tr	0.1528	0.6709	0.8965	0.1145	0.2297
		Title + Category and Subcategory + Abstract	tr	0.1519	0.6739	0.8948	0.1130	0.2319
	Multi-Head Self-Attention + Attention Pooling	Title	tr	0.1544	0.6720	0.9009	0.1182	0.2227
		Title + Category and Subcategory	tr	0.1533	0.6700	0.8896	0.1113	0.2459
		Title + Category and Subcategory + Abstract	tr	0.1472	0.6660	0.8936	0.1091	0.2258
text-embedding-3-large (OpenAI)	Average Pooling	Title	tr	0.0926	0.5706	0.7570	0.0546	0.3053
		Title + Category and Subcategory	tr	0.1040	0.5931	0.8312	0.0663	0.2412
		Title + Category and Subcategory + Abstract	tr	0.1132	0.5981	0.8183	0.0706	0.2855
	Attention Pooling	Title	tr	0.1116	0.5951	0.8329	0.0712	0.2583
		Title + Category and Subcategory	tr	0.1260	0.6176	0.8471	0.0821	0.2713
		Title + Category and Subcategory + Abstract	tr	0.1286	0.6163	0.8494	0.0841	0.2735
	Dense Layer + Attention Pooling	Title	tr	0.1430	0.6590	0.9075	0.1147	0.1898

Generalized Embedding Model	User Encoder	Model Input	Language	F1 Score	AUC	Accuracy	Precision	Recall
text-embedding-3-large (OpenAI)	Dense Layer + Attention Pooling	Title + Category and Subcategory	tr	0.1494	0.6637	0.9048	0.1172	0.2057
		Title + Category and Subcategory + Abstract	tr	0.1466	0.6646	0.9124	0.1214	0.1852
	Multi-Head Self-Attention + Attention Pooling	Title	tr	0.1341	0.6484	0.9151	0.1145	0.1617
		Title + Category and Subcategory	tr	0.1484	0.6588	0.9002	0.1136	0.2140
		Title + Category and Subcategory + Abstract	tr	0.1467	0.6607	0.9062	0.1164	0.1984
multilingual-e5-large-instruct	Average Pooling	Title	vi	0.1460	0.6435	0.8435	0.0938	0.3294
		Title + Category and Subcategory	vi	0.1457	0.6447	0.8487	0.0945	0.3177
		Title + Category and Subcategory + Abstract	vi	0.1434	0.6409	0.8509	0.0935	0.3070
	Attention Pooling	Title	vi	0.1466	0.6438	0.8443	0.0943	0.3291
		Title + Category and Subcategory	vi	0.1522	0.6465	0.8643	0.1020	0.2997
		Title + Category and Subcategory + Abstract	vi	0.1479	0.6431	0.8541	0.0969	0.3114
	Dense Layer + Attention Pooling	Title	vi	0.1525	0.6681	0.8993	0.1159	0.2231
		Title + Category and Subcategory	vi	0.1550	0.6688	0.9024	0.1195	0.2202
		Title + Category and Subcategory + Abstract	vi	0.1561	0.6739	0.9042	0.1215	0.2180
	Multi-Head Self-Attention + Attention Pooling	Title	vi	0.1479	0.6524	0.9056	0.1168	0.2017
		Title + Category and Subcategory	vi	0.1494	0.6546	0.9072	0.1191	0.2006
		Title + Category and Subcategory + Abstract	vi	0.1515	0.6575	0.9040	0.1182	0.2109
Qwen3-Embedding-0.6B	Average Pooling	Title	vi	0.1279	0.6261	0.8590	0.0854	0.2544
		Title + Category and Subcategory	vi	0.1389	0.6392	0.8614	0.0929	0.2752
		Title + Category and Subcategory + Abstract	vi	0.1324	0.6288	0.8588	0.0882	0.2650
	Attention Pooling	Title	vi	0.1324	0.6292	0.8643	0.0894	0.2547
		Title + Category and Subcategory	vi	0.1470	0.6460	0.8655	0.0990	0.2851
		Title + Category and Subcategory + Abstract	vi	0.1404	0.6365	0.8665	0.0951	0.2685
	Dense Layer + Attention Pooling	Title	vi	0.1411	0.6592	0.9084	0.1140	0.1852
		Title + Category and Subcategory	vi	0.1514	0.6632	0.9016	0.1166	0.2161
		Title + Category and Subcategory + Abstract	vi	0.1505	0.6657	0.9014	0.1158	0.2149
	Multi-Head Self-Attention + Attention Pooling	Title	vi	0.1408	0.6598	0.9081	0.1135	0.1854
		Title + Category and Subcategory	vi	0.1512	0.6652	0.9008	0.1159	0.2175
		Title + Category and Subcategory + Abstract	vi	0.1507	0.6663	0.8980	0.1139	0.2226
Qwen3-Embedding-4B	Average Pooling	Title	vi	0.1459	0.6545	0.8516	0.0952	0.3118
		Title + Category and Subcategory	vi	0.1499	0.6621	0.8702	0.1021	0.2818
		Title + Category and Subcategory + Abstract	vi	0.1566	0.6633	0.8702	0.1064	0.2966

Generalized Embedding Model	User Encoder	Model Input	Language	F1 Score	AUC	Accuracy	Precision	Recall
Qwen3-Embedding-4B	Attention Pooling	Title	vi	0.1515	0.6558	0.8654	0.1019	0.2958
		Title + Category and Subcategory	vi	0.1590	0.6650	0.8652	0.1065	0.3137
		Title + Category and Subcategory + Abstract	vi	0.1595	0.6639	0.8656	0.1069	0.3138
	Dense Layer + Attention Pooling	Title	vi	0.1532	0.6737	0.9028	0.1186	0.2165
		Title + Category and Subcategory	vi	0.1562	0.6731	0.8971	0.1171	0.2344
		Title + Category and Subcategory + Abstract	vi	0.1562	0.6764	0.8984	0.1179	0.2315
	Multi-Head Self-Attention + Attention Pooling	Title	vi	0.1537	0.6752	0.9021	0.1185	0.2189
		Title + Category and Subcategory	vi	0.1570	0.6731	0.8979	0.1181	0.2341
		Title + Category and Subcategory + Abstract	vi	0.1558	0.6744	0.8962	0.1163	0.2356
Qwen3-Embedding-8B	Average Pooling	Title	vi	0.1404	0.6422	0.8650	0.0947	0.2714
		Title + Category and Subcategory	vi	0.1307	0.6396	0.8403	0.0839	0.2953
		Title + Category and Subcategory + Abstract	vi	0.1374	0.6437	0.8501	0.0897	0.2939
	Attention Pooling	Title	vi	0.1414	0.6429	0.8618	0.0946	0.2800
		Title + Category and Subcategory	vi	0.1459	0.6492	0.8679	0.0990	0.2776
		Title + Category and Subcategory + Abstract	vi	0.1494	0.6495	0.8673	0.1010	0.2867
	Dense Layer + Attention Pooling	Title	vi	0.1525	0.6720	0.9116	0.1249	0.1956
		Title + Category and Subcategory	vi	0.1553	0.6713	0.8952	0.1155	0.2370
		Title + Category and Subcategory + Abstract	vi	0.1534	0.6733	0.8937	0.1134	0.2369
	Multi-Head Self-Attention + Attention Pooling	Title	vi	0.1541	0.6733	0.9086	0.1234	0.2049
		Title + Category and Subcategory	vi	0.1532	0.6678	0.8909	0.1119	0.2429
		Title + Category and Subcategory + Abstract	vi	0.1531	0.6730	0.8973	0.1151	0.2286
text-embedding-3-large (OpenAI)	Average Pooling	Title	vi	0.1212	0.6091	0.8467	0.0790	0.2602
		Title + Category and Subcategory	vi	0.1224	0.6199	0.8588	0.0819	0.2424
		Title + Category and Subcategory + Abstract	vi	0.1370	0.6269	0.8533	0.0900	0.2864
	Attention Pooling	Title	vi	0.1323	0.6168	0.8674	0.0901	0.2486
		Title + Category and Subcategory	vi	0.1422	0.6299	0.8677	0.0965	0.2698
		Title + Category and Subcategory + Abstract	vi	0.1468	0.6364	0.8677	0.0995	0.2800
	Dense Layer + Attention Pooling	Title	vi	0.1499	0.6657	0.9134	0.1246	0.1879
		Title + Category and Subcategory	vi	0.1539	0.6683	0.9006	0.1177	0.2224
		Title + Category and Subcategory + Abstract	vi	0.1556	0.6738	0.9081	0.1242	0.2084
	Multi-Head Self-Attention + Attention Pooling	Title	vi	0.1520	0.6610	0.9068	0.1206	0.2056
		Title + Category and Subcategory	vi	0.1528	0.6616	0.8937	0.1130	0.2360

Generalized Embedding Model	User Encoder	Model Input	Language	F1 Score	AUC	Accuracy	Precision	Recall
text-embedding-3-large (OpenAI)	Multi-Head Self-Attention + Attention Pooling	Title + Category and Subcategory + Abstract	vi	0.1547	0.6696	0.9055	0.1215	0.2130
multilingual-e5-large-instruct	Average Pooling	Title	zh	0.1346	0.6392	0.8112	0.0827	0.3613
		Title + Category and Subcategory	zh	0.1397	0.6467	0.8505	0.0912	0.2987
		Title + Category and Subcategory + Abstract	zh	0.1422	0.6491	0.8388	0.0907	0.3290
	Attention Pooling	Title	zh	0.1355	0.6401	0.8087	0.0830	0.3690
		Title + Category and Subcategory	zh	0.1455	0.6493	0.8503	0.0947	0.3135
		Title + Category and Subcategory + Abstract	zh	0.1456	0.6498	0.8494	0.0946	0.3156
	Dense Layer + Attention Pooling	Title	zh	0.1367	0.6586	0.9012	0.1060	0.1925
		Title + Category and Subcategory	zh	0.1530	0.6671	0.8815	0.1078	0.2634
		Title + Category and Subcategory + Abstract	zh	0.1520	0.6676	0.9030	0.1179	0.2141
	Multi-Head Self-Attention + Attention Pooling	Title	zh	0.1358	0.6528	0.9019	0.1057	0.1896
		Title + Category and Subcategory	zh	0.1535	0.6642	0.8876	0.1106	0.2508
		Title + Category and Subcategory + Abstract	zh	0.1506	0.6616	0.9062	0.1191	0.2047
Qwen3-Embedding-0.6B	Average Pooling	Title	zh	0.1175	0.6179	0.8163	0.0730	0.3007
		Title + Category and Subcategory	zh	0.1301	0.6379	0.8235	0.0813	0.3247
		Title + Category and Subcategory + Abstract	zh	0.1293	0.6258	0.8562	0.0857	0.2627
	Attention Pooling	Title	zh	0.1201	0.6211	0.8336	0.0765	0.2794
		Title + Category and Subcategory	zh	0.1374	0.6449	0.8306	0.0866	0.3320
		Title + Category and Subcategory + Abstract	zh	0.1349	0.6326	0.8632	0.0908	0.2625
	Dense Layer + Attention Pooling	Title	zh	0.1256	0.6486	0.9008	0.0978	0.1754
		Title + Category and Subcategory	zh	0.1453	0.6624	0.8953	0.1087	0.2189
		Title + Category and Subcategory + Abstract	zh	0.1474	0.6675	0.8908	0.1080	0.2323
	Multi-Head Self-Attention + Attention Pooling	Title	zh	0.1217	0.6476	0.9066	0.0985	0.1592
		Title + Category and Subcategory	zh	0.1443	0.6615	0.8946	0.1077	0.2187
		Title + Category and Subcategory + Abstract	zh	0.1447	0.6673	0.9014	0.1118	0.2053
Qwen3-Embedding-4B	Average Pooling	Title	zh	0.1354	0.6410	0.8322	0.0856	0.3232
		Title + Category and Subcategory	zh	0.1374	0.6483	0.8669	0.0932	0.2608
		Title + Category and Subcategory + Abstract	zh	0.1441	0.6537	0.8645	0.0969	0.2807
	Attention Pooling	Title	zh	0.1395	0.6427	0.8595	0.0929	0.2803
		Title + Category and Subcategory	zh	0.1499	0.6553	0.8624	0.1001	0.2985
		Title + Category and Subcategory + Abstract	zh	0.1504	0.6566	0.8763	0.1043	0.2694
	Dense Layer + Attention Pooling	Title	zh	0.1412	0.6627	0.9056	0.1120	0.1909

Generalized Embedding Model	User Encoder	Model Input	Language	F1 Score	AUC	Accuracy	Precision	Recall
Qwen3-Embedding-4B	Dense Layer + Attention Pooling	Title + Category and Subcategory	zh	0.1567	0.6713	0.8920	0.1148	0.2470
		Title + Category and Subcategory + Abstract	zh	0.1553	0.6738	0.8911	0.1134	0.2465
	Multi-Head Self-Attention + Attention Pooling	Title	zh	0.1417	0.6633	0.8964	0.1068	0.2105
		Title + Category and Subcategory	zh	0.1557	0.6681	0.8960	0.1162	0.2361
		Title + Category and Subcategory + Abstract	zh	0.1550	0.6742	0.8851	0.1105	0.2594
Qwen3-Embedding-8B	Average Pooling	Title	zh	0.1371	0.6467	0.8243	0.0856	0.3434
		Title + Category and Subcategory	zh	0.1194	0.6292	0.8154	0.0740	0.3080
		Title + Category and Subcategory + Abstract	zh	0.1320	0.6444	0.8284	0.0831	0.3211
	Attention Pooling	Title	zh	0.1431	0.6475	0.8443	0.0921	0.3199
		Title + Category and Subcategory	zh	0.1412	0.6482	0.8419	0.0906	0.3199
		Title + Category and Subcategory + Abstract	zh	0.1430	0.6538	0.8410	0.0915	0.3265
	Dense Layer + Attention Pooling	Title	zh	0.1464	0.6661	0.8970	0.1104	0.2172
		Title + Category and Subcategory	zh	0.1569	0.6692	0.8926	0.1152	0.2458
		Title + Category and Subcategory + Abstract	zh	0.1566	0.6737	0.8946	0.1160	0.2406
	Multi-Head Self-Attention + Attention Pooling	Title	zh	0.1474	0.6694	0.8995	0.1124	0.2138
		Title + Category and Subcategory	zh	0.1582	0.6701	0.8920	0.1158	0.2498
		Title + Category and Subcategory + Abstract	zh	0.1566	0.6738	0.8918	0.1146	0.2472
text-embedding-3-large (OpenAI)	Average Pooling	Title	zh	0.1078	0.5939	0.8097	0.0666	0.2829
		Title + Category and Subcategory	zh	0.1193	0.6150	0.8553	0.0792	0.2411
		Title + Category and Subcategory + Abstract	zh	0.1350	0.6332	0.8589	0.0899	0.2709
	Attention Pooling	Title	zh	0.1206	0.6096	0.8516	0.0795	0.2505
		Title + Category and Subcategory	zh	0.1364	0.6318	0.8675	0.0928	0.2576
		Title + Category and Subcategory + Abstract	zh	0.1444	0.6432	0.8705	0.0987	0.2690
	Dense Layer + Attention Pooling	Title	zh	0.1345	0.6611	0.9016	0.1047	0.1883
		Title + Category and Subcategory	zh	0.1529	0.6670	0.8945	0.1135	0.2342
		Title + Category and Subcategory + Abstract	zh	0.1543	0.6743	0.9075	0.1227	0.2077
	Multi-Head Self-Attention + Attention Pooling	Title	zh	0.1328	0.6577	0.9021	0.1038	0.1845
		Title + Category and Subcategory	zh	0.1518	0.6656	0.8954	0.1132	0.2304
		Title + Category and Subcategory + Abstract	zh	0.1541	0.6721	0.9001	0.1175	0.2238

A.2.2 Task 2: All User Classification for given News

The results for training in English only and evaluation in the target language is given in the table below.

Table A.12: Performance of Models on Multilingual Classification of all User for given News trained on English only

Generalized Embedding Model	User Encoder	Model Input	Language	F1 Score	AUC	Accuracy	Precision	Recall
multilingual-e5-large-instruct	Average Pooling	Title	en	0.0389	0.5008	0.8140	0.0404	0.1476
		Title + Category and Subcategory	en	0.0384	0.5019	0.8230	0.0399	0.1354
		Title + Category and Subcategory + Abstract	en	0.0393	0.5002	0.8004	0.0390	0.1526
	Attention Pooling	Title	en	0.0383	0.5019	0.8165	0.0403	0.1471
		Title + Category and Subcategory	en	0.0357	0.5017	0.8362	0.0397	0.1223
		Title + Category and Subcategory + Abstract	en	0.0388	0.5012	0.8077	0.0399	0.1478
	Dense Layer + Attention Pooling	Title	en	0.0327	0.5013	0.8630	0.0388	0.0923
		Title + Category and Subcategory	en	0.0356	0.5014	0.8493	0.0386	0.1047
		Title + Category and Subcategory + Abstract	en	0.0354	0.5010	0.8535	0.0397	0.0997
	Multi-Head Self-Attention + Attention Pooling	Title	en	0.0349	0.5011	0.8623	0.0411	0.0955
		Title + Category and Subcategory	en	0.0342	0.5001	0.8569	0.0376	0.0979
		Title + Category and Subcategory + Abstract	en	0.0369	0.5023	0.8475	0.0400	0.1080
Qwen3-Embedding-0.6B	Average Pooling	Title	en	0.0392	0.5097	0.8168	0.0409	0.1508
		Title + Category and Subcategory	en	0.0420	0.5064	0.7846	0.0401	0.1732
		Title + Category and Subcategory + Abstract	en	0.0414	0.5034	0.8027	0.0436	0.1630
	Attention Pooling	Title	en	0.0390	0.5101	0.8211	0.0416	0.1471
		Title + Category and Subcategory	en	0.0416	0.5060	0.7860	0.0401	0.1727
		Title + Category and Subcategory + Abstract	en	0.0411	0.5031	0.8005	0.0420	0.1637
	Dense Layer + Attention Pooling	Title	en	0.0326	0.5090	0.8546	0.0376	0.0980
		Title + Category and Subcategory	en	0.0383	0.5057	0.8342	0.0416	0.1224
		Title + Category and Subcategory + Abstract	en	0.0364	0.5070	0.8473	0.0407	0.1130
	Multi-Head Self-Attention + Attention Pooling	Title	en	0.0322	0.5073	0.8575	0.0398	0.0971
		Title + Category and Subcategory	en	0.0380	0.5054	0.8430	0.0424	0.1171
		Title + Category and Subcategory + Abstract	en	0.0344	0.5088	0.8587	0.0411	0.1023
Qwen3-Embedding-4B	Average Pooling	Title	en	0.0301	0.5026	0.8454	0.0362	0.1200
		Title + Category and Subcategory	en	0.0378	0.5015	0.8236	0.0427	0.1379

Generalized Embedding Model	User Encoder	Model Input	Language	F1 Score	AUC	Accuracy	Precision	Recall
Qwen3-Embedding-4B	Average Pooling	Title + Category and Subcategory + Abstract	en	0.0350	0.5033	0.8316	0.0415	0.1274
	Attention Pooling	Title	en	0.0297	0.5017	0.8551	0.0364	0.1116
		Title + Category and Subcategory	en	0.0362	0.5014	0.8223	0.0400	0.1368
		Title + Category and Subcategory + Abstract	en	0.0348	0.5019	0.8241	0.0404	0.1337
	Dense Layer + Attention Pooling	Title	en	0.0281	0.4970	0.8669	0.0351	0.0858
		Title + Category and Subcategory	en	0.0323	0.4986	0.8526	0.0371	0.1007
		Title + Category and Subcategory + Abstract	en	0.0350	0.5020	0.8566	0.0412	0.1044
	Multi-Head Self-Attention + Attention Pooling	Title	en	0.0281	0.4974	0.8710	0.0367	0.0845
		Title + Category and Subcategory	en	0.0336	0.4987	0.8550	0.0396	0.1017
		Title + Category and Subcategory + Abstract	en	0.0371	0.5018	0.8489	0.0418	0.1141
Qwen3-Embedding-8B	Average Pooling	Title	en	0.0373	0.5017	0.8335	0.0411	0.1295
		Title + Category and Subcategory	en	0.0476	0.5045	0.7759	0.0449	0.1990
		Title + Category and Subcategory + Abstract	en	0.0408	0.5048	0.8015	0.0401	0.1635
	Attention Pooling	Title	en	0.0349	0.5014	0.8385	0.0388	0.1211
		Title + Category and Subcategory	en	0.0390	0.5048	0.8189	0.0414	0.1410
		Title + Category and Subcategory + Abstract	en	0.0379	0.5055	0.8252	0.0422	0.1327
	Dense Layer + Attention Pooling	Title	en	0.0347	0.5027	0.8579	0.0401	0.0988
		Title + Category and Subcategory	en	0.0396	0.5053	0.8238	0.0394	0.1303
		Title + Category and Subcategory + Abstract	en	0.0395	0.5037	0.8296	0.0397	0.1289
	Multi-Head Self-Attention + Attention Pooling	Title	en	0.0354	0.5025	0.8628	0.0423	0.0977
		Title + Category and Subcategory	en	0.0404	0.5047	0.8327	0.0414	0.1259
		Title + Category and Subcategory + Abstract	en	0.0415	0.5045	0.8349	0.0418	0.1283
text-embedding-3-large (OpenAI)	Average Pooling	Title	en	0.0400	0.5000	0.8071	0.0383	0.1493
		Title + Category and Subcategory	en	0.0428	0.5018	0.8008	0.0400	0.1565
		Title + Category and Subcategory + Abstract	en	0.0413	0.5000	0.8084	0.0395	0.1513
	Attention Pooling	Title	en	0.0374	0.5011	0.8271	0.0380	0.1293
		Title + Category and Subcategory	en	0.0386	0.5032	0.8251	0.0389	0.1309
		Title + Category and Subcategory + Abstract	en	0.0377	0.5013	0.8357	0.0386	0.1223
	Dense Layer + Attention Pooling	Title	en	0.0370	0.4988	0.8622	0.0429	0.0979
		Title + Category and Subcategory	en	0.0415	0.5016	0.8334	0.0437	0.1240
		Title + Category and Subcategory + Abstract	en	0.0386	0.4994	0.8493	0.0420	0.1083
		Title	en	0.0389	0.4994	0.8543	0.0437	0.1062

Generalized Embedding Model	User Encoder	Model Input	Language	F1 Score	AUC	Accuracy	Precision	Recall
text-embedding-3-large (OpenAI)	Multi-Head Self-Attention + Attention Pooling	Title + Category and Subcategory	en	0.0416	0.5000	0.8355	0.0444	0.1243
		Title + Category and Subcategory + Abstract	en	0.0390	0.5009	0.8470	0.0423	0.1115
multilingual-e5-large-instruct	Average Pooling	Title	tr	0.0396	0.5014	0.7962	0.0391	0.1654
		Title + Category and Subcategory	tr	0.0394	0.5009	0.8022	0.0388	0.1541
		Title + Category and Subcategory + Abstract	tr	0.0405	0.4999	0.7783	0.0383	0.1752
	Attention Pooling	Title	tr	0.0367	0.5029	0.8012	0.0376	0.1566
		Title + Category and Subcategory	tr	0.0374	0.5016	0.8061	0.0386	0.1461
		Title + Category and Subcategory + Abstract	tr	0.0401	0.5019	0.7932	0.0404	0.1642
	Dense Layer + Attention Pooling	Title	tr	0.0303	0.5003	0.8741	0.0384	0.0814
		Title + Category and Subcategory	tr	0.0334	0.5015	0.8595	0.0378	0.0955
		Title + Category and Subcategory + Abstract	tr	0.0317	0.5019	0.8624	0.0366	0.0912
	Multi-Head Self-Attention + Attention Pooling	Title	tr	0.0296	0.5004	0.8772	0.0361	0.0784
		Title + Category and Subcategory	tr	0.0340	0.4997	0.8660	0.0397	0.0928
		Title + Category and Subcategory + Abstract	tr	0.0333	0.5013	0.8662	0.0383	0.0922
Qwen3-Embedding-0.6B	Average Pooling	Title	tr	0.0380	0.5034	0.8319	0.0421	0.1325
		Title + Category and Subcategory	tr	0.0420	0.5032	0.8046	0.0412	0.1578
		Title + Category and Subcategory + Abstract	tr	0.0386	0.5011	0.8099	0.0394	0.1441
	Attention Pooling	Title	tr	0.0374	0.5026	0.8343	0.0418	0.1285
		Title + Category and Subcategory	tr	0.0437	0.5032	0.7976	0.0423	0.1670
		Title + Category and Subcategory + Abstract	tr	0.0394	0.5014	0.8088	0.0396	0.1478
	Dense Layer + Attention Pooling	Title	tr	0.0369	0.5008	0.8497	0.0442	0.1079
		Title + Category and Subcategory	tr	0.0409	0.5000	0.8225	0.0431	0.1266
		Title + Category and Subcategory + Abstract	tr	0.0362	0.5019	0.8401	0.0383	0.1113
	Multi-Head Self-Attention + Attention Pooling	Title	tr	0.0366	0.5006	0.8495	0.0431	0.1107
		Title + Category and Subcategory	tr	0.0390	0.5017	0.8325	0.0419	0.1192
		Title + Category and Subcategory + Abstract	tr	0.0381	0.5027	0.8373	0.0399	0.1160
Qwen3-Embedding-4B	Average Pooling	Title	tr	0.0364	0.4997	0.7915	0.0364	0.1665
		Title + Category and Subcategory	tr	0.0382	0.4997	0.8274	0.0432	0.1387
		Title + Category and Subcategory + Abstract	tr	0.0362	0.4994	0.8134	0.0368	0.1404
	Attention Pooling	Title	tr	0.0369	0.5010	0.8052	0.0396	0.1558
		Title + Category and Subcategory	tr	0.0371	0.5009	0.8304	0.0421	0.1358
		Title + Category and Subcategory + Abstract	tr	0.0384	0.5013	0.8059	0.0399	0.1513

Generalized Embedding Model	User Encoder	Model Input	Language	F1 Score	AUC	Accuracy	Precision	Recall
Qwen3-Embedding-4B	Dense Layer + Attention Pooling	Title	tr	0.0318	0.4981	0.8513	0.0379	0.1037
		Title + Category and Subcategory	tr	0.0352	0.4998	0.8379	0.0375	0.1138
		Title + Category and Subcategory + Abstract	tr	0.0365	0.5004	0.8394	0.0404	0.1148
	Multi-Head Self-Attention + Attention Pooling	Title	tr	0.0346	0.4975	0.8533	0.0393	0.1079
		Title + Category and Subcategory	tr	0.0388	0.4991	0.8350	0.0409	0.1218
		Title + Category and Subcategory + Abstract	tr	0.0361	0.4995	0.8456	0.0401	0.1110
Qwen3-Embedding-8B	Average Pooling	Title	tr	0.0347	0.5013	0.8282	0.0392	0.1361
		Title + Category and Subcategory	tr	0.0367	0.5016	0.8324	0.0421	0.1322
		Title + Category and Subcategory + Abstract	tr	0.0392	0.5041	0.8060	0.0393	0.1574
	Attention Pooling	Title	tr	0.0345	0.5014	0.8341	0.0410	0.1311
		Title + Category and Subcategory	tr	0.0342	0.5038	0.8440	0.0421	0.1221
		Title + Category and Subcategory + Abstract	tr	0.0358	0.5034	0.8172	0.0394	0.1428
	Dense Layer + Attention Pooling	Title	tr	0.0342	0.5028	0.8561	0.0400	0.1042
		Title + Category and Subcategory	tr	0.0389	0.5028	0.8256	0.0390	0.1288
		Title + Category and Subcategory + Abstract	tr	0.0389	0.5035	0.8270	0.0398	0.1293
	Multi-Head Self-Attention + Attention Pooling	Title	tr	0.0323	0.5039	0.8640	0.0404	0.0957
		Title + Category and Subcategory	tr	0.0388	0.5015	0.8349	0.0385	0.1228
		Title + Category and Subcategory + Abstract	tr	0.0393	0.5025	0.8352	0.0399	0.1256
text-embedding-3-large (OpenAI)	Average Pooling	Title	tr	0.0460	0.4995	0.7351	0.0397	0.2388
		Title + Category and Subcategory	tr	0.0456	0.5016	0.7654	0.0418	0.1948
		Title + Category and Subcategory + Abstract	tr	0.0455	0.4998	0.7782	0.0426	0.1944
	Attention Pooling	Title	tr	0.0420	0.4991	0.7901	0.0397	0.1773
		Title + Category and Subcategory	tr	0.0427	0.5018	0.7836	0.0396	0.1736
		Title + Category and Subcategory + Abstract	tr	0.0434	0.5004	0.8030	0.0430	0.1641
	Dense Layer + Attention Pooling	Title	tr	0.0376	0.4997	0.8561	0.0416	0.1035
		Title + Category and Subcategory	tr	0.0400	0.5007	0.8354	0.0409	0.1220
		Title + Category and Subcategory + Abstract	tr	0.0393	0.5019	0.8424	0.0407	0.1169
	Multi-Head Self-Attention + Attention Pooling	Title	tr	0.0377	0.4979	0.8633	0.0426	0.1005
		Title + Category and Subcategory	tr	0.0394	0.4986	0.8362	0.0413	0.1187
		Title + Category and Subcategory + Abstract	tr	0.0369	0.4998	0.8579	0.0420	0.1013
multilingual-e5-large-instruct	Average Pooling	Title	vi	0.0395	0.5010	0.8029	0.0420	0.1576
		Title + Category and Subcategory	vi	0.0412	0.5018	0.7933	0.0409	0.1647

Generalized Embedding Model	User Encoder	Model Input	Language	F1 Score	AUC	Accuracy	Precision	Recall
multilingual-e5-large-instruct	Average Pooling	Title + Category and Subcategory + Abstract	vi	0.0400	0.5015	0.7764	0.0395	0.1728
	Attention Pooling	Title	vi	0.0403	0.5016	0.7994	0.0443	0.1636
		Title + Category and Subcategory	vi	0.0372	0.5030	0.8117	0.0408	0.1418
		Title + Category and Subcategory + Abstract	vi	0.0399	0.5011	0.7815	0.0400	0.1697
	Dense Layer + Attention Pooling	Title	vi	0.0271	0.5044	0.8817	0.0381	0.0728
		Title + Category and Subcategory	vi	0.0308	0.5038	0.8663	0.0363	0.0892
		Title + Category and Subcategory + Abstract	vi	0.0323	0.5013	0.8641	0.0389	0.0948
	Multi-Head Self-Attention + Attention Pooling	Title	vi	0.0271	0.5036	0.8839	0.0381	0.0702
		Title + Category and Subcategory	vi	0.0307	0.5021	0.8745	0.0391	0.0833
		Title + Category and Subcategory + Abstract	vi	0.0314	0.5010	0.8645	0.0375	0.0944
Qwen3-Embedding-0.6B	Average Pooling	Title	vi	0.0382	0.5005	0.8325	0.0423	0.1313
		Title + Category and Subcategory	vi	0.0401	0.4994	0.8244	0.0412	0.1400
		Title + Category and Subcategory + Abstract	vi	0.0404	0.4989	0.8019	0.0393	0.1576
	Attention Pooling	Title	vi	0.0395	0.5023	0.8297	0.0432	0.1377
		Title + Category and Subcategory	vi	0.0384	0.5011	0.8197	0.0404	0.1385
		Title + Category and Subcategory + Abstract	vi	0.0400	0.4991	0.7975	0.0386	0.1585
	Dense Layer + Attention Pooling	Title	vi	0.0365	0.4990	0.8442	0.0393	0.1121
		Title + Category and Subcategory	vi	0.0411	0.4994	0.8258	0.0419	0.1307
		Title + Category and Subcategory + Abstract	vi	0.0392	0.4998	0.8355	0.0416	0.1196
	Multi-Head Self-Attention + Attention Pooling	Title	vi	0.0355	0.4994	0.8445	0.0398	0.1129
		Title + Category and Subcategory	vi	0.0403	0.5005	0.8284	0.0421	0.1280
		Title + Category and Subcategory + Abstract	vi	0.0409	0.5006	0.8244	0.0410	0.1319
Qwen3-Embedding-4B	Average Pooling	Title	vi	0.0339	0.5014	0.8365	0.0397	0.1276
		Title + Category and Subcategory	vi	0.0351	0.4987	0.8467	0.0401	0.1125
		Title + Category and Subcategory + Abstract	vi	0.0361	0.5002	0.8260	0.0382	0.1312
	Attention Pooling	Title	vi	0.0313	0.5015	0.8501	0.0395	0.1129
		Title + Category and Subcategory	vi	0.0361	0.5000	0.8402	0.0410	0.1225
		Title + Category and Subcategory + Abstract	vi	0.0367	0.4995	0.8233	0.0416	0.1323
	Dense Layer + Attention Pooling	Title	vi	0.0300	0.4982	0.8604	0.0398	0.0910
		Title + Category and Subcategory	vi	0.0369	0.4998	0.8340	0.0395	0.1193
		Title + Category and Subcategory + Abstract	vi	0.0356	0.4999	0.8405	0.0397	0.1124
		Title	vi	0.0320	0.4981	0.8606	0.0400	0.0986

Generalized Embedding Model	User Encoder	Model Input	Language	F1 Score	AUC	Accuracy	Precision	Recall
Qwen3-Embedding-4B	Multi-Head Self-Attention + Attention Pooling	Title + Category and Subcategory	vi	0.0369	0.5000	0.8426	0.0407	0.1159
		Title + Category and Subcategory + Abstract	vi	0.0345	0.5014	0.8501	0.0391	0.1068
Qwen3-Embedding-8B	Average Pooling	Title	vi	0.0359	0.5004	0.8361	0.0410	0.1253
		Title + Category and Subcategory	vi	0.0405	0.5009	0.8149	0.0425	0.1467
		Title + Category and Subcategory + Abstract	vi	0.0408	0.5030	0.8009	0.0391	0.1598
	Attention Pooling	Title	vi	0.0364	0.5012	0.8329	0.0412	0.1302
		Title + Category and Subcategory	vi	0.0376	0.5019	0.8218	0.0409	0.1344
		Title + Category and Subcategory + Abstract	vi	0.0366	0.5032	0.8175	0.0396	0.1386
	Dense Layer + Attention Pooling	Title	vi	0.0358	0.5005	0.8461	0.0400	0.1142
		Title + Category and Subcategory	vi	0.0399	0.5030	0.8182	0.0382	0.1339
		Title + Category and Subcategory + Abstract	vi	0.0416	0.5035	0.8077	0.0402	0.1452
	Multi-Head Self-Attention + Attention Pooling	Title	vi	0.0334	0.5007	0.8556	0.0394	0.1043
		Title + Category and Subcategory	vi	0.0401	0.5023	0.8257	0.0385	0.1319
		Title + Category and Subcategory + Abstract	vi	0.0398	0.5039	0.8277	0.0400	0.1281
text-embedding-3-large (OpenAI)	Average Pooling	Title	vi	0.0379	0.4989	0.7993	0.0379	0.1587
		Title + Category and Subcategory	vi	0.0425	0.5004	0.7906	0.0401	0.1682
		Title + Category and Subcategory + Abstract	vi	0.0384	0.5001	0.8088	0.0389	0.1471
	Attention Pooling	Title	vi	0.0357	0.4984	0.8163	0.0369	0.1383
		Title + Category and Subcategory	vi	0.0406	0.5022	0.7996	0.0406	0.1537
		Title + Category and Subcategory + Abstract	vi	0.0381	0.5022	0.8236	0.0393	0.1351
	Dense Layer + Attention Pooling	Title	vi	0.0367	0.5009	0.8516	0.0417	0.1065
		Title + Category and Subcategory	vi	0.0387	0.5001	0.8279	0.0404	0.1261
		Title + Category and Subcategory + Abstract	vi	0.0362	0.5008	0.8407	0.0384	0.1153
	Multi-Head Self-Attention + Attention Pooling	Title	vi	0.0339	0.5005	0.8641	0.0400	0.0952
		Title + Category and Subcategory	vi	0.0384	0.5011	0.8313	0.0407	0.1238
		Title + Category and Subcategory + Abstract	vi	0.0352	0.5031	0.8476	0.0394	0.1066
multilingual-e5-large-instruct	Average Pooling	Title	zh	0.0388	0.5035	0.7644	0.0367	0.1961
		Title + Category and Subcategory	zh	0.0404	0.5012	0.7914	0.0399	0.1700
		Title + Category and Subcategory + Abstract	zh	0.0407	0.5013	0.7652	0.0389	0.1870
	Attention Pooling	Title	zh	0.0378	0.5033	0.7585	0.0368	0.1988
		Title + Category and Subcategory	zh	0.0385	0.5015	0.8021	0.0407	0.1586
		Title + Category and Subcategory + Abstract	zh	0.0396	0.5026	0.7807	0.0412	0.1747

Generalized Embedding Model	User Encoder	Model Input	Language	F1 Score	AUC	Accuracy	Precision	Recall
multilingual-e5-large-instruct	Dense Layer + Attention Pooling	Title	zh	0.0300	0.5044	0.8526	0.0358	0.1012
		Title + Category and Subcategory	zh	0.0378	0.5000	0.8339	0.0402	0.1249
		Title + Category and Subcategory + Abstract	zh	0.0356	0.5040	0.8444	0.0406	0.1150
	Multi-Head Self-Attention + Attention Pooling	Title	zh	0.0336	0.5032	0.8526	0.0394	0.1085
		Title + Category and Subcategory	zh	0.0369	0.4996	0.8339	0.0383	0.1223
		Title + Category and Subcategory + Abstract	zh	0.0358	0.5022	0.8435	0.0384	0.1136
Qwen3-Embedding-0.6B	Average Pooling	Title	zh	0.0400	0.5042	0.8047	0.0380	0.1614
		Title + Category and Subcategory	zh	0.0441	0.5039	0.7872	0.0401	0.1756
		Title + Category and Subcategory + Abstract	zh	0.0408	0.5035	0.7983	0.0395	0.1572
	Attention Pooling	Title	zh	0.0405	0.5051	0.7987	0.0381	0.1677
		Title + Category and Subcategory	zh	0.0452	0.5023	0.7800	0.0410	0.1837
		Title + Category and Subcategory + Abstract	zh	0.0411	0.5023	0.7926	0.0392	0.1629
	Dense Layer + Attention Pooling	Title	zh	0.0384	0.5006	0.8353	0.0400	0.1217
		Title + Category and Subcategory	zh	0.0414	0.5001	0.8151	0.0413	0.1378
		Title + Category and Subcategory + Abstract	zh	0.0396	0.5026	0.8192	0.0399	0.1315
	Multi-Head Self-Attention + Attention Pooling	Title	zh	0.0381	0.5025	0.8303	0.0396	0.1255
		Title + Category and Subcategory	zh	0.0421	0.5007	0.8078	0.0415	0.1436
		Title + Category and Subcategory + Abstract	zh	0.0381	0.5040	0.8251	0.0392	0.1255
Qwen3-Embedding-4B	Average Pooling	Title	zh	0.0398	0.5001	0.7976	0.0416	0.1775
		Title + Category and Subcategory	zh	0.0421	0.5017	0.8267	0.0451	0.1450
		Title + Category and Subcategory + Abstract	zh	0.0421	0.4996	0.7979	0.0413	0.1635
	Attention Pooling	Title	zh	0.0369	0.4999	0.8274	0.0421	0.1417
		Title + Category and Subcategory	zh	0.0414	0.5038	0.8346	0.0461	0.1394
		Title + Category and Subcategory + Abstract	zh	0.0402	0.4996	0.8056	0.0412	0.1574
	Dense Layer + Attention Pooling	Title	zh	0.0343	0.4979	0.8563	0.0389	0.1043
		Title + Category and Subcategory	zh	0.0411	0.5009	0.8295	0.0414	0.1334
		Title + Category and Subcategory + Abstract	zh	0.0364	0.4992	0.8451	0.0397	0.1129
	Multi-Head Self-Attention + Attention Pooling	Title	zh	0.0364	0.4972	0.8542	0.0396	0.1084
		Title + Category and Subcategory	zh	0.0404	0.5002	0.8318	0.0401	0.1303
		Title + Category and Subcategory + Abstract	zh	0.0383	0.4988	0.8421	0.0404	0.1215
Qwen3-Embedding-8B	Average Pooling	Title	zh	0.0352	0.5003	0.8122	0.0383	0.1503
		Title + Category and Subcategory	zh	0.0419	0.5016	0.8135	0.0430	0.1658

Generalized Embedding Model	User Encoder	Model Input	Language	F1 Score	AUC	Accuracy	Precision	Recall
Qwen3-Embedding-8B	Average Pooling	Title + Category and Subcategory + Abstract	zh	0.0386	0.5023	0.8220	0.0416	0.1463
	Attention Pooling	Title	zh	0.0324	0.5008	0.8287	0.0372	0.1326
		Title + Category and Subcategory	zh	0.0396	0.5032	0.8304	0.0430	0.1441
		Title + Category and Subcategory + Abstract	zh	0.0351	0.5006	0.8223	0.0396	0.1362
	Dense Layer + Attention Pooling	Title	zh	0.0343	0.5012	0.8505	0.0390	0.1068
		Title + Category and Subcategory	zh	0.0405	0.5023	0.8197	0.0383	0.1357
		Title + Category and Subcategory + Abstract	zh	0.0398	0.5019	0.8146	0.0381	0.1426
	Multi-Head Self-Attention + Attention Pooling	Title	zh	0.0342	0.5007	0.8497	0.0396	0.1084
		Title + Category and Subcategory	zh	0.0409	0.5000	0.8303	0.0392	0.1315
		Title + Category and Subcategory + Abstract	zh	0.0399	0.5018	0.8187	0.0378	0.1392
text-embedding-3-large (OpenAI)	Average Pooling	Title	zh	0.0439	0.5011	0.7649	0.0389	0.2039
		Title + Category and Subcategory	zh	0.0428	0.4998	0.7842	0.0395	0.1700
		Title + Category and Subcategory + Abstract	zh	0.0385	0.5008	0.8051	0.0375	0.1491
	Attention Pooling	Title	zh	0.0405	0.5028	0.7942	0.0383	0.1706
		Title + Category and Subcategory	zh	0.0413	0.5029	0.7992	0.0408	0.1596
		Title + Category and Subcategory + Abstract	zh	0.0380	0.5030	0.8185	0.0381	0.1365
	Dense Layer + Attention Pooling	Title	zh	0.0363	0.5010	0.8478	0.0394	0.1058
		Title + Category and Subcategory	zh	0.0408	0.5003	0.8169	0.0400	0.1385
		Title + Category and Subcategory + Abstract	zh	0.0402	0.5002	0.8250	0.0406	0.1323
	Multi-Head Self-Attention + Attention Pooling	Title	zh	0.0350	0.5016	0.8599	0.0405	0.0976
		Title + Category and Subcategory	zh	0.0392	0.5011	0.8269	0.0400	0.1271
		Title + Category and Subcategory + Abstract	zh	0.0365	0.5003	0.8438	0.0400	0.1142

The results for both training and evaluation in the target language is given in the table below.

Table A.13: Performance of Models on Multilingual Classification of all User for given News trained and evaluated in target language

Generalized Embedding Model	User Encoder	Model Input	Language	F1 Score	AUC	Accuracy	Precision	Recall
multilingual-e5-large-instruct	Average Pooling	Title	en	0.0389	0.5008	0.8140	0.0404	0.1476
		Title + Category and Subcategory	en	0.0384	0.5019	0.8230	0.0399	0.1354
		Title + Category and Subcategory + Abstract	en	0.0393	0.5002	0.8004	0.0390	0.1526

Generalized Embedding Model	User Encoder	Model Input	Language	F1 Score	AUC	Accuracy	Precision	Recall
multilingual-e5-large-instruct	Attention Pooling	Title	en	0.0383	0.5019	0.8165	0.0403	0.1471
		Title + Category and Subcategory	en	0.0357	0.5017	0.8362	0.0397	0.1223
		Title + Category and Subcategory + Abstract	en	0.0388	0.5012	0.8077	0.0399	0.1478
	Dense Layer + Attention Pooling	Title	en	0.0327	0.5013	0.8630	0.0388	0.0923
		Title + Category and Subcategory	en	0.0356	0.5014	0.8493	0.0386	0.1047
		Title + Category and Subcategory + Abstract	en	0.0354	0.5010	0.8535	0.0397	0.0997
	Multi-Head Self-Attention + Attention Pooling	Title	en	0.0349	0.5011	0.8623	0.0411	0.0955
		Title + Category and Subcategory	en	0.0342	0.5001	0.8569	0.0376	0.0979
		Title + Category and Subcategory + Abstract	en	0.0369	0.5023	0.8475	0.0400	0.1080
Qwen3-Embedding-0.6B	Average Pooling	Title	en	0.0392	0.5097	0.8168	0.0409	0.1508
		Title + Category and Subcategory	en	0.0420	0.5064	0.7846	0.0401	0.1732
		Title + Category and Subcategory + Abstract	en	0.0414	0.5034	0.8027	0.0436	0.1630
	Attention Pooling	Title	en	0.0390	0.5101	0.8211	0.0416	0.1471
		Title + Category and Subcategory	en	0.0416	0.5060	0.7860	0.0401	0.1727
		Title + Category and Subcategory + Abstract	en	0.0411	0.5031	0.8005	0.0420	0.1637
	Dense Layer + Attention Pooling	Title	en	0.0326	0.5090	0.8546	0.0376	0.0980
		Title + Category and Subcategory	en	0.0383	0.5057	0.8342	0.0416	0.1224
		Title + Category and Subcategory + Abstract	en	0.0364	0.5070	0.8473	0.0407	0.1130
	Multi-Head Self-Attention + Attention Pooling	Title	en	0.0322	0.5073	0.8575	0.0398	0.0971
		Title + Category and Subcategory	en	0.0380	0.5054	0.8430	0.0424	0.1171
		Title + Category and Subcategory + Abstract	en	0.0344	0.5088	0.8587	0.0411	0.1023
Qwen3-Embedding-4B	Average Pooling	Title	en	0.0301	0.5026	0.8454	0.0362	0.1200
		Title + Category and Subcategory	en	0.0378	0.5015	0.8236	0.0427	0.1379
		Title + Category and Subcategory + Abstract	en	0.0350	0.5033	0.8316	0.0415	0.1274
	Attention Pooling	Title	en	0.0297	0.5017	0.8551	0.0364	0.1116
		Title + Category and Subcategory	en	0.0362	0.5014	0.8223	0.0400	0.1368
		Title + Category and Subcategory + Abstract	en	0.0348	0.5019	0.8241	0.0404	0.1337
	Dense Layer + Attention Pooling	Title	en	0.0281	0.4970	0.8669	0.0351	0.0858
		Title + Category and Subcategory	en	0.0323	0.4986	0.8526	0.0371	0.1007
		Title + Category and Subcategory + Abstract	en	0.0350	0.5020	0.8566	0.0412	0.1044
	Multi-Head Self-Attention + Attention Pooling	Title	en	0.0281	0.4974	0.8710	0.0367	0.0845
		Title + Category and Subcategory	en	0.0336	0.4987	0.8550	0.0396	0.1017

Generalized Embedding Model	User Encoder	Model Input	Language	F1 Score	AUC	Accuracy	Precision	Recall
Qwen3-Embedding-4B	Multi-Head Self-Attention + Attention Pooling	Title + Category and Subcategory + Abstract	en	0.0371	0.5018	0.8489	0.0418	0.1141
Qwen3-Embedding-8B	Average Pooling	Title	en	0.0373	0.5017	0.8335	0.0411	0.1295
		Title + Category and Subcategory	en	0.0476	0.5045	0.7759	0.0449	0.1990
		Title + Category and Subcategory + Abstract	en	0.0408	0.5048	0.8015	0.0401	0.1635
	Attention Pooling	Title	en	0.0349	0.5014	0.8385	0.0388	0.1211
		Title + Category and Subcategory	en	0.0390	0.5048	0.8189	0.0414	0.1410
		Title + Category and Subcategory + Abstract	en	0.0379	0.5055	0.8252	0.0422	0.1327
	Dense Layer + Attention Pooling	Title	en	0.0347	0.5027	0.8579	0.0401	0.0988
		Title + Category and Subcategory	en	0.0396	0.5053	0.8238	0.0394	0.1303
		Title + Category and Subcategory + Abstract	en	0.0395	0.5037	0.8296	0.0397	0.1289
	Multi-Head Self-Attention + Attention Pooling	Title	en	0.0354	0.5025	0.8628	0.0423	0.0977
		Title + Category and Subcategory	en	0.0404	0.5047	0.8327	0.0414	0.1259
		Title + Category and Subcategory + Abstract	en	0.0415	0.5045	0.8349	0.0418	0.1283
text-embedding-3-large (OpenAI)	Average Pooling	Title	en	0.0400	0.5000	0.8071	0.0383	0.1493
		Title + Category and Subcategory	en	0.0428	0.5018	0.8008	0.0400	0.1565
		Title + Category and Subcategory + Abstract	en	0.0413	0.5000	0.8084	0.0395	0.1513
	Attention Pooling	Title	en	0.0374	0.5011	0.8271	0.0380	0.1293
		Title + Category and Subcategory	en	0.0386	0.5032	0.8251	0.0389	0.1309
		Title + Category and Subcategory + Abstract	en	0.0377	0.5013	0.8357	0.0386	0.1223
	Dense Layer + Attention Pooling	Title	en	0.0370	0.4988	0.8622	0.0429	0.0979
		Title + Category and Subcategory	en	0.0415	0.5016	0.8334	0.0437	0.1240
		Title + Category and Subcategory + Abstract	en	0.0386	0.4994	0.8493	0.0420	0.1083
	Multi-Head Self-Attention + Attention Pooling	Title	en	0.0389	0.4994	0.8543	0.0437	0.1062
		Title + Category and Subcategory	en	0.0416	0.5000	0.8355	0.0444	0.1243
		Title + Category and Subcategory + Abstract	en	0.0390	0.5009	0.8470	0.0423	0.1115
multilingual-e5-large-instruct	Average Pooling	Title	tr	0.0396	0.5014	0.7962	0.0391	0.1654
		Title + Category and Subcategory	tr	0.0394	0.5009	0.8022	0.0388	0.1541
		Title + Category and Subcategory + Abstract	tr	0.0405	0.4999	0.7783	0.0383	0.1752
	Attention Pooling	Title	tr	0.0389	0.5010	0.8016	0.0393	0.1595
		Title + Category and Subcategory	tr	0.0375	0.5019	0.8082	0.0383	0.1451
		Title + Category and Subcategory + Abstract	tr	0.0400	0.5022	0.7905	0.0403	0.1667
		Title	tr	0.0327	0.5031	0.8538	0.0397	0.0984

Generalized Embedding Model	User Encoder	Model Input	Language	F1 Score	AUC	Accuracy	Precision	Recall
multilingual-e5-large-instruct	Dense Layer + Attention Pooling	Title + Category and Subcategory	tr	0.0330	0.4999	0.8382	0.0364	0.1085
		Title + Category and Subcategory + Abstract	tr	0.0366	0.5008	0.8339	0.0396	0.1176
	Multi-Head Self-Attention + Attention Pooling	Title	tr	0.0322	0.5016	0.8618	0.0385	0.0934
		Title + Category and Subcategory	tr	0.0354	0.4998	0.8431	0.0393	0.1062
		Title + Category and Subcategory + Abstract	tr	0.0352	0.5012	0.8423	0.0387	0.1085
Qwen3-Embedding-0.6B	Average Pooling	Title	tr	0.0380	0.5034	0.8319	0.0421	0.1325
		Title + Category and Subcategory	tr	0.0420	0.5032	0.8046	0.0412	0.1578
		Title + Category and Subcategory + Abstract	tr	0.0386	0.5011	0.8099	0.0394	0.1441
	Attention Pooling	Title	tr	0.0373	0.5041	0.8376	0.0430	0.1246
		Title + Category and Subcategory	tr	0.0392	0.5031	0.8171	0.0411	0.1400
		Title + Category and Subcategory + Abstract	tr	0.0388	0.5018	0.8082	0.0389	0.1467
	Dense Layer + Attention Pooling	Title	tr	0.0333	0.5002	0.8685	0.0442	0.0878
		Title + Category and Subcategory	tr	0.0346	0.5010	0.8491	0.0401	0.1034
		Title + Category and Subcategory + Abstract	tr	0.0341	0.5030	0.8469	0.0382	0.1039
	Multi-Head Self-Attention + Attention Pooling	Title	tr	0.0330	0.5011	0.8731	0.0454	0.0855
		Title + Category and Subcategory	tr	0.0347	0.5006	0.8548	0.0415	0.0996
		Title + Category and Subcategory + Abstract	tr	0.0341	0.5048	0.8492	0.0381	0.1033
Qwen3-Embedding-4B	Average Pooling	Title	tr	0.0364	0.4997	0.7915	0.0364	0.1665
		Title + Category and Subcategory	tr	0.0382	0.4997	0.8274	0.0432	0.1387
		Title + Category and Subcategory + Abstract	tr	0.0362	0.4994	0.8134	0.0368	0.1404
	Attention Pooling	Title	tr	0.0377	0.5016	0.7983	0.0408	0.1648
		Title + Category and Subcategory	tr	0.0367	0.5000	0.8300	0.0421	0.1338
		Title + Category and Subcategory + Abstract	tr	0.0364	0.4985	0.8135	0.0381	0.1409
	Dense Layer + Attention Pooling	Title	tr	0.0352	0.4975	0.8413	0.0414	0.1113
		Title + Category and Subcategory	tr	0.0346	0.4996	0.8422	0.0379	0.1110
		Title + Category and Subcategory + Abstract	tr	0.0379	0.5000	0.8292	0.0403	0.1230
	Multi-Head Self-Attention + Attention Pooling	Title	tr	0.0352	0.4992	0.8473	0.0405	0.1095
		Title + Category and Subcategory	tr	0.0349	0.4987	0.8511	0.0394	0.1087
		Title + Category and Subcategory + Abstract	tr	0.0370	0.5015	0.8395	0.0411	0.1156
Qwen3-Embedding-8B	Average Pooling	Title	tr	0.0347	0.5013	0.8282	0.0392	0.1361
		Title + Category and Subcategory	tr	0.0367	0.5016	0.8324	0.0421	0.1322
		Title + Category and Subcategory + Abstract	tr	0.0392	0.5041	0.8060	0.0393	0.1574

Generalized Embedding Model	User Encoder	Model Input	Language	F1 Score	AUC	Accuracy	Precision	Recall
Qwen3-Embedding-8B	Attention Pooling	Title	tr	0.0345	0.5010	0.8378	0.0415	0.1288
		Title + Category and Subcategory	tr	0.0349	0.5027	0.8466	0.0423	0.1192
		Title + Category and Subcategory + Abstract	tr	0.0363	0.5042	0.8179	0.0392	0.1421
	Dense Layer + Attention Pooling	Title	tr	0.0334	0.5006	0.8566	0.0405	0.0996
		Title + Category and Subcategory	tr	0.0358	0.5047	0.8445	0.0404	0.1102
		Title + Category and Subcategory + Abstract	tr	0.0380	0.5037	0.8381	0.0406	0.1191
	Multi-Head Self-Attention + Attention Pooling	Title	tr	0.0336	0.5026	0.8572	0.0406	0.1027
		Title + Category and Subcategory	tr	0.0377	0.5035	0.8401	0.0404	0.1178
		Title + Category and Subcategory + Abstract	tr	0.0382	0.5046	0.8385	0.0399	0.1171
text-embedding-3-large (OpenAI)	Average Pooling	Title	tr	0.0460	0.4995	0.7351	0.0397	0.2388
		Title + Category and Subcategory	tr	0.0456	0.5016	0.7654	0.0418	0.1948
		Title + Category and Subcategory + Abstract	tr	0.0455	0.4998	0.7782	0.0426	0.1944
	Attention Pooling	Title	tr	0.0395	0.4990	0.8063	0.0382	0.1532
		Title + Category and Subcategory	tr	0.0440	0.5009	0.7764	0.0404	0.1793
		Title + Category and Subcategory + Abstract	tr	0.0437	0.5012	0.8007	0.0422	0.1645
	Dense Layer + Attention Pooling	Title	tr	0.0372	0.4954	0.8547	0.0403	0.1008
		Title + Category and Subcategory	tr	0.0409	0.4982	0.8376	0.0433	0.1197
		Title + Category and Subcategory + Abstract	tr	0.0357	0.4979	0.8536	0.0410	0.1013
	Multi-Head Self-Attention + Attention Pooling	Title	tr	0.0358	0.4964	0.8686	0.0422	0.0879
		Title + Category and Subcategory	tr	0.0416	0.4973	0.8347	0.0436	0.1245
		Title + Category and Subcategory + Abstract	tr	0.0378	0.4998	0.8467	0.0427	0.1072
multilingual-e5-large-instruct	Average Pooling	Title	vi	0.0395	0.5010	0.8029	0.0420	0.1576
		Title + Category and Subcategory	vi	0.0412	0.5018	0.7933	0.0409	0.1647
		Title + Category and Subcategory + Abstract	vi	0.0400	0.5015	0.7764	0.0395	0.1728
	Attention Pooling	Title	vi	0.0393	0.5011	0.8040	0.0420	0.1569
		Title + Category and Subcategory	vi	0.0380	0.5036	0.8097	0.0412	0.1445
		Title + Category and Subcategory + Abstract	vi	0.0396	0.5009	0.7810	0.0396	0.1700
	Dense Layer + Attention Pooling	Title	vi	0.0358	0.5036	0.8390	0.0396	0.1173
		Title + Category and Subcategory	vi	0.0358	0.5016	0.8362	0.0385	0.1165
		Title + Category and Subcategory + Abstract	vi	0.0346	0.5010	0.8454	0.0388	0.1100
	Multi-Head Self-Attention + Attention Pooling	Title	vi	0.0350	0.5007	0.8547	0.0403	0.1061
		Title + Category and Subcategory	vi	0.0360	0.5026	0.8446	0.0398	0.1107

Generalized Embedding Model	User Encoder	Model Input	Language	F1 Score	AUC	Accuracy	Precision	Recall
multilingual-e5-large-instruct	Multi-Head Self-Attention + Attention Pooling	Title + Category and Subcategory + Abstract	vi	0.0377	0.5027	0.8397	0.0416	0.1182
Qwen3-Embedding-0.6B	Average Pooling	Title	vi	0.0382	0.5005	0.8325	0.0423	0.1313
		Title + Category and Subcategory	vi	0.0401	0.4994	0.8244	0.0412	0.1400
		Title + Category and Subcategory + Abstract	vi	0.0404	0.4989	0.8019	0.0393	0.1576
	Attention Pooling	Title	vi	0.0383	0.5010	0.8354	0.0417	0.1297
		Title + Category and Subcategory	vi	0.0398	0.5000	0.8196	0.0422	0.1418
		Title + Category and Subcategory + Abstract	vi	0.0391	0.5002	0.8017	0.0386	0.1519
	Dense Layer + Attention Pooling	Title	vi	0.0341	0.5009	0.8571	0.0394	0.0991
		Title + Category and Subcategory	vi	0.0360	0.4996	0.8408	0.0396	0.1124
		Title + Category and Subcategory + Abstract	vi	0.0358	0.4990	0.8370	0.0393	0.1137
	Multi-Head Self-Attention + Attention Pooling	Title	vi	0.0331	0.5008	0.8589	0.0393	0.0970
		Title + Category and Subcategory	vi	0.0362	0.5002	0.8423	0.0404	0.1124
		Title + Category and Subcategory + Abstract	vi	0.0366	0.4998	0.8349	0.0384	0.1178
Qwen3-Embedding-4B	Average Pooling	Title	vi	0.0339	0.5014	0.8365	0.0397	0.1276
		Title + Category and Subcategory	vi	0.0351	0.4987	0.8467	0.0401	0.1125
		Title + Category and Subcategory + Abstract	vi	0.0361	0.5002	0.8260	0.0382	0.1312
	Attention Pooling	Title	vi	0.0311	0.5006	0.8475	0.0389	0.1145
		Title + Category and Subcategory	vi	0.0391	0.5003	0.8366	0.0444	0.1307
		Title + Category and Subcategory + Abstract	vi	0.0379	0.4990	0.8206	0.0415	0.1383
	Dense Layer + Attention Pooling	Title	vi	0.0325	0.4977	0.8521	0.0388	0.1016
		Title + Category and Subcategory	vi	0.0348	0.5007	0.8400	0.0370	0.1142
		Title + Category and Subcategory + Abstract	vi	0.0371	0.5010	0.8367	0.0397	0.1187
	Multi-Head Self-Attention + Attention Pooling	Title	vi	0.0336	0.4985	0.8591	0.0414	0.1029
		Title + Category and Subcategory	vi	0.0361	0.5006	0.8495	0.0397	0.1127
		Title + Category and Subcategory + Abstract	vi	0.0359	0.5022	0.8398	0.0389	0.1164
Qwen3-Embedding-8B	Average Pooling	Title	vi	0.0359	0.5004	0.8361	0.0410	0.1253
		Title + Category and Subcategory	vi	0.0405	0.5009	0.8149	0.0425	0.1467
		Title + Category and Subcategory + Abstract	vi	0.0408	0.5030	0.8009	0.0391	0.1598
	Attention Pooling	Title	vi	0.0364	0.5006	0.8339	0.0419	0.1299
		Title + Category and Subcategory	vi	0.0347	0.5002	0.8448	0.0410	0.1158
		Title + Category and Subcategory + Abstract	vi	0.0355	0.5029	0.8256	0.0391	0.1314
		Title	vi	0.0326	0.5004	0.8630	0.0409	0.0943

Generalized Embedding Model	User Encoder	Model Input	Language	F1 Score	AUC	Accuracy	Precision	Recall
Qwen3-Embedding-8B	Dense Layer + Attention Pooling	Title + Category and Subcategory	vi	0.0369	0.5033	0.8373	0.0390	0.1160
		Title + Category and Subcategory + Abstract	vi	0.0390	0.5049	0.8299	0.0404	0.1258
	Multi-Head Self-Attention + Attention Pooling	Title	vi	0.0324	0.5008	0.8651	0.0405	0.0948
		Title + Category and Subcategory	vi	0.0374	0.5048	0.8403	0.0402	0.1156
		Title + Category and Subcategory + Abstract	vi	0.0377	0.5052	0.8448	0.0401	0.1149
text-embedding-3-large (OpenAI)	Average Pooling	Title	vi	0.0379	0.4989	0.7993	0.0379	0.1587
		Title + Category and Subcategory	vi	0.0425	0.5004	0.7906	0.0401	0.1682
		Title + Category and Subcategory + Abstract	vi	0.0384	0.5001	0.8088	0.0389	0.1471
	Attention Pooling	Title	vi	0.0362	0.4987	0.8227	0.0388	0.1352
		Title + Category and Subcategory	vi	0.0406	0.5014	0.8000	0.0402	0.1544
		Title + Category and Subcategory + Abstract	vi	0.0379	0.5014	0.8239	0.0395	0.1331
	Dense Layer + Attention Pooling	Title	vi	0.0371	0.4966	0.8570	0.0435	0.1019
		Title + Category and Subcategory	vi	0.0389	0.4982	0.8238	0.0407	0.1255
		Title + Category and Subcategory + Abstract	vi	0.0356	0.4981	0.8456	0.0401	0.1070
	Multi-Head Self-Attention + Attention Pooling	Title	vi	0.0362	0.4967	0.8519	0.0409	0.1057
		Title + Category and Subcategory	vi	0.0403	0.4981	0.8156	0.0412	0.1326
		Title + Category and Subcategory + Abstract	vi	0.0359	0.4979	0.8425	0.0401	0.1097
multilingual-e5-large-instruct	Average Pooling	Title	zh	0.0388	0.5035	0.7644	0.0367	0.1961
		Title + Category and Subcategory	zh	0.0404	0.5012	0.7914	0.0399	0.1700
		Title + Category and Subcategory + Abstract	zh	0.0407	0.5013	0.7652	0.0389	0.1870
	Attention Pooling	Title	zh	0.0375	0.5033	0.7654	0.0372	0.1929
		Title + Category and Subcategory	zh	0.0394	0.5016	0.7920	0.0406	0.1696
		Title + Category and Subcategory + Abstract	zh	0.0398	0.5017	0.7770	0.0402	0.1762
	Dense Layer + Attention Pooling	Title	zh	0.0323	0.5007	0.8521	0.0391	0.1019
		Title + Category and Subcategory	zh	0.0397	0.4986	0.8031	0.0397	0.1467
		Title + Category and Subcategory + Abstract	zh	0.0364	0.5013	0.8359	0.0409	0.1179
	Multi-Head Self-Attention + Attention Pooling	Title	zh	0.0296	0.5021	0.8605	0.0396	0.0934
		Title + Category and Subcategory	zh	0.0374	0.4995	0.8269	0.0397	0.1312
		Title + Category and Subcategory + Abstract	zh	0.0339	0.5005	0.8482	0.0407	0.1094
Qwen3-Embedding-0.6B	Average Pooling	Title	zh	0.0400	0.5042	0.8047	0.0380	0.1614
		Title + Category and Subcategory	zh	0.0441	0.5039	0.7872	0.0401	0.1756
		Title + Category and Subcategory + Abstract	zh	0.0408	0.5035	0.7983	0.0395	0.1572

Generalized Embedding Model	User Encoder	Model Input	Language	F1 Score	AUC	Accuracy	Precision	Recall
Qwen3-Embedding-0.6B	Attention Pooling	Title	zh	0.0396	0.5041	0.8171	0.0393	0.1487
		Title + Category and Subcategory	zh	0.0455	0.5023	0.7824	0.0417	0.1832
		Title + Category and Subcategory + Abstract	zh	0.0411	0.5023	0.8001	0.0408	0.1583
	Dense Layer + Attention Pooling	Title	zh	0.0338	0.5004	0.8438	0.0374	0.1069
		Title + Category and Subcategory	zh	0.0383	0.4994	0.8308	0.0406	0.1244
		Title + Category and Subcategory + Abstract	zh	0.0377	0.5024	0.8258	0.0400	0.1277
	Multi-Head Self-Attention + Attention Pooling	Title	zh	0.0331	0.5017	0.8570	0.0378	0.1002
		Title + Category and Subcategory	zh	0.0384	0.5019	0.8322	0.0407	0.1247
		Title + Category and Subcategory + Abstract	zh	0.0336	0.5037	0.8433	0.0385	0.1093
Qwen3-Embedding-4B	Average Pooling	Title	zh	0.0398	0.5001	0.7976	0.0416	0.1775
		Title + Category and Subcategory	zh	0.0421	0.5017	0.8267	0.0451	0.1450
		Title + Category and Subcategory + Abstract	zh	0.0421	0.4996	0.7979	0.0413	0.1635
	Attention Pooling	Title	zh	0.0373	0.5008	0.8267	0.0431	0.1418
		Title + Category and Subcategory	zh	0.0439	0.5048	0.8162	0.0471	0.1558
		Title + Category and Subcategory + Abstract	zh	0.0396	0.4999	0.8117	0.0424	0.1487
	Dense Layer + Attention Pooling	Title	zh	0.0332	0.4979	0.8467	0.0405	0.1070
		Title + Category and Subcategory	zh	0.0371	0.5013	0.8255	0.0380	0.1284
		Title + Category and Subcategory + Abstract	zh	0.0373	0.4992	0.8168	0.0378	0.1341
	Multi-Head Self-Attention + Attention Pooling	Title	zh	0.0346	0.4981	0.8375	0.0395	0.1162
		Title + Category and Subcategory	zh	0.0354	0.5021	0.8362	0.0380	0.1190
		Title + Category and Subcategory + Abstract	zh	0.0389	0.5000	0.8146	0.0391	0.1397
Qwen3-Embedding-8B	Average Pooling	Title	zh	0.0352	0.5003	0.8122	0.0383	0.1503
		Title + Category and Subcategory	zh	0.0419	0.5016	0.8135	0.0430	0.1658
		Title + Category and Subcategory + Abstract	zh	0.0386	0.5023	0.8220	0.0416	0.1463
	Attention Pooling	Title	zh	0.0320	0.5010	0.8301	0.0368	0.1276
		Title + Category and Subcategory	zh	0.0403	0.5037	0.8266	0.0438	0.1442
		Title + Category and Subcategory + Abstract	zh	0.0355	0.5012	0.8236	0.0398	0.1349
	Dense Layer + Attention Pooling	Title	zh	0.0357	0.5006	0.8450	0.0418	0.1121
		Title + Category and Subcategory	zh	0.0390	0.5010	0.8311	0.0395	0.1248
		Title + Category and Subcategory + Abstract	zh	0.0376	0.5030	0.8289	0.0379	0.1239
	Multi-Head Self-Attention + Attention Pooling	Title	zh	0.0349	0.5012	0.8539	0.0419	0.1056
		Title + Category and Subcategory	zh	0.0403	0.5010	0.8400	0.0417	0.1219

Generalized Embedding Model	User Encoder	Model Input	Language	F1 Score	AUC	Accuracy	Precision	Recall
Qwen3-Embedding-8B	Multi-Head Self-Attention + Attention Pooling	Title + Category and Subcategory + Abstract	zh	0.0391	0.5030	0.8332	0.0400	0.1233
text-embedding-3-large (OpenAI)	Average Pooling	Title	zh	0.0439	0.5011	0.7649	0.0389	0.2039
		Title + Category and Subcategory	zh	0.0428	0.4998	0.7842	0.0395	0.1700
		Title + Category and Subcategory + Abstract	zh	0.0385	0.5008	0.8051	0.0375	0.1491
	Attention Pooling	Title	zh	0.0392	0.5025	0.8142	0.0383	0.1487
		Title + Category and Subcategory	zh	0.0418	0.5005	0.7958	0.0424	0.1575
		Title + Category and Subcategory + Abstract	zh	0.0385	0.5021	0.8199	0.0395	0.1355
	Dense Layer + Attention Pooling	Title	zh	0.0373	0.4979	0.8445	0.0410	0.1069
		Title + Category and Subcategory	zh	0.0413	0.4999	0.8164	0.0415	0.1375
		Title + Category and Subcategory + Abstract	zh	0.0376	0.5001	0.8444	0.0424	0.1131
	Multi-Head Self-Attention + Attention Pooling	Title	zh	0.0366	0.5010	0.8483	0.0407	0.1033
		Title + Category and Subcategory	zh	0.0402	0.5007	0.8232	0.0414	0.1323
		Title + Category and Subcategory + Abstract	zh	0.0383	0.5010	0.8373	0.0418	0.1210

A.2.3 Task 3: Candidate News Classification for User

The results for training in English only and evaluation in the target language is given in the table below.

Table A.14: Performance of Models on Multilingual Candidate News Classification for User trained on English only

Generalized Embedding Model	User Encoder	Model Input	Language	F1 Score	AUC	Accuracy	Precision	Recall
multilingual-e5-large-instruct	Average Pooling	Title	en	0.1450	0.6589	0.8069	0.1509	0.2973
		Title + Category and Subcategory	en	0.1410	0.6524	0.8226	0.1571	0.2659
		Title + Category and Subcategory + Abstract	en	0.1488	0.6615	0.8199	0.1621	0.2853
	Attention Pooling	Title	en	0.1516	0.6565	0.8029	0.1542	0.3123
		Title + Category and Subcategory	en	0.1498	0.6524	0.8268	0.1658	0.2731
		Title + Category and Subcategory + Abstract	en	0.1568	0.6601	0.8203	0.1680	0.2958
	Dense Layer + Attention Pooling	Title	en	0.1370	0.6516	0.8578	0.1805	0.1982
		Title + Category and Subcategory	en	0.1388	0.6548	0.8546	0.1795	0.2120
		Title + Category and Subcategory + Abstract	en	0.1369	0.6563	0.8561	0.1810	0.2106

Generalized Embedding Model	User Encoder	Model Input	Language	F1 Score	AUC	Accuracy	Precision	Recall
multilingual-e5-large-instruct	Multi-Head Self-Attention + Attention Pooling	Title	en	0.1356	0.6505	0.8582	0.1818	0.1994
		Title + Category and Subcategory	en	0.1299	0.6539	0.8601	0.1819	0.1966
		Title + Category and Subcategory + Abstract	en	0.1336	0.6563	0.8552	0.1773	0.2098
Qwen3-Embedding-0.6B	Average Pooling	Title	en	0.1436	0.6304	0.7854	0.1399	0.3045
		Title + Category and Subcategory	en	0.1346	0.6140	0.7963	0.1384	0.2703
		Title + Category and Subcategory + Abstract	en	0.1348	0.6256	0.7998	0.1413	0.2728
	Attention Pooling	Title	en	0.1455	0.6315	0.7957	0.1449	0.2953
		Title + Category and Subcategory	en	0.1411	0.6217	0.7962	0.1430	0.2808
		Title + Category and Subcategory + Abstract	en	0.1390	0.6312	0.8024	0.1463	0.2777
	Dense Layer + Attention Pooling	Title	en	0.1188	0.6448	0.8507	0.1541	0.1842
		Title + Category and Subcategory	en	0.1315	0.6467	0.8490	0.1632	0.2079
		Title + Category and Subcategory + Abstract	en	0.1276	0.6498	0.8501	0.1671	0.2062
	Multi-Head Self-Attention + Attention Pooling	Title	en	0.1221	0.6465	0.8482	0.1557	0.1916
		Title + Category and Subcategory	en	0.1293	0.6486	0.8506	0.1617	0.2002
		Title + Category and Subcategory + Abstract	en	0.1219	0.6505	0.8551	0.1676	0.1897
Qwen3-Embedding-4B	Average Pooling	Title	en	0.1431	0.6304	0.8113	0.1505	0.2678
		Title + Category and Subcategory	en	0.1242	0.6243	0.8195	0.1394	0.2293
		Title + Category and Subcategory + Abstract	en	0.1231	0.6225	0.8258	0.1415	0.2219
	Attention Pooling	Title	en	0.1433	0.6318	0.8216	0.1568	0.2554
		Title + Category and Subcategory	en	0.1386	0.6306	0.8204	0.1517	0.2496
		Title + Category and Subcategory + Abstract	en	0.1348	0.6284	0.8214	0.1506	0.2461
	Dense Layer + Attention Pooling	Title	en	0.1266	0.6525	0.8535	0.1746	0.1948
		Title + Category and Subcategory	en	0.1364	0.6541	0.8529	0.1753	0.2109
		Title + Category and Subcategory + Abstract	en	0.1320	0.6551	0.8535	0.1770	0.2083
	Multi-Head Self-Attention + Attention Pooling	Title	en	0.1272	0.6552	0.8533	0.1752	0.1975
		Title + Category and Subcategory	en	0.1383	0.6561	0.8513	0.1753	0.2170
		Title + Category and Subcategory + Abstract	en	0.1391	0.6568	0.8453	0.1745	0.2309
Qwen3-Embedding-8B	Average Pooling	Title	en	0.1404	0.6388	0.8124	0.1501	0.2715
		Title + Category and Subcategory	en	0.1554	0.6058	0.7622	0.1417	0.3330
		Title + Category and Subcategory + Abstract	en	0.1458	0.6123	0.7896	0.1422	0.2874
	Attention Pooling	Title	en	0.1420	0.6391	0.8186	0.1543	0.2643
		Title + Category and Subcategory	en	0.1532	0.6238	0.8089	0.1550	0.2770

Generalized Embedding Model	User Encoder	Model Input	Language	F1 Score	AUC	Accuracy	Precision	Recall
Qwen3-Embedding-8B	Attention Pooling	Title + Category and Subcategory + Abstract	en	0.1520	0.6271	0.8088	0.1541	0.2779
	Dense Layer + Attention Pooling	Title	en	0.1312	0.6557	0.8545	0.1732	0.1983
		Title + Category and Subcategory	en	0.1489	0.6539	0.8381	0.1704	0.2488
		Title + Category and Subcategory + Abstract	en	0.1500	0.6557	0.8424	0.1736	0.2463
	Multi-Head Self-Attention + Attention Pooling	Title	en	0.1371	0.6549	0.8506	0.1716	0.2103
		Title + Category and Subcategory	en	0.1474	0.6521	0.8402	0.1682	0.2408
		Title + Category and Subcategory + Abstract	en	0.1497	0.6545	0.8436	0.1729	0.2423
text-embedding-3-large (OpenAI)	Average Pooling	Title	en	0.1150	0.6186	0.8160	0.1357	0.2241
		Title + Category and Subcategory	en	0.1283	0.6205	0.8222	0.1455	0.2353
		Title + Category and Subcategory + Abstract	en	0.1272	0.6308	0.8196	0.1485	0.2444
	Attention Pooling	Title	en	0.1199	0.6251	0.8252	0.1440	0.2195
		Title + Category and Subcategory	en	0.1338	0.6297	0.8337	0.1558	0.2287
		Title + Category and Subcategory + Abstract	en	0.1283	0.6370	0.8351	0.1567	0.2261
	Dense Layer + Attention Pooling	Title	en	0.1314	0.6593	0.8629	0.1850	0.1900
		Title + Category and Subcategory	en	0.1419	0.6575	0.8544	0.1822	0.2207
		Title + Category and Subcategory + Abstract	en	0.1343	0.6588	0.8600	0.1855	0.2026
	Multi-Head Self-Attention + Attention Pooling	Title	en	0.1404	0.6606	0.8551	0.1811	0.2149
		Title + Category and Subcategory	en	0.1409	0.6591	0.8523	0.1810	0.2250
		Title + Category and Subcategory + Abstract	en	0.1369	0.6600	0.8575	0.1855	0.2115
multilingual-e5-large-instruct	Average Pooling	Title	tr	0.1593	0.6469	0.7841	0.1527	0.3440
		Title + Category and Subcategory	tr	0.1543	0.6459	0.8132	0.1620	0.2937
		Title + Category and Subcategory + Abstract	tr	0.1524	0.6520	0.8103	0.1593	0.2976
	Attention Pooling	Title	tr	0.1637	0.6439	0.7833	0.1554	0.3498
		Title + Category and Subcategory	tr	0.1615	0.6447	0.8115	0.1653	0.3066
		Title + Category and Subcategory + Abstract	tr	0.1562	0.6501	0.8154	0.1635	0.2952
	Dense Layer + Attention Pooling	Title	tr	0.1171	0.6402	0.8555	0.1600	0.1699
		Title + Category and Subcategory	tr	0.1297	0.6466	0.8528	0.1698	0.1951
		Title + Category and Subcategory + Abstract	tr	0.1272	0.6428	0.8548	0.1692	0.1896
	Multi-Head Self-Attention + Attention Pooling	Title	tr	0.1195	0.6403	0.8545	0.1608	0.1741
		Title + Category and Subcategory	tr	0.1233	0.6449	0.8530	0.1677	0.1847
		Title + Category and Subcategory + Abstract	tr	0.1242	0.6500	0.8557	0.1684	0.1843
Qwen3-Embedding-0.6B	Average Pooling	Title	tr	0.1284	0.6287	0.8127	0.1408	0.2411

Generalized Embedding Model	User Encoder	Model Input	Language	F1 Score	AUC	Accuracy	Precision	Recall
Qwen3-Embedding-0.6B	Average Pooling	Title + Category and Subcategory	tr	0.1404	0.6365	0.8042	0.1463	0.2767
		Title + Category and Subcategory + Abstract	tr	0.1317	0.6333	0.8181	0.1449	0.2482
	Attention Pooling	Title	tr	0.1292	0.6307	0.8158	0.1427	0.2384
		Title + Category and Subcategory	tr	0.1485	0.6438	0.8040	0.1519	0.2925
		Title + Category and Subcategory + Abstract	tr	0.1384	0.6393	0.8199	0.1512	0.2584
	Dense Layer + Attention Pooling	Title	tr	0.1279	0.6271	0.8405	0.1484	0.2012
		Title + Category and Subcategory	tr	0.1330	0.6297	0.8427	0.1546	0.2121
		Title + Category and Subcategory + Abstract	tr	0.1286	0.6395	0.8500	0.1616	0.2036
	Multi-Head Self-Attention + Attention Pooling	Title	tr	0.1333	0.6315	0.8374	0.1508	0.2124
		Title + Category and Subcategory	tr	0.1284	0.6290	0.8470	0.1535	0.1994
		Title + Category and Subcategory + Abstract	tr	0.1330	0.6388	0.8460	0.1613	0.2131
Qwen3-Embedding-4B	Average Pooling	Title	tr	0.1612	0.6421	0.7806	0.1516	0.3482
		Title + Category and Subcategory	tr	0.1577	0.6433	0.8056	0.1574	0.3016
		Title + Category and Subcategory + Abstract	tr	0.1621	0.6585	0.8087	0.1600	0.3190
	Attention Pooling	Title	tr	0.1605	0.6428	0.7967	0.1561	0.3257
		Title + Category and Subcategory	tr	0.1674	0.6465	0.8120	0.1660	0.3103
		Title + Category and Subcategory + Abstract	tr	0.1675	0.6539	0.8030	0.1595	0.3334
	Dense Layer + Attention Pooling	Title	tr	0.1419	0.6519	0.8464	0.1789	0.2285
		Title + Category and Subcategory	tr	0.1446	0.6503	0.8463	0.1771	0.2321
		Title + Category and Subcategory + Abstract	tr	0.1435	0.6555	0.8450	0.1758	0.2345
	Multi-Head Self-Attention + Attention Pooling	Title	tr	0.1443	0.6560	0.8434	0.1786	0.2370
		Title + Category and Subcategory	tr	0.1506	0.6521	0.8388	0.1740	0.2493
		Title + Category and Subcategory + Abstract	tr	0.1428	0.6575	0.8437	0.1748	0.2352
Qwen3-Embedding-8B	Average Pooling	Title	tr	0.1458	0.6407	0.7823	0.1400	0.3206
		Title + Category and Subcategory	tr	0.1273	0.5991	0.7970	0.1258	0.2551
		Title + Category and Subcategory + Abstract	tr	0.1411	0.6123	0.7837	0.1331	0.2981
	Attention Pooling	Title	tr	0.1463	0.6409	0.7883	0.1426	0.3130
		Title + Category and Subcategory	tr	0.1381	0.6120	0.8074	0.1392	0.2621
		Title + Category and Subcategory + Abstract	tr	0.1502	0.6235	0.7943	0.1435	0.3006
	Dense Layer + Attention Pooling	Title	tr	0.1428	0.6532	0.8422	0.1689	0.2279
		Title + Category and Subcategory	tr	0.1486	0.6478	0.8386	0.1706	0.2474
		Title + Category and Subcategory + Abstract	tr	0.1490	0.6558	0.8371	0.1684	0.2538

Generalized Embedding Model	User Encoder	Model Input	Language	F1 Score	AUC	Accuracy	Precision	Recall
Qwen3-Embedding-8B	Multi-Head Self-Attention + Attention Pooling	Title	tr	0.1435	0.6531	0.8430	0.1708	0.2269
		Title + Category and Subcategory	tr	0.1458	0.6445	0.8439	0.1731	0.2338
		Title + Category and Subcategory + Abstract	tr	0.1439	0.6531	0.8405	0.1674	0.2390
text-embedding-3-large (OpenAI)	Average Pooling	Title	tr	0.1116	0.5954	0.7450	0.1111	0.2845
		Title + Category and Subcategory	tr	0.1127	0.6144	0.8052	0.1252	0.2284
		Title + Category and Subcategory + Abstract	tr	0.1258	0.6175	0.7908	0.1352	0.2682
	Attention Pooling	Title	tr	0.1085	0.6129	0.7883	0.1202	0.2418
		Title + Category and Subcategory	tr	0.1254	0.6305	0.8205	0.1437	0.2348
		Title + Category and Subcategory + Abstract	tr	0.1280	0.6341	0.8133	0.1469	0.2511
	Dense Layer + Attention Pooling	Title	tr	0.1194	0.6417	0.8600	0.1716	0.1758
		Title + Category and Subcategory	tr	0.1354	0.6473	0.8534	0.1755	0.2067
		Title + Category and Subcategory + Abstract	tr	0.1264	0.6567	0.8553	0.1733	0.1997
	Multi-Head Self-Attention + Attention Pooling	Title	tr	0.1253	0.6421	0.8558	0.1704	0.1846
		Title + Category and Subcategory	tr	0.1415	0.6495	0.8484	0.1741	0.2198
		Title + Category and Subcategory + Abstract	tr	0.1236	0.6525	0.8565	0.1713	0.1904
multilingual-e5-large-instruct	Average Pooling	Title	vi	0.1531	0.6436	0.7936	0.1507	0.3215
		Title + Category and Subcategory	vi	0.1528	0.6456	0.8029	0.1558	0.3093
		Title + Category and Subcategory + Abstract	vi	0.1482	0.6434	0.8067	0.1523	0.2976
	Attention Pooling	Title	vi	0.1586	0.6407	0.7872	0.1516	0.3386
		Title + Category and Subcategory	vi	0.1548	0.6436	0.8146	0.1620	0.2951
		Title + Category and Subcategory + Abstract	vi	0.1544	0.6421	0.8077	0.1568	0.3045
	Dense Layer + Attention Pooling	Title	vi	0.1092	0.6248	0.8586	0.1555	0.1519
		Title + Category and Subcategory	vi	0.1225	0.6330	0.8532	0.1631	0.1796
		Title + Category and Subcategory + Abstract	vi	0.1242	0.6330	0.8506	0.1606	0.1862
	Multi-Head Self-Attention + Attention Pooling	Title	vi	0.1083	0.6290	0.8618	0.1607	0.1502
		Title + Category and Subcategory	vi	0.1145	0.6319	0.8554	0.1621	0.1658
		Title + Category and Subcategory + Abstract	vi	0.1200	0.6370	0.8483	0.1557	0.1824
Qwen3-Embedding-0.6B	Average Pooling	Title	vi	0.1256	0.6169	0.8094	0.1377	0.2419
		Title + Category and Subcategory	vi	0.1390	0.6337	0.8148	0.1499	0.2641
		Title + Category and Subcategory + Abstract	vi	0.1295	0.6247	0.8149	0.1405	0.2520
	Attention Pooling	Title	vi	0.1308	0.6187	0.8095	0.1418	0.2496
		Title + Category and Subcategory	vi	0.1455	0.6377	0.8180	0.1558	0.2709

Generalized Embedding Model	User Encoder	Model Input	Language	F1 Score	AUC	Accuracy	Precision	Recall
Qwen3-Embedding-0.6B	Attention Pooling	Title + Category and Subcategory + Abstract	vi	0.1364	0.6306	0.8178	0.1476	0.2611
	Dense Layer + Attention Pooling	Title	vi	0.1255	0.6374	0.8482	0.1567	0.2011
		Title + Category and Subcategory	vi	0.1335	0.6381	0.8483	0.1627	0.2136
		Title + Category and Subcategory + Abstract	vi	0.1280	0.6419	0.8509	0.1641	0.2065
	Multi-Head Self-Attention + Attention Pooling	Title	vi	0.1290	0.6401	0.8444	0.1568	0.2101
		Title + Category and Subcategory	vi	0.1359	0.6376	0.8467	0.1624	0.2157
		Title + Category and Subcategory + Abstract	vi	0.1373	0.6415	0.8408	0.1610	0.2312
Qwen3-Embedding-4B	Average Pooling	Title	vi	0.1478	0.6394	0.7939	0.1471	0.3048
		Title + Category and Subcategory	vi	0.1492	0.6445	0.8138	0.1565	0.2779
		Title + Category and Subcategory + Abstract	vi	0.1517	0.6456	0.8131	0.1559	0.2922
	Attention Pooling	Title	vi	0.1474	0.6399	0.8109	0.1545	0.2819
		Title + Category and Subcategory	vi	0.1618	0.6458	0.8106	0.1623	0.3011
		Title + Category and Subcategory + Abstract	vi	0.1597	0.6425	0.8119	0.1592	0.3034
	Dense Layer + Attention Pooling	Title	vi	0.1269	0.6517	0.8523	0.1735	0.2009
		Title + Category and Subcategory	vi	0.1417	0.6503	0.8443	0.1727	0.2308
		Title + Category and Subcategory + Abstract	vi	0.1383	0.6516	0.8412	0.1693	0.2331
	Multi-Head Self-Attention + Attention Pooling	Title	vi	0.1314	0.6544	0.8478	0.1730	0.2133
		Title + Category and Subcategory	vi	0.1411	0.6508	0.8443	0.1724	0.2296
		Title + Category and Subcategory + Abstract	vi	0.1371	0.6540	0.8432	0.1717	0.2289
Qwen3-Embedding-8B	Average Pooling	Title	vi	0.1342	0.6349	0.8147	0.1481	0.2601
		Title + Category and Subcategory	vi	0.1466	0.6259	0.7970	0.1430	0.2895
		Title + Category and Subcategory + Abstract	vi	0.1479	0.6341	0.8045	0.1466	0.2881
	Attention Pooling	Title	vi	0.1383	0.6343	0.8119	0.1490	0.2692
		Title + Category and Subcategory	vi	0.1587	0.6317	0.7973	0.1508	0.3103
		Title + Category and Subcategory + Abstract	vi	0.1564	0.6350	0.8064	0.1524	0.2973
	Dense Layer + Attention Pooling	Title	vi	0.1463	0.6493	0.8419	0.1692	0.2320
		Title + Category and Subcategory	vi	0.1519	0.6502	0.8387	0.1717	0.2508
		Title + Category and Subcategory + Abstract	vi	0.1556	0.6583	0.8324	0.1686	0.2687
	Multi-Head Self-Attention + Attention Pooling	Title	vi	0.1442	0.6500	0.8443	0.1698	0.2266
		Title + Category and Subcategory	vi	0.1515	0.6463	0.8407	0.1735	0.2467
		Title + Category and Subcategory + Abstract	vi	0.1459	0.6555	0.8407	0.1688	0.2424
text-embedding-3-large (OpenAI)	Average Pooling	Title	vi	0.1240	0.6335	0.8175	0.1460	0.2444

Generalized Embedding Model	User Encoder	Model Input	Language	F1 Score	AUC	Accuracy	Precision	Recall
text-embedding-3-large (OpenAI)	Average Pooling	Title + Category and Subcategory	vi	0.1294	0.6384	0.8278	0.1520	0.2347
		Title + Category and Subcategory + Abstract	vi	0.1437	0.6438	0.8138	0.1534	0.2756
	Attention Pooling	Title	vi	0.1319	0.6392	0.8290	0.1592	0.2449
		Title + Category and Subcategory	vi	0.1457	0.6429	0.8297	0.1660	0.2571
		Title + Category and Subcategory + Abstract	vi	0.1469	0.6532	0.8256	0.1623	0.2668
	Dense Layer + Attention Pooling	Title	vi	0.1447	0.6508	0.8492	0.1755	0.2209
		Title + Category and Subcategory	vi	0.1492	0.6511	0.8469	0.1760	0.2339
		Title + Category and Subcategory + Abstract	vi	0.1474	0.6649	0.8525	0.1838	0.2280
	Multi-Head Self-Attention + Attention Pooling	Title	vi	0.1414	0.6494	0.8480	0.1703	0.2139
		Title + Category and Subcategory	vi	0.1502	0.6542	0.8435	0.1740	0.2402
		Title + Category and Subcategory + Abstract	vi	0.1463	0.6619	0.8503	0.1793	0.2258
multilingual-e5-large-instruct	Average Pooling	Title	zh	0.1504	0.6315	0.7640	0.1380	0.3519
		Title + Category and Subcategory	zh	0.1448	0.6393	0.8031	0.1490	0.2926
		Title + Category and Subcategory + Abstract	zh	0.1526	0.6495	0.7977	0.1551	0.3204
	Attention Pooling	Title	zh	0.1533	0.6300	0.7573	0.1384	0.3637
		Title + Category and Subcategory	zh	0.1499	0.6382	0.8080	0.1539	0.2928
		Title + Category and Subcategory + Abstract	zh	0.1541	0.6466	0.8078	0.1597	0.3066
	Dense Layer + Attention Pooling	Title	zh	0.1172	0.6186	0.8350	0.1357	0.1822
		Title + Category and Subcategory	zh	0.1381	0.6370	0.8338	0.1548	0.2262
		Title + Category and Subcategory + Abstract	zh	0.1423	0.6427	0.8371	0.1604	0.2313
	Multi-Head Self-Attention + Attention Pooling	Title	zh	0.1217	0.6254	0.8364	0.1453	0.1922
		Title + Category and Subcategory	zh	0.1382	0.6410	0.8362	0.1603	0.2246
		Title + Category and Subcategory + Abstract	zh	0.1381	0.6462	0.8377	0.1610	0.2250
Qwen3-Embedding-0.6B	Average Pooling	Title	zh	0.1228	0.6044	0.7647	0.1188	0.2846
		Title + Category and Subcategory	zh	0.1378	0.6316	0.7792	0.1339	0.3104
		Title + Category and Subcategory + Abstract	zh	0.1261	0.6223	0.8124	0.1380	0.2489
	Attention Pooling	Title	zh	0.1254	0.6077	0.7649	0.1208	0.2900
		Title + Category and Subcategory	zh	0.1451	0.6358	0.7825	0.1400	0.3195
		Title + Category and Subcategory + Abstract	zh	0.1308	0.6275	0.8133	0.1419	0.2563
	Dense Layer + Attention Pooling	Title	zh	0.1121	0.6247	0.8407	0.1347	0.1835
		Title + Category and Subcategory	zh	0.1296	0.6396	0.8454	0.1552	0.2101
		Title + Category and Subcategory + Abstract	zh	0.1320	0.6405	0.8405	0.1558	0.2201

Generalized Embedding Model	User Encoder	Model Input	Language	F1 Score	AUC	Accuracy	Precision	Recall
Qwen3-Embedding-0.6B	Multi-Head Self-Attention + Attention Pooling	Title	zh	0.1154	0.6295	0.8361	0.1346	0.1933
		Title + Category and Subcategory	zh	0.1369	0.6403	0.8398	0.1558	0.2255
		Title + Category and Subcategory + Abstract	zh	0.1334	0.6399	0.8404	0.1572	0.2214
Qwen3-Embedding-4B	Average Pooling	Title	zh	0.1415	0.6378	0.7850	0.1392	0.3120
		Title + Category and Subcategory	zh	0.1368	0.6438	0.8245	0.1516	0.2525
		Title + Category and Subcategory + Abstract	zh	0.1422	0.6472	0.8181	0.1516	0.2730
	Attention Pooling	Title	zh	0.1336	0.6381	0.8123	0.1449	0.2615
		Title + Category and Subcategory	zh	0.1465	0.6472	0.8296	0.1617	0.2607
		Title + Category and Subcategory + Abstract	zh	0.1488	0.6457	0.8216	0.1570	0.2752
	Dense Layer + Attention Pooling	Title	zh	0.1143	0.6374	0.8522	0.1596	0.1798
		Title + Category and Subcategory	zh	0.1376	0.6424	0.8487	0.1715	0.2162
		Title + Category and Subcategory + Abstract	zh	0.1334	0.6448	0.8469	0.1685	0.2124
	Multi-Head Self-Attention + Attention Pooling	Title	zh	0.1146	0.6383	0.8480	0.1572	0.1855
		Title + Category and Subcategory	zh	0.1367	0.6444	0.8483	0.1699	0.2159
		Title + Category and Subcategory + Abstract	zh	0.1374	0.6478	0.8422	0.1690	0.2240
Qwen3-Embedding-8B	Average Pooling	Title	zh	0.1479	0.6267	0.7681	0.1371	0.3343
		Title + Category and Subcategory	zh	0.1457	0.6089	0.7680	0.1324	0.3036
		Title + Category and Subcategory + Abstract	zh	0.1530	0.6191	0.7742	0.1403	0.3172
	Attention Pooling	Title	zh	0.1464	0.6266	0.7828	0.1404	0.3125
		Title + Category and Subcategory	zh	0.1589	0.6244	0.7904	0.1489	0.3067
		Title + Category and Subcategory + Abstract	zh	0.1616	0.6258	0.7793	0.1473	0.3285
	Dense Layer + Attention Pooling	Title	zh	0.1423	0.6374	0.8354	0.1596	0.2264
		Title + Category and Subcategory	zh	0.1588	0.6453	0.8358	0.1736	0.2573
		Title + Category and Subcategory + Abstract	zh	0.1611	0.6523	0.8299	0.1692	0.2747
	Multi-Head Self-Attention + Attention Pooling	Title	zh	0.1478	0.6406	0.8299	0.1613	0.2419
		Title + Category and Subcategory	zh	0.1538	0.6395	0.8398	0.1748	0.2436
		Title + Category and Subcategory + Abstract	zh	0.1616	0.6500	0.8284	0.1692	0.2753
text-embedding-3-large (OpenAI)	Average Pooling	Title	zh	0.1182	0.6125	0.7874	0.1255	0.2633
		Title + Category and Subcategory	zh	0.1210	0.6218	0.8232	0.1396	0.2297
		Title + Category and Subcategory + Abstract	zh	0.1363	0.6481	0.8237	0.1549	0.2580
	Attention Pooling	Title	zh	0.1189	0.6235	0.8045	0.1320	0.2479
		Title + Category and Subcategory	zh	0.1324	0.6300	0.8283	0.1513	0.2409

Generalized Embedding Model	User Encoder	Model Input	Language	F1 Score	AUC	Accuracy	Precision	Recall
text-embedding-3-large (OpenAI)	Attention Pooling	Title + Category and Subcategory + Abstract	zh	0.1404	0.6529	0.8267	0.1584	0.2612
	Dense Layer + Attention Pooling	Title	zh	0.1251	0.6385	0.8455	0.1507	0.1920
		Title + Category and Subcategory	zh	0.1460	0.6435	0.8395	0.1642	0.2410
		Title + Category and Subcategory + Abstract	zh	0.1481	0.6554	0.8398	0.1679	0.2465
	Multi-Head Self-Attention + Attention Pooling	Title	zh	0.1247	0.6335	0.8394	0.1435	0.1932
		Title + Category and Subcategory	zh	0.1449	0.6419	0.8366	0.1603	0.2411
		Title + Category and Subcategory + Abstract	zh	0.1445	0.6506	0.8389	0.1625	0.2394

The results for both training and evaluation in the target language is given in the table below.

Table A.15: Performance of Models on Multilingual Candidate News Classification for User trained and evaluated in target language

Generalized Embedding Model	User Encoder	Model Input	Language	F1 Score	AUC	Accuracy	Precision	Recall
multilingual-e5-large-instruct	Average Pooling	Title	en	0.1450	0.6589	0.8069	0.1509	0.2973
		Title + Category and Subcategory	en	0.1410	0.6524	0.8226	0.1571	0.2659
		Title + Category and Subcategory + Abstract	en	0.1488	0.6615	0.8199	0.1621	0.2853
	Attention Pooling	Title	en	0.1516	0.6565	0.8029	0.1542	0.3123
		Title + Category and Subcategory	en	0.1498	0.6524	0.8268	0.1658	0.2731
		Title + Category and Subcategory + Abstract	en	0.1568	0.6601	0.8203	0.1680	0.2958
	Dense Layer + Attention Pooling	Title	en	0.1370	0.6516	0.8578	0.1805	0.1982
		Title + Category and Subcategory	en	0.1388	0.6548	0.8546	0.1795	0.2120
		Title + Category and Subcategory + Abstract	en	0.1369	0.6563	0.8561	0.1810	0.2106
	Multi-Head Self-Attention + Attention Pooling	Title	en	0.1356	0.6505	0.8582	0.1818	0.1994
		Title + Category and Subcategory	en	0.1299	0.6539	0.8601	0.1819	0.1966
		Title + Category and Subcategory + Abstract	en	0.1336	0.6563	0.8552	0.1773	0.2098
Qwen3-Embedding-0.6B	Average Pooling	Title	en	0.1436	0.6304	0.7854	0.1399	0.3045
		Title + Category and Subcategory	en	0.1346	0.6140	0.7963	0.1384	0.2703
		Title + Category and Subcategory + Abstract	en	0.1348	0.6256	0.7998	0.1413	0.2728
	Attention Pooling	Title	en	0.1455	0.6315	0.7957	0.1449	0.2953
		Title + Category and Subcategory	en	0.1411	0.6217	0.7962	0.1430	0.2808
		Title + Category and Subcategory + Abstract	en	0.1390	0.6312	0.8024	0.1463	0.2777

Generalized Embedding Model	User Encoder	Model Input	Language	F1 Score	AUC	Accuracy	Precision	Recall
Qwen3-Embedding-0.6B	Dense Layer + Attention Pooling	Title	en	0.1188	0.6448	0.8507	0.1541	0.1842
		Title + Category and Subcategory	en	0.1315	0.6467	0.8490	0.1632	0.2079
		Title + Category and Subcategory + Abstract	en	0.1276	0.6498	0.8501	0.1671	0.2062
	Multi-Head Self-Attention + Attention Pooling	Title	en	0.1221	0.6465	0.8482	0.1557	0.1916
		Title + Category and Subcategory	en	0.1293	0.6486	0.8506	0.1617	0.2002
		Title + Category and Subcategory + Abstract	en	0.1219	0.6505	0.8551	0.1676	0.1897
Qwen3-Embedding-4B	Average Pooling	Title	en	0.1431	0.6304	0.8113	0.1505	0.2678
		Title + Category and Subcategory	en	0.1242	0.6243	0.8195	0.1394	0.2293
		Title + Category and Subcategory + Abstract	en	0.1231	0.6225	0.8258	0.1415	0.2219
	Attention Pooling	Title	en	0.1433	0.6318	0.8216	0.1568	0.2554
		Title + Category and Subcategory	en	0.1386	0.6306	0.8204	0.1517	0.2496
		Title + Category and Subcategory + Abstract	en	0.1348	0.6284	0.8214	0.1506	0.2461
	Dense Layer + Attention Pooling	Title	en	0.1266	0.6525	0.8535	0.1746	0.1948
		Title + Category and Subcategory	en	0.1364	0.6541	0.8529	0.1753	0.2109
		Title + Category and Subcategory + Abstract	en	0.1320	0.6551	0.8535	0.1770	0.2083
	Multi-Head Self-Attention + Attention Pooling	Title	en	0.1272	0.6552	0.8533	0.1752	0.1975
		Title + Category and Subcategory	en	0.1383	0.6561	0.8513	0.1753	0.2170
		Title + Category and Subcategory + Abstract	en	0.1391	0.6568	0.8453	0.1745	0.2309
Qwen3-Embedding-8B	Average Pooling	Title	en	0.1404	0.6388	0.8124	0.1501	0.2715
		Title + Category and Subcategory	en	0.1554	0.6058	0.7622	0.1417	0.3330
		Title + Category and Subcategory + Abstract	en	0.1458	0.6123	0.7896	0.1422	0.2874
	Attention Pooling	Title	en	0.1420	0.6391	0.8186	0.1543	0.2643
		Title + Category and Subcategory	en	0.1532	0.6238	0.8089	0.1550	0.2770
		Title + Category and Subcategory + Abstract	en	0.1520	0.6271	0.8088	0.1541	0.2779
	Dense Layer + Attention Pooling	Title	en	0.1312	0.6557	0.8545	0.1732	0.1983
		Title + Category and Subcategory	en	0.1489	0.6539	0.8381	0.1704	0.2488
		Title + Category and Subcategory + Abstract	en	0.1500	0.6557	0.8424	0.1736	0.2463
	Multi-Head Self-Attention + Attention Pooling	Title	en	0.1371	0.6549	0.8506	0.1716	0.2103
		Title + Category and Subcategory	en	0.1474	0.6521	0.8402	0.1682	0.2408
		Title + Category and Subcategory + Abstract	en	0.1497	0.6545	0.8436	0.1729	0.2423
text-embedding-3-large (OpenAI)	Average Pooling	Title	en	0.1150	0.6186	0.8160	0.1357	0.2241
		Title + Category and Subcategory	en	0.1283	0.6205	0.8222	0.1455	0.2353

Generalized Embedding Model	User Encoder	Model Input	Language	F1 Score	AUC	Accuracy	Precision	Recall
text-embedding-3-large (OpenAI)	Average Pooling	Title + Category and Subcategory + Abstract	en	0.1272	0.6308	0.8196	0.1485	0.2444
	Attention Pooling	Title	en	0.1199	0.6251	0.8252	0.1440	0.2195
		Title + Category and Subcategory	en	0.1338	0.6297	0.8337	0.1558	0.2287
		Title + Category and Subcategory + Abstract	en	0.1283	0.6370	0.8351	0.1567	0.2261
	Dense Layer + Attention Pooling	Title	en	0.1314	0.6593	0.8629	0.1850	0.1900
		Title + Category and Subcategory	en	0.1419	0.6575	0.8544	0.1822	0.2207
		Title + Category and Subcategory + Abstract	en	0.1343	0.6588	0.8600	0.1855	0.2026
	Multi-Head Self-Attention + Attention Pooling	Title	en	0.1404	0.6606	0.8551	0.1811	0.2149
		Title + Category and Subcategory	en	0.1409	0.6591	0.8523	0.1810	0.2250
		Title + Category and Subcategory + Abstract	en	0.1369	0.6600	0.8575	0.1855	0.2115
multilingual-e5-large-instruct	Average Pooling	Title	tr	0.1593	0.6469	0.7841	0.1527	0.3440
		Title + Category and Subcategory	tr	0.1543	0.6459	0.8132	0.1620	0.2937
		Title + Category and Subcategory + Abstract	tr	0.1524	0.6520	0.8103	0.1593	0.2976
	Attention Pooling	Title	tr	0.1596	0.6465	0.7866	0.1537	0.3404
		Title + Category and Subcategory	tr	0.1602	0.6447	0.8129	0.1651	0.3027
		Title + Category and Subcategory + Abstract	tr	0.1554	0.6504	0.8142	0.1623	0.2959
	Dense Layer + Attention Pooling	Title	tr	0.1404	0.6505	0.8459	0.1725	0.2202
		Title + Category and Subcategory	tr	0.1402	0.6545	0.8469	0.1750	0.2257
		Title + Category and Subcategory + Abstract	tr	0.1414	0.6547	0.8439	0.1731	0.2328
	Multi-Head Self-Attention + Attention Pooling	Title	tr	0.1302	0.6513	0.8516	0.1745	0.2031
		Title + Category and Subcategory	tr	0.1322	0.6550	0.8527	0.1781	0.2109
		Title + Category and Subcategory + Abstract	tr	0.1302	0.6557	0.8507	0.1758	0.2114
Qwen3-Embedding-0.6B	Average Pooling	Title	tr	0.1284	0.6287	0.8127	0.1408	0.2411
		Title + Category and Subcategory	tr	0.1404	0.6365	0.8042	0.1463	0.2767
		Title + Category and Subcategory + Abstract	tr	0.1317	0.6333	0.8181	0.1449	0.2482
	Attention Pooling	Title	tr	0.1304	0.6330	0.8185	0.1457	0.2384
		Title + Category and Subcategory	tr	0.1456	0.6446	0.8168	0.1554	0.2716
		Title + Category and Subcategory + Abstract	tr	0.1425	0.6400	0.8181	0.1532	0.2670
	Dense Layer + Attention Pooling	Title	tr	0.1136	0.6370	0.8576	0.1543	0.1632
		Title + Category and Subcategory	tr	0.1300	0.6471	0.8551	0.1684	0.1964
		Title + Category and Subcategory + Abstract	tr	0.1350	0.6516	0.8538	0.1723	0.2056
		Title	tr	0.1118	0.6383	0.8579	0.1534	0.1595

Generalized Embedding Model	User Encoder	Model Input	Language	F1 Score	AUC	Accuracy	Precision	Recall
Qwen3-Embedding-0.6B	Multi-Head Self-Attention + Attention Pooling	Title + Category and Subcategory	tr	0.1287	0.6479	0.8565	0.1678	0.1911
		Title + Category and Subcategory + Abstract	tr	0.1358	0.6514	0.8529	0.1700	0.2052
Qwen3-Embedding-4B	Average Pooling	Title	tr	0.1612	0.6421	0.7806	0.1516	0.3482
		Title + Category and Subcategory	tr	0.1577	0.6433	0.8056	0.1574	0.3016
		Title + Category and Subcategory + Abstract	tr	0.1621	0.6585	0.8087	0.1600	0.3190
	Attention Pooling	Title	tr	0.1647	0.6413	0.7855	0.1545	0.3470
		Title + Category and Subcategory	tr	0.1704	0.6450	0.8101	0.1668	0.3166
		Title + Category and Subcategory + Abstract	tr	0.1641	0.6576	0.8095	0.1614	0.3201
	Dense Layer + Attention Pooling	Title	tr	0.1497	0.6566	0.8433	0.1756	0.2407
		Title + Category and Subcategory	tr	0.1532	0.6606	0.8445	0.1809	0.2467
		Title + Category and Subcategory + Abstract	tr	0.1495	0.6609	0.8449	0.1768	0.2435
	Multi-Head Self-Attention + Attention Pooling	Title	tr	0.1485	0.6583	0.8434	0.1760	0.2415
		Title + Category and Subcategory	tr	0.1521	0.6602	0.8450	0.1814	0.2447
		Title + Category and Subcategory + Abstract	tr	0.1461	0.6604	0.8471	0.1777	0.2373
Qwen3-Embedding-8B	Average Pooling	Title	tr	0.1458	0.6407	0.7823	0.1400	0.3206
		Title + Category and Subcategory	tr	0.1273	0.5991	0.7970	0.1258	0.2551
		Title + Category and Subcategory + Abstract	tr	0.1411	0.6123	0.7837	0.1331	0.2981
	Attention Pooling	Title	tr	0.1458	0.6400	0.7894	0.1425	0.3098
		Title + Category and Subcategory	tr	0.1386	0.6140	0.8109	0.1416	0.2588
		Title + Category and Subcategory + Abstract	tr	0.1504	0.6242	0.7959	0.1452	0.2995
	Dense Layer + Attention Pooling	Title	tr	0.1318	0.6513	0.8494	0.1708	0.2077
		Title + Category and Subcategory	tr	0.1406	0.6533	0.8443	0.1736	0.2323
		Title + Category and Subcategory + Abstract	tr	0.1402	0.6570	0.8439	0.1709	0.2325
	Multi-Head Self-Attention + Attention Pooling	Title	tr	0.1391	0.6533	0.8436	0.1714	0.2259
		Title + Category and Subcategory	tr	0.1477	0.6509	0.8372	0.1697	0.2481
		Title + Category and Subcategory + Abstract	tr	0.1362	0.6544	0.8442	0.1690	0.2267
text-embedding-3-large (OpenAI)	Average Pooling	Title	tr	0.1116	0.5954	0.7450	0.1111	0.2845
		Title + Category and Subcategory	tr	0.1127	0.6144	0.8052	0.1252	0.2284
		Title + Category and Subcategory + Abstract	tr	0.1258	0.6175	0.7908	0.1352	0.2682
	Attention Pooling	Title	tr	0.1132	0.6172	0.8010	0.1309	0.2402
		Title + Category and Subcategory	tr	0.1359	0.6340	0.8140	0.1501	0.2601
		Title + Category and Subcategory + Abstract	tr	0.1308	0.6374	0.8134	0.1497	0.2578

Generalized Embedding Model	User Encoder	Model Input	Language	F1 Score	AUC	Accuracy	Precision	Recall
text-embedding-3-large (OpenAI)	Dense Layer + Attention Pooling	Title	tr	0.1341	0.6484	0.8552	0.1718	0.1928
		Title + Category and Subcategory	tr	0.1365	0.6530	0.8554	0.1785	0.2087
		Title + Category and Subcategory + Abstract	tr	0.1269	0.6541	0.8617	0.1811	0.1873
	Multi-Head Self-Attention + Attention Pooling	Title	tr	0.1173	0.6478	0.8640	0.1753	0.1632
		Title + Category and Subcategory	tr	0.1398	0.6523	0.8519	0.1774	0.2155
		Title + Category and Subcategory + Abstract	tr	0.1311	0.6553	0.8579	0.1798	0.2005
multilingual-e5-large-instruct	Average Pooling	Title	vi	0.1531	0.6436	0.7936	0.1507	0.3215
		Title + Category and Subcategory	vi	0.1528	0.6456	0.8029	0.1558	0.3093
		Title + Category and Subcategory + Abstract	vi	0.1482	0.6434	0.8067	0.1523	0.2976
	Attention Pooling	Title	vi	0.1534	0.6432	0.7941	0.1510	0.3213
		Title + Category and Subcategory	vi	0.1534	0.6443	0.8143	0.1611	0.2939
		Title + Category and Subcategory + Abstract	vi	0.1540	0.6420	0.8082	0.1567	0.3033
	Dense Layer + Attention Pooling	Title	vi	0.1406	0.6470	0.8425	0.1712	0.2268
		Title + Category and Subcategory	vi	0.1382	0.6523	0.8469	0.1752	0.2236
		Title + Category and Subcategory + Abstract	vi	0.1345	0.6520	0.8459	0.1730	0.2216
	Multi-Head Self-Attention + Attention Pooling	Title	vi	0.1287	0.6461	0.8500	0.1736	0.2050
		Title + Category and Subcategory	vi	0.1279	0.6504	0.8537	0.1777	0.2035
		Title + Category and Subcategory + Abstract	vi	0.1303	0.6520	0.8492	0.1748	0.2131
Qwen3-Embedding-0.6B	Average Pooling	Title	vi	0.1256	0.6169	0.8094	0.1377	0.2419
		Title + Category and Subcategory	vi	0.1390	0.6337	0.8148	0.1499	0.2641
		Title + Category and Subcategory + Abstract	vi	0.1295	0.6247	0.8149	0.1405	0.2520
	Attention Pooling	Title	vi	0.1288	0.6195	0.8138	0.1428	0.2421
		Title + Category and Subcategory	vi	0.1465	0.6379	0.8165	0.1563	0.2754
		Title + Category and Subcategory + Abstract	vi	0.1357	0.6309	0.8205	0.1494	0.2568
	Dense Layer + Attention Pooling	Title	vi	0.1213	0.6376	0.8543	0.1617	0.1846
		Title + Category and Subcategory	vi	0.1347	0.6449	0.8492	0.1686	0.2159
		Title + Category and Subcategory + Abstract	vi	0.1337	0.6493	0.8501	0.1695	0.2140
	Multi-Head Self-Attention + Attention Pooling	Title	vi	0.1228	0.6398	0.8534	0.1609	0.1855
		Title + Category and Subcategory	vi	0.1368	0.6459	0.8483	0.1678	0.2183
		Title + Category and Subcategory + Abstract	vi	0.1381	0.6497	0.8472	0.1695	0.2229
Qwen3-Embedding-4B	Average Pooling	Title	vi	0.1478	0.6394	0.7939	0.1471	0.3048
		Title + Category and Subcategory	vi	0.1492	0.6445	0.8138	0.1565	0.2779

Generalized Embedding Model	User Encoder	Model Input	Language	F1 Score	AUC	Accuracy	Precision	Recall
Qwen3-Embedding-4B	Average Pooling	Title + Category and Subcategory + Abstract	vi	0.1517	0.6456	0.8131	0.1559	0.2922
	Attention Pooling	Title	vi	0.1484	0.6393	0.8063	0.1532	0.2894
		Title + Category and Subcategory	vi	0.1637	0.6456	0.8063	0.1614	0.3102
		Title + Category and Subcategory + Abstract	vi	0.1585	0.6445	0.8073	0.1568	0.3102
	Dense Layer + Attention Pooling	Title	vi	0.1361	0.6550	0.8490	0.1736	0.2162
		Title + Category and Subcategory	vi	0.1442	0.6560	0.8457	0.1769	0.2341
		Title + Category and Subcategory + Abstract	vi	0.1419	0.6584	0.8461	0.1740	0.2322
	Multi-Head Self-Attention + Attention Pooling	Title	vi	0.1378	0.6570	0.8482	0.1748	0.2195
		Title + Category and Subcategory	vi	0.1448	0.6563	0.8460	0.1777	0.2345
		Title + Category and Subcategory + Abstract	vi	0.1433	0.6588	0.8442	0.1742	0.2376
Qwen3-Embedding-8B	Average Pooling	Title	vi	0.1342	0.6349	0.8147	0.1481	0.2601
		Title + Category and Subcategory	vi	0.1466	0.6259	0.7970	0.1430	0.2895
		Title + Category and Subcategory + Abstract	vi	0.1479	0.6341	0.8045	0.1466	0.2881
	Attention Pooling	Title	vi	0.1371	0.6347	0.8110	0.1482	0.2689
		Title + Category and Subcategory	vi	0.1519	0.6319	0.8158	0.1541	0.2754
		Title + Category and Subcategory + Abstract	vi	0.1536	0.6348	0.8130	0.1540	0.2839
	Dense Layer + Attention Pooling	Title	vi	0.1316	0.6531	0.8564	0.1774	0.1970
		Title + Category and Subcategory	vi	0.1453	0.6547	0.8441	0.1755	0.2390
		Title + Category and Subcategory + Abstract	vi	0.1426	0.6575	0.8432	0.1718	0.2382
	Multi-Head Self-Attention + Attention Pooling	Title	vi	0.1362	0.6535	0.8532	0.1763	0.2076
		Title + Category and Subcategory	vi	0.1456	0.6522	0.8399	0.1732	0.2453
		Title + Category and Subcategory + Abstract	vi	0.1417	0.6554	0.8456	0.1714	0.2292
text-embedding-3-large (OpenAI)	Average Pooling	Title	vi	0.1240	0.6335	0.8175	0.1460	0.2444
		Title + Category and Subcategory	vi	0.1294	0.6384	0.8278	0.1520	0.2347
		Title + Category and Subcategory + Abstract	vi	0.1437	0.6438	0.8138	0.1534	0.2756
	Attention Pooling	Title	vi	0.1273	0.6373	0.8310	0.1566	0.2349
		Title + Category and Subcategory	vi	0.1490	0.6407	0.8276	0.1666	0.2643
		Title + Category and Subcategory + Abstract	vi	0.1470	0.6527	0.8237	0.1612	0.2705
	Dense Layer + Attention Pooling	Title	vi	0.1341	0.6473	0.8591	0.1750	0.1919
		Title + Category and Subcategory	vi	0.1433	0.6536	0.8490	0.1755	0.2266
		Title + Category and Subcategory + Abstract	vi	0.1388	0.6562	0.8550	0.1785	0.2126
		Title	vi	0.1402	0.6479	0.8527	0.1733	0.2116

Generalized Embedding Model	User Encoder	Model Input	Language	F1 Score	AUC	Accuracy	Precision	Recall
text-embedding-3-large (OpenAI)	Multi-Head Self-Attention + Attention Pooling	Title + Category and Subcategory	vi	0.1466	0.6538	0.8438	0.1742	0.2409
		Title + Category and Subcategory + Abstract	vi	0.1404	0.6581	0.8535	0.1786	0.2176
multilingual-e5-large-instruct	Average Pooling	Title	zh	0.1504	0.6315	0.7640	0.1380	0.3519
		Title + Category and Subcategory	zh	0.1448	0.6393	0.8031	0.1490	0.2926
		Title + Category and Subcategory + Abstract	zh	0.1526	0.6495	0.7977	0.1551	0.3204
	Attention Pooling	Title	zh	0.1528	0.6307	0.7600	0.1386	0.3601
		Title + Category and Subcategory	zh	0.1527	0.6385	0.8007	0.1525	0.3085
		Title + Category and Subcategory + Abstract	zh	0.1529	0.6475	0.8061	0.1582	0.3078
	Dense Layer + Attention Pooling	Title	zh	0.1209	0.6386	0.8448	0.1547	0.1950
		Title + Category and Subcategory	zh	0.1483	0.6526	0.8274	0.1652	0.2684
		Title + Category and Subcategory + Abstract	zh	0.1311	0.6510	0.8458	0.1688	0.2192
	Multi-Head Self-Attention + Attention Pooling	Title	zh	0.1203	0.6403	0.8452	0.1562	0.1924
		Title + Category and Subcategory	zh	0.1462	0.6526	0.8322	0.1670	0.2559
		Title + Category and Subcategory + Abstract	zh	0.1285	0.6525	0.8488	0.1719	0.2101
Qwen3-Embedding-0.6B	Average Pooling	Title	zh	0.1228	0.6044	0.7647	0.1188	0.2846
		Title + Category and Subcategory	zh	0.1378	0.6316	0.7792	0.1339	0.3104
		Title + Category and Subcategory + Abstract	zh	0.1261	0.6223	0.8124	0.1380	0.2489
	Attention Pooling	Title	zh	0.1212	0.6078	0.7810	0.1224	0.2644
		Title + Category and Subcategory	zh	0.1459	0.6358	0.7837	0.1411	0.3189
		Title + Category and Subcategory + Abstract	zh	0.1305	0.6277	0.8172	0.1438	0.2500
	Dense Layer + Attention Pooling	Title	zh	0.1128	0.6324	0.8506	0.1473	0.1734
		Title + Category and Subcategory	zh	0.1336	0.6449	0.8460	0.1657	0.2203
		Title + Category and Subcategory + Abstract	zh	0.1387	0.6491	0.8406	0.1623	0.2329
	Multi-Head Self-Attention + Attention Pooling	Title	zh	0.1067	0.6325	0.8559	0.1482	0.1583
		Title + Category and Subcategory	zh	0.1343	0.6459	0.8454	0.1657	0.2214
		Title + Category and Subcategory + Abstract	zh	0.1293	0.6509	0.8501	0.1643	0.2069
Qwen3-Embedding-4B	Average Pooling	Title	zh	0.1415	0.6378	0.7850	0.1392	0.3120
		Title + Category and Subcategory	zh	0.1368	0.6438	0.8245	0.1516	0.2525
		Title + Category and Subcategory + Abstract	zh	0.1422	0.6472	0.8181	0.1516	0.2730
	Attention Pooling	Title	zh	0.1364	0.6378	0.8089	0.1454	0.2706
		Title + Category and Subcategory	zh	0.1542	0.6472	0.8161	0.1588	0.2913
		Title + Category and Subcategory + Abstract	zh	0.1448	0.6466	0.8261	0.1568	0.2638

Generalized Embedding Model	User Encoder	Model Input	Language	F1 Score	AUC	Accuracy	Precision	Recall
Qwen3-Embedding-4B	Dense Layer + Attention Pooling	Title	zh	0.1263	0.6478	0.8524	0.1669	0.1916
		Title + Category and Subcategory	zh	0.1510	0.6567	0.8422	0.1768	0.2492
		Title + Category and Subcategory + Abstract	zh	0.1490	0.6562	0.8395	0.1724	0.2499
	Multi-Head Self-Attention + Attention Pooling	Title	zh	0.1339	0.6495	0.8443	0.1651	0.2116
		Title + Category and Subcategory	zh	0.1472	0.6566	0.8458	0.1771	0.2389
		Title + Category and Subcategory + Abstract	zh	0.1553	0.6556	0.8338	0.1691	0.2622
Qwen3-Embedding-8B	Average Pooling	Title	zh	0.1479	0.6267	0.7681	0.1371	0.3343
		Title + Category and Subcategory	zh	0.1457	0.6089	0.7680	0.1324	0.3036
		Title + Category and Subcategory + Abstract	zh	0.1530	0.6191	0.7742	0.1403	0.3172
	Attention Pooling	Title	zh	0.1469	0.6263	0.7855	0.1426	0.3108
		Title + Category and Subcategory	zh	0.1628	0.6274	0.7889	0.1518	0.3184
		Title + Category and Subcategory + Abstract	zh	0.1609	0.6269	0.7821	0.1477	0.3243
	Dense Layer + Attention Pooling	Title	zh	0.1356	0.6478	0.8435	0.1655	0.2176
		Title + Category and Subcategory	zh	0.1532	0.6527	0.8428	0.1751	0.2479
		Title + Category and Subcategory + Abstract	zh	0.1465	0.6569	0.8437	0.1729	0.2418
	Multi-Head Self-Attention + Attention Pooling	Title	zh	0.1373	0.6489	0.8448	0.1657	0.2139
		Title + Category and Subcategory	zh	0.1570	0.6518	0.8410	0.1743	0.2516
		Title + Category and Subcategory + Abstract	zh	0.1491	0.6556	0.8401	0.1709	0.2487
text-embedding-3-large (OpenAI)	Average Pooling	Title	zh	0.1182	0.6125	0.7874	0.1255	0.2633
		Title + Category and Subcategory	zh	0.1210	0.6218	0.8232	0.1396	0.2297
		Title + Category and Subcategory + Abstract	zh	0.1363	0.6481	0.8237	0.1549	0.2580
	Attention Pooling	Title	zh	0.1177	0.6244	0.8147	0.1356	0.2332
		Title + Category and Subcategory	zh	0.1370	0.6322	0.8266	0.1533	0.2496
		Title + Category and Subcategory + Abstract	zh	0.1399	0.6535	0.8290	0.1592	0.2571
	Dense Layer + Attention Pooling	Title	zh	0.1240	0.6451	0.8510	0.1571	0.1882
		Title + Category and Subcategory	zh	0.1449	0.6534	0.8449	0.1731	0.2385
		Title + Category and Subcategory + Abstract	zh	0.1364	0.6588	0.8552	0.1789	0.2109
	Multi-Head Self-Attention + Attention Pooling	Title	zh	0.1234	0.6459	0.8519	0.1565	0.1841
		Title + Category and Subcategory	zh	0.1439	0.6542	0.8463	0.1737	0.2335
		Title + Category and Subcategory + Abstract	zh	0.1422	0.6596	0.8490	0.1755	0.2270

A.2.4 Task 4: Candidate News Ranking for User

The results for training in English only and evaluation in the target language is given in the table below.

Table A.16: Performance of Models on Multilingual Candidate News Ranking for User trained on English only

Generalized Embedding Model	User Encoder	Model Input	Language	AUC	MRR	NDCG@5	NDCG@10
multilingual-e5-large-instruct	Average Pooling	Title	en	0.6589	0.3141	0.3483	0.4071
		Title + Category and Subcategory	en	0.6524	0.3168	0.3493	0.4088
		Title + Category and Subcategory + Abstract	en	0.6615	0.3231	0.3584	0.4165
	Attention Pooling	Title	en	0.6565	0.3169	0.3520	0.4104
		Title + Category and Subcategory	en	0.6524	0.3207	0.3549	0.4141
		Title + Category and Subcategory + Abstract	en	0.6601	0.3256	0.3622	0.4201
	Dense Layer + Attention Pooling	Title	en	0.6516	0.3197	0.3518	0.4130
		Title + Category and Subcategory	en	0.6548	0.3231	0.3562	0.4162
		Title + Category and Subcategory + Abstract	en	0.6563	0.3239	0.3570	0.4175
	Multi-Head Self-Attention + Attention Pooling	Title	en	0.6505	0.3217	0.3530	0.4138
		Title + Category and Subcategory	en	0.6539	0.3244	0.3564	0.4159
		Title + Category and Subcategory + Abstract	en	0.6563	0.3241	0.3568	0.4174
Qwen3-Embedding-0.6B	Average Pooling	Title	en	0.6304	0.3036	0.3340	0.3923
		Title + Category and Subcategory	en	0.6140	0.2925	0.3157	0.3757
		Title + Category and Subcategory + Abstract	en	0.6256	0.3030	0.3296	0.3892
	Attention Pooling	Title	en	0.6315	0.3042	0.3357	0.3940
		Title + Category and Subcategory	en	0.6217	0.2968	0.3220	0.3823
		Title + Category and Subcategory + Abstract	en	0.6312	0.3064	0.3346	0.3942
	Dense Layer + Attention Pooling	Title	en	0.6448	0.3011	0.3318	0.3943
		Title + Category and Subcategory	en	0.6467	0.3106	0.3416	0.4023
		Title + Category and Subcategory + Abstract	en	0.6498	0.3139	0.3453	0.4062
	Multi-Head Self-Attention + Attention Pooling	Title	en	0.6465	0.3020	0.3327	0.3949
		Title + Category and Subcategory	en	0.6486	0.3108	0.3421	0.4028
		Title + Category and Subcategory + Abstract	en	0.6505	0.3137	0.3450	0.4061
Qwen3-Embedding-4B	Average Pooling	Title	en	0.6304	0.3034	0.3333	0.3934
		Title + Category and Subcategory	en	0.6243	0.2973	0.3224	0.3834

Generalized Embedding Model	User Encoder	Model Input	Language	AUC	MRR	NDCG@5	NDCG@10
Qwen3-Embedding-4B	Average Pooling	Title + Category and Subcategory + Abstract	en	0.6225	0.2993	0.3262	0.3867
	Attention Pooling	Title	en	0.6318	0.3064	0.3371	0.3971
		Title + Category and Subcategory	en	0.6306	0.3046	0.3326	0.3924
		Title + Category and Subcategory + Abstract	en	0.6284	0.3048	0.3341	0.3935
	Dense Layer + Attention Pooling	Title	en	0.6525	0.3176	0.3500	0.4101
		Title + Category and Subcategory	en	0.6541	0.3206	0.3535	0.4129
		Title + Category and Subcategory + Abstract	en	0.6551	0.3230	0.3554	0.4154
	Multi-Head Self-Attention + Attention Pooling	Title	en	0.6552	0.3195	0.3510	0.4114
		Title + Category and Subcategory	en	0.6561	0.3228	0.3554	0.4145
		Title + Category and Subcategory + Abstract	en	0.6568	0.3240	0.3561	0.4161
Qwen3-Embedding-8B	Average Pooling	Title	en	0.6388	0.3105	0.3374	0.3974
		Title + Category and Subcategory	en	0.6058	0.2916	0.3098	0.3714
		Title + Category and Subcategory + Abstract	en	0.6123	0.2913	0.3088	0.3720
	Attention Pooling	Title	en	0.6391	0.3111	0.3393	0.3994
		Title + Category and Subcategory	en	0.6238	0.3030	0.3267	0.3868
		Title + Category and Subcategory + Abstract	en	0.6271	0.3025	0.3271	0.3874
	Dense Layer + Attention Pooling	Title	en	0.6557	0.3195	0.3517	0.4124
		Title + Category and Subcategory	en	0.6539	0.3249	0.3573	0.4160
		Title + Category and Subcategory + Abstract	en	0.6557	0.3230	0.3558	0.4163
	Multi-Head Self-Attention + Attention Pooling	Title	en	0.6549	0.3179	0.3494	0.4105
		Title + Category and Subcategory	en	0.6521	0.3231	0.3554	0.4144
		Title + Category and Subcategory + Abstract	en	0.6545	0.3224	0.3551	0.4155
text-embedding-3-large (OpenAI)	Average Pooling	Title	en	0.6186	0.2957	0.3161	0.3777
		Title + Category and Subcategory	en	0.6205	0.2964	0.3172	0.3796
		Title + Category and Subcategory + Abstract	en	0.6308	0.3084	0.3328	0.3929
	Attention Pooling	Title	en	0.6251	0.3019	0.3257	0.3863
		Title + Category and Subcategory	en	0.6297	0.3060	0.3319	0.3924
		Title + Category and Subcategory + Abstract	en	0.6370	0.3143	0.3418	0.4018
	Dense Layer + Attention Pooling	Title	en	0.6593	0.3266	0.3606	0.4210
		Title + Category and Subcategory	en	0.6575	0.3283	0.3620	0.4218
		Title + Category and Subcategory + Abstract	en	0.6588	0.3278	0.3622	0.4222
		Title	en	0.6606	0.3275	0.3614	0.4218

Generalized Embedding Model	User Encoder	Model Input	Language	AUC	MRR	NDCG@5	NDCG@10
text-embedding-3-large (OpenAI)	Multi-Head Self-Attention + Attention Pooling	Title + Category and Subcategory	en	0.6591	0.3297	0.3634	0.4228
		Title + Category and Subcategory + Abstract	en	0.6600	0.3291	0.3636	0.4233
multilingual-e5-large-instruct	Average Pooling	Title	tr	0.6469	0.3158	0.3482	0.4066
		Title + Category and Subcategory	tr	0.6459	0.3144	0.3470	0.4058
		Title + Category and Subcategory + Abstract	tr	0.6520	0.3155	0.3500	0.4085
	Attention Pooling	Title	tr	0.6439	0.3160	0.3492	0.4076
		Title + Category and Subcategory	tr	0.6447	0.3169	0.3507	0.4093
		Title + Category and Subcategory + Abstract	tr	0.6501	0.3170	0.3522	0.4108
	Dense Layer + Attention Pooling	Title	tr	0.6402	0.3087	0.3395	0.4010
		Title + Category and Subcategory	tr	0.6466	0.3171	0.3494	0.4093
		Title + Category and Subcategory + Abstract	tr	0.6428	0.3130	0.3438	0.4046
	Multi-Head Self-Attention + Attention Pooling	Title	tr	0.6403	0.3117	0.3421	0.4027
		Title + Category and Subcategory	tr	0.6449	0.3158	0.3468	0.4068
		Title + Category and Subcategory + Abstract	tr	0.6500	0.3183	0.3497	0.4097
Qwen3-Embedding-0.6B	Average Pooling	Title	tr	0.6287	0.2959	0.3194	0.3811
		Title + Category and Subcategory	tr	0.6365	0.3027	0.3261	0.3880
		Title + Category and Subcategory + Abstract	tr	0.6333	0.3026	0.3292	0.3893
	Attention Pooling	Title	tr	0.6307	0.2982	0.3228	0.3848
		Title + Category and Subcategory	tr	0.6438	0.3104	0.3364	0.3977
		Title + Category and Subcategory + Abstract	tr	0.6393	0.3088	0.3377	0.3973
	Dense Layer + Attention Pooling	Title	tr	0.6271	0.2999	0.3269	0.3882
		Title + Category and Subcategory	tr	0.6297	0.3065	0.3337	0.3946
		Title + Category and Subcategory + Abstract	tr	0.6395	0.3114	0.3420	0.4014
	Multi-Head Self-Attention + Attention Pooling	Title	tr	0.6315	0.3025	0.3302	0.3914
		Title + Category and Subcategory	tr	0.6290	0.3051	0.3323	0.3933
		Title + Category and Subcategory + Abstract	tr	0.6388	0.3115	0.3419	0.4015
Qwen3-Embedding-4B	Average Pooling	Title	tr	0.6421	0.3125	0.3440	0.4029
		Title + Category and Subcategory	tr	0.6433	0.3099	0.3407	0.4010
		Title + Category and Subcategory + Abstract	tr	0.6585	0.3244	0.3617	0.4184
	Attention Pooling	Title	tr	0.6428	0.3145	0.3470	0.4060
		Title + Category and Subcategory	tr	0.6465	0.3165	0.3498	0.4091
		Title + Category and Subcategory + Abstract	tr	0.6539	0.3240	0.3621	0.4190

Generalized Embedding Model	User Encoder	Model Input	Language	AUC	MRR	NDCG@5	NDCG@10
Qwen3-Embedding-4B	Dense Layer + Attention Pooling	Title	tr	0.6519	0.3222	0.3556	0.4142
		Title + Category and Subcategory	tr	0.6503	0.3232	0.3552	0.4140
		Title + Category and Subcategory + Abstract	tr	0.6555	0.3247	0.3577	0.4167
	Multi-Head Self-Attention + Attention Pooling	Title	tr	0.6560	0.3247	0.3584	0.4163
		Title + Category and Subcategory	tr	0.6521	0.3231	0.3552	0.4135
		Title + Category and Subcategory + Abstract	tr	0.6575	0.3259	0.3593	0.4175
Qwen3-Embedding-8B	Average Pooling	Title	tr	0.6407	0.3064	0.3368	0.3972
		Title + Category and Subcategory	tr	0.5991	0.2804	0.3037	0.3653
		Title + Category and Subcategory + Abstract	tr	0.6123	0.2905	0.3147	0.3781
	Attention Pooling	Title	tr	0.6409	0.3074	0.3382	0.3988
		Title + Category and Subcategory	tr	0.6120	0.2935	0.3208	0.3810
		Title + Category and Subcategory + Abstract	tr	0.6235	0.2996	0.3276	0.3896
	Dense Layer + Attention Pooling	Title	tr	0.6532	0.3188	0.3518	0.4126
		Title + Category and Subcategory	tr	0.6478	0.3207	0.3531	0.4127
		Title + Category and Subcategory + Abstract	tr	0.6558	0.3218	0.3567	0.4162
	Multi-Head Self-Attention + Attention Pooling	Title	tr	0.6531	0.3186	0.3516	0.4124
		Title + Category and Subcategory	tr	0.6445	0.3203	0.3522	0.4111
		Title + Category and Subcategory + Abstract	tr	0.6531	0.3203	0.3544	0.4141
text-embedding-3-large (OpenAI)	Average Pooling	Title	tr	0.5954	0.2699	0.2863	0.3488
		Title + Category and Subcategory	tr	0.6144	0.2857	0.3035	0.3658
		Title + Category and Subcategory + Abstract	tr	0.6175	0.2994	0.3216	0.3817
	Attention Pooling	Title	tr	0.6129	0.2863	0.3059	0.3675
		Title + Category and Subcategory	tr	0.6305	0.3025	0.3246	0.3854
		Title + Category and Subcategory + Abstract	tr	0.6341	0.3110	0.3377	0.3961
	Dense Layer + Attention Pooling	Title	tr	0.6417	0.3182	0.3475	0.4079
		Title + Category and Subcategory	tr	0.6473	0.3236	0.3540	0.4136
		Title + Category and Subcategory + Abstract	tr	0.6567	0.3262	0.3595	0.4186
	Multi-Head Self-Attention + Attention Pooling	Title	tr	0.6421	0.3149	0.3448	0.4058
		Title + Category and Subcategory	tr	0.6495	0.3222	0.3536	0.4135
		Title + Category and Subcategory + Abstract	tr	0.6525	0.3219	0.3549	0.4152
multilingual-e5-large-instruct	Average Pooling	Title	vi	0.6436	0.3121	0.3469	0.4045
		Title + Category and Subcategory	vi	0.6456	0.3147	0.3490	0.4067

Generalized Embedding Model	User Encoder	Model Input	Language	AUC	MRR	NDCG@5	NDCG@10
multilingual-e5-large-instruct	Average Pooling	Title + Category and Subcategory + Abstract	vi	0.6434	0.3106	0.3459	0.4029
	Attention Pooling	Title	vi	0.6407	0.3112	0.3464	0.4042
		Title + Category and Subcategory	vi	0.6436	0.3158	0.3509	0.4087
		Title + Category and Subcategory + Abstract	vi	0.6421	0.3129	0.3488	0.4059
	Dense Layer + Attention Pooling	Title	vi	0.6248	0.2992	0.3276	0.3895
		Title + Category and Subcategory	vi	0.6330	0.3063	0.3371	0.3975
		Title + Category and Subcategory + Abstract	vi	0.6330	0.3025	0.3325	0.3938
	Multi-Head Self-Attention + Attention Pooling	Title	vi	0.6290	0.3060	0.3347	0.3951
		Title + Category and Subcategory	vi	0.6319	0.3078	0.3372	0.3979
		Title + Category and Subcategory + Abstract	vi	0.6370	0.3045	0.3360	0.3973
Qwen3-Embedding-0.6B	Average Pooling	Title	vi	0.6169	0.2927	0.3171	0.3800
		Title + Category and Subcategory	vi	0.6337	0.3061	0.3340	0.3954
		Title + Category and Subcategory + Abstract	vi	0.6247	0.2997	0.3270	0.3881
	Attention Pooling	Title	vi	0.6187	0.2946	0.3203	0.3828
		Title + Category and Subcategory	vi	0.6377	0.3107	0.3412	0.4016
		Title + Category and Subcategory + Abstract	vi	0.6306	0.3049	0.3345	0.3950
	Dense Layer + Attention Pooling	Title	vi	0.6374	0.3066	0.3366	0.3974
		Title + Category and Subcategory	vi	0.6381	0.3123	0.3422	0.4020
		Title + Category and Subcategory + Abstract	vi	0.6419	0.3153	0.3470	0.4062
	Multi-Head Self-Attention + Attention Pooling	Title	vi	0.6401	0.3070	0.3371	0.3984
		Title + Category and Subcategory	vi	0.6376	0.3116	0.3414	0.4012
		Title + Category and Subcategory + Abstract	vi	0.6415	0.3153	0.3462	0.4057
Qwen3-Embedding-4B	Average Pooling	Title	vi	0.6394	0.3124	0.3427	0.4023
		Title + Category and Subcategory	vi	0.6445	0.3104	0.3412	0.4019
		Title + Category and Subcategory + Abstract	vi	0.6456	0.3127	0.3474	0.4061
	Attention Pooling	Title	vi	0.6399	0.3140	0.3454	0.4053
		Title + Category and Subcategory	vi	0.6458	0.3159	0.3484	0.4086
		Title + Category and Subcategory + Abstract	vi	0.6425	0.3141	0.3500	0.4076
	Dense Layer + Attention Pooling	Title	vi	0.6517	0.3196	0.3529	0.4120
		Title + Category and Subcategory	vi	0.6503	0.3210	0.3534	0.4128
		Title + Category and Subcategory + Abstract	vi	0.6516	0.3213	0.3547	0.4135
		Title	vi	0.6544	0.3219	0.3545	0.4138

Generalized Embedding Model	User Encoder	Model Input	Language	AUC	MRR	NDCG@5	NDCG@10
Qwen3-Embedding-4B	Multi-Head Self-Attention + Attention Pooling	Title + Category and Subcategory	vi	0.6508	0.3219	0.3534	0.4128
		Title + Category and Subcategory + Abstract	vi	0.6540	0.3236	0.3568	0.4153
Qwen3-Embedding-8B	Average Pooling	Title	vi	0.6349	0.3087	0.3374	0.3978
		Title + Category and Subcategory	vi	0.6259	0.2965	0.3215	0.3834
		Title + Category and Subcategory + Abstract	vi	0.6341	0.3028	0.3285	0.3909
	Attention Pooling	Title	vi	0.6343	0.3083	0.3376	0.3981
		Title + Category and Subcategory	vi	0.6317	0.3063	0.3337	0.3955
		Title + Category and Subcategory + Abstract	vi	0.6350	0.3083	0.3370	0.3985
	Dense Layer + Attention Pooling	Title	vi	0.6493	0.3163	0.3488	0.4102
		Title + Category and Subcategory	vi	0.6502	0.3233	0.3561	0.4153
		Title + Category and Subcategory + Abstract	vi	0.6583	0.3240	0.3590	0.4188
	Multi-Head Self-Attention + Attention Pooling	Title	vi	0.6500	0.3169	0.3495	0.4106
		Title + Category and Subcategory	vi	0.6463	0.3226	0.3546	0.4137
		Title + Category and Subcategory + Abstract	vi	0.6555	0.3221	0.3563	0.4165
text-embedding-3-large (OpenAI)	Average Pooling	Title	vi	0.6335	0.3110	0.3383	0.3947
		Title + Category and Subcategory	vi	0.6384	0.3107	0.3353	0.3940
		Title + Category and Subcategory + Abstract	vi	0.6438	0.3154	0.3450	0.4035
	Attention Pooling	Title	vi	0.6392	0.3187	0.3493	0.4052
		Title + Category and Subcategory	vi	0.6429	0.3213	0.3497	0.4071
		Title + Category and Subcategory + Abstract	vi	0.6532	0.3237	0.3560	0.4135
	Dense Layer + Attention Pooling	Title	vi	0.6508	0.3215	0.3550	0.4149
		Title + Category and Subcategory	vi	0.6511	0.3246	0.3579	0.4171
		Title + Category and Subcategory + Abstract	vi	0.6649	0.3312	0.3677	0.4256
	Multi-Head Self-Attention + Attention Pooling	Title	vi	0.6494	0.3171	0.3498	0.4111
		Title + Category and Subcategory	vi	0.6542	0.3240	0.3582	0.4175
		Title + Category and Subcategory + Abstract	vi	0.6619	0.3272	0.3624	0.4221
multilingual-e5-large-instruct	Average Pooling	Title	zh	0.6315	0.2993	0.3318	0.3906
		Title + Category and Subcategory	zh	0.6393	0.3074	0.3393	0.3980
		Title + Category and Subcategory + Abstract	zh	0.6495	0.3175	0.3511	0.4085
	Attention Pooling	Title	zh	0.6300	0.3001	0.3328	0.3915
		Title + Category and Subcategory	zh	0.6382	0.3094	0.3424	0.4009
		Title + Category and Subcategory + Abstract	zh	0.6466	0.3179	0.3523	0.4096

Generalized Embedding Model	User Encoder	Model Input	Language	AUC	MRR	NDCG@5	NDCG@10
multilingual-e5-large-instruct	Dense Layer + Attention Pooling	Title	zh	0.6186	0.2845	0.3117	0.3759
		Title + Category and Subcategory	zh	0.6370	0.3029	0.3355	0.3969
		Title + Category and Subcategory + Abstract	zh	0.6427	0.3067	0.3393	0.4003
	Multi-Head Self-Attention + Attention Pooling	Title	zh	0.6254	0.2955	0.3234	0.3860
		Title + Category and Subcategory	zh	0.6410	0.3102	0.3421	0.4026
		Title + Category and Subcategory + Abstract	zh	0.6462	0.3103	0.3441	0.4048
Qwen3-Embedding-0.6B	Average Pooling	Title	zh	0.6044	0.2827	0.3055	0.3676
		Title + Category and Subcategory	zh	0.6316	0.3027	0.3308	0.3910
		Title + Category and Subcategory + Abstract	zh	0.6223	0.3013	0.3275	0.3883
	Attention Pooling	Title	zh	0.6077	0.2849	0.3088	0.3707
		Title + Category and Subcategory	zh	0.6358	0.3068	0.3376	0.3968
		Title + Category and Subcategory + Abstract	zh	0.6275	0.3045	0.3332	0.3933
	Dense Layer + Attention Pooling	Title	zh	0.6247	0.2938	0.3208	0.3825
		Title + Category and Subcategory	zh	0.6396	0.3074	0.3362	0.3979
		Title + Category and Subcategory + Abstract	zh	0.6405	0.3061	0.3368	0.3978
	Multi-Head Self-Attention + Attention Pooling	Title	zh	0.6295	0.2939	0.3213	0.3834
		Title + Category and Subcategory	zh	0.6403	0.3076	0.3367	0.3982
		Title + Category and Subcategory + Abstract	zh	0.6399	0.3071	0.3376	0.3986
Qwen3-Embedding-4B	Average Pooling	Title	zh	0.6378	0.3040	0.3340	0.3931
		Title + Category and Subcategory	zh	0.6438	0.3086	0.3371	0.3972
		Title + Category and Subcategory + Abstract	zh	0.6472	0.3120	0.3451	0.4036
	Attention Pooling	Title	zh	0.6381	0.3055	0.3362	0.3955
		Title + Category and Subcategory	zh	0.6472	0.3147	0.3460	0.4053
		Title + Category and Subcategory + Abstract	zh	0.6457	0.3147	0.3487	0.4071
	Dense Layer + Attention Pooling	Title	zh	0.6374	0.3089	0.3374	0.3976
		Title + Category and Subcategory	zh	0.6424	0.3185	0.3480	0.4074
		Title + Category and Subcategory + Abstract	zh	0.6448	0.3152	0.3471	0.4064
	Multi-Head Self-Attention + Attention Pooling	Title	zh	0.6383	0.3091	0.3365	0.3970
		Title + Category and Subcategory	zh	0.6444	0.3196	0.3488	0.4075
		Title + Category and Subcategory + Abstract	zh	0.6478	0.3182	0.3498	0.4090
Qwen3-Embedding-8B	Average Pooling	Title	zh	0.6267	0.2928	0.3192	0.3829
		Title + Category and Subcategory	zh	0.6089	0.2806	0.3030	0.3657

Generalized Embedding Model	User Encoder	Model Input	Language	AUC	MRR	NDCG@5	NDCG@10
Qwen3-Embedding-8B	Average Pooling	Title + Category and Subcategory + Abstract	zh	0.6191	0.2890	0.3117	0.3765
	Attention Pooling	Title	zh	0.6266	0.2945	0.3222	0.3849
		Title + Category and Subcategory	zh	0.6244	0.2966	0.3242	0.3847
		Title + Category and Subcategory + Abstract	zh	0.6258	0.2971	0.3240	0.3874
	Dense Layer + Attention Pooling	Title	zh	0.6374	0.3052	0.3359	0.3979
		Title + Category and Subcategory	zh	0.6453	0.3191	0.3507	0.4101
		Title + Category and Subcategory + Abstract	zh	0.6523	0.3187	0.3520	0.4128
	Multi-Head Self-Attention + Attention Pooling	Title	zh	0.6406	0.3077	0.3382	0.4002
		Title + Category and Subcategory	zh	0.6395	0.3178	0.3481	0.4072
		Title + Category and Subcategory + Abstract	zh	0.6500	0.3180	0.3511	0.4117
text-embedding-3-large (OpenAI)	Average Pooling	Title	zh	0.6125	0.2893	0.3117	0.3726
		Title + Category and Subcategory	zh	0.6218	0.2966	0.3190	0.3790
		Title + Category and Subcategory + Abstract	zh	0.6481	0.3162	0.3442	0.4041
	Attention Pooling	Title	zh	0.6235	0.2985	0.3244	0.3834
		Title + Category and Subcategory	zh	0.6300	0.3062	0.3325	0.3913
		Title + Category and Subcategory + Abstract	zh	0.6529	0.3212	0.3520	0.4110
	Dense Layer + Attention Pooling	Title	zh	0.6385	0.3035	0.3330	0.3949
		Title + Category and Subcategory	zh	0.6435	0.3143	0.3463	0.4061
		Title + Category and Subcategory + Abstract	zh	0.6554	0.3211	0.3543	0.4146
	Multi-Head Self-Attention + Attention Pooling	Title	zh	0.6335	0.2971	0.3252	0.3888
		Title + Category and Subcategory	zh	0.6419	0.3109	0.3417	0.4026
		Title + Category and Subcategory + Abstract	zh	0.6506	0.3145	0.3479	0.4090

The results for both training and evaluation in the target language is given in the table below.

Table A.17: Performance of Models on Multilingual Candidate News Ranking for User trained and evaluated in target language

Generalized Embedding Model	User Encoder	Model Input	Language	AUC	MRR	NDCG@5	NDCG@10
multilingual-e5-large-instruct	Average Pooling	Title	en	0.6589	0.3141	0.3483	0.4071
		Title + Category and Subcategory	en	0.6524	0.3168	0.3493	0.4088

Generalized Embedding Model	User Encoder	Model Input	Language	AUC	MRR	NDCG@5	NDCG@10
multilingual-e5-large-instruct	Average Pooling	Title + Category and Subcategory + Abstract	en	0.6615	0.3231	0.3584	0.4165
	Attention Pooling	Title	en	0.6565	0.3169	0.3520	0.4104
		Title + Category and Subcategory	en	0.6524	0.3207	0.3549	0.4141
		Title + Category and Subcategory + Abstract	en	0.6601	0.3256	0.3622	0.4201
	Dense Layer + Attention Pooling	Title	en	0.6516	0.3197	0.3518	0.4130
		Title + Category and Subcategory	en	0.6548	0.3231	0.3562	0.4162
		Title + Category and Subcategory + Abstract	en	0.6563	0.3239	0.3570	0.4175
	Multi-Head Self-Attention + Attention Pooling	Title	en	0.6505	0.3217	0.3530	0.4138
		Title + Category and Subcategory	en	0.6539	0.3244	0.3564	0.4159
		Title + Category and Subcategory + Abstract	en	0.6563	0.3241	0.3568	0.4174
Qwen3-Embedding-0.6B	Average Pooling	Title	en	0.6304	0.3036	0.3340	0.3923
		Title + Category and Subcategory	en	0.6140	0.2925	0.3157	0.3757
		Title + Category and Subcategory + Abstract	en	0.6256	0.3030	0.3296	0.3892
	Attention Pooling	Title	en	0.6315	0.3042	0.3357	0.3940
		Title + Category and Subcategory	en	0.6217	0.2968	0.3220	0.3823
		Title + Category and Subcategory + Abstract	en	0.6312	0.3064	0.3346	0.3942
	Dense Layer + Attention Pooling	Title	en	0.6448	0.3011	0.3318	0.3943
		Title + Category and Subcategory	en	0.6467	0.3106	0.3416	0.4023
		Title + Category and Subcategory + Abstract	en	0.6498	0.3139	0.3453	0.4062
	Multi-Head Self-Attention + Attention Pooling	Title	en	0.6465	0.3020	0.3327	0.3949
		Title + Category and Subcategory	en	0.6486	0.3108	0.3421	0.4028
		Title + Category and Subcategory + Abstract	en	0.6505	0.3137	0.3450	0.4061
Qwen3-Embedding-4B	Average Pooling	Title	en	0.6304	0.3034	0.3333	0.3934
		Title + Category and Subcategory	en	0.6243	0.2973	0.3224	0.3834
		Title + Category and Subcategory + Abstract	en	0.6225	0.2993	0.3262	0.3867
	Attention Pooling	Title	en	0.6318	0.3064	0.3371	0.3971

Generalized Embedding Model	User Encoder	Model Input	Language	AUC	MRR	NDCG@5	NDCG@10
Qwen3-Embedding-4B	Attention Pooling	Title + Category and Subcategory	en	0.6306	0.3046	0.3326	0.3924
		Title + Category and Subcategory + Abstract	en	0.6284	0.3048	0.3341	0.3935
	Dense Layer + Attention Pooling	Title	en	0.6525	0.3176	0.3500	0.4101
		Title + Category and Subcategory	en	0.6541	0.3206	0.3535	0.4129
		Title + Category and Subcategory + Abstract	en	0.6551	0.3230	0.3554	0.4154
	Multi-Head Self-Attention + Attention Pooling	Title	en	0.6552	0.3195	0.3510	0.4114
		Title + Category and Subcategory	en	0.6561	0.3228	0.3554	0.4145
		Title + Category and Subcategory + Abstract	en	0.6568	0.3240	0.3561	0.4161
Qwen3-Embedding-8B	Average Pooling	Title	en	0.6388	0.3105	0.3374	0.3974
		Title + Category and Subcategory	en	0.6058	0.2916	0.3098	0.3714
		Title + Category and Subcategory + Abstract	en	0.6123	0.2913	0.3088	0.3720
	Attention Pooling	Title	en	0.6391	0.3111	0.3393	0.3994
		Title + Category and Subcategory	en	0.6238	0.3030	0.3267	0.3868
		Title + Category and Subcategory + Abstract	en	0.6271	0.3025	0.3271	0.3874
	Dense Layer + Attention Pooling	Title	en	0.6557	0.3195	0.3517	0.4124
		Title + Category and Subcategory	en	0.6539	0.3249	0.3573	0.4160
		Title + Category and Subcategory + Abstract	en	0.6557	0.3230	0.3558	0.4163
	Multi-Head Self-Attention + Attention Pooling	Title	en	0.6549	0.3179	0.3494	0.4105
		Title + Category and Subcategory	en	0.6521	0.3231	0.3554	0.4144
		Title + Category and Subcategory + Abstract	en	0.6545	0.3224	0.3551	0.4155
text-embedding-3-large (OpenAI)	Average Pooling	Title	en	0.6186	0.2957	0.3161	0.3777
		Title + Category and Subcategory	en	0.6205	0.2964	0.3172	0.3796
		Title + Category and Subcategory + Abstract	en	0.6308	0.3084	0.3328	0.3929
	Attention Pooling	Title	en	0.6251	0.3019	0.3257	0.3863
		Title + Category and Subcategory	en	0.6297	0.3060	0.3319	0.3924

Generalized Embedding Model	User Encoder	Model Input	Language	AUC	MRR	NDCG@5	NDCG@10
text-embedding-3-large (OpenAI)	Attention Pooling	Title + Category and Subcategory + Abstract	en	0.6370	0.3143	0.3418	0.4018
	Dense Layer + Attention Pooling	Title	en	0.6593	0.3266	0.3606	0.4210
		Title + Category and Subcategory	en	0.6575	0.3283	0.3620	0.4218
		Title + Category and Subcategory + Abstract	en	0.6588	0.3278	0.3622	0.4222
	Multi-Head Self-Attention + Attention Pooling	Title	en	0.6606	0.3275	0.3614	0.4218
		Title + Category and Subcategory	en	0.6591	0.3297	0.3634	0.4228
		Title + Category and Subcategory + Abstract	en	0.6600	0.3291	0.3636	0.4233
multilingual-e5-large-instruct	Average Pooling	Title	tr	0.6469	0.3158	0.3482	0.4066
		Title + Category and Subcategory	tr	0.6459	0.3144	0.3470	0.4058
		Title + Category and Subcategory + Abstract	tr	0.6520	0.3155	0.3500	0.4085
	Attention Pooling	Title	tr	0.6465	0.3159	0.3485	0.4069
		Title + Category and Subcategory	tr	0.6447	0.3167	0.3504	0.4091
		Title + Category and Subcategory + Abstract	tr	0.6504	0.3168	0.3520	0.4106
	Dense Layer + Attention Pooling	Title	tr	0.6505	0.3146	0.3457	0.4079
		Title + Category and Subcategory	tr	0.6545	0.3196	0.3524	0.4131
		Title + Category and Subcategory + Abstract	tr	0.6547	0.3197	0.3535	0.4136
	Multi-Head Self-Attention + Attention Pooling	Title	tr	0.6513	0.3173	0.3487	0.4099
		Title + Category and Subcategory	tr	0.6550	0.3220	0.3540	0.4143
		Title + Category and Subcategory + Abstract	tr	0.6557	0.3216	0.3547	0.4145
Qwen3-Embedding-0.6B	Average Pooling	Title	tr	0.6287	0.2959	0.3194	0.3811
		Title + Category and Subcategory	tr	0.6365	0.3027	0.3261	0.3880
		Title + Category and Subcategory + Abstract	tr	0.6333	0.3026	0.3292	0.3893
	Attention Pooling	Title	tr	0.6330	0.3011	0.3278	0.3892
		Title + Category and Subcategory	tr	0.6446	0.3125	0.3404	0.4012
		Title + Category and Subcategory + Abstract	tr	0.6400	0.3106	0.3406	0.4003
	Dense Layer + Attention Pooling	Title	tr	0.6370	0.3044	0.3344	0.3961

Generalized Embedding Model	User Encoder	Model Input	Language	AUC	MRR	NDCG@5	NDCG@10
Qwen3-Embedding-0.6B	Dense Layer + Attention Pooling	Title + Category and Subcategory	tr	0.6471	0.3153	0.3474	0.4079
		Title + Category and Subcategory + Abstract	tr	0.6516	0.3185	0.3526	0.4122
	Multi-Head Self-Attention + Attention Pooling	Title	tr	0.6383	0.3049	0.3347	0.3962
		Title + Category and Subcategory	tr	0.6479	0.3154	0.3479	0.4079
		Title + Category and Subcategory + Abstract	tr	0.6514	0.3181	0.3521	0.4119
Qwen3-Embedding-4B	Average Pooling	Title	tr	0.6421	0.3125	0.3440	0.4029
		Title + Category and Subcategory	tr	0.6433	0.3099	0.3407	0.4010
		Title + Category and Subcategory + Abstract	tr	0.6585	0.3244	0.3617	0.4184
	Attention Pooling	Title	tr	0.6413	0.3140	0.3467	0.4056
		Title + Category and Subcategory	tr	0.6450	0.3175	0.3512	0.4101
		Title + Category and Subcategory + Abstract	tr	0.6576	0.3249	0.3628	0.4194
	Dense Layer + Attention Pooling	Title	tr	0.6566	0.3219	0.3558	0.4154
		Title + Category and Subcategory	tr	0.6606	0.3272	0.3617	0.4208
		Title + Category and Subcategory + Abstract	tr	0.6609	0.3261	0.3610	0.4204
	Multi-Head Self-Attention + Attention Pooling	Title	tr	0.6583	0.3233	0.3575	0.4168
		Title + Category and Subcategory	tr	0.6602	0.3276	0.3616	0.4208
		Title + Category and Subcategory + Abstract	tr	0.6604	0.3271	0.3612	0.4208
Qwen3-Embedding-8B	Average Pooling	Title	tr	0.6407	0.3064	0.3368	0.3972
		Title + Category and Subcategory	tr	0.5991	0.2804	0.3037	0.3653
		Title + Category and Subcategory + Abstract	tr	0.6123	0.2905	0.3147	0.3781
	Attention Pooling	Title	tr	0.6400	0.3068	0.3375	0.3982
		Title + Category and Subcategory	tr	0.6140	0.2954	0.3231	0.3832
		Title + Category and Subcategory + Abstract	tr	0.6242	0.3004	0.3284	0.3903
	Dense Layer + Attention Pooling	Title	tr	0.6513	0.3160	0.3483	0.4090
		Title + Category and Subcategory	tr	0.6533	0.3211	0.3545	0.4143

Generalized Embedding Model	User Encoder	Model Input	Language	AUC	MRR	NDCG@5	NDCG@10
Qwen3-Embedding-8B	Dense Layer + Attention Pooling	Title + Category and Subcategory + Abstract	tr	0.6570	0.3210	0.3556	0.4152
	Multi-Head Self-Attention + Attention Pooling	Title	tr	0.6533	0.3170	0.3495	0.4104
		Title + Category and Subcategory	tr	0.6509	0.3199	0.3531	0.4132
		Title + Category and Subcategory + Abstract	tr	0.6544	0.3214	0.3549	0.4143
text-embedding-3-large (OpenAI)	Average Pooling	Title	tr	0.5954	0.2699	0.2863	0.3488
		Title + Category and Subcategory	tr	0.6144	0.2857	0.3035	0.3658
		Title + Category and Subcategory + Abstract	tr	0.6175	0.2994	0.3216	0.3817
	Attention Pooling	Title	tr	0.6172	0.2934	0.3148	0.3760
		Title + Category and Subcategory	tr	0.6340	0.3075	0.3314	0.3914
		Title + Category and Subcategory + Abstract	tr	0.6374	0.3133	0.3412	0.3998
	Dense Layer + Attention Pooling	Title	tr	0.6484	0.3146	0.3464	0.4084
		Title + Category and Subcategory	tr	0.6530	0.3223	0.3554	0.4152
		Title + Category and Subcategory + Abstract	tr	0.6541	0.3216	0.3557	0.4160
	Multi-Head Self-Attention + Attention Pooling	Title	tr	0.6478	0.3155	0.3468	0.4087
		Title + Category and Subcategory	tr	0.6523	0.3223	0.3550	0.4151
		Title + Category and Subcategory + Abstract	tr	0.6553	0.3233	0.3573	0.4172
multilingual-e5-large-instruct	Average Pooling	Title	vi	0.6436	0.3121	0.3469	0.4045
		Title + Category and Subcategory	vi	0.6456	0.3147	0.3490	0.4067
		Title + Category and Subcategory + Abstract	vi	0.6434	0.3106	0.3459	0.4029
	Attention Pooling	Title	vi	0.6432	0.3120	0.3468	0.4046
		Title + Category and Subcategory	vi	0.6443	0.3158	0.3508	0.4085
		Title + Category and Subcategory + Abstract	vi	0.6420	0.3127	0.3487	0.4057
	Dense Layer + Attention Pooling	Title	vi	0.6470	0.3145	0.3451	0.4069
		Title + Category and Subcategory	vi	0.6523	0.3190	0.3509	0.4117
		Title + Category and Subcategory + Abstract	vi	0.6520	0.3177	0.3501	0.4115
		Title	vi	0.6461	0.3166	0.3465	0.4077

Generalized Embedding Model	User Encoder	Model Input	Language	AUC	MRR	NDCG@5	NDCG@10
multilingual-e5-large-instruct	Multi-Head Self-Attention + Attention Pooling	Title + Category and Subcategory	vi	0.6504	0.3203	0.3514	0.4113
		Title + Category and Subcategory + Abstract	vi	0.6520	0.3181	0.3500	0.4107
Qwen3-Embedding-0.6B	Average Pooling	Title	vi	0.6169	0.2927	0.3171	0.3800
		Title + Category and Subcategory	vi	0.6337	0.3061	0.3340	0.3954
		Title + Category and Subcategory + Abstract	vi	0.6247	0.2997	0.3270	0.3881
	Attention Pooling	Title	vi	0.6195	0.2961	0.3220	0.3847
		Title + Category and Subcategory	vi	0.6379	0.3117	0.3422	0.4025
		Title + Category and Subcategory + Abstract	vi	0.6309	0.3060	0.3359	0.3961
	Dense Layer + Attention Pooling	Title	vi	0.6376	0.3060	0.3360	0.3972
		Title + Category and Subcategory	vi	0.6449	0.3148	0.3466	0.4068
		Title + Category and Subcategory + Abstract	vi	0.6493	0.3174	0.3503	0.4103
	Multi-Head Self-Attention + Attention Pooling	Title	vi	0.6398	0.3070	0.3371	0.3986
		Title + Category and Subcategory	vi	0.6459	0.3150	0.3471	0.4070
		Title + Category and Subcategory + Abstract	vi	0.6497	0.3178	0.3506	0.4104
Qwen3-Embedding-4B	Average Pooling	Title	vi	0.6394	0.3124	0.3427	0.4023
		Title + Category and Subcategory	vi	0.6445	0.3104	0.3412	0.4019
		Title + Category and Subcategory + Abstract	vi	0.6456	0.3127	0.3474	0.4061
	Attention Pooling	Title	vi	0.6393	0.3140	0.3455	0.4051
		Title + Category and Subcategory	vi	0.6456	0.3160	0.3490	0.4088
		Title + Category and Subcategory + Abstract	vi	0.6445	0.3139	0.3494	0.4078
	Dense Layer + Attention Pooling	Title	vi	0.6550	0.3202	0.3535	0.4132
		Title + Category and Subcategory	vi	0.6560	0.3244	0.3580	0.4172
		Title + Category and Subcategory + Abstract	vi	0.6584	0.3249	0.3593	0.4184
	Multi-Head Self-Attention + Attention Pooling	Title	vi	0.6570	0.3215	0.3543	0.4144
		Title + Category and Subcategory	vi	0.6563	0.3243	0.3574	0.4168

Generalized Embedding Model	User Encoder	Model Input	Language	AUC	MRR	NDCG@5	NDCG@10
Qwen3-Embedding-4B	Multi-Head Self-Attention + Attention Pooling	Title + Category and Subcategory + Abstract	vi	0.6588	0.3257	0.3594	0.4188
Qwen3-Embedding-8B	Average Pooling	Title	vi	0.6349	0.3087	0.3374	0.3978
		Title + Category and Subcategory	vi	0.6259	0.2965	0.3215	0.3834
		Title + Category and Subcategory + Abstract	vi	0.6341	0.3028	0.3285	0.3909
	Attention Pooling	Title	vi	0.6347	0.3090	0.3380	0.3983
		Title + Category and Subcategory	vi	0.6319	0.3067	0.3344	0.3959
		Title + Category and Subcategory + Abstract	vi	0.6348	0.3083	0.3371	0.3985
	Dense Layer + Attention Pooling	Title	vi	0.6531	0.3193	0.3515	0.4123
		Title + Category and Subcategory	vi	0.6547	0.3257	0.3585	0.4183
		Title + Category and Subcategory + Abstract	vi	0.6575	0.3235	0.3578	0.4173
	Multi-Head Self-Attention + Attention Pooling	Title	vi	0.6535	0.3190	0.3509	0.4118
		Title + Category and Subcategory	vi	0.6522	0.3248	0.3569	0.4167
		Title + Category and Subcategory + Abstract	vi	0.6554	0.3229	0.3569	0.4166
text-embedding-3-large (OpenAI)	Average Pooling	Title	vi	0.6335	0.3110	0.3383	0.3947
		Title + Category and Subcategory	vi	0.6384	0.3107	0.3353	0.3940
		Title + Category and Subcategory + Abstract	vi	0.6438	0.3154	0.3450	0.4035
	Attention Pooling	Title	vi	0.6373	0.3162	0.3459	0.4023
		Title + Category and Subcategory	vi	0.6407	0.3211	0.3498	0.4072
		Title + Category and Subcategory + Abstract	vi	0.6527	0.3235	0.3558	0.4133
	Dense Layer + Attention Pooling	Title	vi	0.6473	0.3153	0.3468	0.4086
		Title + Category and Subcategory	vi	0.6536	0.3227	0.3556	0.4157
		Title + Category and Subcategory + Abstract	vi	0.6562	0.3227	0.3574	0.4176
	Multi-Head Self-Attention + Attention Pooling	Title	vi	0.6479	0.3153	0.3463	0.4077
		Title + Category and Subcategory	vi	0.6538	0.3237	0.3560	0.4160
		Title + Category and Subcategory + Abstract	vi	0.6581	0.3238	0.3579	0.4182
multilingual-e5-large-instruct	Average Pooling	Title	zh	0.6315	0.2993	0.3318	0.3906

Generalized Embedding Model	User Encoder	Model Input	Language	AUC	MRR	NDCG@5	NDCG@10
multilingual-e5-large-instruct	Average Pooling	Title + Category and Subcategory	zh	0.6393	0.3074	0.3393	0.3980
		Title + Category and Subcategory + Abstract	zh	0.6495	0.3175	0.3511	0.4085
	Attention Pooling	Title	zh	0.6307	0.3011	0.3341	0.3925
		Title + Category and Subcategory	zh	0.6385	0.3099	0.3431	0.4017
		Title + Category and Subcategory + Abstract	zh	0.6475	0.3183	0.3526	0.4097
	Dense Layer + Attention Pooling	Title	zh	0.6386	0.3017	0.3299	0.3927
		Title + Category and Subcategory	zh	0.6526	0.3181	0.3502	0.4105
		Title + Category and Subcategory + Abstract	zh	0.6510	0.3154	0.3478	0.4086
	Multi-Head Self-Attention + Attention Pooling	Title	zh	0.6403	0.3015	0.3298	0.3928
		Title + Category and Subcategory	zh	0.6526	0.3163	0.3483	0.4092
		Title + Category and Subcategory + Abstract	zh	0.6525	0.3149	0.3476	0.4086
Qwen3-Embedding-0.6B	Average Pooling	Title	zh	0.6044	0.2827	0.3055	0.3676
		Title + Category and Subcategory	zh	0.6316	0.3027	0.3308	0.3910
		Title + Category and Subcategory + Abstract	zh	0.6223	0.3013	0.3275	0.3883
	Attention Pooling	Title	zh	0.6078	0.2847	0.3086	0.3707
		Title + Category and Subcategory	zh	0.6358	0.3076	0.3387	0.3978
		Title + Category and Subcategory + Abstract	zh	0.6277	0.3048	0.3337	0.3936
	Dense Layer + Attention Pooling	Title	zh	0.6324	0.2977	0.3243	0.3867
		Title + Category and Subcategory	zh	0.6449	0.3117	0.3415	0.4024
		Title + Category and Subcategory + Abstract	zh	0.6491	0.3107	0.3424	0.4036
	Multi-Head Self-Attention + Attention Pooling	Title	zh	0.6325	0.2973	0.3241	0.3861
		Title + Category and Subcategory	zh	0.6459	0.3121	0.3417	0.4028
		Title + Category and Subcategory + Abstract	zh	0.6509	0.3115	0.3433	0.4046
Qwen3-Embedding-4B	Average Pooling	Title	zh	0.6378	0.3040	0.3340	0.3931
		Title + Category and Subcategory	zh	0.6438	0.3086	0.3371	0.3972

Generalized Embedding Model	User Encoder	Model Input	Language	AUC	MRR	NDCG@5	NDCG@10
Qwen3-Embedding-4B	Average Pooling	Title + Category and Subcategory + Abstract	zh	0.6472	0.3120	0.3451	0.4036
	Attention Pooling	Title	zh	0.6378	0.3069	0.3379	0.3973
		Title + Category and Subcategory	zh	0.6472	0.3160	0.3477	0.4068
		Title + Category and Subcategory + Abstract	zh	0.6466	0.3153	0.3495	0.4077
	Dense Layer + Attention Pooling	Title	zh	0.6478	0.3113	0.3422	0.4026
		Title + Category and Subcategory	zh	0.6567	0.3257	0.3588	0.4175
		Title + Category and Subcategory + Abstract	zh	0.6562	0.3209	0.3545	0.4143
	Multi-Head Self-Attention + Attention Pooling	Title	zh	0.6495	0.3116	0.3418	0.4029
		Title + Category and Subcategory	zh	0.6566	0.3251	0.3575	0.4162
		Title + Category and Subcategory + Abstract	zh	0.6556	0.3203	0.3537	0.4133
Qwen3-Embedding-8B	Average Pooling	Title	zh	0.6267	0.2928	0.3192	0.3829
		Title + Category and Subcategory	zh	0.6089	0.2806	0.3030	0.3657
		Title + Category and Subcategory + Abstract	zh	0.6191	0.2890	0.3117	0.3765
	Attention Pooling	Title	zh	0.6263	0.2960	0.3243	0.3870
		Title + Category and Subcategory	zh	0.6274	0.3015	0.3307	0.3901
		Title + Category and Subcategory + Abstract	zh	0.6269	0.2990	0.3263	0.3890
	Dense Layer + Attention Pooling	Title	zh	0.6478	0.3127	0.3442	0.4039
		Title + Category and Subcategory	zh	0.6527	0.3229	0.3552	0.4150
		Title + Category and Subcategory + Abstract	zh	0.6569	0.3222	0.3563	0.4157
	Multi-Head Self-Attention + Attention Pooling	Title	zh	0.6489	0.3137	0.3451	0.4054
		Title + Category and Subcategory	zh	0.6518	0.3222	0.3544	0.4147
		Title + Category and Subcategory + Abstract	zh	0.6556	0.3211	0.3548	0.4149
text-embedding-3-large (OpenAI)	Average Pooling	Title	zh	0.6125	0.2893	0.3117	0.3726
		Title + Category and Subcategory	zh	0.6218	0.2966	0.3190	0.3790
		Title + Category and Subcategory + Abstract	zh	0.6481	0.3162	0.3442	0.4041
	Attention Pooling	Title	zh	0.6244	0.2997	0.3263	0.3855

Generalized Embedding Model	User Encoder	Model Input	Language	AUC	MRR	NDCG@5	NDCG@10
text-embedding-3-large (OpenAI)	Attention Pooling	Title + Category and Subcategory	zh	0.6322	0.3092	0.3365	0.3947
		Title + Category and Subcategory + Abstract	zh	0.6535	0.3222	0.3527	0.4121
	Dense Layer + Attention Pooling	Title	zh	0.6451	0.3083	0.3382	0.3999
		Title + Category and Subcategory	zh	0.6534	0.3234	0.3554	0.4149
		Title + Category and Subcategory + Abstract	zh	0.6588	0.3231	0.3575	0.4173
	Multi-Head Self-Attention + Attention Pooling	Title	zh	0.6459	0.3085	0.3386	0.4001
		Title + Category and Subcategory	zh	0.6542	0.3228	0.3550	0.4145
		Title + Category and Subcategory + Abstract	zh	0.6596	0.3229	0.3572	0.4169