An Automated Intelligent E-learning Approach for Ensuring Learning Outcomes

Hossain Ashik Mahmud Chowdury
Student Id: 012 091 018

A thesis in the Department of Computer Science and Engineering presented in partial fulfillment of the requirements for the Degree of Master of Science in Computer Science and Engineering

United International University
Dhaka, Bangladesh

February 2018

Declaration

I, Hossain Ashik Mahmud Chowdury, declare that this thesis titled, "**An Automated Intelligent E-learning Approach for Ensuring Learning Outcomes**" and the work presented in it are my own. I confirm that:

- This work was done mainly while in candidature for a MSc degree at United International University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at United International University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

_____

Name: Hossain Ashik Mahmud Chowdury
Student ID: 012 091 018
Department: Computer Science and Engineering

Certificate

I do hereby declare that the research works embodied in this thesis entitled "**An Automated Intelligent E-learning Approach for Ensuring Learning Outcomes**" is the outcome of an original work carried out by Hossain Ashik Mahmud Chowdury under my supervision.

I further certify that the dissertation meets the requirements and the standard for the degree of MSc in Computer Science and Engineering.

_____

Suman Ahmmed
Assistant Professor
Department of Computer Science and Engineering, United International University
Director, Center for Development of IT Professionals (CDIP)

Abstract

E-learning is an innovative and effective way to perceive knowledge in the modern world. The web-based interactive system has become one of the most common needs in our generation. Thus E-learning has achieved enormous popularity in the world since this decade. Most of the universities are using web-based systems to provide interactive systems to students which can make a faster communication to grow up the knowledge in every academic section. Many advantages come out with the adoption of e-learning systems majorly paperless environment, low cost, freedom of time and geographic locations, web materials are accessible in any order, in any devices even in smartphones. In advanced countries, e-learning systems play a major role in an economy to raise productive output in the industries without having paid a huge amount for personal staff that is locating physically and also provides big advantages to peoples of developed countries who cannot attend physically courses either afford experts in a professional domain. This is an immense achievement of e-learning based systems to promote education online.

In this dissertation, it has been proposed some helpful and efficient technique to design a web-based intelligent e-learning system for academic institutions like university and high school. The learning contents may be presented in different advantageous ways to teach any topic efficiently. Contents can be presented in text, video, audio or in form of animation. Apart from that, also a model is prepared to ensure the learning outcome. Evaluation part of an e-learning system is very critical. Most of the e-learning website use multiple-choice questions for evaluation which cannot guarantee learning outcome because a student may get good marks by with random selection. Here, some proficient ways are proposed which would enforce a student to learn. In artificial intelligence based games player follows several levels where players go to the final level step by step. In this proposed model follows the same to teach a student, that means system also examines the student step by step. That is why a student can hardly pass only by selecting the answer randomly. This

system has analyzed the design of proposed e-learning system consists of use case diagrams, sequence diagrams, class diagrams and database diagrams. As an intelligent system, this system is going to apply some machine learning techniques like "Multiple Linear Regression Algorithm" which not only guide us to define most effective exam approach but also will ensure student outcome.

This dissertation is organized as - chapter 1 includes the introductory portion of the necessary background. Chapter 2 comes with literature review and chapter 3 with the methodology details of proposed system. In-depth analysis of proposed system is prepared in chapter 4 and finally concluding in chapter 5.

Acknowledgement

At the very beginning, I would like to thank Almighty Allah for giving me the opportunity to do this thesis and for blessing me with some great people who helped me a lot in my personal and professional life. Also, I would like to thank my supervisor Asst. Prof. Suman Ahmmed for his continuous support and effort to bringing this project into the light. I would also love to thank Dr. Keshav Dahal of University of Bradford, UK for being a supervisor of previous versions of this thesis. I am really like to acknowledge my professor Dr. Alamgir Hossain of University of Bradford, UK who helped me to select this topic for my MSc program. I would especially like to express my deepest gratitude to my Prof. Dr. Chowdhury Mofizur Rahman of United International University, BD who guided me to be a part of Erasmus Mundus Student Exchange Program where I got an opportunity to start this thesis work.

Last of all, I also like to thank my family and friends, especially my friend and colleague Mr. Sifatur Rahim who gave me several ideas to prepare this thesis.

Table of Contents

LIST OF TABLES

LIST OF FIGURES

Chapter 1

1. Introduction

1.1 Background

E-Learning is actually all kind of electronic supported teaching and learning system whose goal is to build of knowledge with different references to individual experience, practice and knowledge of the learner. IT and communication system serve as a media to implement this learning system even without network [1].

There are many definitions of e-learning like academic papers on the subject, but most of them focus on the same kind of features. Like the European e-learning Action Plan definition as an example: “the use of new multimedia technologies and the Internet to improve the quality of learning by facilitating access to resources and services as well as remote exchanges and collaboration”. Here it is said that this IT system is act like a vehicles for education services and resources, and as the conduits for collaboration and communication. If we may be so bold as to focus on the essentials, however, our preferred definition would view e-learning simply as: “online access to learning resources, anywhere and anytime” [2]

E-learning is a system which transfer knowledge and skills by computer and network. It’s refers to using of electronic processes and applications for learning and teaching. It’s processes and applications include Web-based or computer-based learning, virtual classrooms and digital collaboration. Content is delivered via audio or video tape, satellite TV, the Internet, intranet/extranet, DVD-ROM and CD-ROM. It can be self paced or instructor led and includes media in the different form like text, streaming video, animation, audio and image [3].

Acronyms like CBT (*Computer-Based Training*), IBT (*Internet-Based Training*) or WBT (*Web-Based Training*) have been used as synonyms to E-Learning. Today

one can still find these terms being used, along with variations of e-learning such as elearning, Elearning, and eLearning. [3]

Early E-Learning systems, based on Computer-Based Learning/Training often attempted to replicate autocratic teaching styles whereby the role of the e-learning system was assumed to be for transferring knowledge, as opposed to systems developed later based on Computer Supported Collaborative Learning (CSCL), which encouraged the shared development of knowledge.

As early as 1993, William D. Graziadei [4] described an online computer-delivered lecture, tutorial and assessment project using electronic Mail, two VAX Notes conferences and Gopher/Lynx [5] together with several software programs that allowed students and instructor to create a Virtual Instructional Classroom Environment in Science (VICES) in Research, Education, Service & Teaching (REST) [6]. In 1997 Graziadei, W.D., et al.,[28] published an article entitled "Building Asynchronous and Synchronous Teaching-Learning Environments: Exploring a Course/Classroom Management System Solution". [7] They described a process at the State University of New York (SUNY) of evaluating products and developing an overall strategy for technology-based course development and management in teaching-learning. The product(s) had to be easy to use and maintain, portable, replicable, scalable, and immediately affordable, and they had to have a high probability of success with long-term cost-effectiveness. Today many technologies can be, and are, used in e-Learning, from blogs to collaborative software, ePortfolios, and virtual classrooms. Most eLearning situations use combinations of these techniques.

1.2 Motivation

E-Learning can provide benefits for the organizations and individuals involved. The benefits are shown step by step.

a. Improved resource utilisation
    1. Travel time and accommodation savings

2. Procurement savings
3. Cost-effective micro training courses
4. Reduced time for updating
5. Better utilisation of existing technology
6. Economies of scale possible
7. Just in time (JIT)
8. Scalability
9. Rationalisation of offerings

b. Geographical reach
   1. Flexibility – any place , any time
   2. Provides access for people in remote locations
   3. Better support of field staff at customer premises
   4. Rapid transfer of knowledge including best practice
   5. More training to more people

c. Effectiveness of delivery
   1. Faster roll-out and employee development
   2. Reduced time to competency
   3. More focused training and development
   4. Consistency of information
   5. Lean training to support lean organisation

d. Organisational benefits
   1. Getting better overall performance from technology
   2. Support for staff at workplace – JIT training improves problem-solving capability
   3. Helps to attract good quality talent
   4. Demonstrates real commitment to employee development
   5. Cost-effective delivery of promised lifelong learning
   6. Demonstrating of effectiveness of product

e. Enhanced learning
   1. Self-paced learning
   2. Increased self-management capability

3. Enhanced individual choice by providing an alternative approach to training and development

Further benefits include enhanced business responsiveness and community building. The latter can be across the supply chain or within the local community by offering outreach programmes. [8]

The worldwide e-learning industry is estimated to be worth over thirty-eight (38) billion euro according to conservative estimates; although in the European Union only about 20% of e-learning products are produced within the common market. [9] Developments in internet and multimedia technologies are the basic enabler of e-learning, with consulting, content, technologies, services and support being identified as the five key sectors of the e-learning industry. [10]

By 2006, 3.5 million students were participating in on-line learning at institutions of higher education in the United States. According to the Sloan Foundation reports,[11][12] there has been an increase of around 12-14 per cent per year on average in enrolments for fully online learning over the five years 2004-2009 in the US post-secondary system, compared with an average of approximately 2 percent increase per year in enrolments overall. Allen and Seamen (2009) [11] claim that almost a quarter of all students in post-secondary education were taking fully online courses in 2008, and a report by Ambient Insight Research [13] suggests that in 2009, 44 per cent of post-secondary students in the USA were taking some or all of their courses online, and projected that this figure would rise to 81 per cent by 2014. Thus it can be seen that e-learning is moving rapidly from the margins to being a predominant form of post-secondary education, at least in the USA.

Many higher educations, for-profit institutions, now offer online classes. By contrast, only about half of private, non-profit schools offer them. The Sloan report, based on a poll of academic leaders, says that students generally appear to be at least as satisfied with their online classes as they are with traditional ones. Private institutions may become more involved with online presentations as the cost of

instituting such a system decreases. Properly trained staff must also be hired to work with students online. These staff members need to understand the content area, and also be highly trained in the use of the computer and Internet. Online education is rapidly increasing, and online doctoral programs have even developed at leading research universities. [14]

I have ten years working experience as Big Data Analyst in Software Company. So, I like to do research on data mining. There have several applications to implement data mining. Nowadays, e-learning is very popular and interesting field to implement data mining. Many research groups are working on intelligent e-learning system using data mining. They are trying to ensure students learning outcome through intelligent e-learning system. In my MSc research my target is to find some good data mining technique which will be helpful for intelligent e-learning to ensure learning outcome of a student.

1.3 Aims & Objectives

The goal of this thesis is to design a automated intelligent e-learning system for university. This automated intelligent system will be a common platform. Many universities will be able to use this system. The learning content like different subject will be presented in different way. Content will be presented in text, video, audio, animation format and also some useful link. Without that, in the system, an algorithm will be followed which will ensure learning outcome. Evaluation part of this e-learning system is very critical. Here system will use multiple-choice questions for evaluation. Also here some technique will be proposed which will force the student to learn the content. Proposed system will follow Artificial Intelligence Gaming technique where player follows several levels. Players go to the final level step by step. In this model, proposed system followed this technique to train a student step by step. Also system will examine the student step by step. That is why, student will not pass only by selecting the answer randomly.

Following the key objectives are fulfilled in this project in order to achieve the goal:

1. literature review to justify the investigation
2. Discuss about existing technique and tools
3. Design a model to teach a student about particular subject
4. Design an algorithm to evaluate a student to ensure learning outcome
5. Design a Flowchart for the system.
6. Design Use Case Diagrams of the system.
7. Design Sequence Diagrams of the system.
8. Design Class Diagram of the system.
9. Design Database Diagram of the system.
10. Generate sample data.
11. Examin sample data using proposed method.
12. Discuss about proposed system outcome.

1.4 Outline of the Report

The report is organized as follows: Chapter 2 presents the literature review and related work, Chapter 3 shows the theory and methodology of the intelligent web-based e-learning system model with its system design (like flowchart, algorithm, use case, sequence, class, database diagram), Chapter 4 presents Investigation and Result and then conclusions and future work are given in Chapter 5.

Chapter 2

2. Literature Review and Related Works

2.1 Literature Review

E-learning is a technique which delivers instruction on a computer via internet or CD. The content of e-learning may be in text, streaming video or audio format. Nowadays, e-learning is very useful for the organizations and individuals.

In a research by the U.S. Department of Education, found that online learning for higher education performed better than face-to-face courses [15]. Recognized expert can share their knowledge in a border way. E-learning is very flexible and convenient way for learners. Student can learn their courses at home. E-learning is very cost effective for education institutions. It is less expensive than face-to-face education.

It is becoming a dominant delivery method in many workplace learning settings for many organizations. Organizations are finding the potential of e-learning to train to employees. But they also face some barriers to e-learning, such as situational, organizational and technical barriers [16]. Also Managers are concerned about technology requirements and cost for implementing e-learning [17]. Furthermore, there is excess of emerging technologies which have workplace learning implications. Thus, using a perfect e-storm with countless emerging technologies, organization can significantly reduce budgets [18].

Several surveys were conducted on the current status and future trends of e-learning in workplace learning system [19]. Where we find that, e-learning and blended forms of e-learning are on the rise in many organization's training environments. There is also a demand for authentic learning opportunities and simulated experiences and e-learning can provide that scope.

Another survey was conducted by Yongmei Bentley, Anjali Shegunshi and Mike Scannell [20]. In this report they investigated the distance learning support systems of a UK University's overseas MBA program. This program is provided to several countries around the world. The university program provide online and face-to-face both learning technique. The researcher aim was to evaluate the learning support mechanisms that are used to deliver the program overseas and to determine their impact on the learning experience for the students of MBA. The outcomes of this research helped to improve the learning support system and also enhanced the quality of MBA program. And this also helped to provide guidelines for other Higher Educational Institutions that offer or intend to offer blended learning courses globally.

Chih-Ming Chen presented a personalized e-learning system with self-regulated learning assisted mechanisms that will help learners to enhance their self-regulated learning abilities [23]. He proposed four self-regulated learning types. His experimental results show that, his method help to improve performance of personalized e-learning system.

Chun, Gwo, Yungho proposed to apply an algorithm of Apriori for concept map to develop an intelligent concept diagnostic (ICDS) [26]. Teachers can diagnose the learning barriers and misconception. Using their algorithm they reached at best Remedial-Instruction Path (RIP).

For e-learning, there have several challenges. One of them is to ensure positive learning outcome. Institutes need to ensure that students are actually learned from online. Online testing is usually in the form of multiple-choice questions, without any essay type of learning assessment. Major reasons for employing multiple-choice tasks in e-learning include ease of implementation and ease of managing learner's responses. To address this limitation in online assessment of learning outcome Minoru Nakayama, Hiroh Yamamoto and Rowena Santiago [28] are investigated an automatic assessment system as a natural language processing tool for conducting essay type tests in online learning. This essay testing also examines the relationship between learner characteristics and

learner performance. In the research students were enrolled in two-unit courses which were taught by the same professor as follows: hybrid learning course at bachelor's level, fully online course at bachelor's level, and hybrid learning course at masters level. They found that expert evaluation of the essay is focused on logical structure more than other factors. When comparing the score of experts' assessment between hybrid learning and fully online courses at the bachelor's level, they did not found significant differences. In their research they found that in fully online learning, as well as in hybrid learning, learning performance can be measured using essay tests without the need for a face-to-face session to conduct this type of assessment.

In a book, Lluís Vicent and Xavier Gumara [29] presented how data mining techniques can be used in instructional design, in learning and in the assessment of the students, to reduce, interpret and classify the information, factor and cluster analysis. There, two examples of data mining applications in e-learning are presented and proved to give knowledge to the teacher. They used two examples. In the first example, factor analysis has been used in order to set the minimum group of tools for developing an online degree in the European Higher Education Area (EHEA). With this analysis it can be understood that a complementarity between face to face classes and online work was needed. In the second example, a complementary application to questionnaire tools from LMS has been implemented and demonstrated to expand reporting options from open source LMS. The main power of the system is to provide an automatic classification of students according to their performance on different competences. They showed, factor analysis is a data mining technique that helped us to reduce a large number of variables in two different problems, both related to learning.

On the other hand, though in e-learning and other educational multimedia environments, the evaluation of interaction and behaviour is particularly crucial, Data mining, a non-intrusive, objective analysis technology, can be proposed as the central evaluation technology for the analysis of the usage of computer-based educational environments and in particular of the interaction with educational

content. Pahl, Claus [30] shows us the benefits of data mining technology in discovering latent knowledge make it a convincing solution for the observation-based evaluation of learner behaviour in feature-rich educational environments in particular.

Gaming can be another technique to teach student through e-learning. In a book, Donald Clark [31] shows how gaming technique can be used to in e-learning platform. In his paper, he mentioned, what do existing games teach us and how can games accelerate learning.

In evaluation part of e-learning instructor need to generate question set.. But also question set can be generated using ontology technique.

2.2 Evaluation of Some Popular e-learning Tools

Nowadays, there have several tools for e-learning system. Some of them are discussed below.

SmartBuilder: SmartBuilder combines authoring tools like Flash and a rapid development timeline. It does not require scripting. And with its point-and-click interface, SmartBuilder has a learning curve that's measured in days vs months. It has content management features to collaboration with other users.

SmartBuilder is good to develop Product training, Compliance training, Safety training, IT training, Equipment training, Soft skills simulations, Interactive exercises, Games and Case studies [34]

iSpring: iSpring provides PowerPoint to Flash solutions for trade shows, web conferencing, advertising, e-learning and rich media content creation. It is easy and fast to create Flash banners in PowerPoint with iSpring. PowerPoint offers different effects and facilities to create animated content. It is easy to make user's PowerPoint texts or images modification. User just needs to gather all these things and let iSpring to output them into Flash banner. [35]

Easy Authoring launched Learning Management System: Easy Authoring launched Learning Management System hosting service allow virtual learning environment in a very easy way. It is easy for installation, setup, hosting, maintenance, support and training. Additionally, they provide customize learning platform with user corporate colors and logo for corporate branding.[36]

Highlights of Learning Management System

- It comes at a relevant point in time as the training and education industry has begun to focus on informal and blended learning models.
- It is packaged with an impressive array of features, and is highly configurable and extensible.
- Moodle is used worldwide by many colleges, universities, corporate, government organizations, etc.
- It is an easy and cost-effective way of delivering online learning, offline learning and blended learning in an institute.

Engage: Campus is a tool for managing knowledge resources and delivering them consistently throughout an organization. The learning management features enable organizations to instantly deploy training initiatives through its course catalogue, flexible reporting system.[37]

Mercury eLearning creation tool: Mercury is a Flash application that builds rich, SCORM-compliant eLearning from simple XML instructions. Any eLearning content can be defined quickly and easily, including the resource's language(s), navigation layout, number of pages, content structure, sessions, activities, assessments, videos, photos, audio, animations and text.[38]

Skillcast: Skillcast is an online e-learning authoring software that allows user to make courses without design or programming experience. A basic computer skill is good enough to build any content. user can upload content as is from PowerPoint slide decks, or use pre-built page templates available to user in Skillcast's integrated Asset Library. User can also select from thousands of graphics, configurable

interactivities and multimedia extensions. The interface of Skillcast Editor is as user friendly as PowerPoint, but with all the features user need to build high-impact e-learning courses.[39]

Moodle: Moodle is an Open Source Course Management System (CMS), also known as a Learning Management System (LMS) or a Virtual Learning Environment (VLE). It has become very popular among educators around the world as a tool for creating online dynamic web sites for their students. [40]

The focus of the Moodle project is always on giving educators the best tools to manage and promote learning, but there are many ways to use Moodle:

- Moodle has features that allow it to scale to very large deployments and hundreds of thousands of students, yet it can also be used for a primary school or an education hobbyist.
- Many institutions use it as their platform to conduct fully online courses, while some use it simply to augment face-to-face courses (known as blended learning).
- Many users love to use the activity modules (such as forums, databases and wikis) to build richly collaborative communities of learning around their subject matter (in the social constructionist tradition), while others prefer to use Moodle as a way to deliver content to students (such as standard SCORM packages) and assess learning using assignments or quizzes.

2.3 Some Important Terms

2.3.1 Data Mining

Data mining is a process to select, explore and model large amount of data for discovering unknown relationships and patterns which actually give a useful and clear result to the data analyst [43]. Coined in the mid-1990s, the term data mining has today become a synonym for 'Knowledge Discovery in Databases' which, as proposed by Fayyad et al. [44], emphasized the data analysis process rather than the

use of specific analysis methods. Different problems of data mining are most of time solved by using

Data mining problems are often solved by using a combination of different approaches taken from computer science, including multi-dimensional databases, machine learning, soft computing and data visualization, and from statistics, including hypothesis testing, clustering, classification and regression techniques.

The real challenge of data mining is selecting right appropriate and combination of these techniques to reliably and efficiently solve a problem. [45]

2.3.2 Machine Learning Algorithm

"Machine learning is a field of computer science that gives computer systems the ability to "learn" (i.e. step by step performance improvement of a task) with data, without complex programming. "Machine learning" name was coined in 1959 by Arthur Samuel. machine learning can explores and construct of algorithms that can learn from and make predictions on data using pattern recognition and computational learning theory of artificial intelligence. These algorithms can overcome from static program instructions by making data-driven predictions or decisions, by building a model from sample data. Machine learning is solution where designing and programming explicit algorithms with good performance is difficult or infeasible, like applications include email filtering, detection of network intruders or malicious insiders working towards a data breach, optical character recognition (OCR), learning to rank, and computer vision."[46]

2.3.3 Multiple Linear Regression Model

Multiple linear regression is the most common form of linear regression analysis. The multiple linear regression is used to build the relationship between one continuous dependent variable and two or more independent variables for

predictive analysis. The independent variables can be continuous or categorical .[47]

If we consider Y is a dependent variable and there have N number of independent variables then Y will be as following according to Multiple Linear Regression.

$$Y = b0 + b1*X1 + b2*X2 + b3*X3 + \ldots\ldots + bn*Xn$$

Chapter 3

3. Methodology and Design

3.1 Theory

There is no standard algorithm to judge the online student after test via e-learning system. At e-learning web site student can select random result on test and can pass also. If any student passes after online test, it does not mean that he/she really learned from web. May be student can pass only selecting random result. So, at e-learning, online testing cannot ensure students learning outcome. To ensure this, system need any specific algorithm to use in an intelligent e-learning system. This algorithm will force the student to learn via online and also ensure actual learning outcome of that student. The learning content can be text, video, audio, animation. The intelligent system will act individually with the student and will learn and update the testing and learning methods automatically. This system will use machine learning algorithm to train itself. The Multiple Linear Regression Algorithm will be used to judgment student exam result. Also statistical significance based on P-value will be used to define best method for an exam.

3.2 System Flow and Algorithm

The system can be act like an Artificial Intelligent Game. There will be several stage of learning. For a specific subject, course tutor will provide some learning content in text, video, audio or animation format. Then he/she will prepare 15-20 question set for different level. Also the tutor will define marks distribution for the questions. Different levels activities are described below.

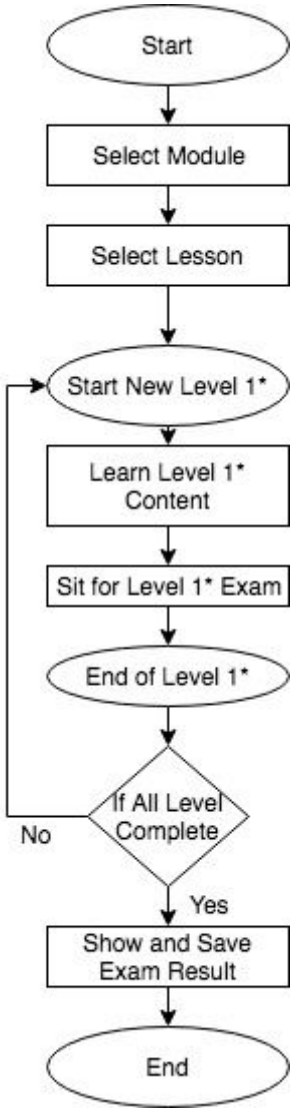

Figure.1 Flow chart for evaluation system

Step 1: In this step student will select module from his/her course curriculum.

Step 2: At second step student will select a lesson (chapter) of the module.

Step 3: Now new level will start. At first face of the level, student will learn about the topic from any kind of content like text, video, audio or animation.

Step 4: In this step student will sit for an exam. Here system will provide 20 MCQ type questions with four options. Student will select correct answer as per his/her knowledge. After completing all answer he/she will submit the exam form.

Step 5: Now system will check, are all levels completed or not? If NO than student will be redirected to step 3 again for new level. This cycle will continue until all level completed. If no more level exists then student will go to final step.

Step Final: At final step system will show the result of the exam and also will save the case history to the database.

At each level student will get same lesson as different content. That's mean lesson content will be dynamic at each level but will teach same information. Also in each level exam question will be same but question and answer both's presenting language will be different. Like at level one, question can be "What is computer?". But at level two question will be like "What is definition of computer?". An Machine Learning Algorithm will define how many level should have for an exam.

3.3 Why Machine Learning

This proposed system would use Multiple Linear Regression Algorithm of Machine Learning for couple of reasons.

Reason 1: To identify most statistically significant variable to design an exam method.

Reason 2: To judge student exam result.

3.4 Why Multi Level Exam

In real life people learn things from different perspective. If people learn any content in different way like text, video, audio, animated mode or with real life example then it become very easy to understand and also people can remember the topic. Also it is known that at gaming world people go through different level. Which not only ensure his/her expertise but also make the game exciting for the player. So, this system is following same approach here. At different level system will present same topic from different perspective. Which will not only make the learning process exciting but also ensure student learning outcome by complex intelligent judgement process.

3.5 How to Define Best Exam Configuration

First of all it has been selected three criterias to design an exam process.

max_exam_level: It will define how many level should have for an exam process.

min_avg_result: Let's think that a student will answer same MCQ question in five level. So, he/she may get total 0-5 numbers for that particular question. If he/she get 3 then his/her average number will be (¾)*100 = 60% for that question. Now if minimum average number is defined as 60 to take the answers as correct then end of the exam the actual result will be 1 for that question. If he/she get less average number then system will consider the result as 0.

min_pass_marks: Let's think that a lesson will contain total 20 questions. So, after the exam a student may get total 0-20 numbers according to the logic of min_avg_result. Here min_pass_marks will judge by another level of marking. If his/her total marks is more than or equal to the min_pass_marks then he/she will be considered as pass in the exam, otherwise fail.

At starting system doesn't know, are this variables actually can dominate the exam result or not. And also system doesn't know what combination of this variable will be best to examine a student. So, after the sample data examination by machine learning system will find these two questions answer.

3.6 How to Judge Student Exam Result

After millions of exam record generation, system will build a Multiple Linear Regression Model. By that model system will be able to judge new student's exam result. System will compare test dataset with training dataset. If system see student's exam result is almost same for the same kind of module, lesson and exam method then system can ensure that student really has gone through actual learning process without any random selection.

3.7 Use Case Diagram of the System

Here Use Case diagram of the system is presented. Different user's functionality is shown in Different Use Case. System has Use Case for three actors. They are:

a. Student
b. Teacher
c. Admin

Following Use Case contain the whole systems functionality for above three actors. Student has two functions. One is Learning Process. Second is Exam Process. Teacher has Module Handling Process, Q/A (Question/Answer) Creation Process. Admin has User Handling Process and Administration Process. Each user has some common functionality. Like login and Logout.

a. Learning Process: Through learning process, student will learn new topic/module.
b. Exam Process: Through exam process, student will sit for exam to judge his/her learning outcome.
c. Module Handling Process: Through module handling process, teachers will add/ update/ remove lesson or material for a module.
d. Q/A Creation Process: Through Q/A creation process, teachers will add/update/remove Question/Answer/Explanation of a lesson.

e. User Handling Process: Through user handling process, Admin will create/update/remove student and teacher account.

f. Administration Process: Through administration process, admin will assign modules to student or teacher.

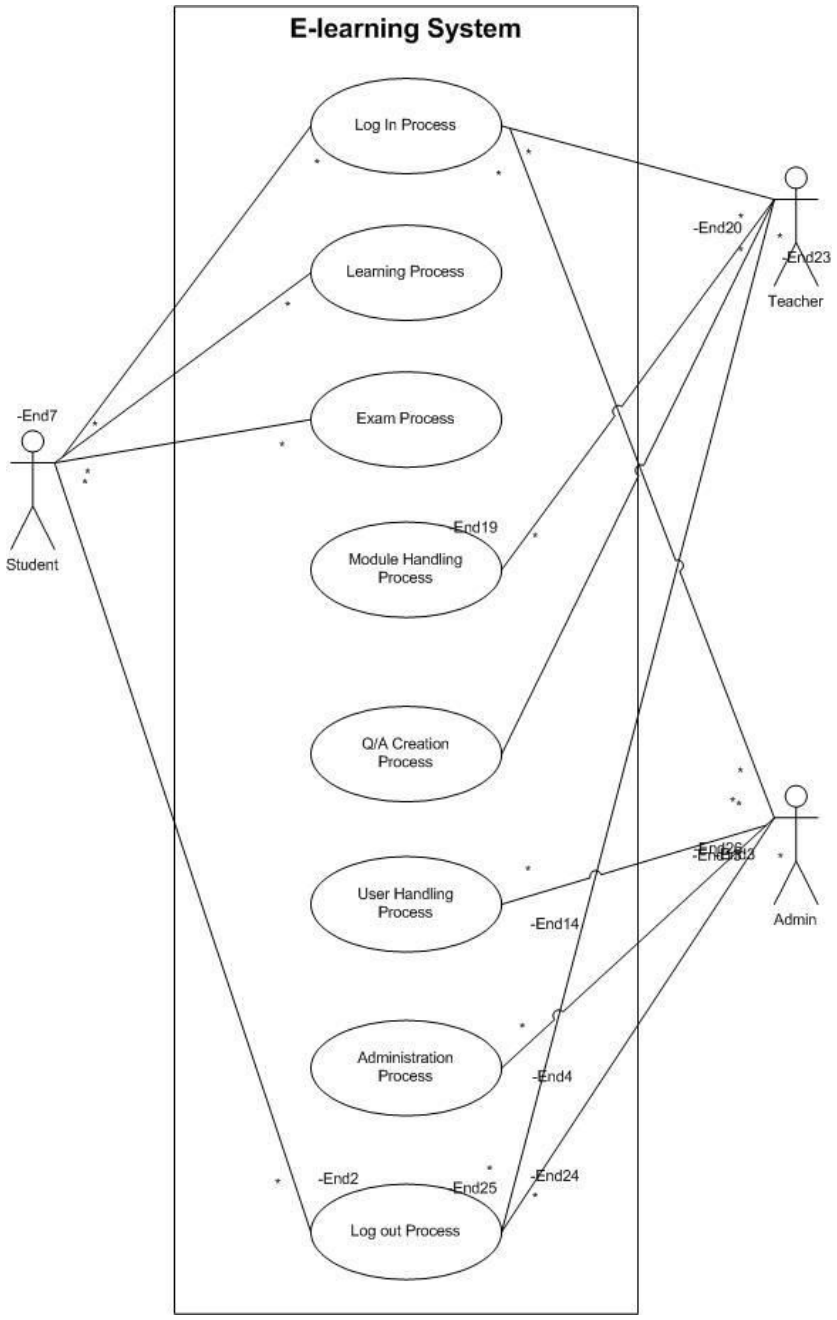

Figure 2 : Whole e-learning system Use Case.

Following Use Case contain learning functionality of a student. Student has seven functions.

a. Login: Student login to the system using this function.
b. Module Selection: Student selects a module about which he wants to learn.
c. Component Selection: After select the module, student selects the learning Component type. Like text, audio, video, animation, Image, reference link etc.
d. Exam Process: Student sits for exam through this process.
e. Save Level Position: If student want to leave learning process now and start again from current level then they can save current learning level position.
f. Start New Level: Student can start new/next level through this process after finish a level exam.
g. Logout: Student logout from the system.

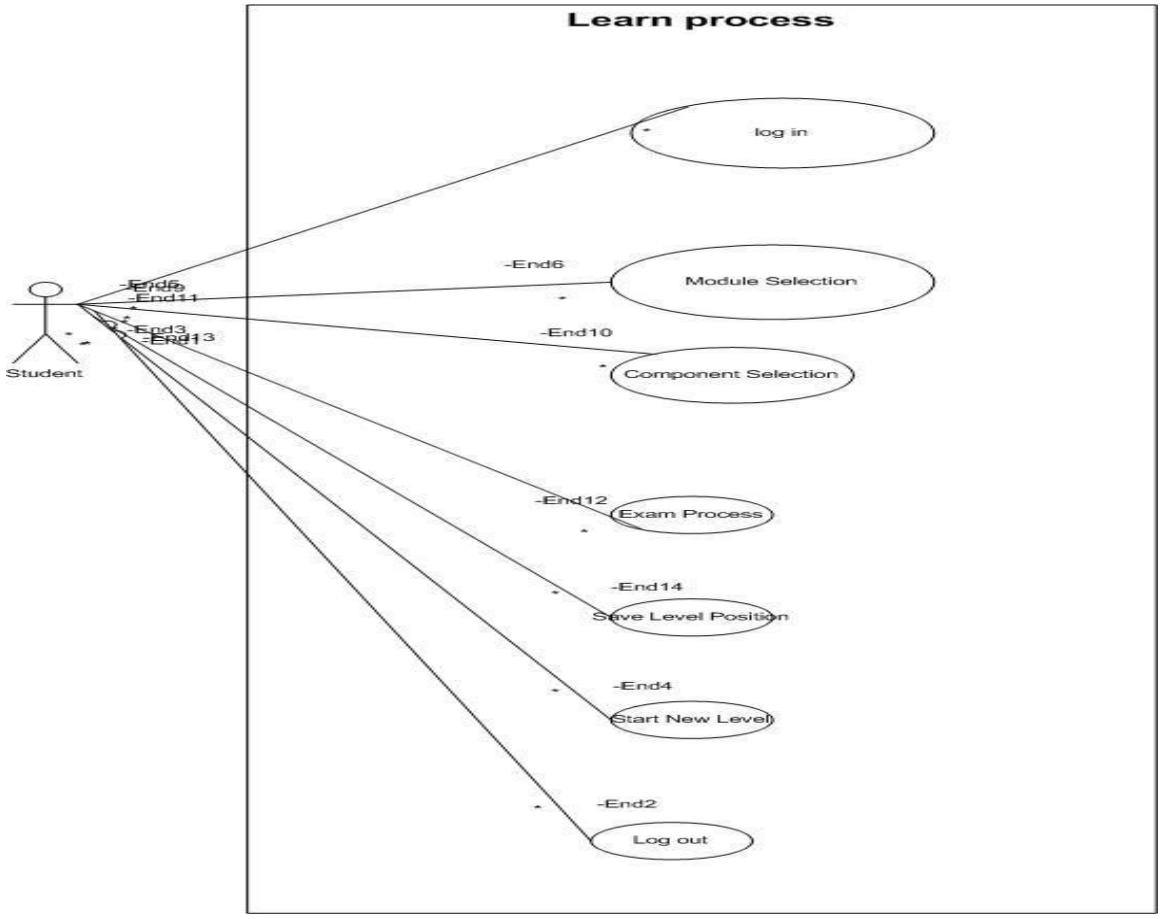

Figure 3 : Student Learning Process Use Case

Following Use Case contain the Exam process of a student. There have two functions.

a. Test Start: Start a new exam.
b. Answer Submission: Student selects the answer and submits for result.

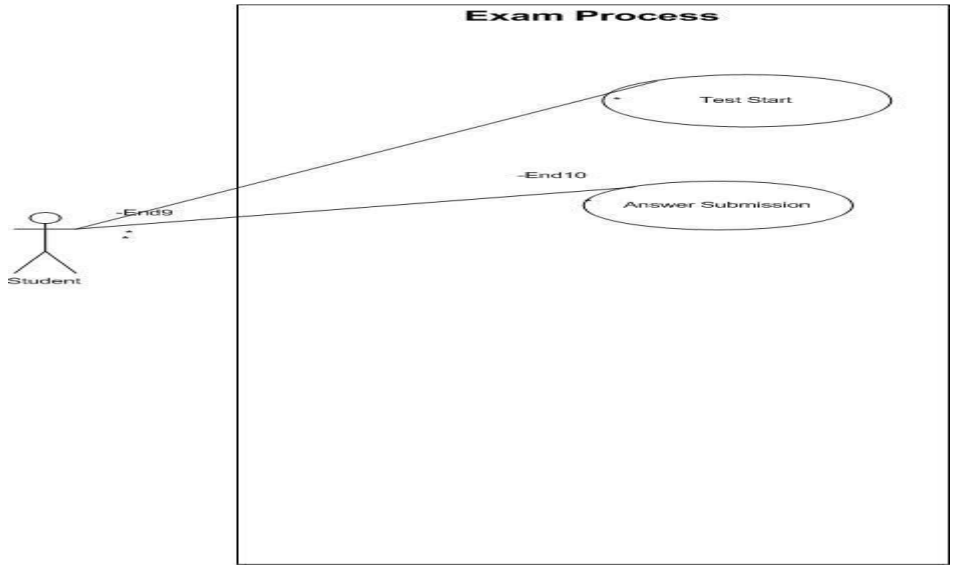

Figure 4 : Student Exam Process Use Case.

Following Use Case contain the Module Handling Process of a teacher. In this process a teacher has nine functions. There are:

a. Module Selection: Select a Module to add/update/remove Lesson/Material.
b. Lesson Addition: Add a new Lesson for the Module.
c. Lesson Selection: Select a lesson to add/update/remove Material.
d. Component Selection: Select a component type to upload module materials.
e. Material Addition: Upload new material for the lesson.
f. Lesson Update: Update existing Lesson information.
g. Material Update: Update existing Material.
h. Lesson removal: Remove/Delete existing Lesson.
i. Material removal: Remove/Delete existing Material.

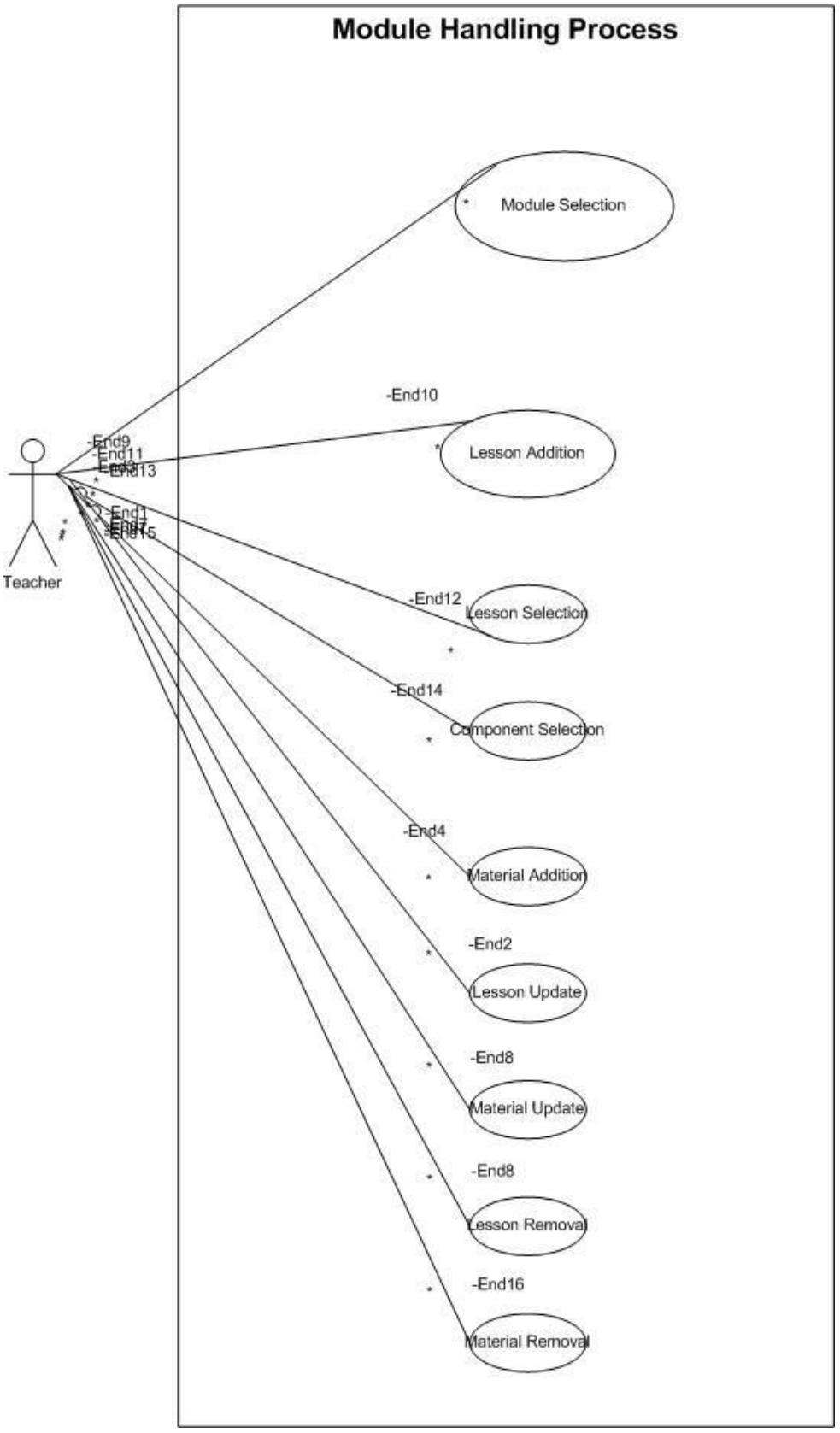

Figure 5 : Teacher Module Handling Process Use Case

Following Use Case contain the Q/A (Question/Answer) Creation Process of a teacher. There have nine functionalities. They are:

a. Module Selection: Select a Module to add/update/remove Question/Answer.
b. Lesson Selection: After select module he/she select a Lesson to add/update/remove Question/Answer.
c. Question Addition: Add several new questions for the lesson.
d. Answer Addition: Add several wrong answers and a correct answer for each question (it is a multiple chose exam).
e. Explanation Addition: Add little explanation of the correct answer. This will be shown at next level if any student selects wrong answer in exam for this question.
f. Question Update: Update existing Question.
g. Answer Update: Update existing Answer.
h. Explanation Update: Update existing Explanation.
i. Question Remove: Remove existing Question.
j. Answer Remove: Remove existing Answer.
k. Explanation Remove: Remove existing Explanation.

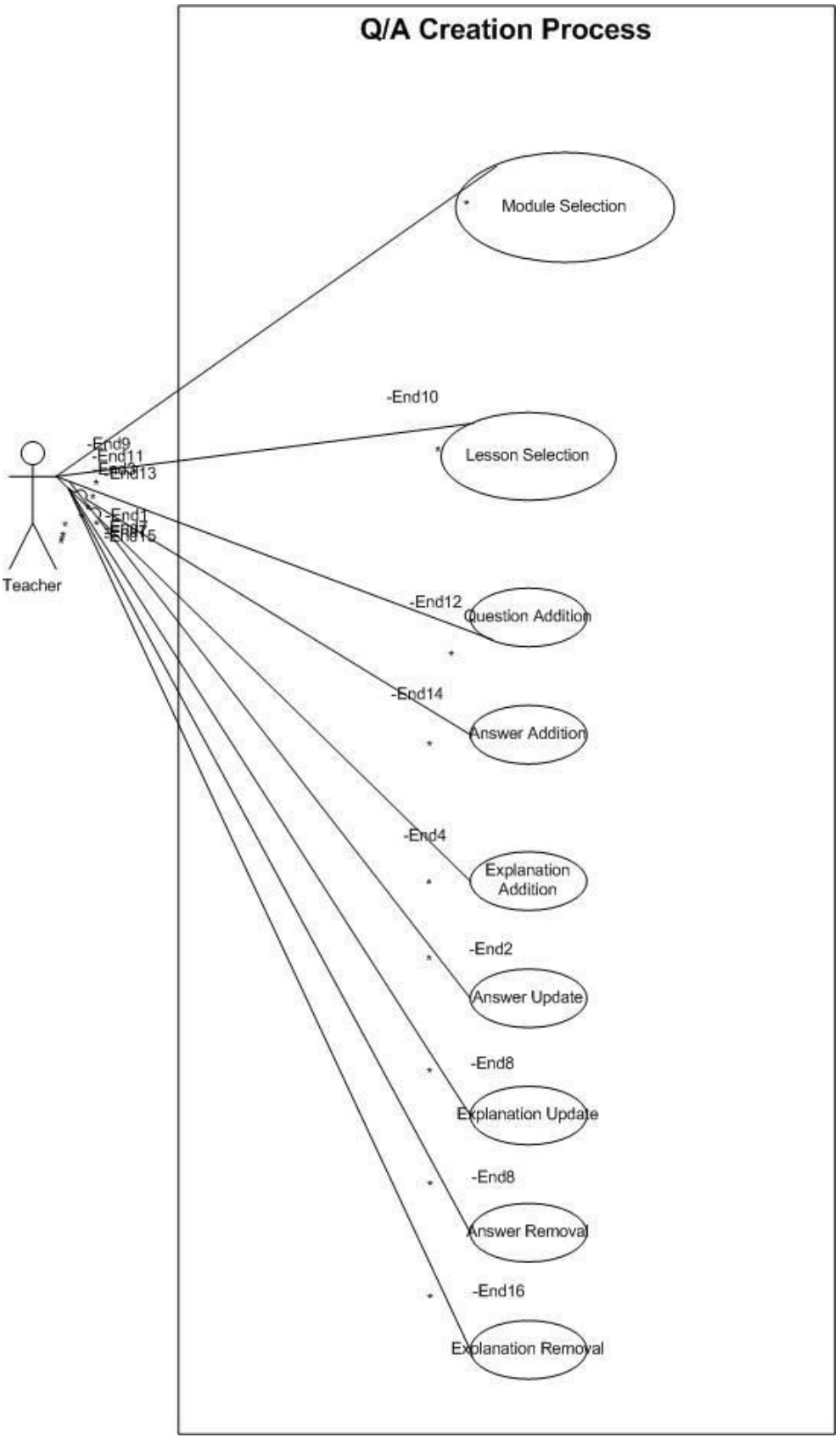

Figure 6 : Use Case for Q/A Creation Process for a teacher

Following Use case contain the User Handling Process for an Admin. There have six functionalities. They are:

a. Student Account Creation: Admin create new student account for the system.
b. Teacher Account Creation: Admin create new teacher account for the system.
c. Student Account Update: Admin update existing student account information.
d. Teacher Account Update: Admin update existing teacher account information.
e. Student Account Remove: Admin remove existing student account information.
f. Teacher Account Remove: Admin remove existing teacher account information.

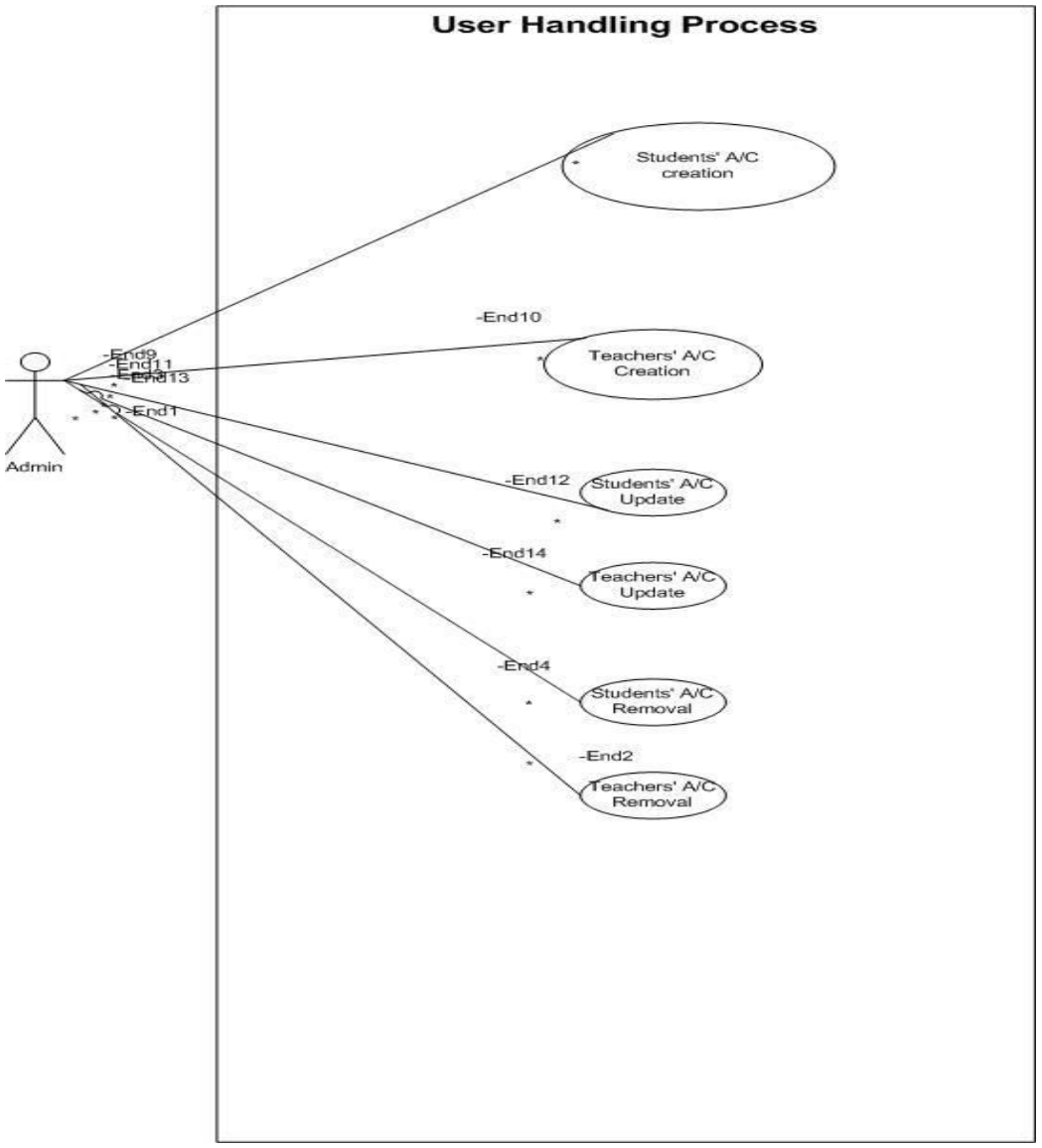

Figure 7 : Use Case for Admin User Handling Process

Following Use Case contain the Administration Process of an Admin. There have seven functionalities. They are:

a. Module Creation: Create a new module.
b. Module Allocation for Student: Admin assign a module to a student.
c. Module Allocation for Teacher: Admin assign a module to a teacher.
d. Module Update for Student: Admin update assigned module list for a student.
e. Module Update for Teacher: Admin update assigned module list for a teacher.
f. Module Remove for Student: Admin Remove assigned module for a student.

g. Module Remove for Teacher: Admin Remove assigned module for a teacher.

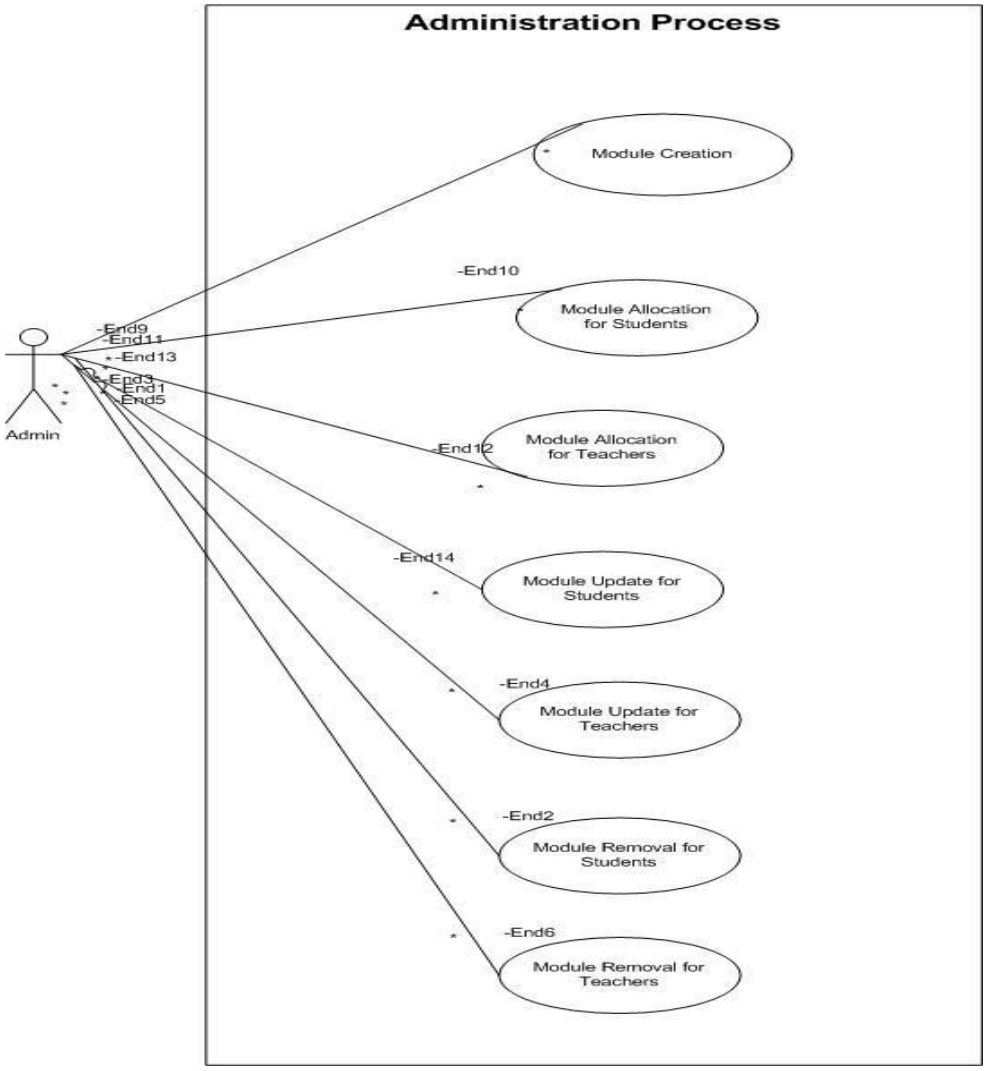

Figure 8 : Use Case for Admin Administration Process

3.8 Sequence Diagrams of the System

Now, there are explained the sequence diagrams of the system. There are nine sequence diagrams. Students have one sequence diagram. Teachers have five sequence diagrams and Admin have three sequence diagrams.

Following sequence diagram is for a student. The sequence diagram is explained step by step:

a. Student login to the system and get main page.
b. Form “main page” student select the module about which module he/she want to learn.
c. From the “component list page”, student selects the component type. The component types are text, animation, audio, video, image, reference link.
d. After selecting component type the materials of that module will be displayed at “display component page”. The sequence of the lessons of the module will be maintained automatically.
e. After learn about the lesson student can sit for the exam. If student start exam, he/she will be redirect to “test page”.
f. When student will submit the exam then he/she will get the result.
g. Now student can start again a new level/lesson from “component list page”. After each exam level/lesson will be updated automatically.
h. Or student can save the level to start learning again later.
i. Students can logout from the system.

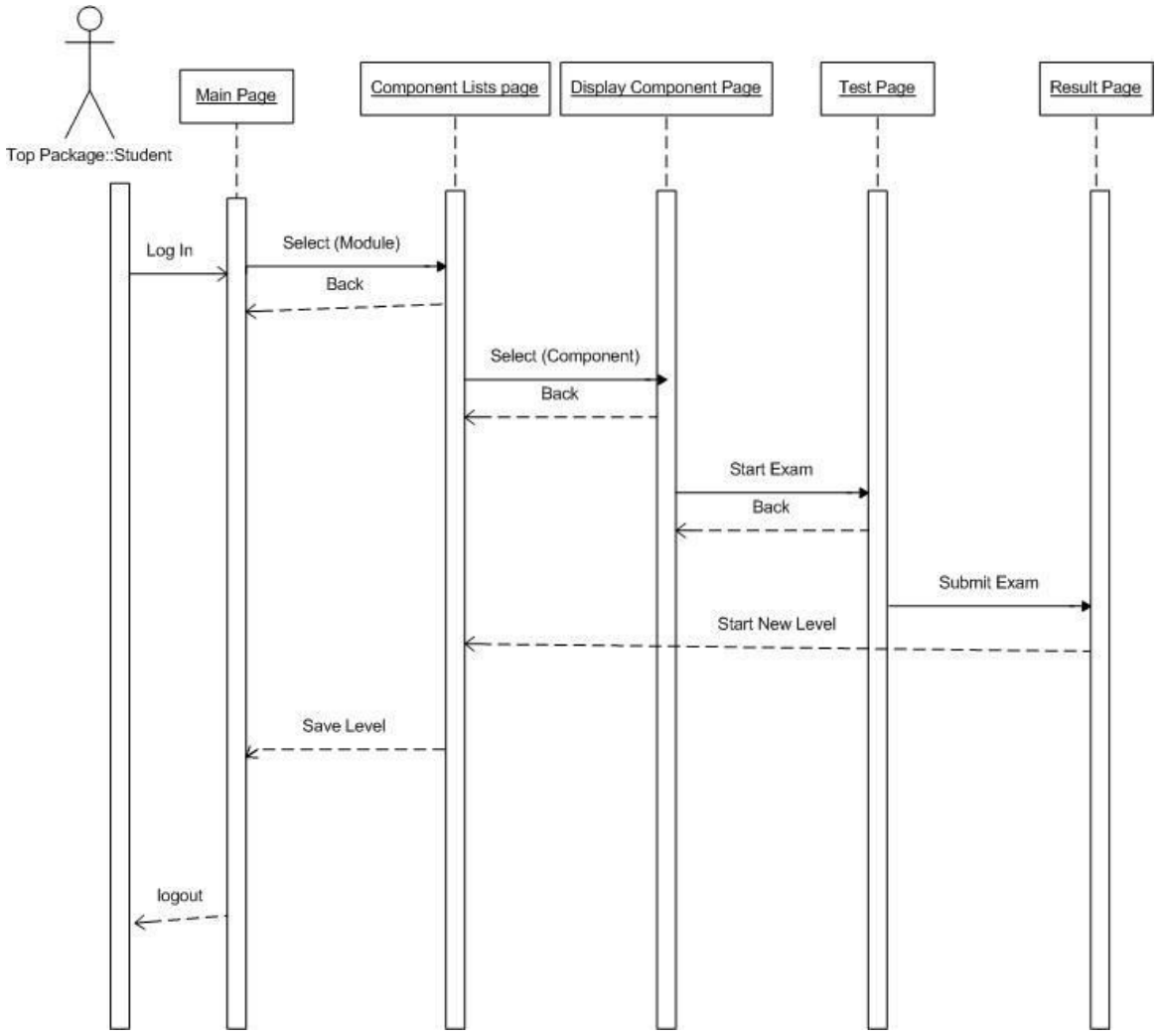

Figure 9 : Sequence Diagram of student learning process

Following sequence diagram is for teacher to add lesson and material for a module. The sequence diagram is explained step by step:

a. Teacher login to the system and get “main page”.
b. From main page he/she can select module/topic to add lesson/material.
c. At “Add Lesson Page”, teacher will give input about basic information of the lesson.
d. To input material, now teacher have to select component type.
e. At “add material page”, teacher will upload materials.
f. “Display preview page” will show the preview of lesson and materials.
g. After save command, system will show confirmation page.

h. Now, teacher can add another lesson from main page or logout from the system.

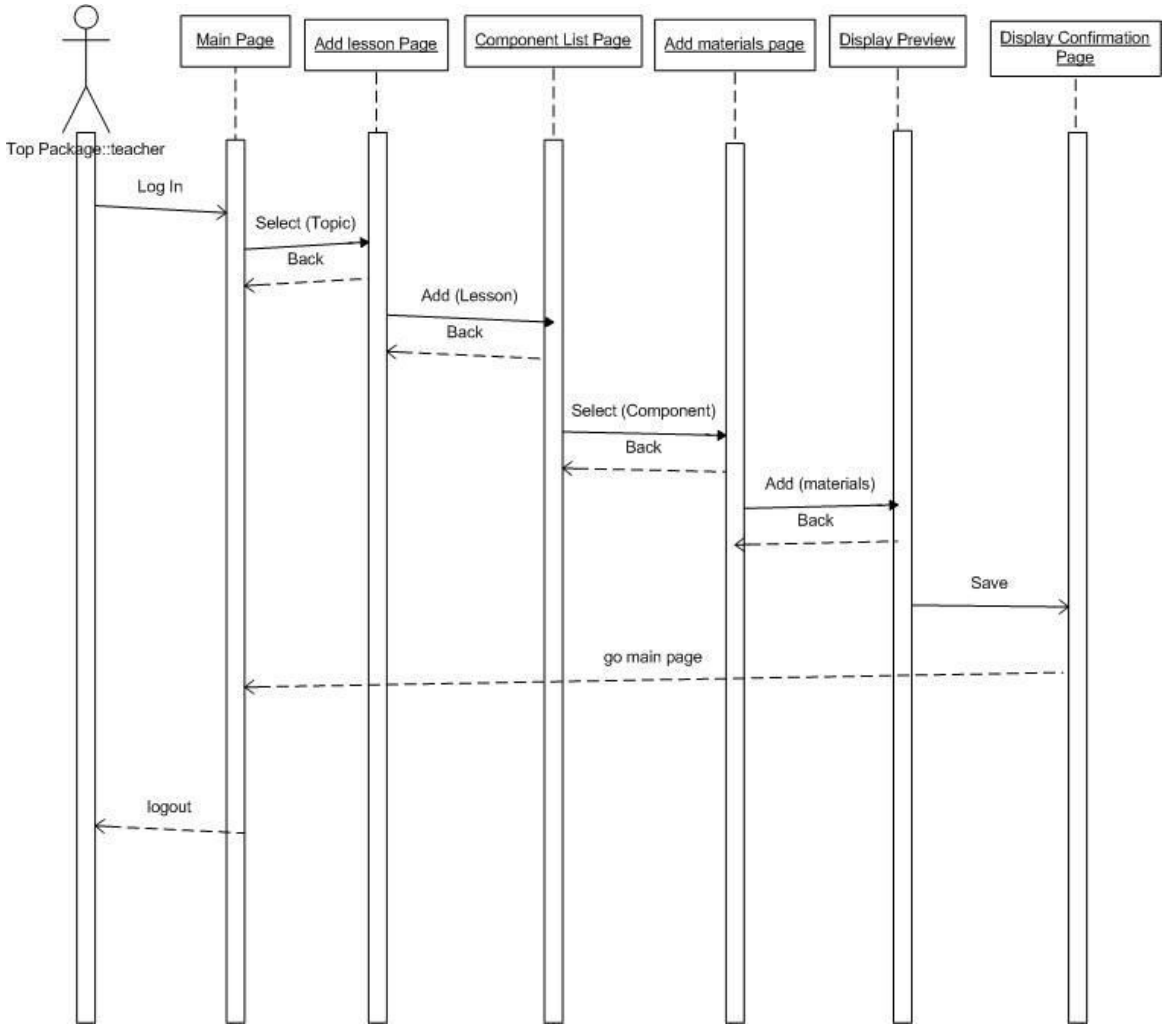

Figure 10 : Sequence Diagram to add lesson and material by teacher

Following sequence diagram is for teacher to update lesson for a module. The sequence diagram is explained step by step:

a. Teacher login to the system and get “main page”.
b. To update lesson, teacher will to “lesson list page” to select the lesson.
c. Form “lesson list page” he/she will select the lesson.
d. At “lesson update page”, teacher will update the lesson information.
e. “Display preview page” will show preview of the lesson information.
f. “Display Confirmation page” will show confirmation message.

g. Now, teacher can update another lesson from main page or logout from the system.

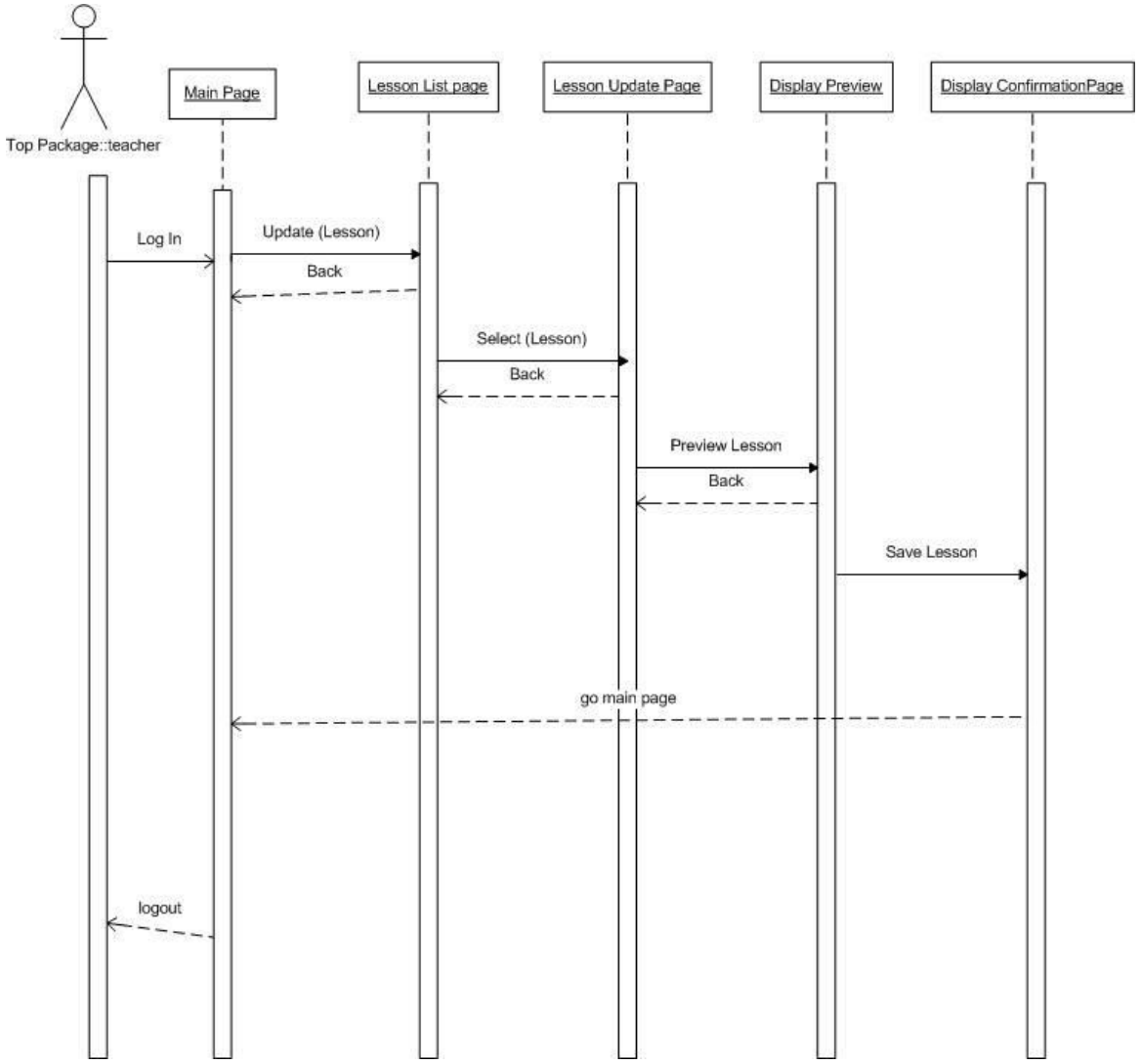

Figure 11: Sequence Diagram to update lesson by teacher

Following sequence diagram is for teacher to update material for a module. The sequence diagram is explained step by step:

a. Teacher login to the system and get “main page”.
b. To “update material”, teacher will go to “lesson list page” to select the lesson.
c. Form “lesson list page” he/she will select the lesson.
d. Now teacher will select component type to update the material of that component.

e. At “material update page”, teacher will update the material information.
f. “Display preview page” will show preview of the material information.
g. “Display Confirmation page” will show confirmation message.
h. Now, teacher can update another material from main page or logout from the system.

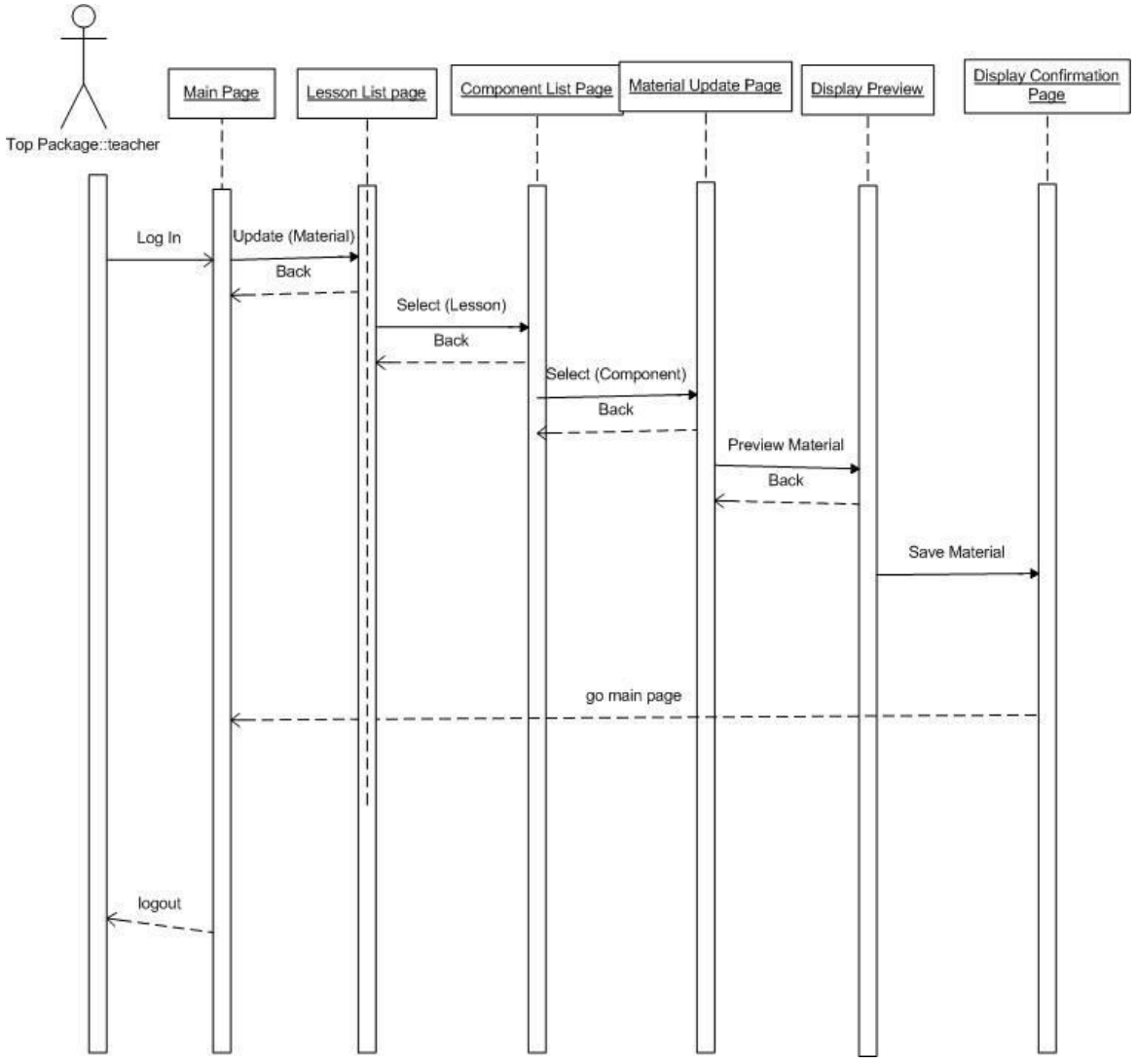

Figure 12: Sequence Diagram to update material by teacher

Following sequence diagram is for teacher to add Q/A (Question/Answer) for a module. The sequence diagram is explained step by step:

a. Teacher login to the system and get “main page”.
b. To “add Q/A”, teacher will go to “lesson list page” to select the lesson.
c. Form “lesson list page” he/she will select the lesson.
d. Now teacher will input one question, four wrong answers, one correct answer and brief explanation of the correct answer at “Add Q/A page”.
e. “Display preview page” will show preview of the Q/A information.
f. “Display Confirmation page” will show confirmation message.
g. Now, teacher can add another Q/A from main page or logout from the system.

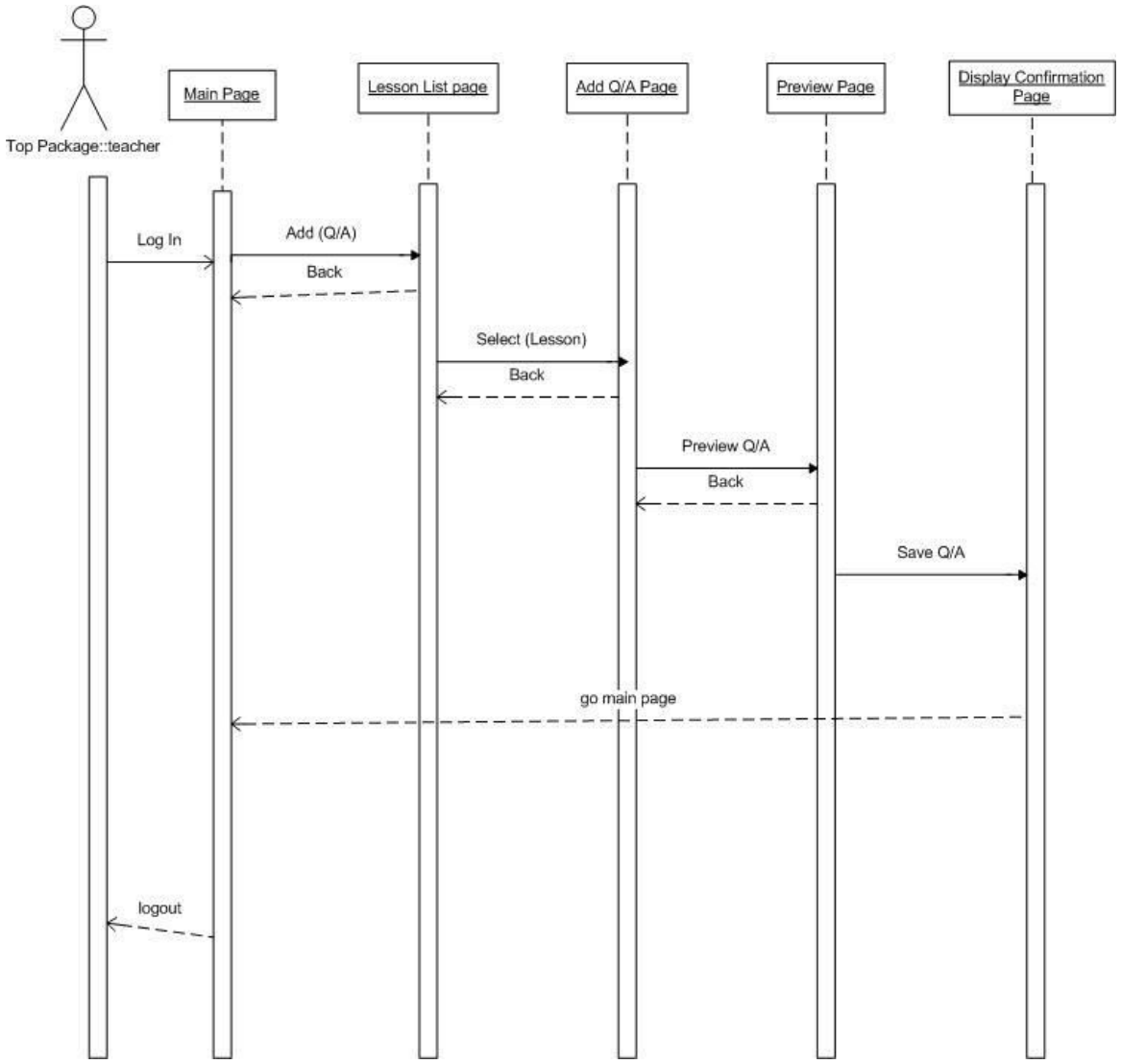

Figure 13 : Sequence Diagram to add Q/A by teacher.

Following sequence diagram is for teacher to update Q/A (Question/Answer) for a module. The sequence diagram is explained step by step:

a. Teacher login to the system and get “main page”.
b. To “update Q/A”, teacher will go to “lesson list page” to select the lesson.
c. Form “lesson list page” he/she will select the lesson.
d. Now teacher will update existing question, answers and brief explanation of the correct answer at “Update Q/A page”.
e. “Display preview page” will show preview of the Q/A information.
f. “Display Confirmation page” will show confirmation message.

g. Now, teacher can update another Q/A from main page or logout from the system.

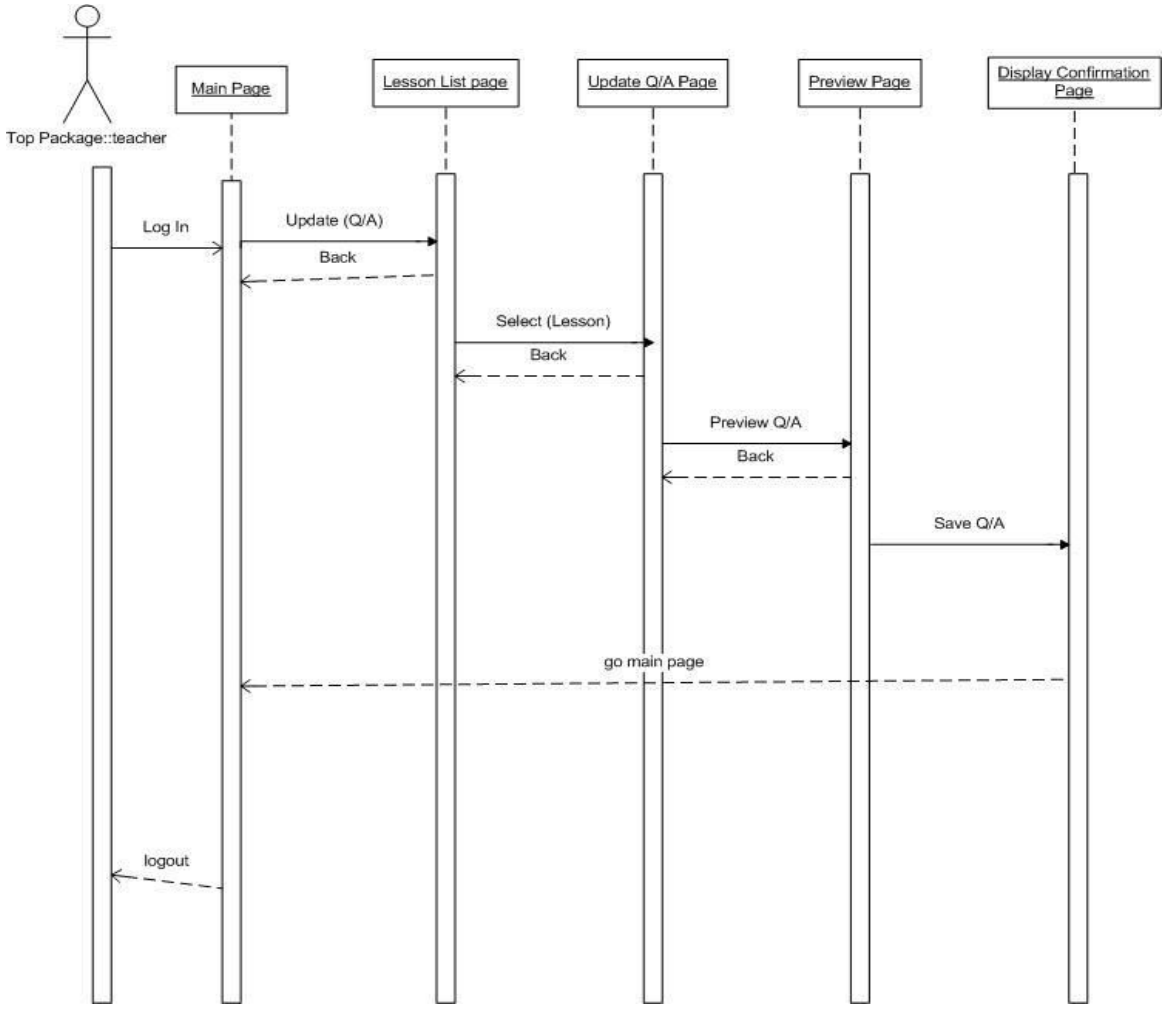

Figure 14 : Sequence Diagram to update Q/A by teacher.

Following sequence diagram is for Admin to create users for system. The sequence diagram is explained step by step:

a. Admin login to the system and get “main page”.
b. To “add user”, admin will go to “Create User Page”.
c. Admin input basic information of user (like student or teacher).
d. “Display preview page” will show preview of the information.
e. “Display Confirmation page” will show confirmation message.

f. Now, admin can create another user from main page or logout from the system.

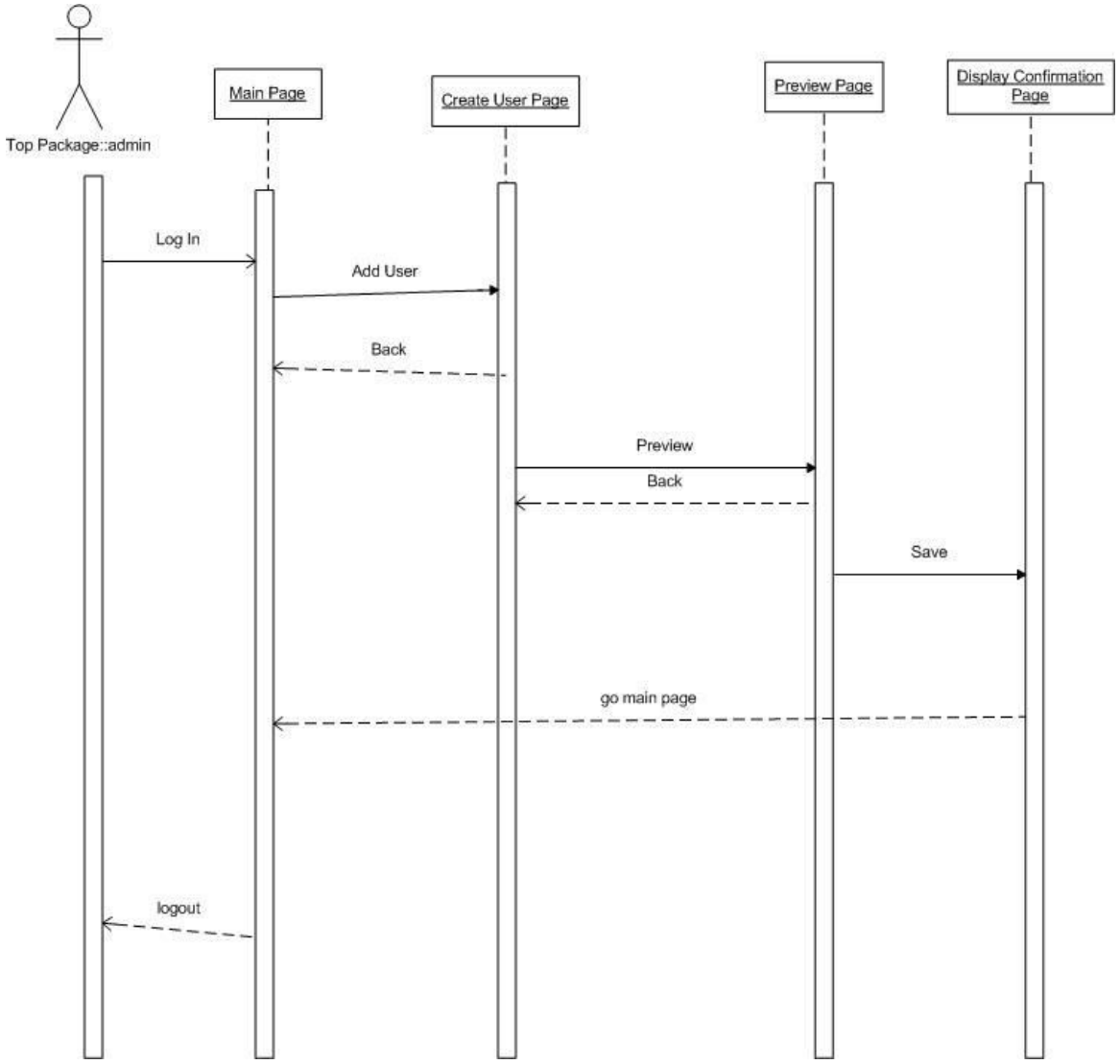

Figure 15 : Sequence Diagram to create a user by admin.

Following sequence diagram is for Admin to update users for system. The sequence diagram is explained step by step:

a. Admin login to the system and get “main page”.
b. To “update user”, admin will go to “User List Page”.
c. Now select user.
d. Admin update basic information of user (like student or teacher).

e. “Display preview page” will show preview of the information.

f. “Display Confirmation page” will show confirmation message.

g. Now, admin can update another user from main page or logout from the system.

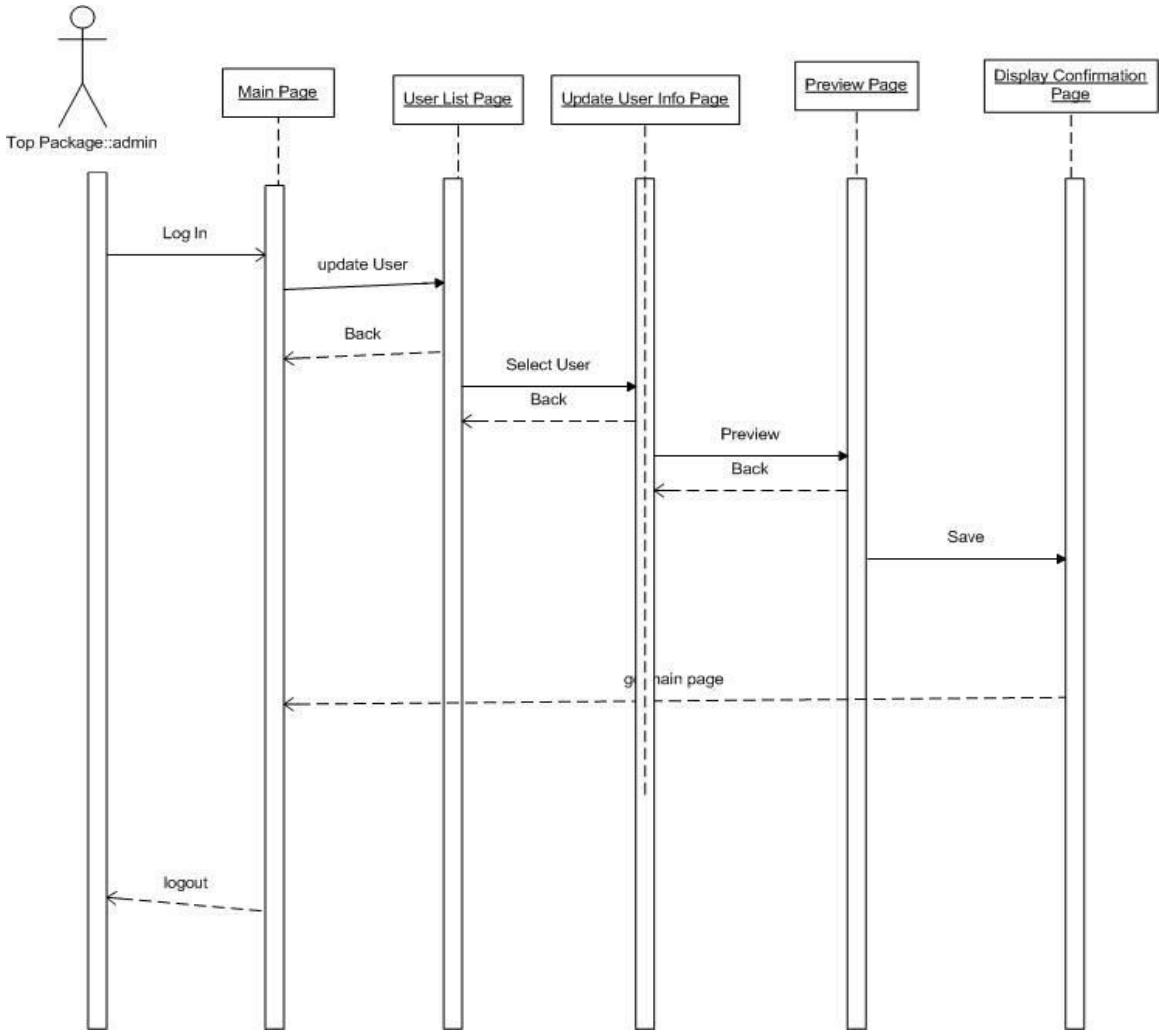

Figure 16 : Sequence Diagram to update a user information by admin.

Following sequence diagram is for Admin to assign/allocate users to modules. The sequence diagram is explained step by step:

a. Admin login to the system and get “main page”.

b. To “assign module”, admin will go to “Module List Page”.

c. Now select module and go to “User List Page”.

d. Now select User for that module.

e. “Display preview page” will show preview of the information.

f. “Display Confirmation page” will show confirmation message.

g. Now, admin can assign another user from main page or logout from the system.

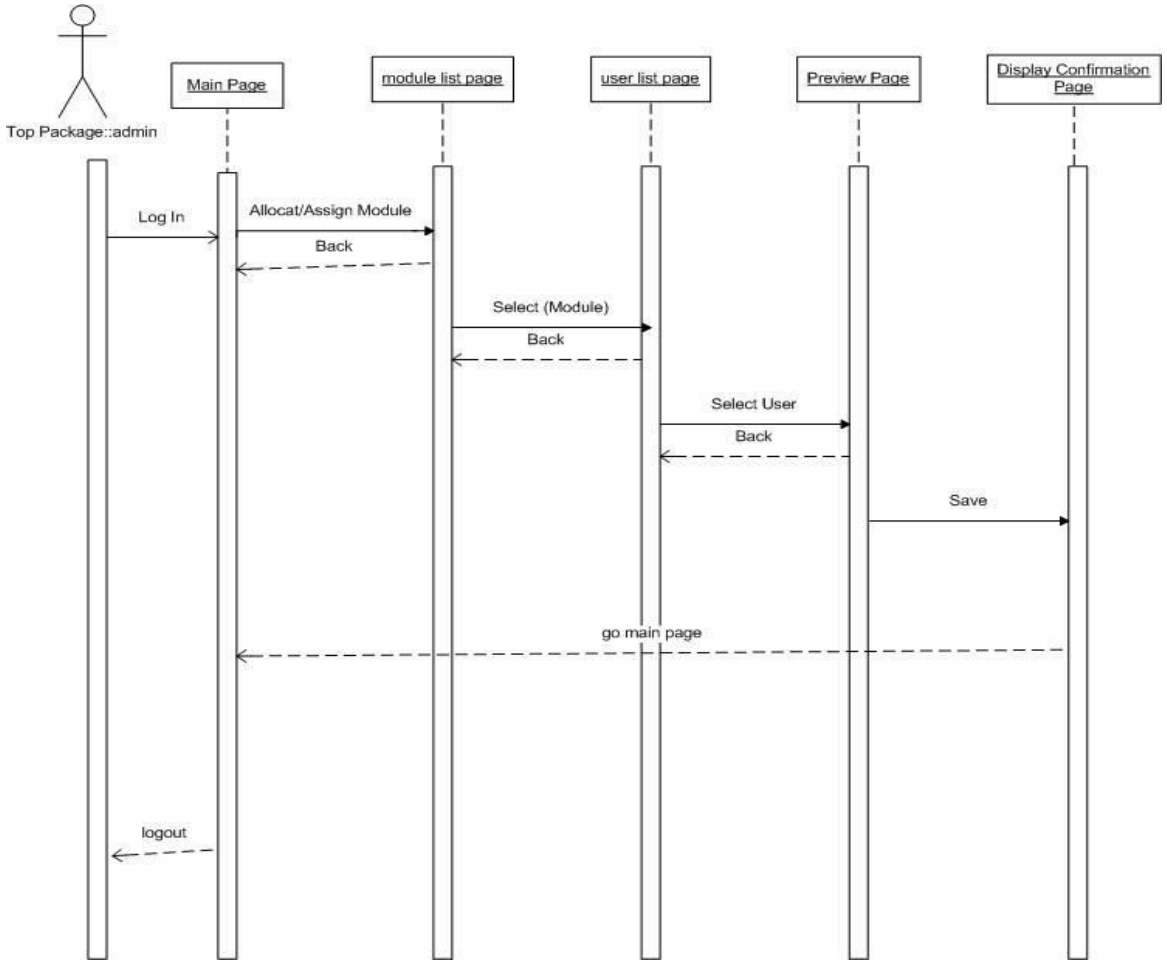

Figure 17 : Sequence Diagram to assign a user to a module by admin.

3.9 Class Diagram of the System

Following class diagram represent the whole e-learning system. This diagram contains 19 classes. The classes are:

a. Student: contain the basic information of a student.

b. Teacher: contain the basic information of a teacher.
c. Module: contain the basic information of a module.
d. Lesson: contain the basic information of a lesson.
e. Material: contain different learning material of a lesson.
f. Attachment: contain material's attachment. Like image, video, audio, animation.
g. QA: contain information of question, answer, explanation.
h. Admin: contain the basic information of an admin
i. User: contain the basic information of a system user. Like user type is student or teacher or admin.
j. StudentModule: contain basic information of a student's assigned module.
k. StudentLesson: contain basic information of a student's assigned module's lesson.
l. TeacherModule: contain basic information of a teacher's assigned module.
m. Date: contain basic action date and actor information. Most of class extend this class.
n. AdminUserHandlingProcess: contain basic methods of admin user handling process.
o. AdminModuleHandlingProcess: contain basic methods of admin module handling process.
p. StudentLearnProcess: contain basic methods of student learning process.
q. StudentExamProcess: contain basic methods of student exam process.
r. TeachersQACreationProcess: contain basic methods of teacher Q/A creation process.
s. TeachersModuleHandlingProcess: contain basic methods of teacher module handling process.

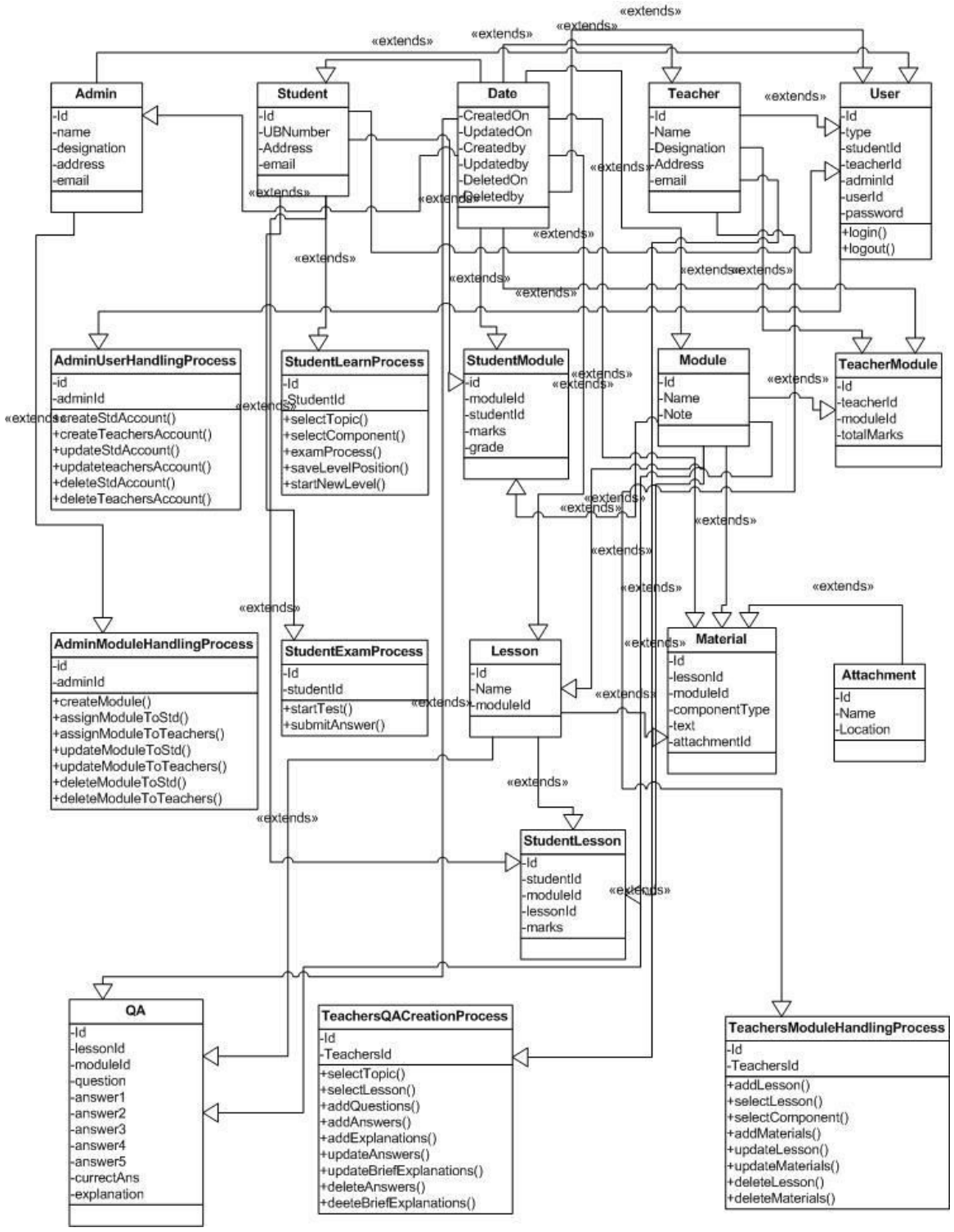

Figure 18 : Class diagram of e-learning system.

3.10 Database Diagram of the System

Following database diagram represent the whole e-learning system's database. This diagram contains 13 tables. The tables are:

a. Student: this table contain the basic information of a student.
b. Teacher: this table contain the basic information of a teacher.
c. Module: this table contain the basic information of a module.
d. Lesson: this table contain the basic information of a lesson.
e. Material: this table contain different learning material of a lesson.
f. Attachment: this table contain material's attachment information. Like image, video, audio, animation and file location of these materials.
g. QA: this table contain information of question, answer, explanation.
h. Admin: this table contain the basic information of an admin
i. User: this table contain the basic information of a system user. Like user type is student or teacher or admin.
j. StudentModule: this table contain basic information of a student's assigned module.
k. StudentLesson: this table contain basic information of a student's assigned module's lesson.
l. TeacherModule: this table contain basic information of a teacher's assigned module.
m. Date: this table contain basic action date and actor information.

All IDs are primary key of corresponding table.

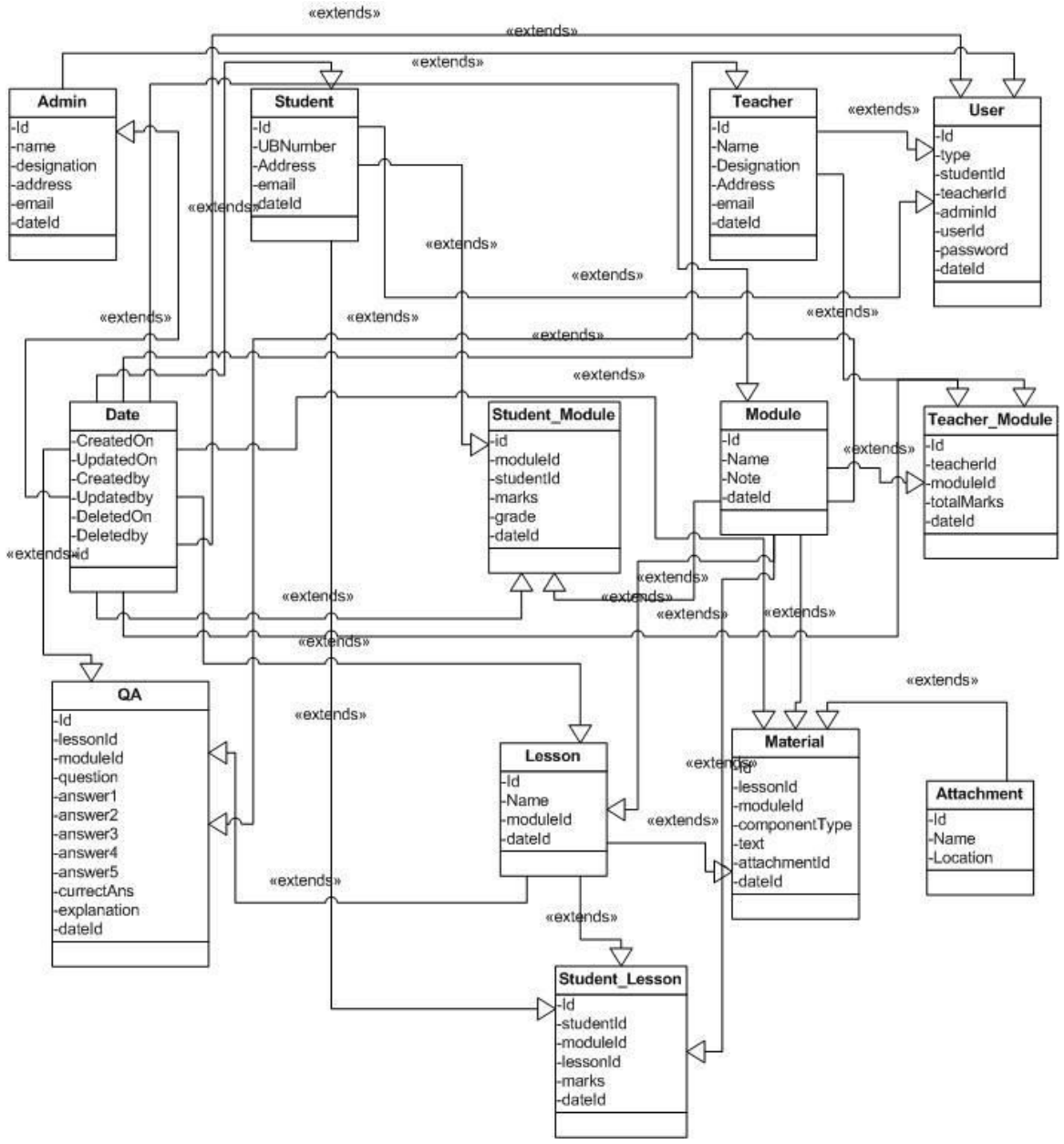

Figure 19 : Database diagram of e-learning system.

Chapter 4

4. Investigation and Result

4.1 Data Processing and Investigation

For investigation almost 34 million sample data are generated in a PostgreSQL Server. First of all 100 sample student information are inserted into **student** table. Then inserted 5 modules into **module** table. For each **module,** 10 **lessons** are inserted where for each **lesson** 1 **question_set** is created. Next 20 **question_group** are inserted for each **question_set**. This **question_group** will contain basic questions. Now for each **question_group,** 5 **question** are generated which actually will represent same question of that particular **question_group** but in different language. Like if a **question_text** is "What is Array?" in a **question_group** then other **questions** will be like "What is the definition of Array?" or "What do you mean by Array" etc. Each **question** of a particular **question_group** will represent each **level** of an exam. Not only the **question** but also the **answers** of each MCQ will be in different language and will mean same **answer**.

After creating sample student, module, lesson, question_set, question_group and question, exam method(exam_config) are generated based on different value of max_exam_level, min_avg_result and min_pass_marks. Here it has been considered 4 (2,3,4,5) max_exam_level, 6 (70,60,50,45,40,35) min_avg_result and 4 (14,12,10,7) min_pass_marks. So, total 4*6*4 = 96 exam method and all are inserted into table **exam_config**.

max_exam_level	min_avg_result	min_pass_marks
5	70	14
5	70	12
5	70	10
5	70	7
5	60	14
5	50	12
5	45	14
5	40	7
5	35	14

Table 1: Sample Data of exam_config Table

Now a SQL script is prepared which will generate exam result for each student, question and exam_config. Here this script is slightly supervised the answer selection to increase pass marks for max level exam.

Example Answer Query:

```
(ARRAY[q.correct_answer,1,2,3,4])[floor(random()*(6 -
q.exam_level_number))+1]
```

Here it's shown that query is selected answer randomly from currect_answer and other 4 options. But greater level has more probability to select correct answer rather then less level.

This SQL script is generated 33600000 data for each possible combination of tables student, module, lesson, question_set, question_group, question and exam_config.

All SQL scripts are included at appendix A.

4.2 Results and Findings

Now sample data will examined in three different steps to get following question's answer.

1. Do max_exam_level, min_avg_result, min_pass_marks have statistical significance on exam result?
2. Which exam method can provide better learning outcome?
3. How system can ensure learning outcome by a particular student's exam result?

4.2.1 Exam Method Variable Analysis

First of all system have calculated total_pass for all student_exam. It's shown that if system increase max_exam_level and decrease min_avg_result, min_pass_marks then total_pass increase.

To examine this result by Multiple Linear Regression Model, system need all positive directional value for better understanding. So, min_avg_result and min_pass_marks are inverted as max_avg_wrong and max_fail_marks accordingly.

Now proposed system has examined the result using R machine learning language. After building **regressor** model on result data it is found that P-Values (which represent statistical significance of a independent variable) of all variables are very very low (*** for all variables). That means proposed system has selected right variables to build a method for exam which can manipulate the student exam result.

```
Residuals:
     Min       1Q   Median       3Q      Max
-3150.70  -916.34    53.82  1030.98  2634.62

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    -7139.67     881.24  -8.102 3.07e-12 ***
max_exam_level   551.09     126.07   4.371 3.41e-05 ***
max_avg_wrong    102.02      11.56   8.826 1.02e-13 ***
max_fail_marks   296.52      53.65   5.527 3.36e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1324 on 87 degrees of freedom
Multiple R-squared:  0.5964,    Adjusted R-squared:  0.5824
F-statistic: 42.85 on 3 and 87 DF,  p-value: < 2.2e-16
```

Figure 20: Compare Statistical Significance based on P-Value

4.2.2 Identify Best Exam Method

In this step, it is shown that the total_pass values are changed according to the max_fail_marks, max_avg_wrong and max_exam_level values. If max_fail_marks, max_avg_wrong and max_exam_level values are increased then total_pass values also increase. But higher total_pass doesn't mean better model. Because if total_pass is almost 5000 between total 5000 exam then it means that this exam method is to much easy. On the other hand is total_pass is under 2000 then it means this exam method is extremely hard.

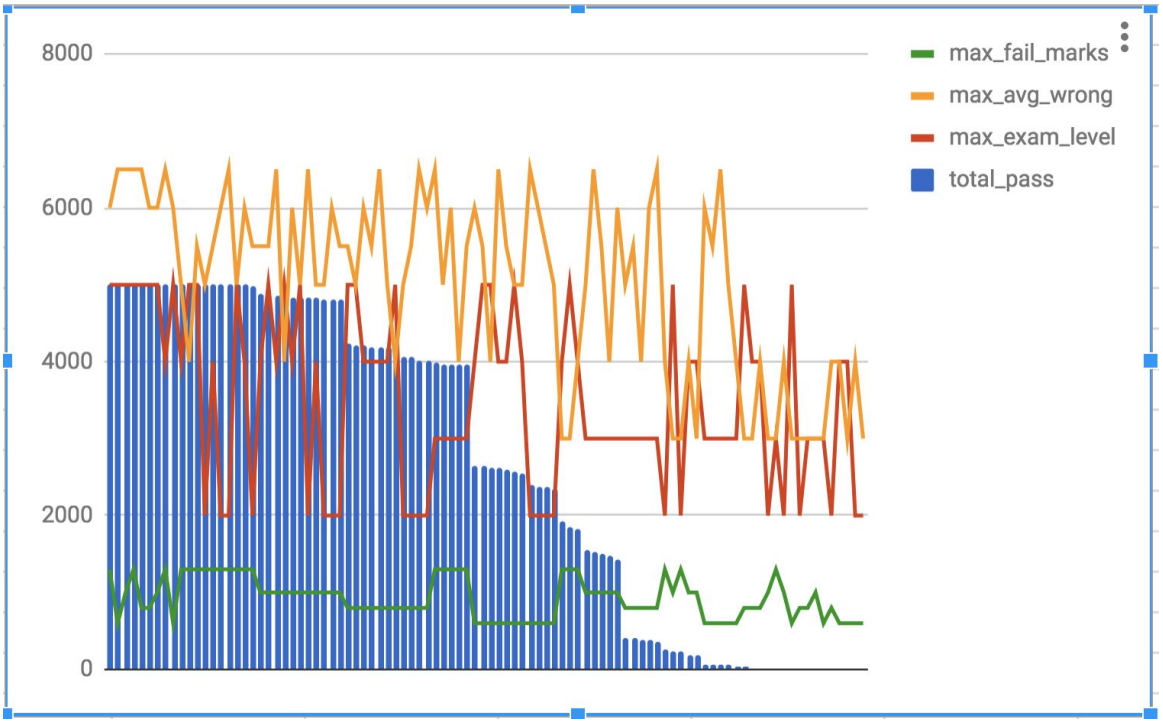

Figure 21: Comparison between total_pass and other values

So, system should select an acceptable range for total_pass like 3000 to 45000. Though investigation is done on sample data, it may change for real life data.

total_pass	max_exam_level	min_avg_result	min_pass_marks
3955	3	45	7
3958	3	60	7
3968	3	40	7
3974	3	50	7
3998	3	35	7
4012	2	40	12
4016	2	35	12
4058	2	45	12
4070	2	50	12
4127	5	60	12
4190	4	35	12
4190	4	50	12
4191	4	45	12
4200	4	40	12
4218	5	50	12
4226	5	45	12

Table 2: Best Exam Configuration for Better Learning Process.

4.2.3 Ensure Learning Outcome

In this step, it has been selected all final_level_marks for each student's exam with module_id, lesson_id and exam_config_id. Here final_level_marks is dependent variable and module_id, lesson_id and exam_config_id are independent variables. So, system has trained a Multiple Linear Regression Model using these data and tested on some sample value. It has found that this model can predict what should be the final_level_marks for a student if he/she sit for same kind of exam like (module, lesson and exam_config). So, in real life, if system get final_level_marks of a student,

is far different than predicted final_level_marks by system for same exam scenario then system can predict that this student learning outcome is not acceptable.

Here are some comparison between test dataset and predicted dataset.

A	B	C	D	E	F
			Test Data	**Predicted Data**	
module_id	**lesson_id**	**exam_config_id**	**final_level_marks**	**final_level_marks**	**Difference**
5	20	4	5	4	1
4	39	21	11	14	3
2	12	72	12	14	2
3	43	95	17	19	2
5	35	86	15	14	1
4	24	81	14	13	1
3	18	85	12	14	2
2	42	11	13	14	1
5	25	92	18	19	1
3	8	3	4	3	1
5	5	65	14	13	1
5	15	78	12	14	2
4	19	16	18	14	4
3	13	14	14	13	1
4	24	68	13	14	1
4	24	28	1	2	1
4	14	2	8	3	5
3	3	56	7	6	1

Table 3: Comparison Between Test Dataset and Predicted Dataset

It is shown that the difference between Test Data and Predicted Data is very low. So, it prove that this proposed system is well trained to ensure student learning outcome.

Chapter 5

5. Conclusion and Future Work Plan

The proposed technique will ensure a positive learning outcome for the students who are learning any module using web based e-learning system. This structure can be used as dynamic system for different course module. Here, it has been presented Flowchart, Algorithm, Use Case diagram, Sequence diagram, Class diagram, Database diagram and Machine Learning Algorithm to develop the system. My future plan is to develop this system and test in a real environment. After testing the system, I can improve the system. After completing this e-learning system, this will be used in different educational institution to reduce traditional face-to-face class teaching technique. Also I can introduce dynamic exam method for each module and lesson and Ontology to generate questions automatically. There is also a scope to evaluate and identifying best exam method by Backward Elimination approach of Multiple Linear Regression Model in R.

Duration of this research project is twelve months which has started 25th of January 2010. At first semester a lots of literature review are made on my chosen area. Some new ideas have been generated and designed the system template. At second semester (1st July 2010 to 5th September 2010) Flowchart, Algorithm, Use Case Diagram, Sequence Diagram, Class Diagram and Database Diagram of this system are designed. At last semester (Dec 2017 to Feb 2018) this thesis is revised and improved the process design, generated sample data, examined the sample data by Machine Learning Algorithm like Multiple Linear Regression Model.

References

[1]. Tavangarian D., Leypold M., Nölting K., Röser M.,(2004). “Is e-Learning the Solution for Individual Learning?” Journal of e-Learning, 2004.

[2]. Bryn Holmes, John Gardner, “E-learning: concepts and practice” ,2006

[3]. http://en.wikipedia.org/wiki/E-learning, Visited last at 15-June-2010

[4]. Graziadei, W.D., 1993. Virtual Instructional Classroom Environment in Science (VICES) in Research, Education, Service & Teaching (REST)

[5]. Graziadei, W. D., Mosaic & Gopher/Lynx, 1993.

[6]. William D. Graziadei, Sharon Gallagher,Ronald N. Brown,Joseph Sasiadek, “Building Asynchronous and Synchronous Teaching-Learning Environments: Exploring a Course/Classroom Management System Solution”

[7]. Graziadei, W. D., et al., “Building Asynchronous and Synchronous Teaching-Learning Environments: Exploring a Course/Classroom Management System Solution”. 1997.

[8]. David Birchall and Damian Woolfall, “Corporate e-learning: Delivering business benefits” , 2003

[9]. EC (2000). Communication from the Commission: E-Learning - Designing "Tejas at Niit" tomorrow’s education. Brussels: European Commission

[10]. Nagy, A. (2005). The Impact of E-Learning, in: Bruck, P.A.; Buchholz, A.; Karssen, Z.; Zerfass, A. (Eds). E-Content: Technologies and Perspectives for the European Market. Berlin: Springer-Verlag, pp.79-96

[11]. Allen, I. E. and Seaman, J. (2008) Staying the Course: Online Education in the United States, 2008 Needham MA: Sloan Consortium

[12]. Allen, I.E. and Seaman, J. (2003) Sizing the Opportunity: The Quality and Extent of Online Education in the United States, 2002 and 2003 Wellesley, MA: The Sloan Consortium

[13]. Ambient Insight Research (2009) US Self-paced e-Learning Market Monroe WA: Ambient Insight Research

[14]. Hebert, D. G. (2007). "Five Challenges and Solutions in Online Music Teacher Education". Research and Issues in Music Education 5 (1). http://www.stthomas.edu/rimeonline/vol5/hebert.htm.

[15]. Means, B.; Toyama, Y.; Murphy, R.; Bakia, M.; Jones, K. (2009), "Evaluation of Evidence-Based Practices in Online Learning: A Meta-Analysis and Review of Online Learning Studies", retrieved 20 August 2009.

[16]. Mungania, P. (2003). The seven e-learning barriers facing employees. Retrieved November 18, 2004, from http://www.masie.com/researchgrants/2003/Mungania_Exec_Summary.pdf

[17]. Ellis, R. (2004, November). Learning Circuits e-learning trends 2004. Learning Circuits.

[18]. Bonk, C. J. (2004). The perfect e-storm: Emerging technologies, enhanced pedagogy, enormous learner demand, and erased budgets. London, UK: The Observatory on Borderless Higher Education.

[19]. Kim, K. J., Bonk, C. J., & Zeng, T. (2005, June). "Surveying the future of workplace e-learning: The rise of blending, interactivity, and authentic learning." E-Learn Magazine.

[20]. Bentley, Y, Shegunshi, A and Scannell, M (2010) "Evaluating the Impact of Distance Learning Support Systems on the Learning Experience of MBA Students in a Global Context" Electronic Journal of e-Learning Volume 8 Issue 2 2010, (pp51 - 62), available online at www.ejel.org

[23]. Chih-Ming Chen, "Personalized E-learning system with self-regulated learning assisted mechanisms for promoting learning performance", 2008.

[26]. Chun-Hsiung Lee *, Gwo-Guang Lee, Yungho Leu, “Application of automatically constructed concept map of learning to conceptual diagnosis of e-learning”. 2007

[28]. Nakayama, M, Yamamoto, H and Santiago, R. (2010) “The Role of Essay Tests Assessment in e-Learning: A Japanese Case Study” Electronic Journal of e-Learning Volume 8 Issue 2 2010, (pp173 - 178), available online at www.ejel.org

[29]. Lluís Vicent and Xavier Gumara, chapter 8 “Data Mining for Instructional Design, Learning and Assessment” , ” E-learning experiences and future”, Book edited by: Safeeullah Soomro, ISBN: 978-953-307-092-6, Publisher: INTECH, Publishing date: April 2010

[30]. Pahl, Claus, " Data mining technology for the evaluation of learning content interaction" International Journal on E-Learning (October 1, 2004)

[31]. Donald Clark “Games and e-learning ” http://www.caspianlearning.co.uk/Whtp_caspian_games_1.1.pdf, Visited last at 15-June-2010

[34]. http://www.suddenlysmart.com/smartbuilder.htm, Visited last at 15-June-2010

[35]. http://www.ispringsolutions.com/, Visited last at 15-June-2010

[36]. http://www.easyauthoring.com/lms.shtml, Visited last at 15-June-2010

[37]. http://www.engagesoftware.com/Products/Modules/Engage_Campus.aspx, Visited last at 15-June-2010

[38]. http://www.madewithmercury.com/, Visited last at 15-June-2010

[39]. http://www.skillcast.com/?gclid=CK3H1qm5oKICFQE8lAodClHswg, Visited last at 15-June-2010

[40]. http://moodle.org/, Visited last at 15-June-2010

[43] P. Giudici, Applied Data Mining Statistical Methods for Business and Industry, Wiley & Sons, 2003.

[44]. U. Fayyad, G. Piatetsky-Shapiro, P. Smyth, Data mining and knowledge discovery in databases, Commun. ACM 39 (1996) 24–26.

[45]. Riccardo Bellazzi, Blaz Zupan, "Predictive data mining in clinical medicine: Current issues and guidelines" international journal of medical informatics 7 7 (2 0 0 8) 81–97

[46]. https://en.wikipedia.org/wiki/Machine_learning Visited last at 26-Feb-2018

[47] http://www.statisticssolutions.com/what-is-multiple-linear-regression/ Visited last at 26-Feb-2018

Appendix A

a. Server Configuration

We have used a MacBook Pro for this thesis work with following configuration

Model Name: MacBook Pro (13-inch, 2017)

Processor Name: Intel Core i5

Processor Speed: 2.3 GHz

Number of Processors: 1

Total Number of Cores: 2

L2 Cache (per Core): 256 KB

L3 Cache: 4 MB

Memory: 8 GB

OS: macOS High Sierra

b. Database Preparation Scripts

Here are all scripts for database preparation.

```
CREATE SCHEMA IF NOT EXISTS elearndb AUTHORIZATION postgres;

CREATE USER eladmin WITH PASSWORD 'shuvo123'; --  CONNECTION LIMIT
UNLIMITED;

GRANT ALL PRIVILEGES ON SCHEMA elearndb TO eladmin;
GRANT ALL PRIVILEGES ON ALL TABLES IN SCHEMA elearndb TO eladmin;
GRANT EXECUTE ON ALL FUNCTIONS IN SCHEMA elearndb TO eladmin;

GRANT ALL PRIVILEGES ON DATABASE postgres TO eladmin;

CREATE TABLE "elearndb"."student" (
  "id" int4 NOT NULL,
  "student_name" varchar(128) COLLATE "pg_catalog"."default" NOT NULL,
  "roll_number" int4 NOT NULL,
  "create_date" timestamp(0),
  "update_date" timestamp(0)
)
;
ALTER TABLE "elearndb"."student" OWNER TO "eladmin";
ALTER TABLE "elearndb"."student" ADD CONSTRAINT "student_pkey" PRIMARY KEY
```

```
("id");

CREATE TABLE "elearndb"."exam_level" (
  "id" int8 NOT NULL,
  "student_exam_id" int4,
  "exam_level_number" int4,
  "question_group_id" int4,
  "question_id" int4,
  "answer" int4,
  "result" int4,
  "total_marks" int4,
  "create_date" timestamp(0),
  "update_date" timestamp(0)
)
;
ALTER TABLE "elearndb"."exam_level" OWNER TO "eladmin";
ALTER TABLE "elearndb"."exam_level" ADD CONSTRAINT "uk_se_lvl" UNIQUE
("student_exam_id", "exam_level_number", "question_group_id");
ALTER TABLE "elearndb"."exam_level" ADD CONSTRAINT "exam_level_pkey" PRIMARY
KEY ("id");
ALTER TABLE "elearndb"."exam_level" ADD CONSTRAINT "fk_el_q" FOREIGN KEY
("question_id") REFERENCES "elearndb"."question" ("id") ON DELETE NO ACTION
ON UPDATE NO ACTION;
ALTER TABLE "elearndb"."exam_level" ADD CONSTRAINT "fk_el_qg" FOREIGN KEY
("question_group_id") REFERENCES "elearndb"."question_group" ("id") ON
DELETE NO ACTION ON UPDATE NO ACTION;
ALTER TABLE "elearndb"."exam_level" ADD CONSTRAINT "fk_el_se" FOREIGN KEY
("student_exam_id") REFERENCES "elearndb"."student_exam" ("id") ON DELETE NO
ACTION ON UPDATE NO ACTION;

CREATE TABLE "elearndb"."exam_config" (
  "id" int4 NOT NULL,
  "min_exam_level" int4,
  "max_exam_level" int4,
  "min_avg_result" numeric(32,2),
  "create_date" timestamp(0),
  "update_date" timestamp(0),
  "min_pass_marks" int4
)
;
ALTER TABLE "elearndb"."exam_config" OWNER TO "eladmin";
ALTER TABLE "elearndb"."exam_config" ADD CONSTRAINT "exam_config_pkey"
PRIMARY KEY ("id");

CREATE TABLE "elearndb"."lesson" (
  "id" int4 NOT NULL,
  "lesson_name" varchar(128) COLLATE "pg_catalog"."default",
  "module_id" int4,
  "lesson_content_1" varchar(256) COLLATE "pg_catalog"."default",
  "create_date" timestamp(0),
```

```
  "update_date" timestamp(0),
  "lesson_content_2" varchar(256) COLLATE "pg_catalog"."default",
  "lesson_content_3" varchar(256) COLLATE "pg_catalog"."default"
)
;
ALTER TABLE "elearndb"."lesson" OWNER TO "eladmin";
ALTER TABLE "elearndb"."lesson" ADD CONSTRAINT "lesson_pkey" PRIMARY KEY
("id");
ALTER TABLE "elearndb"."lesson" ADD CONSTRAINT "fk_lesson_module" FOREIGN
KEY ("module_id") REFERENCES "elearndb"."module" ("id") ON DELETE NO ACTION
ON UPDATE NO ACTION;

CREATE TABLE "elearndb"."module" (
  "id" int4 NOT NULL,
  "module_name" varchar(128) COLLATE "pg_catalog"."default" NOT NULL,
  "create_date" timestamp(0),
  "update_date" timestamp(0)
)
;
ALTER TABLE "elearndb"."module" OWNER TO "eladmin";
ALTER TABLE "elearndb"."module" ADD CONSTRAINT "module_pkey" PRIMARY KEY
("id");

CREATE TABLE "elearndb"."question_group" (
  "id" int4 NOT NULL,
  "question_text" varchar(256) COLLATE "pg_catalog"."default",
  "question_set_id" int4,
  "create_date" timestamp(0),
  "update_date" timestamp(0)
)
;
ALTER TABLE "elearndb"."question_group" OWNER TO "eladmin";
ALTER TABLE "elearndb"."question_group" ADD CONSTRAINT "question_group_pkey"
PRIMARY KEY ("id");
ALTER TABLE "elearndb"."question_group" ADD CONSTRAINT "fk_qg_qs" FOREIGN
KEY ("question_set_id") REFERENCES "elearndb"."question_set" ("id") ON
DELETE NO ACTION ON UPDATE NO ACTION;

CREATE TABLE "elearndb"."question_set" (
  "id" int4 NOT NULL,
  "set_name" varchar(128) COLLATE "pg_catalog"."default",
  "lesson_id" int4,
  "create_date" timestamp(0),
  "update_date" timestamp(0)
)
;
ALTER TABLE "elearndb"."question_set" OWNER TO "eladmin";
ALTER TABLE "elearndb"."question_set" ADD CONSTRAINT "question_set_pkey"
PRIMARY KEY ("id");
ALTER TABLE "elearndb"."question_set" ADD CONSTRAINT "fk_set_lesson" FOREIGN
```

```
KEY ("lesson_id") REFERENCES "elearndb"."lesson" ("id") ON DELETE NO ACTION
ON UPDATE NO ACTION;

CREATE TABLE "elearndb"."question" (
  "id" int4 NOT NULL,
  "question_text" varchar(256) COLLATE "pg_catalog"."default",
  "question_group_id" int4,
  "ans_option_a" varchar(256) COLLATE "pg_catalog"."default",
  "ans_option_b" varchar(256) COLLATE "pg_catalog"."default",
  "ans_option_c" varchar(256) COLLATE "pg_catalog"."default",
  "ans_option_d" varchar(256) COLLATE "pg_catalog"."default",
  "correct_answer" int4,
  "create_date" timestamp(0),
  "update_date" timestamp(0),
  "exam_level_number" int4
)
;
ALTER TABLE "elearndb"."question" OWNER TO "eladmin";
ALTER TABLE "elearndb"."question" ADD CONSTRAINT "question_pkey" PRIMARY KEY
("id");
ALTER TABLE "elearndb"."question" ADD CONSTRAINT "fk_q_qg" FOREIGN KEY
("question_group_id") REFERENCES "elearndb"."question_group" ("id") ON
DELETE NO ACTION ON UPDATE NO ACTION;

CREATE TABLE "elearndb"."student_exam_outcome" (
  "id" int4 NOT NULL,
  "student_exam_id" int4,
  "final_level_marks" int4,
  "exam_result" int4,
  "create_date" timestamp(0)
)
;
ALTER TABLE "elearndb"."student_exam_outcome" OWNER TO "eladmin";
ALTER TABLE "elearndb"."student_exam_outcome" ADD CONSTRAINT
"student_exam_outcome_pkey" PRIMARY KEY ("id");
ALTER TABLE "elearndb"."student_exam_outcome" ADD CONSTRAINT "fk_exo_se"
FOREIGN KEY ("student_exam_id") REFERENCES "elearndb"."student_exam" ("id")
ON DELETE NO ACTION ON UPDATE NO ACTION;

CREATE TABLE "elearndb"."student_exam" (
  "id" int4 NOT NULL,
  "student_id" int4 NOT NULL,
  "question_set_id" int4 NOT NULL,
  "exam_config_id" int4 NOT NULL,
  "create_date" timestamp(0)
)
;
ALTER TABLE "elearndb"."student_exam" OWNER TO "eladmin";
ALTER TABLE "elearndb"."student_exam" ADD CONSTRAINT "uk_st_q_ec" UNIQUE
("student_id", "question_set_id", "exam_config_id");
```

```
ALTER TABLE "elearndb"."student_exam" ADD CONSTRAINT "student_exam_pkey"
PRIMARY KEY ("id");
ALTER TABLE "elearndb"."student_exam" ADD CONSTRAINT "fk_se_ec" FOREIGN KEY
("exam_config_id") REFERENCES "elearndb"."exam_config" ("id") ON DELETE NO
ACTION ON UPDATE NO ACTION;
ALTER TABLE "elearndb"."student_exam" ADD CONSTRAINT "fk_se_qs" FOREIGN KEY
("question_set_id") REFERENCES "elearndb"."question_set" ("id") ON DELETE NO
ACTION ON UPDATE NO ACTION;
ALTER TABLE "elearndb"."student_exam" ADD CONSTRAINT "fk_se_st" FOREIGN KEY
("student_id") REFERENCES "elearndb"."student" ("id") ON DELETE NO ACTION ON
UPDATE NO ACTION;
```

c. Sample Data Generation

Here are all scripts to generate sample data

```
-- student
INSERT INTO elearndb.student
SELECT generate_series(1,100) AS id
, 'Student_Name_' || LPAD(generate_series(1,100)::text, 3, '0')
, generate_series(1,100)
, now()
, now()
;

-- module
INSERT INTO elearndb.module
SELECT generate_series(1,5) AS id
, 'Module_Name_' || LPAD(generate_series(1,5)::text, 3, '0')
, now()
, now()
;

-- lesson
INSERT INTO elearndb.lesson
SELECT row_number() over () AS id
, 'Lesson_Name_' || LPAD(row_number()over()::text, 3, '0')
, m.id
, 'Lesson_Content_ABC_' || LPAD(row_number()over()::text, 3, '0')
, now()
, now()
, 'Lesson_Content_CDR_' || LPAD(row_number()over()::text, 3, '0')
, 'Lesson_Content_OLE_' || LPAD(row_number()over()::text, 3, '0')
FROM generate_series(1,10) s, elearndb.module m
;

-- question_set
INSERT INTO elearndb.question_set
SELECT l.id AS id
```

```
, 'Set_Name_' || LPAD(l.id::text, 3, '0')
, l.id
, now()
, now()
FROM elearndb.lesson l
;

-- question_group
INSERT INTO elearndb.question_group
SELECT row_number()over() AS id
, 'Question Group Text ' || LPAD(row_number()over()::text, 3, '0')
, qs.id
, now()
, now()
FROM generate_series(1,20) s, elearndb.question_set qs
;

-- question
INSERT INTO elearndb.question
SELECT row_number()over() AS id
, 'Question Text ' || LPAD(row_number()over()::text, 3, '0')
, qg.id
, 'Answer Option A ' || LPAD(row_number()over()::text, 3, '0')
, 'Answer Option B ' || LPAD(row_number()over()::text, 3, '0')
, 'Answer Option C ' || LPAD(row_number()over()::text, 3, '0')
, 'Answer Option D ' || LPAD(row_number()over()::text, 3, '0')
, floor(random() * (4-1+1) + 1)::int
, now()
, now()
, q
FROM generate_series(1,5) q, elearndb.question_group qg
;

-- exam_config
INSERT INTO elearndb.exam_config
SELECT row_number()over() AS id
, 2 min_exam_level
, mxel
, s.miar
, now()
, now()
, h.mipm
FROM generate_series(2,5) mxel
, (VALUES (70),(60),(50),(45),(40),(35)) s(miar)
, (VALUES (14),(12),(10),(7)) h(mipm)
;

-- student_exam
INSERT INTO elearndb.student_exam
SELECT row_number()over() AS id
```

```
, s.id
, qs.id
, ec.id
, now()
FROM elearndb.student s, elearndb.question_set qs, elearndb.exam_config ec
;

-- exam_level
INSERT INTO elearndb.exam_level
SELECT (LPAD(tb.student_id::text, 3, '0')||LPAD(row_number()over()::text, 8,
'0'))::int8 id
, tb.student_exam_id
, tb.exam_level_number
, tb.question_group_id
, tb.question_id
, tb.answer
, (CASE WHEN tb.answer = tb.correct_answer THEN 1 ELSE 0 END) result
, SUM(CASE WHEN tb.answer = tb.correct_answer THEN 1 ELSE 0 END)
            OVER (PARTITION BY tb.student_exam_id, tb.question_group_id
            ORDER BY tb.student_exam_id, tb.question_group_id,
tb.exam_level_number) total_marks
, now() create_date
, now() update_date
FROM (
            SELECT se.id student_exam_id
            , s.id student_id
            , q.exam_level_number
            , qg.id question_group_id
            , q.id question_id
            , (ARRAY[q.correct_answer,1,2,3,4])[floor(random()*(6 -
q.exam_level_number))+1] answer
            , q.correct_answer
            FROM elearndb.student_exam se
            INNER JOIN elearndb.student s ON s.id = se.student_id
            INNER JOIN elearndb.exam_config ec ON ec.id = se.exam_config_id
            INNER JOIN elearndb.question_set qs ON qs.id =
se.question_set_id
            INNER JOIN elearndb.lesson l ON l.id = qs.lesson_id
            INNER JOIN elearndb.question_group qg ON qg.question_set_id =
qs.id
            INNER JOIN elearndb.question q ON q.question_group_id = qg.id
                        AND q.exam_level_number <= ec.max_exam_level
            ORDER BY se.id, qg.id, q.exam_level_number
            ) tb
ORDER BY tb.student_exam_id, tb.question_group_id, tb.exam_level_number
;

-- student_exam_outcome
INSERT INTO elearndb.student_exam_outcome
SELECT row_number()over() id
```

```
, tb.student_exam_id
, SUM(tb.marks) final_level_marks
, (CASE WHEN SUM(tb.marks) >= tb.min_pass_marks THEN 1 ELSE 0 END)
exam_result
, now()
FROM (
              SELECT el.student_exam_id
              , el.question_group_id
              , ec.max_exam_level
              , ec.min_avg_result
              , ec.min_pass_marks
              , (CASE WHEN floor((max(el.total_marks)::float /
max(el.exam_level_number)::float) * 100) >= ec.min_avg_result
                           THEN 1 ELSE 0 END) marks
              FROM elearndb.exam_level el
              INNER JOIN elearndb.student_exam se ON se.id =
el.student_exam_id
              INNER JOIN elearndb.exam_config ec ON ec.id = se.exam_config_id
              GROUP BY el.student_exam_id
              , el.question_group_id
              , ec.max_exam_level
              , ec.min_avg_result
              , ec.min_pass_marks
              ) tb
GROUP BY tb.student_exam_id
, tb.min_pass_marks
;
```

d. R-Script for Exam Config Investigation

SQL Script to select total_pass count for each exam_config

```
SELECT ec.max_exam_level max_exam_level
, (100 - ec.min_avg_result) max_avg_wrong
, (20 - ec.min_pass_marks) max_fail_marks
, SUM(o.exam_result) total_pass
FROM elearndb.student_exam_outcome o
INNER JOIN elearndb.student_exam se ON se.id = o.student_exam_id
INNER JOIN elearndb.exam_config ec ON ec.id = se.exam_config_id
GROUP BY 1, 2, 3
ORDER BY 1, 2, 3
;
```

R-Script to ensure exam config variables statistical significance.

```
# Import data
dataset = read.csv('config_2.csv')

# Splitting the dataset into the training set and test set
set.seed(123)
split = sample.split(dataset$total_pass, SplitRatio = 0.95)
training_set = subset(dataset, split == TRUE)
test_set = subset(dataset, split == FALSE)

# Fitting MLR to the training set
regressor = lm(formula = total_pass ~ .,
               data = training_set)
summary(regressor)
```

e. SQL-Script to define best exam config

SQL Script to select top most exam config

```
SELECT SUM(o.exam_result) total_pass
, ec.max_exam_level
, ec.min_avg_result
, ec.min_pass_marks
FROM elearndb.student_exam_outcome o
INNER JOIN elearndb.student_exam se ON se.id = o.student_exam_id
INNER JOIN elearndb.exam_config ec ON ec.id = se.exam_config_id
GROUP BY ec.max_exam_level, ec.min_avg_result, ec.min_pass_marks
HAVING SUM(o.exam_result) >= 3000 AND SUM(o.exam_result) < 4500
ORDER BY 1, 2, 3, 4
;
```

f. R-Script for Student Learning Outcome Investigation

SQL Script to select module_id, lesson_id, exam_config_id, final_level_marks

```
SELECT m.id module_id
, l.id lesson_id
, ec.id exam_config_id
, o.final_level_marks
FROM elearndb.student_exam_outcome o
INNER JOIN elearndb.student_exam se ON se.id = o.student_exam_id
INNER JOIN elearndb.exam_config ec ON ec.id = se.exam_config_id
INNER JOIN elearndb.question_set qs ON qs.id = se.question_set_id
INNER JOIN elearndb.lesson l ON l.id = qs.lesson_id
INNER JOIN elearndb."module" m ON m.id = l.module_id
;
```

R-Script to ensure Student Learning Outcome.

```
# Import data
dataset = read.csv('exam.csv')

# Encoding categorical data
dataset$module_id <- as.factor(dataset$module_id)
dataset$lesson_id <- as.factor(dataset$lesson_id)
dataset$exam_config_id <- as.factor(dataset$exam_config_id)

# Splitting the dataset into the training set and test set
set.seed(123)
split = sample.split(dataset$final_level_marks, SplitRatio = 0.95)
training_set = subset(dataset, split == TRUE)
test_set = subset(dataset, split == FALSE)

# Fitting MLR to the training set
regressor = lm(formula = final_level_marks ~ .,
               data = training_set)

summary(regressor)

# Predicting the Test set result
y_pred = predict(regressor, newdata = test_set)
y_pred
```

****** ******* ******* THE END ******* ******* ******