Deanship of Graduate Studies
Al-Quds University

A smart Framework for identifying the training needs of Palestinian Teachers using Data Mining Techniques

Nidaa "Mohammad Ghazi" Hasan Farhaneh

M.Sc. Thesis

Jerusalem – Palestine

1442/ 2020

A smart Framework for identifying the training needs of Palestinian Teachers using Data Mining Techniques

Prepared By:
Nidaa "Mohammad Ghazi" Hasan Farhaneh

B.Sc.: Computer Science – Al-Quds University – Palestine

Supervisor: Dr. Rashid Jayousi

A Thesis Submitted in Partial Fulfillment of Requirements for the Degree of Master in Computer Science /Department of Computer Science/ Faculty of Graduate Studies /Al-Quds University.

1442/ 2020

Al-Quds University
Deanship of Graduate Studies
Computer Science Department

Thesis Approval

A smart Framework for identifying the training needs of Palestinian Teachers using Data Mining Techniques

Prepared By: Nidaa "Mohammad Ghazi" Hasan Farhaneh

Registration No: 21710109

Supervisor: Dr. Rashid Jayousi

Master thesis submitted and accepted Date: 15/ 01/2021

The names and signatures of the examining committee members are as follows:

1-	**Head of Committee: Dr. Rashid Jayousi**	**Signature: ……..**
2-	**Internal Examiner: Dr. Radwan Qasrawi**	**Signature: ……**
3-	**External Examiner: Dr. Samah Ereqat**	**Signature: ……**

Jerusalem – Palestine

1442/ 2020

Declaration

I certify that this thesis, submitted for the degree of Master, is the result of my research, except where otherwise acknowledged, and that this thesis (or any part of the same) has not been submitted for any higher degree or to any other university or institution .

Signature: 

Nidaa "Mohammad Ghazi" Hasan Farhaneh

Date: 27 / 12 /2020

Dedication

إلى من جاء بالحق معلماً.. إلى النبي محمداً "صلى الله عليه وسلم"
إلى من زرع فيّ الصبر والعزيمة .. إلى روح والدي الغالي رحمه الله تعالى
إلى والدتي الحنونة أطال الله في عمرها وأمدها بوافر الصحة والعافية
إلى زوجي وشريك حياتي أكرم أكرمه الله وحفظه لي ذخراً وسنداً
إلى مهج القلب أبنائي وبناتي أمل وعمر وعلي ومالك وماريا رعاهم الله
إلى إخواني وأخواتي الغوالي حفظهم الله جميعاً
أهديكم بحثي هذا

Acknowledgments

I extend my sincere thanks and gratitude to the supervisor of my post; **Dr. Rashid Jayousi**; who did not skimp on his advice, guidance, and extensive knowledge, was also proud of the members of the discussion committee represented by the distinguished Ladies and gentlemen: **Dr. Samah Ereqat** as an external examiner and **Dr. Radwan Qasrawi** as an internal examiner, thanking them for the valuable instructions and guidance they provided to me.

I am also pleased to present a bouquet of thanks and gratitude to the distinguished gentlemen in the Ministry of Education and the Ramallah and Al-Bireh District, who provided me with study data and information on teachers and training courses that were passed, and for the facilities, they provided to me to conduct this study, and to everyone who has encouraged and supported me, I offer the highest signs of thanks and gratitude

List of Contents

List of Appendices

List of Tables

List of Figures:

List of Abbreviation:

Abbreviation	Meaning
API	Application Program Interface
RestAPI	Representational state transfer
FE	Front End
BE	Back End
ADAM	Algorithm Development And Mining
ETL	Extract, Transform, Load
ML	Machine Learning
NumPy	Numeric Python
Jhipster	java hipster
FN	False Negative
TP	True Positive
FP	False Positive
TN	True Negative
AI	Artificial Intelligent

Abstract

This study aims to establish a Smart Framework that works as a system based on searching for data related to teachers of the Arabic Language, Science, Mathematics and English Language studies within specific criteria and variables, to determine the training needs by predictions through framework using data mining AI algorithm, which is necessary to enhance teacher's performance in the classroom environment. The researcher took a sample of public school teachers in the academic year 2014/2015 AD and determined the following variables: academic qualification, specialization, stages that he teaches, an annual calendar, training courses, and their number of hours in content areas and teaching methods.

The search was conducted using Angular as a JavaScript tool for the client-side (FE) in the framework and the spring boot as a server-side (BE) connected to a database MySQL and Flask Python server which is responsible for machine learning of deep neural network (DNN) as (AI) algorithm, all servers connected via (API) in Jhipster environment. Results of applying the prototype showed high effectiveness in identifying training needs to improve teachers' performance in an organized and interactive way.

The researcher recommends the Ministry of Education to use this intelligent technology tool for its effectiveness in predicting needed training courses for a teacher to enhance his performance, and also to organize teacher's data in an oriented way.

Chapter 1: Introduction

1.1 Introduction:

The modern world is witnessing an accelerated development in information technology, the amount of data reaching different institutions is increasing day by day, it is no longer feasible to store such data without trying to organize and benefit from it, but the traditional statistical methods are unable to control it due to its abundance, complexity and dramatic change. This has led researchers to consider appropriate methods and mechanisms for data mining, whose algorithms can be used to extract useful information that can improve performance and efficiency, such as business, commerce, medical and other fields.

Many studies in the world have been interested in the use of data mining in the field of education, some of them explored the knowledge in universities (Al_Janaei, et al., 2011), and some were aimed at predicting the appropriate specialization for graduate students based on criteria and data related to their success in some subjects (Ajab, et al., 2014), including some research on Influential factors in evaluating teacher performance (Kumar, Saurabh, 2013), and other research related to the pedagogical aspect, which benefits all elements of the educational process; data mining helps both the teacher and the learner to improve their performance, and decision-makers in the educational institution invest in improving the rules governing the educational learning process and improving the human and material resources needed to develop them (Khazim, et al., 2017).

The educational researchers have tended to study the methods and programs that affect the professional development of the teacher and improve his performance, and its relationship to the level of academic achievement of students, and the appropriate techniques to determine training courses appropriate to the needs of professional teachers, based on information provided from the performance evaluation process.

The teacher performance evaluation contributes to improving the educational process because it has a great role in introducing the teacher to important aspects, that help him in developing his skills and improving his performance, by leading him to instructions for appropriate areas of training. (Bichi, 2017)

What distinguishes this study is its interest in focusing on enhancing teacher's performance and not evaluating his performance considering the evaluation as a variable, by establishing a smart technology tool which is a framework, works automatically to run technology for data mining and use of AI algorithm, Deep Neural Network (DNN), through an easy-to-use interface that serves as a link between the user and the system, to produce a smart predictive result related to training needs to develop teacher's performance.

Keywords: Framework; Teacher Performance Assessment; Data Mining.

1.2 Problem Statement:

Academic and pedagogical qualification of teachers is one of the most important priorities of the Palestinian Ministry of Education, and is a prerequisite for the success of the educational process because educational qualification and training are important in improving the quality of education and raising the academic achievement of students in different subjects. As time goes by, the educational institution has a huge data for teachers, containing their specializations, achievements, courses, and training programs that they have joined, aside from this data, a decision is often impose training on teachers without relying on clear criteria showing the need for it, whether if it will enhance teacher's performance or not, and repeatedly for the same content, leading to negative attitudes towards training without progress in his development, as well as the loss of time and exhaustion of effort and money, feasibility the problem of the study is, therefore:

"Design a technology tool which is a Smart Framework to identify teacher's training needs, using AI algorithm to extract knowledge that can be used to explore influential patterns in teacher's performance enhancement, to support the development of teaching and learning."

1.3 Research Questions:

This study seeks to answer the following questions:

- How can Data Mining techniques be used to arrive at influential patterns related to teacher performance enhancement in Palestine?
- What is the effectiveness of establishing the Framework as an intermediary between the user and the data mining system in enhancing the performance of Palestinian teachers and identifying their training needs?

1.4 Objectives of the Study:

This study aims to design a Smart Framework that acts as a smart technology tool based on data mining using an AI algorithm, it starts from an interface that acts as a link between the user and the system and is used by decision-makers in the Palestinian Ministry of Education, to enhance teacher's performance by smart prediction of the system for training needs, to develop his professional performance at the highest quality and lowest cost.

1.5 Research Significance:

The need for an appropriate smart system based on Data Mining techniques is increasing, with a rapid increase in the amount of data for Palestinian teachers, to arrive at patterns that demonstrate the correlation between several factors related to teacher performance enhancement and training needs. The importance of this study is that it is one of the first studies that seek to provide a smart framework technology tool to take advantage of the large and scattered data for teachers, configure and analyze it, using data mining AI algorithms, to reach these patterns, and provide the ability to explore the training needs of teachers to help make the right decisions when assigning courses and training programs that the teacher should attend.

1.6 Research Contribution:

The study distinguished in the following contributions:

1. Some studies were mainly aimed at evaluating the teacher's performance through multiple models and methods as in (Hemaid, El-Halees, 2015) study, while this study

considered the teacher's evaluation as one of the criteria upon which to determine his training needs to improve his performance.

2. Development of a procedural application tool (framework) to identify the appropriate training courses necessary to improve the performance of the teacher, while many studies such as (Hervie, Winful, 2018) were limited to highlighting the importance of educational training in improving the performance of the teacher.

3. Given the importance of school education and its role in the renaissance and development of society in all aspects of life; this study directed its tools to school teachers who are the basis of the educational-learning process and the center of strength in it. While we find that previous studies in this context have been directed to the university academic community as (Oancea et al., 2013).

1.7 Research Methodology:

The methodology used in this study is applied analytical, by building a Smart Framework based on data mining tools and algorithms, this approach consists of several phases, from building the model by AI algorithms using Deep Neural Network (DNN), and designing the Framework to writing software to explore patterns influencing teacher's performance enhancement and identifying their training needs, and ending with applying a group of teachers to test the Smart prototype of the Framework.

1.8 Research Limitations:

1) The study was limited to male and female teachers in Palestine for the main subjects (Arabic, Science, Mathematics, and English).
2) The study was conducted on the data collected in the academic years 2014/2015, 1009 teacher records were collected, this is what can be provided.
3) The following criteria adopted as possible influences in determining the teacher's training needs: Qualifications, Evaluation, specialization, stages (the teaching stages), the courses a teacher passed in both the field of specialization (content) and teaching methods (pedagogy), number of hours of each course, the sum of hours of all courses taken.
4) The study is limited to the topic of designing a smart framework based on Data Mining techniques and algorithms specifically neural networks in exploring patterns that affect training needs and develop their professional performance.

1.9 Study population & The Sample:

The study population consists of male and female teachers of the Palestinian Ministry of Education for the main subjects (Arabic, Science, Mathematics, and English). The sample of the study consisted of the teachers of the Directorate of Education in Ramallah and Al-Bireh for the same subjects in the academic years 2014-2015, 1009 teacher records were collected. It was selected as a purposive sample to facilitate the researcher to collect data from the place of residence because of the difficulty of moving between cities due to the closures caused by the Covid-19 virus.

Chapter 2: Background and Literature Review

2.1 Background:

2.1.1. Data Mining:

Data mining emerged in the late 1980s, the process of analyzing a large amount of data that is usually large and scattered to find patterns and models with logical relationships that summarize the data in a new way to become useful and understandable to the owner of that data. Prospecting techniques focus on forecasting and future exploration, which helps to make sound decisions at high speed and on time, which makes them different from classical statistical methods, where such data is difficult to control. (Ajab et al., 2014)

Data mining involves multiple methods based on the type of algorithm that is adopted in proportion to the purpose of the mining and the nature of the data. The following is a list of some of these methods:

1. Decision Trees and Rules.
2. Nonlinear Regression and Classification Methods.
3. Example-based Methods.
4. Probabilistic Graphical Dependency Models.
5. Relational Learning Models.

The most common tasks performed by the data mining system are clusters, aggregation, classification, prediction, data series analysis, and external analysis. (Padhy, et al., 2012)

The model in figure (2.1) shows the sequence of operations through which data mining works:

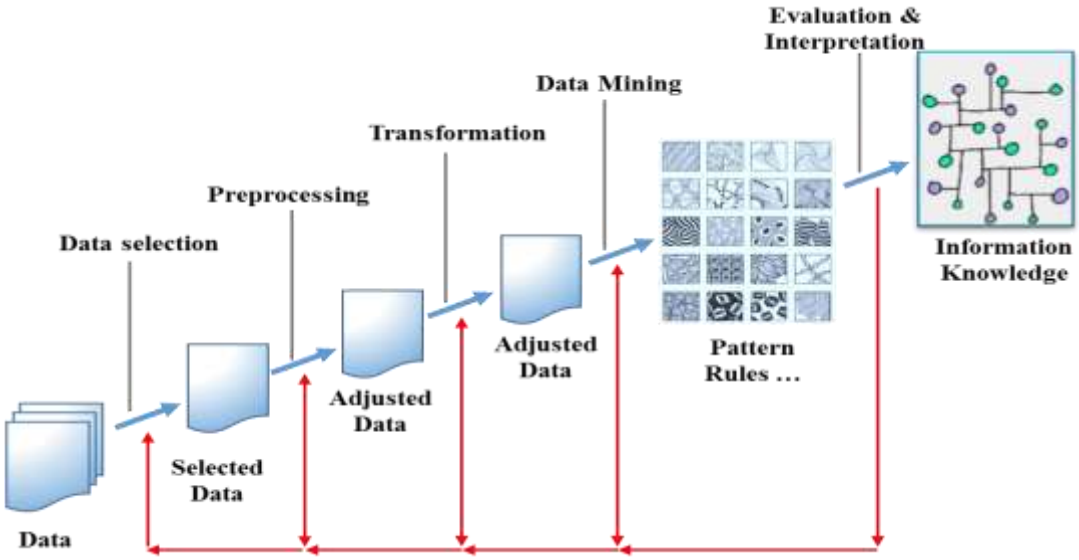


Figure 2.1: Data mining model

Many studies have dealt with data mining because this topic is important in the areas of planning, economics, health, artificial intelligence, education, and other fields. The study carried out by (Edmond, et al., 2004) sought to use data mining to discover patterns of behavior of web users, the results suggest that you can use ways to optimize and customize websites to improve user engagement. The study (Figueiredo, et al., 2005) concerned with the characterization of electricity consumers through a proposed framework based on exploration techniques in user data to find relevant knowledge of how and when consumers use electricity. The quality of the proposed framework has been clarified by applying it to a case-series of real data provided by the Portuguese Electricity Distribution Company, the results being satisfactory given the limitations of the available database.

In the field of trade and business data mining techniques were used extensively, to achieve the greatest possible success in the promotion and gain Customer Relationship Management CRM (Bahari, Elayidom, 2015), as well as in the detection of financial fraud as revealed in the study (Ngai, et al., 2011), and used six algorithms of data mining which are: classification, regression, aggregation, prediction, external detection, and visualization. This is to detect four categories of financial fraud: bank fraud, insurance fraud, securities, and commodity fraud, and other related financial fraud.

2.1.2. Neural Network:

Neural networks are a very old computational tool for language processing, and it is a simplified model for human neurons, and more recently it is considered a network of small computing units, each of which takes a vector for the input values and produces a single output value, which is often deep (it has many layers DNNs) to be an effective tool for solving large-scale problems and providing enough data to automatically learn features.

Some people may confuse the difference between the two methods of functioning of neural networks (NN) and the work of the personal computer (PC), given that there is a close relationship between the two methods in the field of education. A digital computer usually uses one or more central processing units (CPUs), while NN consists of many processing elements (called neurons), and the computer performs sequential operations, while NN performs high-intensity parallel processors, and the computer was designed to address computational problems, While NN has been developed for cognitive task processes, such as understanding language or pattern recognition, these tasks are still out of reach of modern computers. And if neural network learning is understood, teachers will be able to teach their students differently, which is more efficient and effective. Especially with the current acceleration of science and information technology. (Francis T.S., 2002)

A study (Jurafsky, Martin, 2019) presented two classes of neural networks that have been applied in modeling language so that probabilities are assigned to word sequences and predict upcoming words. Many other aspects of neural models have also been introduced, such as recurrent neural networks, decoding models, etc. Deep Neural Networks (DNNs) have also shown remarkable performance in complex machine learning tasks, such as image classification or speech recognition.

In the study (Nurettin, Ilker, 2015) a new approach was applied in the analysis of educational qualitative data, in which a model of neural networks was proposed to achieve qualitative results from a huge amount of categorical data, and in this direction, the Cascading Neural Network Model of Back Propagation (CFBPN) was developed. To analyze categorical data to determine students' attitudes, data were collected using the conceptual comprehension test that includes open questions. The results of this study indicated that the use of the CFBPN model in analyzing data from educational research examining attitudes, behaviors, or beliefs

may help in obtaining more detailed information about the analyzed data and thus about the characteristics of the participants involved.

2.1.3. Smart Framework:

A framework in general is a conceptual structure designed to act as a guide that extends the structure into something useful. In computer systems, a framework is an architecture that contains layers of software that can be built and linked together, and some computer systems frameworks also include programming interfaces for a set of functions within the system. (Rouse, 2015)

The framework tool provides a set of pre-written codes for part of the application, and this concept exists in several programming languages, including Java and PHP, which explains that it helps to create web applications or software components and other uses, and this method has been used to save time on the programmer, Because it establishes several software libraries, each one specialized in a field and does not have to reprogram it, and the person who uses the framework has to use these libraries according to his need and the style in which he works, to develop and innovate in other matters not previously existing. (Nagawi, 2018)

Frameworks contributed to the education field, in an advanced way, the study (Elhoseny, et al., 2018) proposed a framework for smoothly adapt the traditional e-learning systems to be suitable for smart city applications. Some learning platforms such as e-learning systems must change their methodologies for data processing to be smarter, due to the huge data size on the internet.

The study (Dˇzeroski, 2006) dealt with an ambitious task represented in formulating a general framework for data extraction and discussed the requirements that such a framework should fulfill, such as elegant handling with different types of data, models, and patterns. It discussed data mining languages, what they should support, and what It should include it from designing and implementing algorithms in the data extraction process, as well as configuring them in multi-stage knowledge discovery scenarios, ranging from developing

basic concepts to using these concepts to formulate data extraction tasks and design general algorithms to achieve this.

In the study (Haruechaiyasak, et al., 2015), a new framework was proposed that relies on data mining algorithms to build a web page recommendation system, which is considered as a middleware with a user interface that automatically and intelligently creates a list of information that suits an individual's needs, and in which two methods are used to filter Recommended information; The first method is through content-based filtering by analyzing the information content. The second method was through collaborative filtering by referring to user behaviors and other reactions.

The new challenge is to analyze the use of the web to manage and discover media patterns from different types of data stored in structured or unstructured databases for system monitoring and decision making. In the study (WU et al., 2004), a new integrated framework for data storage and data extraction was introduced for managing websites and detecting patterns for analyzing user behavior. This framework was distinguished by combining multi-dimensional web databases to support online analytical processing to improve web services. The model used suggested some statistical indexes and practical solutions to intelligently discover interesting patterns for website optimization and customization. The results show that this integrated data storage and mining model is effective for application in practical web applications such as sports websites, for example.

Big data includes large size, high speed, and a variety of expandable data, whether organized or unorganized, and analyzing and examining this data plays a big role in making better strategic decisions about it, and among these data is what is related to e-learning and its development, as it has recently become It is widespread all over the world, and the study (Nikolovska et al., 2018) presented research in this area, focusing on analyzing the database of the e-learning platform Moodle, to better understand the teaching and learning processes to reach teachers and learners alike to be effective Two major in improving the

educational process. To facilitate this process, tools for processing and analyzing big data were used, and a framework was developed based on that.

The framework was used in areas other than in educational fields. In the study (Chen et al., 2004) a framework was proposed for extracting multimedia data in analyzing video footage of football goals, extracting visual and silent features, cleaning up noise, and excluding irrelevant data. The classification was done in this study using a decision tree model, and the experimental results showed the effectiveness of the proposed work to extract football goals data.

In the field of trade and economics, the researcher (Ladas, et al., 2014) prepared a framework for modeling and analyzing consumer debt based on Data Mining techniques to explore the psychological factors of the consumer as a problem of a complex nature, and the results confirmed the strong influence of psychological factors in modeling consumer debt, and a new approach was proposed. By taking into account the psychological characteristics of consumers.

The data mining framework creates a close relationship with customers and manages the relationship between organizations and customers. Data Mining has gained wide popularity in recent years in many customer relationship management (CRM) applications, and the classification model is a useful technique in this field; in a study (Bahari, Elayidom, 2015), a data mining model was used to predict customer behavior to enhance retention decision-making processes. An effective framework has been proposed, and two classification models, Naïve Bayes and Neural Networks have been studied to show that the accuracy of the neural network is relatively better.

In the field of improving agricultural production, the study (Vagh et al., 2010) proposed a framework for managing and analyzing agricultural data that includes the stages of storing, processing, and updating data periodically, forecasting crops and planting and harvesting strategies, which enables farmers to develop plans according to seasonal and climatic changes and to reach the best Practices in agricultural production.

In industry, the study (Sajadfar, Ma, 2015) presented an informational framework for cost estimation supported by data extraction algorithms, using commonly available manufacturing process data associated with ERP systems. The proposed method combines linear regression techniques and data mining techniques and enhances the strengths of both. And creates a mechanism for discovering cost advantages. The study results indicate that the combined method is flexible and effective for determining costs. When comparing results between experimental prediction and five different data extraction algorithms, it appears that the ANN algorithm is the most accurate in this area.

In the world of crime and forensic research, a study (Al-Janabi, 2011) proposed a framework for analyzing and detecting crime and criminal data using decision tree algorithms for classifying data and Simple K Means for data collection. He prepared this study important in assisting specialists in discovering patterns and trends, setting expectations, finding potential relationships and interpretations, mapping criminal networks, identifying potential suspects, and the classification is based mainly on classifying crimes according to type, location, time, and other characteristics. This is to find relationships between various crimes and criminal aspects that have some previously unknown common characteristics. The results of both classifications and aggregation were used to predict trends and behavior of specific objects (crimes and criminals), and data on both crimes and criminals were collected from police departments data set available on the Internet to create and test the proposed framework, then this data was pre-processed to obtain clean data. And accurate using various pre-processing techniques (cleaning, missing values, and discrepancy removal), crimes and criminals were grouped into groups according to their important traits, and WEKA mining software and Microsoft Excel were used to analyze the given data.

2.1.4. Framework Prototype:

Prototypes can be created to reveal both technical and contextual design considerations, which ultimately determine whether a developed product will be relevant to the stakeholders' needs and expectations. (Cooper, 2019)

Prototypes are fundamental tools used throughout design processes. During early design stages, including problem definition and concept development, prototypes can support stakeholder engagement, which is considered critical for success. The study explored Front-End prototyping strategies for stakeholder engagement. (Rodriguez-Calero, et al., 2020)

2.1.5. Front-End and Back-End Concepts:

The front end can broadly be defined as including background research, needs finding, problem scoping and definition, requirements or attribute elicitation, specifications development, concept generation, and concept development (Ulrich, Eppinger, 2016)

The front end of any website is defined as the screen that enables users to interact with the site, by choosing links and navigating through the pages of the site, and it includes drop-down menus, scroll slides, fonts, and colors, and is prepared using a mixture of programming languages such as HTML, CSS, and JavaScript. While the backend of a website consists of a server, application, and database, the interface developer maintains the technology that powers these components. (Wales, 2014)

The figure below illustrates the concepts:

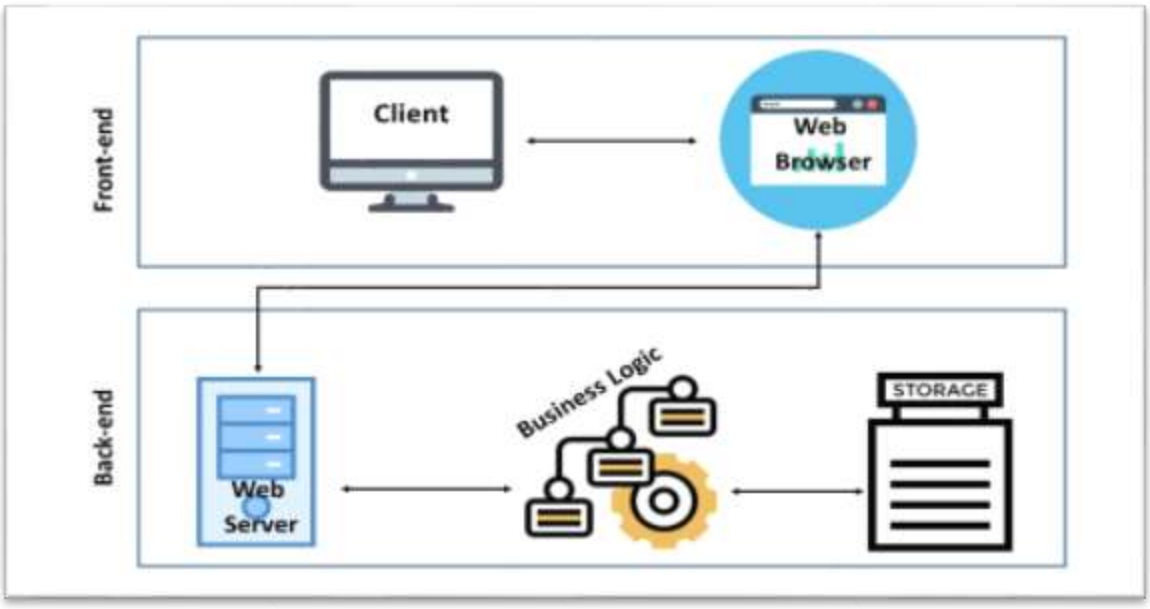


Figure 2.2: Front-end/ Back-end

2.1.6. Teacher Performance Evaluation:

Enhancing teacher's performance contributes to developing the educational process because it has a great role in introducing the teacher to important aspects that help him in developing his skills and improving his performance. To provide notes and guidance and an accurate description of the effectiveness of teaching, its strengths, and areas for its development, followed by instructions for appropriate areas of training, support, and professional development opportunities. (Bichi, 2017).

2.2 Literature Review:

Renewal and development has become a feature of this era, and keeping pace with this era has become an imperative need, and this needs educators and teachers who have a tendency to renew and develop, capable of developing the personality of the learner in all its aspects, which reflects a new vision of the roles of the teacher, beyond just being indoctrinated and transferring information To a facilitator of the teaching and learning process, knowing and understanding the methods of uncovering students' talents, nurturing and developing them, managing the educational process well by practicing planning skills according to students' diverse abilities and preparations, and having the ability to manage class learning and implement it effectively, broadly to see and practice every new thing in the field of education, teaching methods and educational technology in general, and in its academic specialization in particular, creativity consists in designing innovative educational methods, and in initiating to find solutions and proposals for issues or problems facing it, which pushes towards improving learning outcomes of its students and encouraging them to creativity. (Al-Burai, 2012)

Evaluating the teacher's performance is the first and important step in improving his performance, to monitor his strengths and weaknesses. Many studies have spoken in this regard, methods differ from one study to another in the way of evaluation, the study (Olcum, Titrek, 2015) consider the principal of the school plays an important and profound role in the success of the educational process, its development, and its progress with the support, motivation, and support it provides. Therefore, authority was given to use several specific models and

criteria in the evaluation of the teacher. Studies have shown that this is a common interest for both researchers, educational policymakers, and decision-makers the relationship between the decision-making methods of school principals and the levels of job satisfaction of teachers was examined through the Satisfaction Survey (MSQ) and the General Decision-Making Questionnaire (GDMSQ), and the study showed that the rational decision-making method (RDMS) It is the most used, However, because it increases the job satisfaction of the teachers more than other methods, while the managers sometimes resort to using the intuitive method (IDMS) and the method of decision-making (DMS), they rarely use the spontaneous decision-making method (ADMS) because of its effect in reducing Teacher job satisfaction.

Some studies have neglected the role of the school principal as an authorized party in the teacher performance evaluation as a study (Bichi, 2017), which gave this to other parties through a model that addresses four axes: the educational supervisor evaluation, student evaluation, peer evaluation, and self-evaluation. This model provided useful information on teacher effectiveness, which may serve as a basis for decision-making to improve performance and its development.

As for the study (Baby, 2018), it was considered that the student is the only source of information related to the learning environment, and he is the strongest evaluator according to the standards of the quality of educational content, the teaching method, the effectiveness of evaluation and the level of satisfaction. In this study, a model based on data mining technology was included, including a questionnaire that answers on her behalf, one of its results was the clarity of the effectiveness of this model in predicting the performance of teachers. Student feedback was also used in the study (Kumar, Saurabh, 2018) of aspects related to teacher performance metrics such as presentation style, content arrangement, voice modulation, explanatory strength, student attendance, and interaction in the lesson. While the study (Silva, et al., 2017) shows that students' assessment of teachers ’abilities was not accurate by adopting specialization criteria, the number of years of experience, and gender as independent variables in the study, the

reason for this was due to the low size of the study sample, which affected the credibility of the results and the possibility of generalization.

Some studies dealt with the use of data mining techniques and algorithms to predict the quality of teacher performance, potential, and capabilities, to enable decision-makers to make their decisions and understand specific patterns of teacher motivation, growth, and retreat, as in the study (Kumar, Saurabh, 2018), where a Likert questionnaire included Criteria and variables to test many features in the field of teaching methods, and the use of the model that resulted from data mining has resulted in some of the features being effective in predicting teacher performance, and the strongest impact was the ability of the teacher to arrange the content, followed by the student's achievement results. While the study of (Hemaid, El-Halees, 2015) focused on the use of data mining to study factors associated with evaluating teacher performance through specialization criteria, experience, and the number of years of service in the field of education, to determine the necessary courses for teachers in need to improve their performance.

The professional development of the teacher is a major task of the educational institution, represented by the educational supervisor and the principal of the school. The interest in the teacher and its development and qualification is a reflection of the importance of the role he plays in the educational process, and the consequent comprehensive revival in society, in terms of cultural, social, and economic aspects, for the development of civilization depends to achieve A competent teacher with a stable personality able to give and take, improves engagement with the student generation and understands their needs and learning styles, who can guide, guide and facilitate learning for them. (Abu-Samra, 2013)

The educational researchers have tended to study the methods and programs that affect the professional development of the teacher and improve his performance, and its relationship to the level of academic achievement of students, and the appropriate techniques to determine training courses appropriate to the needs of professional teachers, based on information provided from the performance evaluation process. In a study (Mahgoub, Yassir, 2014) a training program was

designed to improve the quality of the educational process by developing the teacher's skills and knowledge in teaching methods and activating teaching aids and other skills. The program was applied to an experimental group of teachers and was used to assess its effect. The questionnaire tool consisted of ten criteria, and it was found that there were statistically significant differences between the experimental group and the control group due to the training program used.

Another study(Hervie, Winful, 2018) support the previous study, where a random sample of teachers was chosen during the service and trained and developed to improve their performance, and the results indicated a strong relationship between the level of teachers 'performance and the level of achievement of their students, and that training programs It was effective in enhancing teachers 'capabilities, and poor teacher performance was due to a lack of frequent training, a lack of learning tools, lack of incentives, weak motivation, and inappropriate supervision.

In the same context, several studies discussed the relationship between the academic achievement of students as a dependent variable and the development of the professional teacher as an independent variable outlined by systematic reviews such as study (Sodiya, Quliyeva, 2019), which showed the importance of dealing with this research topic, which indicates a clear impact of improving teacher performance and professional development in Increase students' academic achievement, using several criteria that were most noticeable in these studies: class management, teaching and learning approach, teacher personal characteristics, cooperative learning facilitation, and ability to control and control within the classroom. While the systematic review in (Leeuwen, Janssen, 2019) study examined the relationship between teacher orientation strategies for students and their interactions during cooperative learning and the results of this interaction, the study concluded that the teacher character plays a critical and important role in supporting class interaction and stimulating learning.

The study (Suarez, Toro, 2018) was distinguished by that it examined the evaluation of the education process itself, through a model based on improving teacher performance and its link to student achievement, school director evaluation, and self-evaluation used questionnaires and diagnostic interviews were conducted with principals and teachers, and took field observations To

verify the available information, the results have shown that improving the performance of the teacher in the school is one of the strongest influences in improving the quality of education, by mastering planning skills and motivating students to learn, and it was found that training evaluation lacks follow-up and continuous evaluation, and that poor teacher preparation leads to low levels performance.

Data mining algorithms were also used in the teacher performance evaluation process as a first step to improving performance. In the study (Hemaid, El-Halees, 2015), specialization, experience, and the number of years of service in teaching were approved as criteria for an evaluation to predict the necessary courses to improve performance, and several training sessions were held For the targeted teachers, after implementing the activities of the experiment in their classes with the support and encouragement of the trainers, it was found that there was a noticeable improvement in the competence of the teacher, with a state of feeling comfortable and satisfied.

In the field of education, educational researchers have been interested in the subject of data mining, to help decision-makers in the educational institution to invest in improving the human and material resources necessary to develop the educational learning process. In a study (khazem et al., 2016), researchers were interested in the feasibility of applying some data mining techniques to test the study plans applied in the semester-based universities, where the previous course is a cognitive requirement for the next course, to provide indicators that help the management of the institution to evaluate Study plans to ensure the delivery of expected knowledge to the student, which contributes to improving the quality of the generalization and outputs.

In the study (Bhullar, 2012) the use of data mining through data classification and extraction was investigated using the Weka tool, and certain algorithms were used to predict student outcomes, helping teachers to provide them with academic assistance, and to act before they fail, to provide the confidence gained from knowledge. The numbers of students who are likely to succeed or fail, assist the students themselves in making important decisions for future work, and see the students' domain and interests.

The fields of use of neural networks varied with its many technologies, and one of the most important and used fields is the academic educational field, and one example of this is presented by the study (Mahapatra, Khan, 2007) which adopted the idea that the diverse nature of the requirements of the technical education system (TES) increases the difficulty of access. To the desired quality, and hence identifying the appropriate minimum quality elements for all stakeholders will aid in system design and improvement of customer satisfaction. For this purpose, a measurement tool known as EduQUAL was developed and an integrative approach using neural networks to assess the quality of service was proposed. Four neural network models based on the reverse propagation algorithm were used to predict the quality of education. The study also showed that the P-E gap model is the best, and sensitivity analysis of the best model for each stakeholder was performed to assess the robustness of the model, and finally, areas for improvement were suggested to the administrators of institutions.

A study (Abu-Naser, 2012) showed that using artificial neural networks technology and expert systems to gain knowledge of the learner model in the intelligent teaching system for linear programming (LP-ITS), to be able to determine the level of academic performance of learners to determine the level of difficulty for linear programming problems in preparation for solving them. . It was found from this study that the accuracy of predicting learners' performance was very high, indicating that the artificial neural network is skilled enough to make appropriate predictions.

The problem of low student achievement is one of the biggest problems in higher education, and many students leave their universities and give up their studies. This is for various reasons, including poor knowledge in the field of study, very low grades, and the inability to pass the exams, in addition to the lack of financial resources at times, and anticipating student results is a very important matter for managing universities that want to avoid this problem. The study (Oancea et al., 2013) investigated the use of a neural network to predict the results of students through their scores in the first year of the study. For this purpose, a sample of 1000 students from "Nicolae Titulescu" University in Bucharest was used, 800 were used to train the network, and 200 were used to test the network. The neural

network was a multi-layer preceptor (MLP) with one input layer, two hidden layers, and one output layer and trained with a version of the flexible back-propagation algorithm. The results of the network's training have been very positive, and the ability to anticipate student outcomes has proven very beneficial for the university administration to take early action to avoid the phenomenon of dropping out.

Studies on the design and implementation of the data mining framework have been numerous and in many fields; In the field of education and learning improvement, a study (Khader et al., 2016) presented a framework and implementation processes for data extraction techniques to improve students' level in higher education, the framework proposed courses for students according to their discovered skills and provided useful advice and information for the general education system to make more effective and efficient decisions on time .

To enhance teacher evaluation in the interest of students, their families, and their communities; Educators need feedback on their performance to help them determine how to improve their practices and develop schools as professional learning communities. In this regard, the study (Santiago, Benavides, 2009) proposes a conceptual framework for analyzing teacher evaluation, detailing the main components and aspects of a comprehensive teacher evaluation model, which must be taken into account when designing a teacher evaluation model, and how to use evaluation results to improve actual practices according to the framework. Conceptually developed in the study.

Another study (Ibrahim, et al., 2018) aimed to develop a framework based on data mining to analyze student assessment notes that would be obtained from social media sites based on textual comments and the notes included in them. The study consisted of three phases: The first phase was building a model that automatically detects students 'attitudes using sentiment analysis methods. The second stage consisted of building a model that automatically classifies the evaluation issues, and the third stage examines the correlation between the topic (the problem) and students' performance. Various common algorithms have been used to classify the text and analyze student feedback to enhance the learning process.

With the increasing importance of distance learning and in conjunction with the increase in interest in developing data mining tools in learning environments, especially in training course management systems known as "Moodle"; Many tools have been developed that usually require the user's knowledge of data mining techniques, and they also require time to obtain mining results from those tools, which created a problem for the user. The study (Gonçalves et al., 2017) provided a structure that uses data mining to obtain quick results illustrated with graphs, and the proposed tool was applied to a real data set for pedagogical students and based on the proposed framework, a tool was developed to display the results of data mining with a few clicks For Moodle users who wanted to have them for everyday use, users were able to provide faster feedback on how students are progressing in the distance learning courses.

Chapter 3: Methodology

The methodology used in this study is applied analytical, by preparing the system based on data mining tools and algorithms, this approach consists of several phases, from building the model by AI algorithms and designing the framework to writing software to explore patterns influencing teacher performance assessment and identifying their training needs, and ending with applying a group of teachers to test the framework.

3.1 Study Population & the Sample:

The study population consisted of male and female teachers of the Palestinian Ministry of Education for the main subjects (Arabic, Science, Mathematics, and English). The sample of the study consisted of the teachers of the Directorate of Education in Ramallah and Al-Bireh for the same subjects in the academic years 2014-2015, a data of 1009 teachers were collected. It was selected as a purposive sample to facilitate the researcher to collect data from the place of residence because of the difficulty of moving between cities due to the closures caused by the Covid-19 virus.

3.2 Steps of Building the Smart Framework:

After determined the variables that needed in the model which were (specialization, Qualification, Evaluation, Courses, Hours, the sum of hours, Type, Suggested Courses) it was collected, unfortunately, it wasn’t easy, there was no database in the ministry of education or any of its facilities, several places were visited to get the scattered data collected, and enter it in an organized manner in an excel file, 1009 teacher's records were collected.

3.2.1 Data Preprocessing:

- **Coding the data**: for easy preprocessing, string values converted to numeric values by coding it, variables coded were (courses, qualification, specialization, stage), as shown in table (3.1):

Table 3.1: Coded Data (string values converted to numeric values)

No.	Specialization	Evaluation	Qual-ification	Stage	Courses	Type	Hours	Sum
1010	0	76	2	0	089	p	15	55
1010	0	76	2	0	025	p	25	55
1010	0	76	2	0	024	c	15	55
1011	1	80	1	0	056	c	85	306
1011	1	80	1	0	021	p	18	306

Python language used to process the data flexibly, the following is performed:

- **Performing Two files**: Two files in excel were prepared, one named "suggested_courses" contains the number of teachers and the suggested courses, the other file consists of the rest of the variables, named "teachers_courses".

3.2.2 Normalization of "Evaluation" and "Sum" variables:

Normalization is a scaling technique or a mapping technique or a preprocessing stage (Patro, Saho, 2015).

Normalization is required when we are dealing with attributes on a different scale when multiple attributes are there but attributes have values on different scales, this may lead to poor data models while performing data mining operations. So they are normalized to bring all the attributes on the same scale.

In the data collected two variables normalized ('Evaluation', 'sum') using three methods of normalization process (Standard, Min-Max, Robust), every method affected the model results differently, so compartment was done before adopting one of them.

3.2.2.1 Standard Scalar:

The Standard Scaler is another popular scaler, for each feature, the Standard Scaler scales the values such that the mean is 0 and the standard deviation is 1. (Huilgol, 2020)

$$x_{Scaled} = \frac{x - \mu}{\sigma} \quad (1)$$

Standard deviation is the square root of the variance, the formula is: (Ayeni, 2014)

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \mu)^2}{n-1}} \quad (2)$$

Where x_i: the individual value of matric, μ: mean of matric, n: number of values.

3.2.2.2 Min-Max Scalar:

Min-max is a data normalization technique, it helps to normalize the data and scale it between 0 and 1, and it is not very well efficient in handling the outliers.

The formula for calculating the scaled value is: (Huilgol, 2020)

$$x_{Scaled} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (3)$$

3.2.2.3 Robust Scalar:

It scales features using statistics that are robust to outliers. This method removes the median and scales the data in the range between the 1st quartile and 3rd quartile. i.e., in between the 25th quartile and 75th quartile range. This method uses the interquartile range so that it is robust to outliers, with the formula: (Huilgol, 2020)

$$x_{Scaled} = \frac{x - Q1}{Q3 - Q1} \quad (4)$$

After the variables normalized “evaluation” and “sum”, their new values (normalized) reentered in the “teacher” array and an excel file named “normalized teacher”.

Table 3.2: Normalized values for “Evaluation” and “sum”

Number	Specialization	Evaluation	Qualification	Stage	Sum
0	1	0.44	2	1	-0.84
1	0	-0.84	2	0	-1.02
2	0	0.69	0	0	-0.37
3	1	-1.35	0	1	0.41
4	1	0.18	1	0	2.68

3.2.3 Data Mapping:

step1: mapping is to give variables new values, for easier handling of data and getting rid of empty spaces in the model, mapping is made for courses for each specializing by storing each in an array named “unique_courses” dimensions [specialization][courses], after counting each from an excel file “teachers_courses”.

Table 3.3: Number of Courses per Specialization

Specialization	Number of courses
0	69
1	67
2	86
3	71

Table 3.4: Unique Courses per Specialization

Specialization	Courses
0	[89 58 45 23 10 83 49 64 18 49 83 ……...]
1	[84 85 52 76 10 82 56 57 70 21 33 ……...]
2	[03 33 76 08 18 82 09 78 85 10 48 ……...]
3	[10 42 85 76 80 25 44 21 89 82 58 ……..….]

Step2: mapping is to store the array “unique_courses” in another array named “courses_inverse” but inversely, that is the course number in “unique_courses” is an index

in “courses_inverse”, and index in “unique_courses” is content in “courses_inverse”. Dimensions [specialization][index_of_courses] [4,91], where:

4: number of specialization.
91: the largest number of the courses numbers. (Max)

The set of numbers below shows the course inverse array for English specialization “0”.

```
array([ 0., 52., 0.,13., 45.,  0., 62., 34., 8., 20., 16., 55.,  0., 67., 0., 68., 53.,  0.,  38., 37.
        , 46., 11., 42., 17., 3.,  1., 0., 63., 60.,  0., 21., 44., 57., 47., 66., 64., 33., 0., 58.,
        24., 51., 22., 48., 49., 23., 18., 0., 0., 65., 14., 41., 35., 6., 50., 29., 43., 61., 0., 1
        9., 31., 26.,  0., 32.,  2., 12., 27.,  9., 28., 56., 0.,  4.,  0.,  0., 0., 0., 40., 7., 0., 10.,
        25., 59., 0., 5., 15., 0., 30., 0., 36., 0., 0., 39., 54.])
```

Figure 3.1: Course Inverse for English Specialization

Step3: the inverse arrays for the four specializations merged in a one-dimensional array named “courses_absolute_address” to get rid of the spaces not used, as follows:

- A large number of courses of the whole specializations is 86, so every specialization had the same length which is 86, in that way the courses compressed without any spaces the array length is (86*4).

- “teachers_encoded_coursers” is an array that combines between “teachers_indexes” and “courses_absolute_address”, it filled with “1” for taken courses for each teacher and “0” for not taken courses.

- **Generating the Input (Vector) “X” in the DNN Model:**

The input of the model assigned as “X”, an “x_original” generated from the array “teachers_encoded_coursers” of 329 entries which represent the taken courses besides of eleven entries for specialization, qualification, stage, evaluation, and sum represented as one-hot encoding which is a representation of categorical variables as binary vectors, these categorical values are first mapped to integer values. Each integer value is then represented as a binary vector that is all 0s, except the index of the integer which is marked as 1, for 1009 teacher as following: (Cerda, et al., 2018)

- **Specialization**: it took 4 entries.

Table 3.5: One Hot Encoding Per Specialization

Specialization	One Hot Encoding	
English	**0**	**1000**
Science	1	0100
Arabic	2	0010
Math	3	0001

- **Qualification:** it took 4 entries.

Table 3.6: One Hot Encoding Per Qualification

Specialization	One Hot Encoding	
Diploma	**0**	**1000**
Bach.	1	0100
Master	2	0010
PhD	3	0001

- **Stage:** one entry.

Table 3.7: One Hot Encoding Per Stage

Stage	One Hot Encoding	
elementary	0	10
Secondary	1	01

- **Evaluation**: one entry, the normalized value.

- **Sum:** one entry, the normalized value.

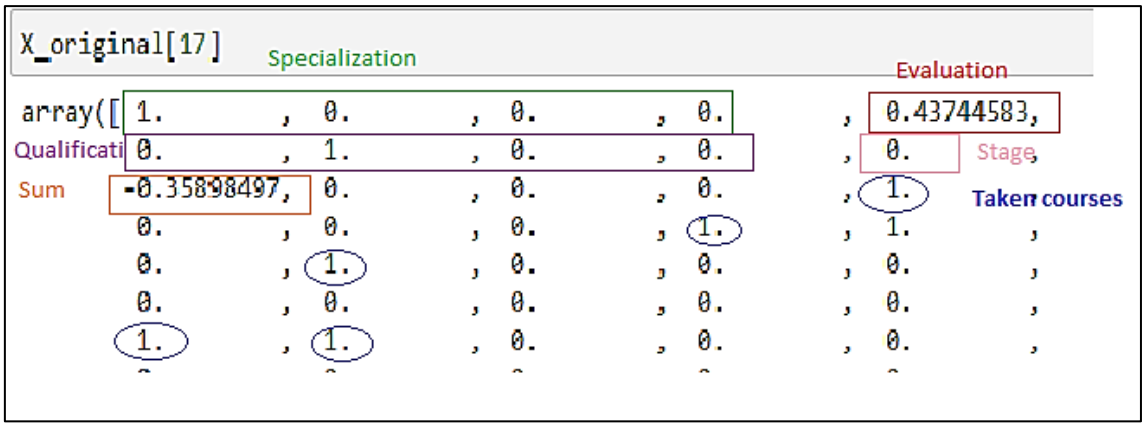


Figure 3.2: The Input vector X

The rest of the entries are the taken courses according to the teacher's specialization.

- **Generating the Label "Y" (suggested courses) :**

A suggested courses is a prerequisite for training the DNN model, as known values to calculate the difference between the predicted values and the actual data which is "Y".

The label assigned as "Y", a "Y_original" generated from the array "teachers_encoded_courses_y" which represents the suggested courses as following:

"sgst_courses" and the variable 'specialization' in "teacher" merged according to "No" in an array called:"teachers_courses_suggested". Each specialization suggested course is counted.

Table 3.8: Number of Suggested Courses per Specialization

Specialization	number of Suggested courses
0	21
1	20
2	25
3	18

- A unique_courses_y array contains the "teachers_courses_suggested" group by specialization and suggestion, then an" inverse_courses_y "array established by swapping the index and the suggested courses the same way in generating "X".

- To get rid of the spaces in the inverse_courses, the courses compressed in a one-dimensional array called “courses_abolute_address_y”, finally it was combined with the teacher's index to get up with “teachers_encoded_courses_y” which equal the input “Y”, Teachers_encoded_courses_y dimensions are (1009, 93).

Table 3.9: Unique Suggested Courses per Specialization

Specialization	Courses
0	[76, 82, 78, 10, 8, 89, 25, 65, 58, 45, 56, 85……...]
1	[82, 10, 56, 8, 80, 78, 20, 25, 85, 76, 44, 83……...]
2	[10, 85, 7, 82, 78, 76, 33, 3, 8, 44, 21, 18……...]
3	[82, 8, 20, 78, 70, 76, 85, 80, 89, 10, 42, 21 ……..]

An example of suggested courses(Y) for a teacher:

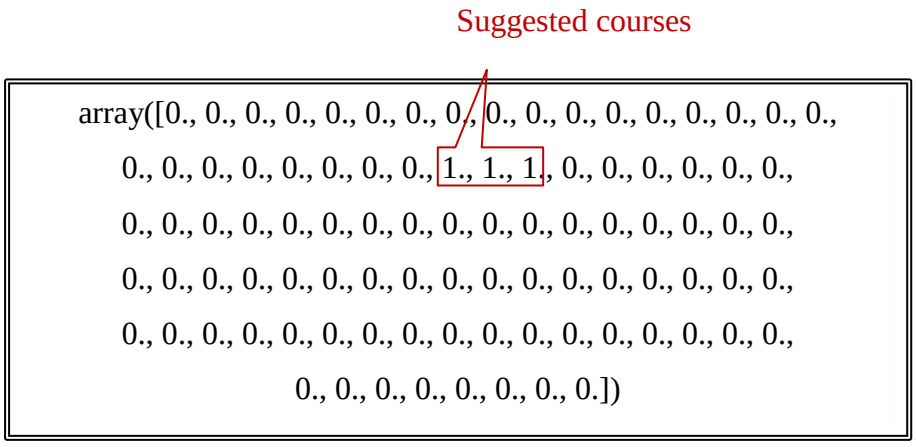


Figure 3.3: The Label Y (suggested courses for a teacher)

The processes ended up with the input X with dimensions (1009, 340)

And the suggested courses the label Y with dimensions (1009, 93)

3.3 The Deep Neural Network (DNN) Model:

A mining model was created by applying an algorithm to data, it is a set of data, statistics, and patterns that can be applied to new data to generate predictions and make inferences about relationships, and in this study deep Neural network AI algorithm was applied.

To build the model, data is needed to train it and test it, so the data divided into train and test data.

Since there were two inputs as shown above X, Y, a train data took values for x, from 0→ 850 entries, for test data from 850 to the end of data which equals 159. For input Y a train data took values from 0→850, for test data from 850 to the end as in X to be integrated.

3.4 Building the (DNN) Model:

The Deep Neural Network(DNN) model was composed of three hidden layers with 200 nodes in each, besides the output layer.

```
x = Dense (200, activation='relu')(inputs)
x = Dropout (0.4)(x)
x = Dense (200, activation='relu')(x)
x = Dropout (0.4)(x)
x = Dense (200, activation='relu')(x)
x = Dropout (0.4)(x)
Output = Dense (Y.shape[1], activation='sigmoid')(x)
Model = Model (inputs, output, name='teacher-model')
```

Figure 3.4: Statements of Building DNN Model

Dropout: to minimize the overfitting, so the model can deal with test data.

Dense: is a layer with 200 nodes in this model.

Relu: a function used to convert the negative values to zero and decimal numbers remain the same. Usually, it is used in the first layers.

Sigmoid: a function used in the last layer to make sure all values between 0 and 1.

For compiling the model a "binary_crossentropy" loss function was used, its formula:

$$H_p(q) = -\frac{1}{N}\sum_{i=1}^{N} y_i \cdot \log(p(y_i)) + (1 - y_i) \cdot \log(1 - p(y_i)) \quad (5)$$

The loss function tries every time to reduce the error.

The optimizer used in the model is “Adam”, "Adam is defined as an optimization algorithm that can be used instead of the classical stochastic gradient descent procedure to update network weights iterative based in training data" (Brownlee, 2017).

The model was trained according to accuracy.

```
Model.compile (loss='binary_crossentropy', optimizer='adam', metrics=['accuracy'])
```

Calculating Metrics:

- Accuracy is one metric for evaluating classification models, Formally, accuracy has the following definition:

$$Accuracy = \frac{number\ of\ \ correct\ \ predictions}{Total\ \ number\ of\ \ \ predictions} \quad (6)$$

For binary classification, accuracy can also be calculated in terms of positives and negatives as follows:

$$\frac{TP + TN}{TP\ + TN\ + FN\ + FP} \quad (7)$$

(Where TP = True Positives, TN = True Negatives, FP = False Positives, and FN = False Negatives)

- Precision trying to answer this question:

What proportion of positive identifications was correct?

$$Precision = \frac{TP}{TP\ + FP} \quad (8)$$

- Recall trying to answer this question:

What proportion of actual positives was identified correctly?

$$Recall = \frac{TP}{TP\ + FN} \quad (9)$$

3.5 The Smart Framework:

Figure (3.5) represent the architecture of the proposed system, which composed of three basic components:

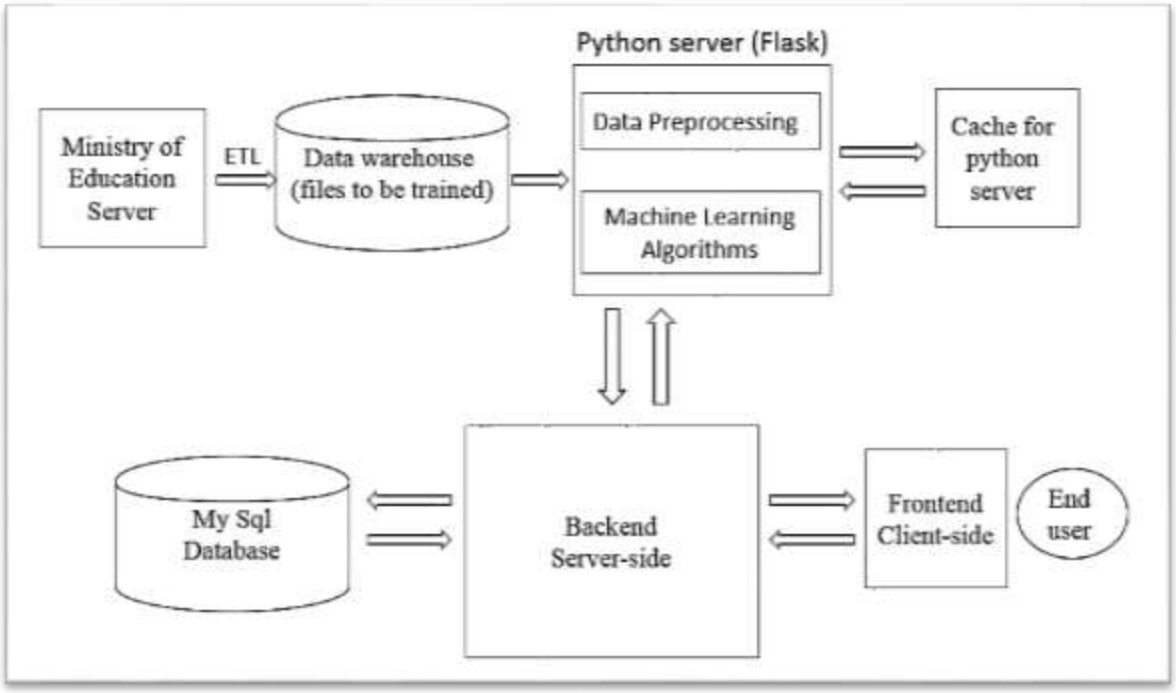


Figure 3.5: The smart Framework Architecture

Note that the figure consists of three main components:

- The first component is the Frontend or named “client-side” using the Angular javascript app tool.it communicates with the spring boot server via API to read and write data.
- The second component is the Backend server or named “server-side” using spring boot, it’s connected to the database to read and write data from it. It communicates with each Angular and python server.
- The third component is the python server “Flask”, its responsible for preprocessing the incoming data by reading it from the warehouse, to improve the efficiency of data mining algorithms, and then train it to get the required valid model. It communicates with the spring boot server via API to read and write data.

Other components:

- The server of the Ministry of Education connected with the data warehouse by ETL (Extract, Transform, Load) to send a list of teacher's information as files to be trained.
- The database contains the predicted results of the training and testing data about teacher's suggested courses, in addition to metrics of the trained model.
- The end-user can access the database to view the teacher's suggested courses and can view the machine learning page and ask for metrics when the data trained, depending on if the end-user is an admin or not.

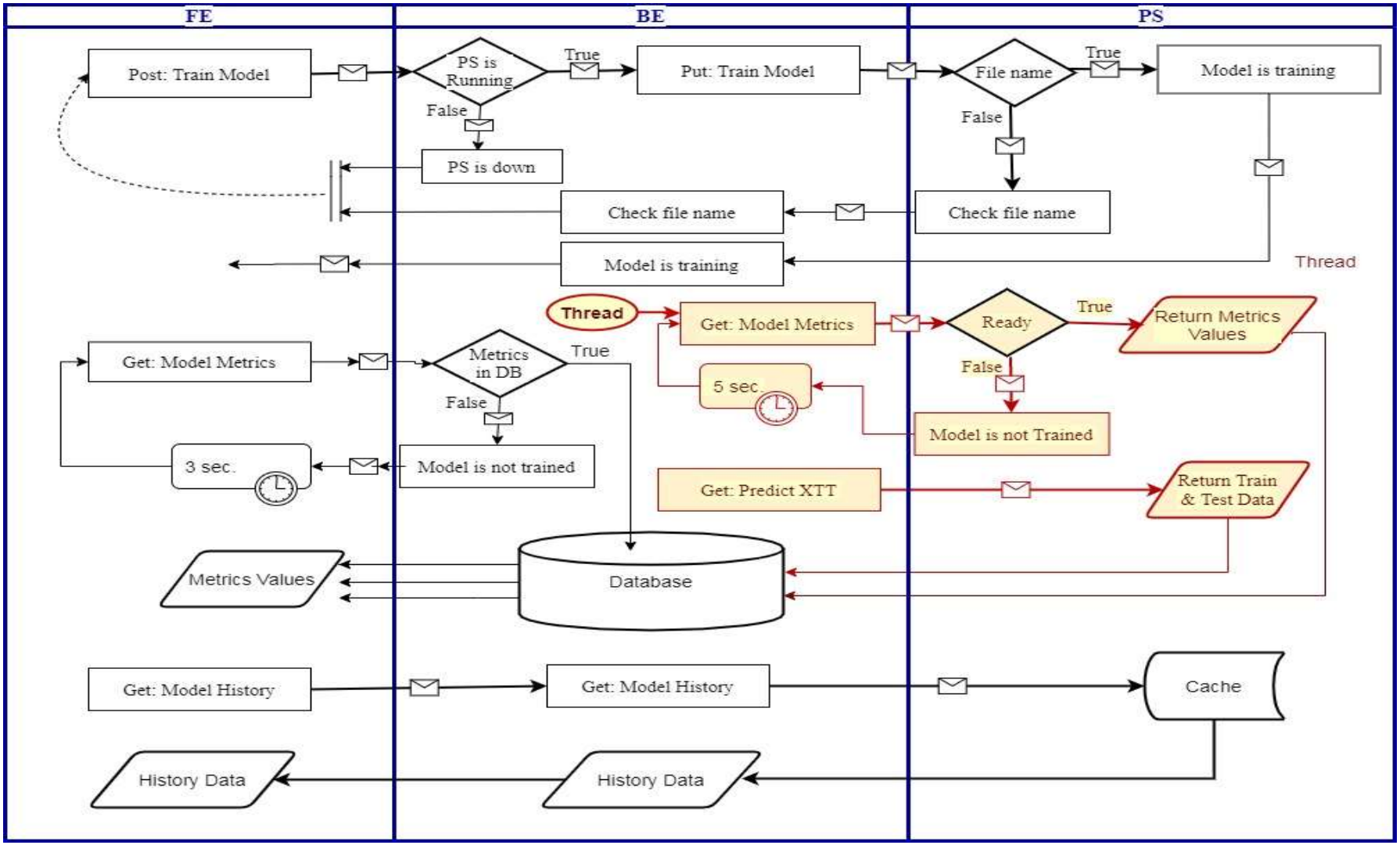


Figure 3.6: Sequential Diagram of Smart Framework

3.6 Description of the System:

According to the sequential diagram in figure 3.6 the requests were between the three main components as follow:

- When the frontend sends a request to the backend to train the model it sends “POST: train MLModel “.
- The backend sends a request to a python server “Put: train MLModel “.

  Response:

  On success: "message": "Model is training"

  This message will be delivered to BE, and in turn, it will send it to FE.

  On failure: "message": "Python server is down."

  or "message": "check file name"

(Message might vary with error cause variation).

- In case of success which is "Model is training" the BE server will create a thread responsible for sending requests to the python server to get the metrics and the predicted values as following:
- The thread will send a request to the python server “GET: model Metrics “.

  Response:

  On failure: "message": "Model is not trained yet"

  In this case, the thread will keep sending the same request every five seconds until succeeded.

  On success: metrics will be returned to the Backend server and saved in the database.
- Another request by the thread will be sent to the python server “POST: predictXTT”

  Response: return the train and test data and save them in the database.

  The role of the thread is ended at this point.
- During previous operations, the frontend sends a request to the BE server for metrics

  “GET: model Metrics “, BE checks the DB if there is any data there.

Response:

If the data does not exist it sends a message: "Model is not trained”

The FB keeps sending the same request every three seconds until succeeded.

On success (if the data exist): metrics returned to FE.

Another request from the FE was sent to the BE server asking for the history, “GET: model History “.

Since history is not stored in the database, the BE server sends a request to the python server: “GET: model History “.

Response:

The python server checks the cache memory-related and load the history data and deliver it to BE, and in turn, it sends it to FE to view it as curves for “accuracy” and “loss”.

If history does not exist: "message": "Model is not trained".

3.7 Software Development Environment:

The system was developed using the jhipster environment as a handy application generator that created a spring boot (java part) and the Angular (hipster part) application. Angular was considered as a platform, and framework for building single-page client applications (interface) named Front End. The server side was spring boot Back End, it considered as an open-source java-based framework to create a microservice.

Spring boot connected to the database to write and read data. Flask server python was used as a web framework that provides useful tools and features that make creating web applications in Python easier. It gave flexibility and was a more accessible framework, Neural Network AI algorithm done in flask server using TensorFlow as an open-source software library for machine learning and inference of Deep Neural Networks. NumPy is used also as a Python library used for working with arrays, functions, and matrices, another library named Pandas used for data analysis. It provided highly optimized performance with back-end source code. To transfer data between all servers a tool was used named API it is a software intermediary that allows two applications or software programs to communicate with each other.

It is spelled out the proper way to write a program requesting services from an operating system or other application.

A RESTful API is an architectural style for an application program interface that uses HTTP requests to access and use data. Using these operations:

The API operations used for the Smart Framework shown in appendix A.

Table 3.10: API Operations

Operation	Description
Get	Returns data based on the parameters you provide when you call the operation.
Delete	Deletes objects.
Post	Creates new objects.
Put	Updates existing objects.

The status codes used in the system:

Table 3.11: Status Codes of the System

Status Code	Name	Use case
200	OK	A success GET or PUT request
404	Not Found	Unsuccessful request due to invalid parameter in URL
401	Unauthorized	Requesting a restricted URL with invalid credentials
500	Internal server error	Problem with the server or DB server.

3.8 Smart Prototype Design:

A smart prototype was implemented to simulate the suggested system with the mentioned components:

- The first component is the Frontend or named “client-side “.
- The second component is the Backend server or named “server-side”, it’s connected to the database to read and write data from it.
- The third component is the python server “Flask”.

Other components:

- The database (MySQL) contains the predicted results of the training and testing data.
- The end-user can access the database to view the teacher’s suggested courses and could view the machine learning page and ask for metrics when the data trained depending if the end-user is an admin or not.

Chapter 4: Experimental Results and Analysis:

The current study seeks to design a framework based on data mining to extract knowledge that can be used by decision-makers to identify the training needs of the teacher to support his academic and professional development, and then improve the teaching and learning process. A sample of government teachers in Palestine was taken for the main subjects (Arabic Language, Science, Mathematics and English Language) in the academic year 2014/2015 AD, and the researcher approved the following criteria for the teacher: academic qualification, specialization, stages he teaches, annual evaluation, courses, and its number of hours he has passed in each of the fields of specialization and teaching methods. Which is expected to be influential in determining its development training needs, and has adopted appropriate software tools for this type of applied research.

4.1 Normalization Methods Compartment Results:

At the stage of normalization, the three methods were applied in the model; since the data are shuffled every running time results will be different, every method affected the training model results differently, so three trials for every method have been made, then the best results choose for each, and finally compared the final results of the three methods and adopted the best. The results were as following:

After applying the Standard Scalar three times; it gives the results in a table (4.1) as follows:

Table 4.1: Metrics using Standard Scalar

Standard	Trial (1)	Trial(2)	Trial(3)
Accuracy	0.974	0.974	0.974
precision	0.72	0.74	0.73
recall	0.73	0.75	0.74
TP	525	543	537
FP	199	191	195
TN	13878	13872	13874
FN	185	181	181

As the results show trial (2) is the best because the precision TP and recall are the highest. FN is the lowest which is the best.

When the Min-Max Scalar was applied; the results in table (4.2) were got:

Table 4.2: Metrics Using Min-Max Scalar

Min-Max	Trial (1)	Trial(2)	Trial(3)
Accuracy	0.971	0.974	0.973
precision	0.70	0.73	0.72
recall	0.71	0.75	0.74
TP	510	534	534
FP	215	200	206
TN	13859	13881	13867
FN	203	172	180

As the results show trial (2) is the best, because the precision, TP, and recall are the highest. FN is the lowest which is the best.

Table (4.3) shows the results when the Robust Scalar was applied:

Table 4.3: Metrics Using Robust Scalar

Robust	Trial (1)	Trial(2)	Trial(3)
Accuracy	0.972	0.975	0.972
precision	0.71	0.74	0.72
recall	0.71	0.75	0.70
TP	508	541	505
FP	205	189	196
TN	13870	13878	13875
FN	204	179	211

As the results show trial (2) is the best, because the precision, TP, and recall are the highest. FN is the lowest which is the best.

After making a comparison of the three methods with their final results, the standard Scalar was adopted and used in normalizing the data as shown in table (4.4).

Table 4.4: Compartment of Three Normalization Methods

methods / **metrics**	**Standard**	**Min-Max**	**Robust**
Accuracy	0.975	0.975	0.975
precision	0.74	0.73	0.74
recall	0.75	0.75	0.75
TP	543	534	541
FP	191	200	189
TN	13872	13881	13878
FN	181	172	179

4.2 Metrics Results

The metrics results of this model after training shows the success of this model, accuracy is the fraction of predictions the model got right. The accuracy of the trained model is 0.9757219. The curve in figure 4.1 shows the values of the accuracy increasing as the epoch's increases, which means higher predictions.

According to table 4.4, the precision is approximately 0.74, which means when it predicts courses it is correct 74% of the time, and the recall is 0.75 which means it correctly identifies 75% of all courses.

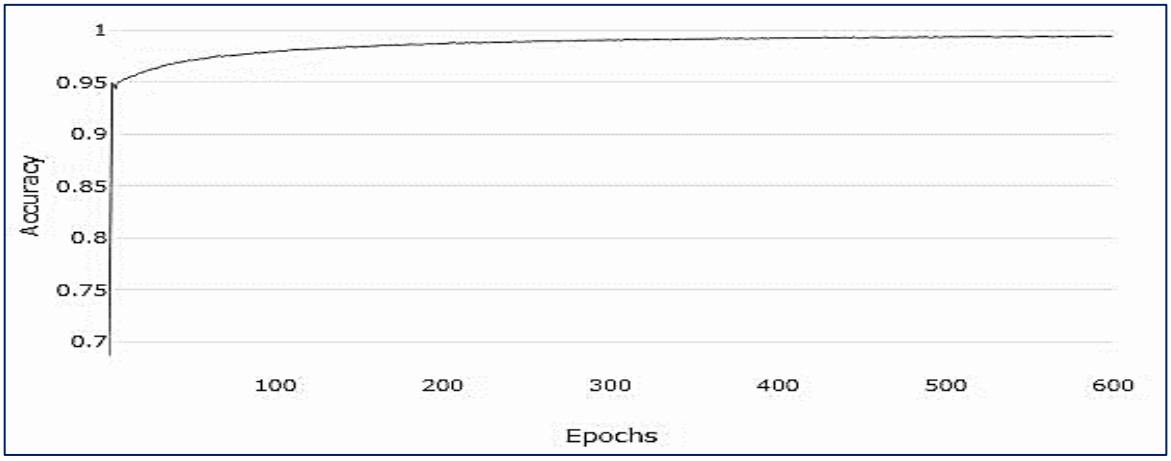


Figure 4.1 Data Accuracy of DNN Model

The loss function tries every time to reduce the error, to express by this the proximity or distance of the model to the desired results, the lower the function value, the higher prediction. That is: $Y \approx Y^{\wedge}$, Y: train, $Y^{\wedge}$: predictive (test)

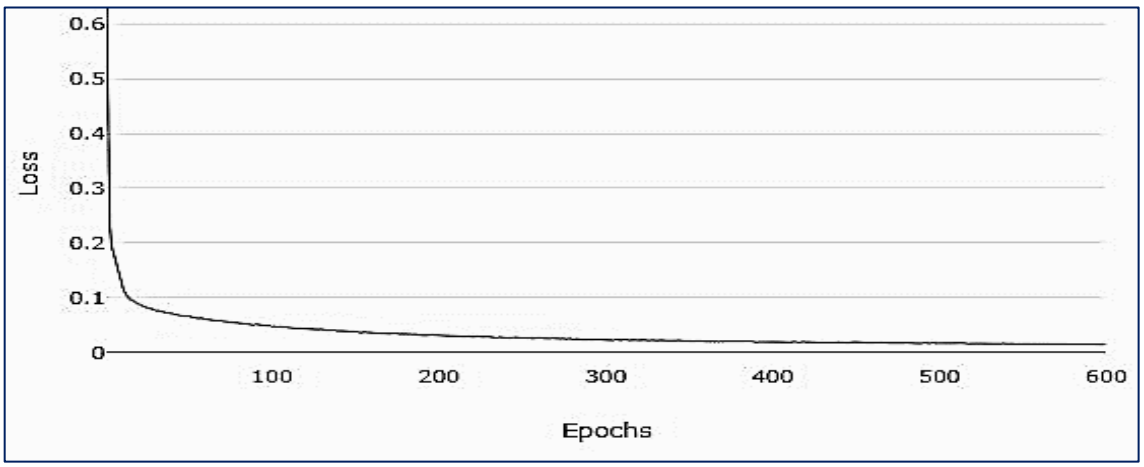


Figure 4.2 Binary Crossentropy Loss Function for DNN Model

As shown in figure (4.2) according to the model results, the values of the loss function are decreasing which means higher prediction.

After training the model, test data (159) entries were predicted, the difference between the prediction data(y^) and the actual data (suggested courses y_test) was:

(predictions - y_test) =372 fault from the prediction data size (159*93 =14787), which means (14787-372 =14415 is correct),

Base on the results of test data (prediction) in table 4.4 and after adopting the Standard Scalar, the accuracy was 0.97484 which means higher predictions:

According to equation (6), accuracy = (14415/ 14787) = 0.97484

For example, Prediction results for teacher number (521) is:

```
{'specialization': 1, 'course': 82},
{'specialization': 1, 'course': 56},
{'specialization': 1, 'course': 85}],
```

Actual data (y-test) for the same teacher is:

```
{'specialization': 1, 'course': 82},
{'specialization': 1, 'course': 56},
{'specialization': 1, 'course': 76}],
```

The model missed only one course.

As a result of preparing the data and building the model, the plotting of the model as shown in figure (4.3):

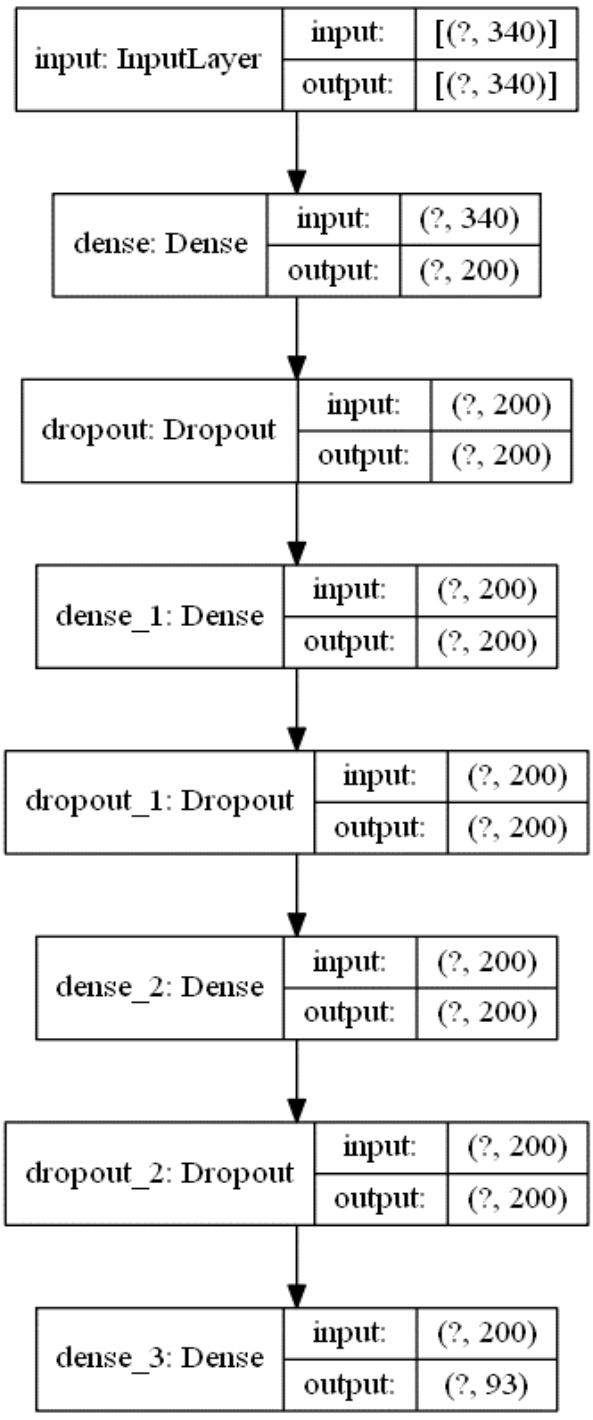


Figure 4.3: Neural Networks (DNN) Model

The input has 340 entries with a variable number of teachers, to begin with, then three hidden layers with 200 entries as the number of nodes, finally output is 93 labels with a variable number of teachers.

4.3 System Mechanism:

The way the system works; as a result of building the Smart Framework Prototype is as follows: (for figures see appendix B)

- To let the framework run successfully, the python server should be run first, then the Back-end server.
- Open the Database (my SQL) to read and write data.
- Type http://localhost:8080 in the browser to get to the system.
- Two ways to access the system through the main page:
    1. **Sign in**: as an administrator account, he can enter the system and have all the permissions. But if it is a user account, he can log in with the permissions specified by the admin; asking for the user name and password.
    2. **Sign up or Register**: Any user can create an account and wait until approval is given from the administrator through the Administration tab, and the last one can specify the user's permissions.
- The menu of the main page has four tabs:
    1. **Home**: to sign in or sign up when opening the system, and a blank page with a welcome message for the entered user.
    2. **System**: it contains a list of four titles which are:
        - **Courses**: view courses for the main subjects (Maths, Science, Arabic, English) with its code, name of the course, type (content, pedagogy), and no. of hours for each course. (no. of all =293).
        - **Teachers**: when data brought from DB after training the model, it viewed the prediction courses for teachers besides the variables (Number, Specialization, Evaluation, Qualification, Stage, Sum of hours, Suggested courses) another two

columns are added, the creation date to check when was the prediction done or search by the date, the other column name “Is predicted” with two options “true” for the training data because the suggested courses are known, and “False” for the test data to predict. All the columns are filtered to search for specific data.

- **Model Training**: to build a model and train it, write the file's name correctly with its extensions, one for the model training and the other for model labels.

  After finishing training, it viewed metrics of the model (confusion matrix, accuracy, Recall, Precision), and plotted curves of “accuracy” and “loss” for the model in a very clear, neat, and interactive way. Model and all data can be cleared and deleted through this page by an icon name “Delete”.

- **Data Prediction**: This page is for predicting data through the built model, by typing the file name in the space provided. When the data predicted to succeed, it is stored in the teacher table.

3. **Administration:** views all users' information including admins of the system, controlling the activation of each by admin, and the date of account creation, editing, deleting, and defining everyone’s role.

4. **Account:** a list of three titles when the user signs in:

   - **Settings:** to change user information (Fname, Lname, E-mail)

   - **Password:** for changing user password, by writing the old one and confirming the new one.

   - **Sign out:** to leave the system.

Chapter 5: Discussion and Conclusion

5.1 Discussion:

Some studies were mainly aimed at evaluating the teacher's performance through multiple models and methods as in (Hemaid, El-Halees, 2015) study, while this study was distinguished by considering the teacher's evaluation as one of the criteria upon which to determine his training needs to improve his performance.

What distinguished this study is the development of a procedural application tool (framework) to identify the appropriate training courses necessary to improve the performance of the teacher, while many studies as (Hervie, Winful, 2018) were limited to highlighting the importance of educational training in improving the performance of the teacher.

Given the importance of school education and its role in the renaissance and development of society in all aspects of life; this study directed its tools to school teachers who are the basis of the educational-learning process and the center of strength in it. While we find that previous studies in this context have been directed to the university academic community as (Oancea et al., 2013).

Initially, the validity of the model checked by its accuracy, and it was 0.9997 after training, it increases with the epochs increasing, the binary_crossentropy used as a loss function by calculating the cost which is the difference between the predictive values (predictive courses y^) and the actual values (suggested courses y-train), it gives higher predictions by trying to reduce the error with epochs increasing, it ends up with 0.0024.

After training the model, test data (159) entries were predicted, the difference between the prediction data(y^) and the actual data (suggested courses y_test) was 372 fault from the prediction data size 14787, which means 14415 is correct.

Base on the results of test data (prediction) in table 4.4 and after adopting the Standard Scalar, the accuracy was 0.97484 which means higher predictions:

The framework was tested by 300 teachers giving the required results which are the predicted courses of each.

Of course, accuracy is not enough to determine the success of the model, precision and recall metrics support the validation of the model, the result of precision was 0.73978 which means when it predicts courses it is correct 0.74% of the time, and recall was 0.75 which means it correctly identifies 75% of all predicted courses. When these two metrics are high the F1 score is high. If the data were more than used, these values will be even higher.

Adam's choice as an optimizer gave the model training more speed, also to Prevent the Neural network from Overfitting a layer was added in the model named dropout, it uses a technique where randomly selected neurons are ignored during training in every phase. They are "dropped-out" randomly, the value chosen is 0.4, which means it dropped 40% of the neurons in every phase (dropped 80 neurons from 200), which gave the model effectiveness.

The selection of the variants was successful to get the required results of the framework that approved its efficiency and applicability in determining the required courses for a teacher to enhance his performance and then improve his evaluation.

5.2 Conclusion:

- The results of applying the system as a Smart Prototype of the Framework showed high effectiveness in identifying the necessary training needs for the teachers to improve their academic performance in an organized and interactive way.
- The AI algorithm of Deep Neural Networks(DNN), was effective to arrive at influential patterns related to teacher performance enhancement in Palestine.
- A teacher will not be confused about taking unnecessary courses or duplicate one, besides reducing the teacher's usual negative attitudes towards training.
- The Ministry of education also reduces the expense of holding courses and will have teacher's data stored in an oriented organized way.

5.3 Future Work:

- The researcher recommends the Ministry of Education to use this intelligent technology tool for its effectiveness in predicting needed training courses for a teacher to enhance his performance, and also to organize teacher's data in an oriented way

- As well as classifying teacher's data in the framework according to the governorate in which they teach, so that each directorate is allowed to enter its page of this system.

References:

- Abdul-Hay, E. M. (2017). **One of the most important uses of teaching competencies,** available at https://www.new-educ.com/category/studies/page/8, entry date: 19/01/2020.

- Abu-Naser S. (2012). **Predicting Learners Performance Using Artificial Neural Networks in Linear Programming Intelligent Tutoring System**, International Journal of Artificial Intelligence & Applications (IJAIA), Vol.3, No.2, Gaza, Palestine.

- Abu_Samra, M., Moamer, M. (2013). **The Role of Educational Supervision in Supporting the New Teacher in Palestine**, An-Najah University Research Journal, Vol. 27(2).

- Ajab, A., Basheer, A., Mohammad, M., (2014), **Evaluation of student academic performance using data mining techniques**, Sudan University of Science and Technology, College of Computer Science and Information Technology, Sudan.

- Ayeni, A. W. (2014). **Empirics of Standard Deviation,** available at https://www.researchgate.net/publication/264276808_Empirics_of_Standard_Deviation, entry date: 18/01/2021.

- Baby A. M. (2018). **Pedagogue Performance Assessment (PPA) using Data mining Techniques**, IOP Conf. Series: Materials Science and Engineering 396.

- Bahari, F., Elayidom, S. (2015). **An Efficient CRM-Data Mining Framework for the Prediction of Customer Behavior**, International Conference on Information and Communication Technologies (ICICT), Procedia Computer Science, Vol. 46, p (725–731).

- Bhullar, M. S., Iaeng, M., & Kaur, A. (2012). **Use of Data Mining in Education Sector**, WCECS, San Francisco, USA, Vol. 1, p (978 -988).

- Bichi, A. A. (2017). **Evaluation of Teacher Performance in Schools: Implication for Sustainable Development Goals**, Northwest Journal of Educational Studies, Vol. 2, No.1.

- Brownlee, j. (2017). **Gentle Introduction to the Adam Optimization Algorithm for Deep Learning**, available at https://machinelearningmastery.com/adam-optimization-algorithm-for-deep-learning/, date of entry: 01/11/2020.

- Al-Burai, A. M. (2012). **The Reality of Teacher Preparation and Qualification Institutions in Yemen "An Analytical Study"**, Journal of the College of Basic Education, Yemen, Issue (76).

- Cerda, P., Varoquaux. G., Kégl, B. (2018). **Similarity encoding for learning with dirty categorical variables**, available at: https://link.springer.com/article/10.1007/s10994-018-5724-2, date of entry: 18/01/2021.

- Chen, S. C., Shyu, M. L., Chen, M., & Zhang, C. (2004). **A decision tree-based multimodal data mining framework for soccer goal detection,** IEEE International Conference on Multimedia and Expo, Taipei, Taiwan.

- Cooper, R. G. (2019), **The drivers of success in new-product development**, Industrial Marketing Management, Vol 76, Pages 36-47.

- Džeroski, S. (2006). **Towards a General Framework for Data Mining**, International Workshop on Knowledge Discovery in Inductive Databases, Knowledge Discovery in Inductive Databases. KDID 2006. Lecture Notes in Computer Science, vol 4747. Springer, Berlin, Heidelberg.

- Edmond, H., Michael, K., & Huang, J. Z. (2004). **A data Warehousing and Data Mining Framework for web usage management**, Communications in Information Systems, Vol. 4, No. 4, pp. 301-324.

- Elhoseny, H. Elhoseny, M., Riad, A. M., & Elhoseny, A. E. (2018). **A Framework for Big Data Analysis in Smart Cities**, available at https://link.springer.com/chapter/10.1007/978-3-319-74690-6_40, date of entry: 29/12/2020.

- Figueiredo, V., Rodrigues, F., & Vale, Z. (2005), **An Electric Energy Consumer Characterization Framework Based on Data Mining Techniques**, VOL. 20, NO. 2.

- Francis T.S., (2002). **Optical Neural Network and Education**, available at https://spie.org/ETOP/2001/234_1.pdf, date of entry: 08/10/2020.

- Gonçalves, F. D., Maciel, M. A., & Rodrigues, R. L. (2017). **Development of a Data Mining Education Framework for Data Visualization in Distance Learning Environments,** Retrieved September 30, 2020, available at http://www.ecomp.poli.br/amam/wp-content/uploads/2019/11/2017_Development-of-a-Data-Mining-Education-Framework-for-Visualization-of-Data-in-Distance-Learning-Environments.pdf.

- Haruechaiyasak, C., Shyu, M., & Chen, S. (2015). **A Data Mining Framework for Building a Web-Page Recommender System**, National Electronics and Computer Technology Center (NECTEC).

- Hemaid R. Kh., El-Halees A. M. (2015). **Improving Teacher Performance using Data Mining**, International Journal of Advanced Research in Computer and Communication Engineering, Vol. 4, Issue2.

- Hervie D. M., Winful E. C. (2018). **Enhancing Teachers' Performance through Training and Development in Ghana Education Service (A Case Study of Ebenezer Senior High School)**, Journal of Human Resource Management, 6(1): 1-8.

- Huilgol P., (2020). **Feature Transformation and Scaling Techniques to Boost Your Model Performance**, available at:

https://www.analyticsvidhya.com/blog/2020/07/types-of-feature-transformation-and-scaling/, date of entry: 18/01/2021.

- Ibrahim, Z. M., Badr-El-Den, M., & Haig, E. (2018). **A Data Mining Framework for Analyzing Students' Feedback of Assessment,** 13th European Conference on Technology Enhanced Learning Doctoral Consortium Leeds, United Kingdom.

- Jais I. K. M., Ismail A. R., Nisa S. Q. (2019). **Adam Optimization Algorithm for Wide and Deep Neural Network**, Knowledge Engineering and Data Science 2(1):41, p41-46.

- Al_Janaie, A., Al_Hadad, H., Albar, A., Zahrawi, A. (2011). **Exploring some patterns affecting the academic performance of students at the University of Science and Technology using data mining techniques,** Journal of Science and Technology, Vol. (16), Issue (1).

- Jurafsky D., Martin J. H., (2019). **Neural Networks and Neural Language Models**, Available at https://web.stanford.edu/~jurafsky/slp3/7.pdf, date of entry: 08/10/2020.

- Kadhim B., Al-Janabi, S. (2011). **A Proposed Framework for Analyzing Crime Data Set Using Decision Tree and Simple K-Means Mining Algorithms**, Journal of Kufa for Mathematics and Computer, Vol.1, No.3, pp.8- 24.

- Khazim, R., Rahhal, Y., Rahmah, N. (2017). **The use of data mining techniques to improve the quality of the educational process and the delivery of knowledge, an applied case: The Higher Institute of Business Administration,** Tishreen University Journal of Science and Studies, Engineering Sciences Series, Vol. (39), Issue (6).

- Khedr, A., Abdo, A., & Amar, A. (2016). **A Proposed Framework for Data Mining Techniques in Learning System**, International Journal of Advanced Research in Computer Science and Software Engineering, V. 6, Issue 3. Cairo, Egypt.

- Kumar, A., Saurabh (2013). **Evaluation of Teacher's Performance: A Data Mining Approach,** International Journal of Computer Science and Mobile Computing (IJCSMC), Vol. 2, Issue. 12, pg. 359– 369.

- Ladas, A., Ferguson, E., Aickelin, U., & Garibaldi J. (2014). **A Data Mining framework to model Consumer Indebtedness with Psychological Factors,** International Conference on Data Mining Workshop, Shenzhen, China.

- Laura G., Courtney B., & Olivia L. (2008). **Approaches to Evaluating Teacher Effectiveness: A Research Synthesis**, National Comprehensive Center for Teacher Quality (NCCTQ).

- Leeuwen A., Janssen J. (2019). **A systematic review of teacher guidance during collaborative learning in primary and secondary education**, Educational Research Review 27 71–89.

- Mahapatra S. S., Khan M. S., (2007). **A neural network approach for assessing quality in technical education: an empirical study**, International Journal of Productivity and Quality Management (IJPQM), Vol. 2, No. 3.

- Mahgoub Y. M. (2014). **Development of Teacher Performance and its Impact on Enhancing on the Quality of the Educational Process**, Pensee Journal, Vol 76, No. 2.

- Margaret, R. (2015). **Framework**, Available at https://whatis.techtarget.com/definition/framework, date of entry: 22/07/2020.

- Nagawi, A. (2018), **What is simply a Framework?** Available at: https://www.geekgenes.com/2018/04/what-is-framework, entry date 29/9/2020.

- Ngai, E.W.T., Hu, Y., Wong, Y.H., Chen, Y., & Sun, X, (2011). **The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature**, Decision Support Systems, Vol. 50, P (559–569).

- Nikolovska, A., Spasov, S., Velinov, A., & Zdravev, Z. (2018). **Framework for Big Data Analytics of Moodle Data Using Hadoop in the Cloud,** Conference: Eighth International Scientific Conference Computer Science'2018, Kavala, Greece.

- Oancea B., Dragoescu R., & Ciucu S. (2013), **Predicting students' results in higher education using a neural network,** Published in International Conference on Applied Information and Communication Technologies (AICT2013) (April 2013): pp. 190-193.

- Ola, A. F, Pallaniappan, S. (2013). **A data mining model for evaluation of instructors' performance in higher institutions of learning using machine learning algorithms,** International Journal of Conceptions on Computing and Information Technology, Vol. 1, Issue 2, ISSN: 2345 – 9808.

- Olcum D., Titrek O. (2015). **The Effect of School Administrators' Decision-Making Styles on Teacher Job Satisfaction**, Social and Behavioral Sciences, 1936 – 1946.

- Padhy, N., Mishra, P., & Panigrahi, R. (2012). **The Survey of Data Mining Applications and Feature Scope**, International Journal of Computer Science, Engineering and Information Technology (IJCSEIT), Vol. 2, No.3.

- Patro, G. k., Saho, K. K. (2015). **Normalization: A Preprocessing Stage**, Available at https://arxiv.org/ftp/arxiv/papers/1503/1503.06462.pdf, date of entry: 22/07/2020.

- Rodriguez-Calero, I., Coulentianos, M., Daly, S., Burridge, J., Sienko, K. (2020), **Prototyping strategies for stakeholder engagement during front-end design: Design practitioners' approaches in the medical device industry,** Design Studies, Vol.71, No. C.

- Rouse, M. (2015). **Framework**, Available at: https://whatis.techtarget.com/definition/framework, date of entry: 29/09/2020.

- Sajadfar, N., Ma, y. (2015). **A hybrid cost estimation framework based on feature-oriented data mining approach,** Advanced Engineering Informatics, 29 (3).

- Santiago, P. & Benavides, F. (2009). **Teacher Evaluation: A Conceptual Framework and Examples of Country Practices**, paper presented at the OECD-Mexico Workshop towards a Teacher Evaluation Framework in Mexico.

- Silva N. J., Lobo C. C., Pereira C. S., & Durão N. S. (2017). **Evaluation of Teaching Performance in Higher Education: Students Perspectives and Teaching Management Indicators**, Proceedings of INTED2017 Conference, Valencia, Spain.

- Sodiya O., Quliyeva A. (2019). **A Systematic Review: Relationship Between Teacher's Professional Development and Students' Academic Achievement**, Conference: 4th ADA International Education Conference.

- Suarez E. G., Toro R. M. (2018). **Teaching Performance Evaluation Model: Preparation for Student Learning within the Framework for Teacher Good Performance**, Propósitos y Representaciones, Vol. 6, N 2: pp.431-452.

- Syed, A. (2015). **Open-source data mining tools: an analytical evaluation study,** Taibah University Journal of Arts and Humanities, Egypt, Issue (10), p. (791).

- Ulrich, K. T., Eppinger, S. D. (2016), **Concept Development: The Front-End Process,** Product Design and Development, 6th ed., pp. 16.

- UNESCO (2014). **The general strategic plan of the Regional Center for Quality and Excellence in Education (2015-2020),** United Nations Educational, Scientific and Cultural Organization, UNESCO, Kingdom of Saudi Arabia, Riyadh, available at https://rcqe.org/wp-content/uploads/2016/10/Strategy2015V2-arabic.pdf, date of entry: 29/09/2020.

- Vagh, Y., Armstrong, L., & Diepeveen, D. A. (2010). **Application of a data mining framework for the identification of agricultural production areas in WA**. Proceedings of the 14th Pacific-Asia Conference on Knowledge Discovery and Data Mining. (pp. 11-22). Hyderabad, India. Edith Cowan University.

- Wales, M. (2014). **3 Web Dev Careers Decoded: Front-End vs Back-End vs Full Stack**, Retrieved September 29, 2020, Available at: https://blog.udacity.com/2014/12/front-end-vs-back-end-vs-full-stack-web-developers.html, date of entry: 30/12/2020.

- Wu, E. H., Ng, M. K., & Huang, J. Z. (2004)**. A Data Warehousing and Data Mining Framework for Web Usage Management**, Communications in Information & Systems, Vol. 4, No.4, p 301-324.

Appendix A

API operations used in the System:

The API operations used for the Smart Framework Prototype are as following:

❖ POST: /API/trainMLModel
http://localhost:8080/api/trainMLModel?trainingFile=train_test_set.xlsx&labelsFile=suggested_courses.xlsx

Takes in the path two files:

- Training File
- Labels File

Response:

- On Success: status 200 with the JSON in the body:

```
{
   "message": "Model is training"
}
```

- On Failure: status 404 with the JSON in the body: (Message might vary with error cause variation).

```
{
   "message": "Python server is down."
}
```

❖ POST: /API/predictData_

http://localhost:8080/api/predictData?newDataFile=teachers_test_300.xlsx

Takes in the path a file: newDataFile

Response:

- On Success: status 200 with the JSON in the body:

```
{
   "message": "Prediction succeeded."
}
```

- On Failure: status 404 with the JSON in the body: (Message might vary with error cause variation).

  {

  "message": "Model is not trained yet"

  }

❖ **GET**: /api/modelMetrics

http://localhost:8080/api/modelMetrics

Response:

- On Success: status 200 with the json in the body:

  {

  "tp": 539.0,

  "tn": 13881.0,

  "fp": 205.0,

  "fn": 162.0,

  "accuracy": 0.9751809,

  "precision": 0.7244624,

  "recall": 0.7689016

  }

- On Failure: status 404 with the JSON in the body: (Message might vary with error cause variation).

  {

  "message": "Model is not trained yet"

  }

❖ **GET**: /api/modelHistory

http://localhost:8080/api/modelHistory

takes no parameters

Response:

- On Success: status 200 with the json in the body:

  {

  "loss": [

```
      0.6526795765932869,
      0.41481562572367053,
      0.23371745032422683,
      0.19550187447491815,
      0.17901217586853924,
      .
      .
     0.015425456797375399,
      0.015075195635504583,
      0.015300996601581573
   ],
   "accuracy": [
      0.6541935205459595,
      0.8864010572433472,
      0.9452624320983887,
      0.9481719732284546,
      0.9448449611663818,
     .
     .
      0.9944846034049988,
      0.9943327903747559,
      0.9941303730010986,
      0.9941049218177795,
      0.9940923452377319,
      0.9940291047096252,
      0.9940164685249329
   ]
}
```


- On Failure: status 404 with the JSON in the body: (Message might vary with error cause variation).


```
{
   "message": "Model is not trained yet"
}
```

❖ **DELETE:** /api/deleteModelAndTeachers

http://localhost:8080/api/deleteModelAndTeachers

takes no parameters

Response:

- On Success: status 200 with the JSON in the body:

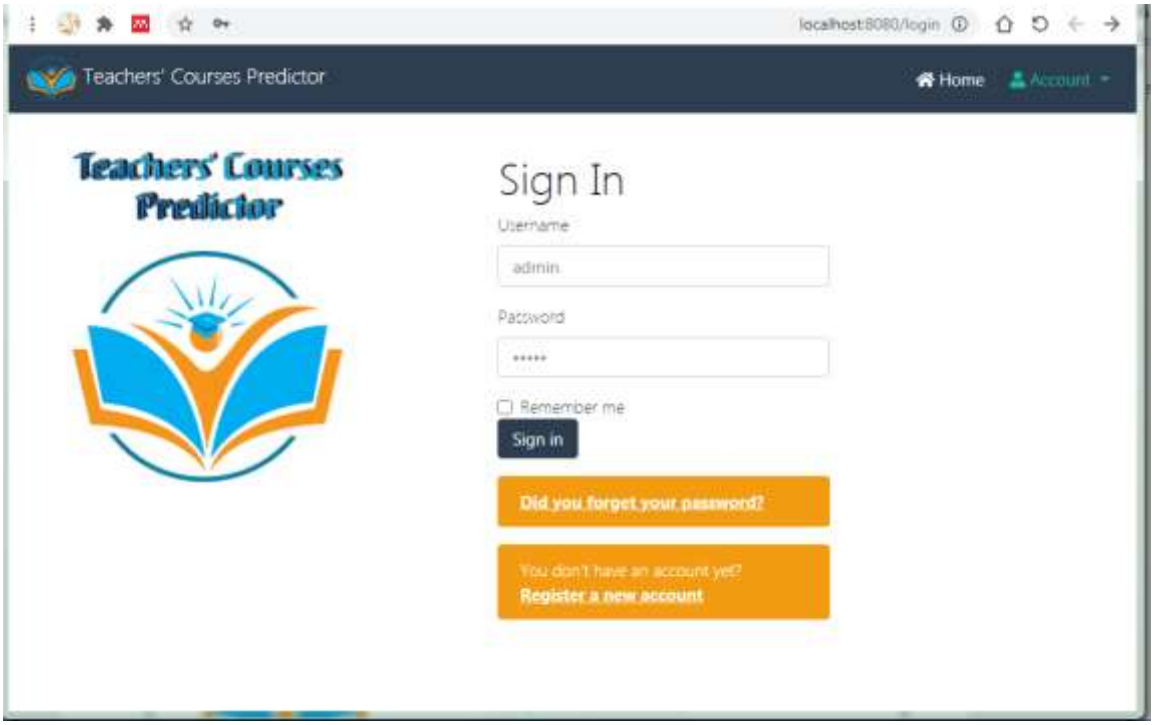
```
{
   "message": "Succeeded."
}
```

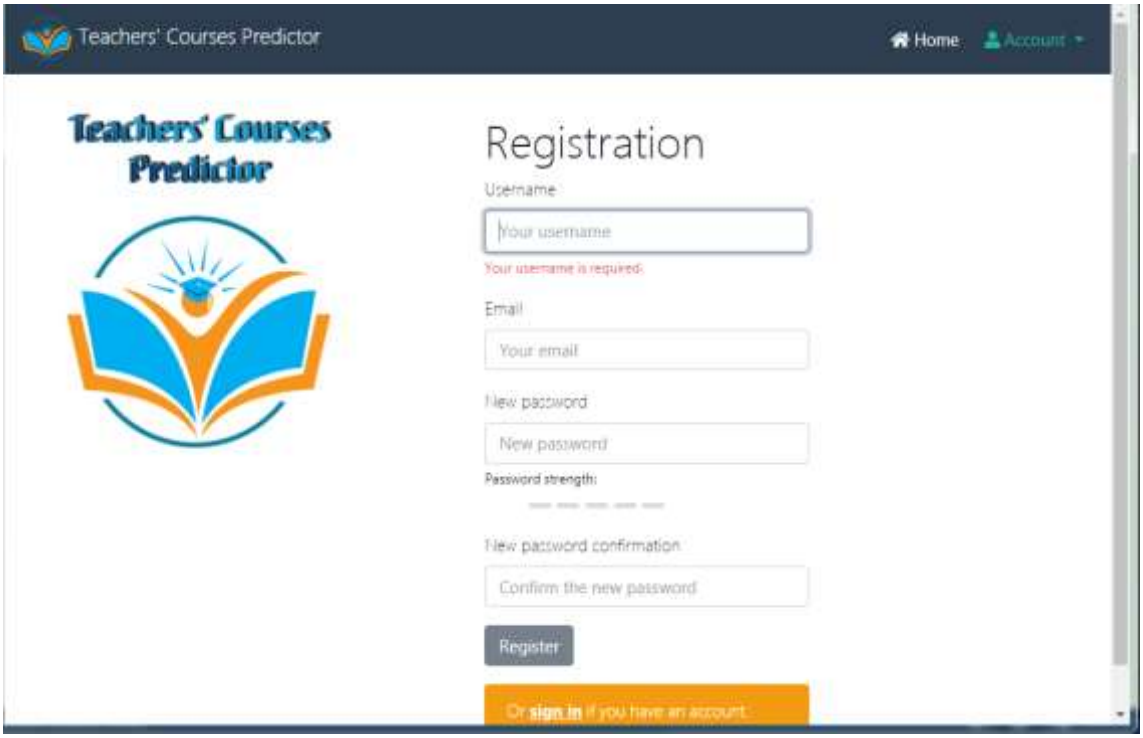
- On Failure: status 404 with the JSON in the body: (Message might vary with error cause variation).

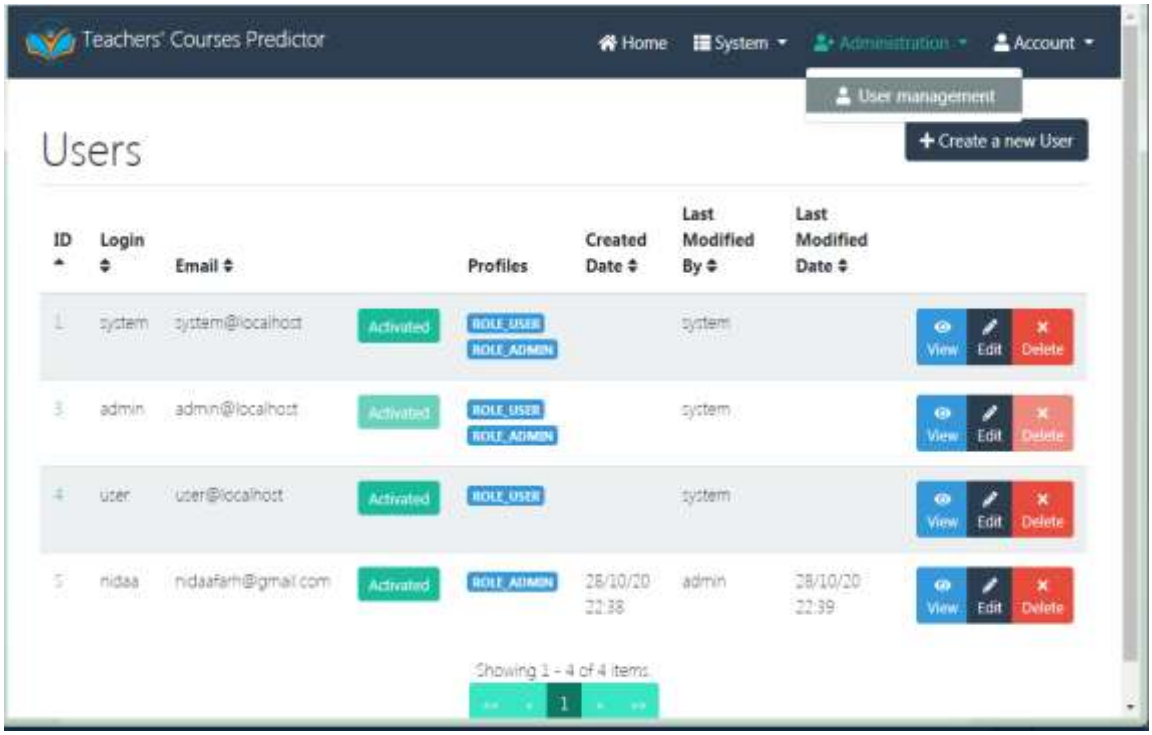
```
{
   "message": "Python server is down."
}
```

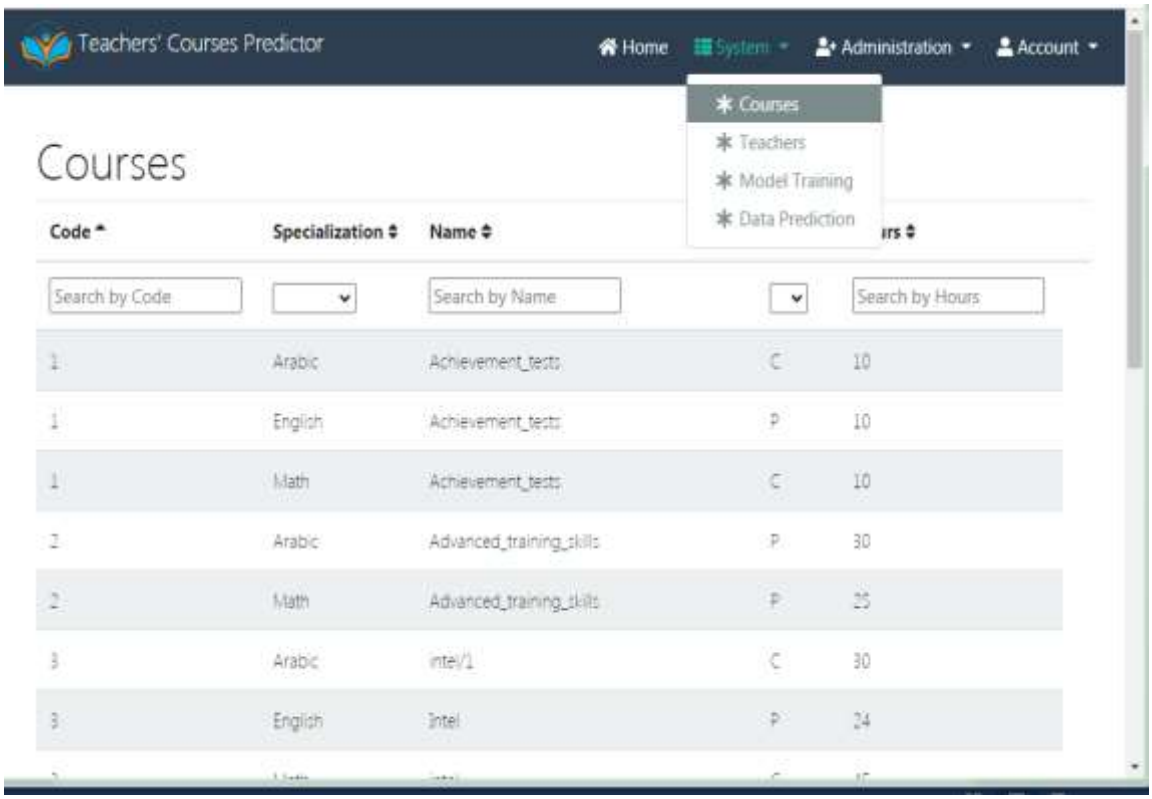
Appendix B

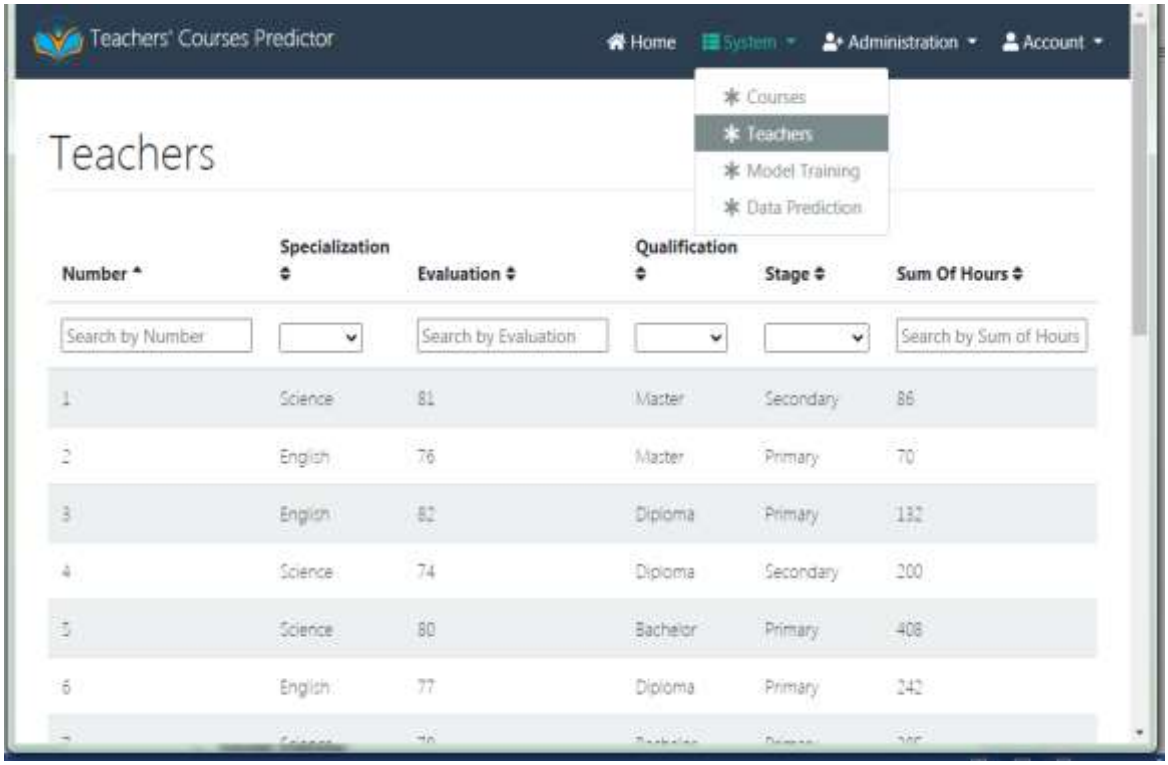
Smart Prototype Design:

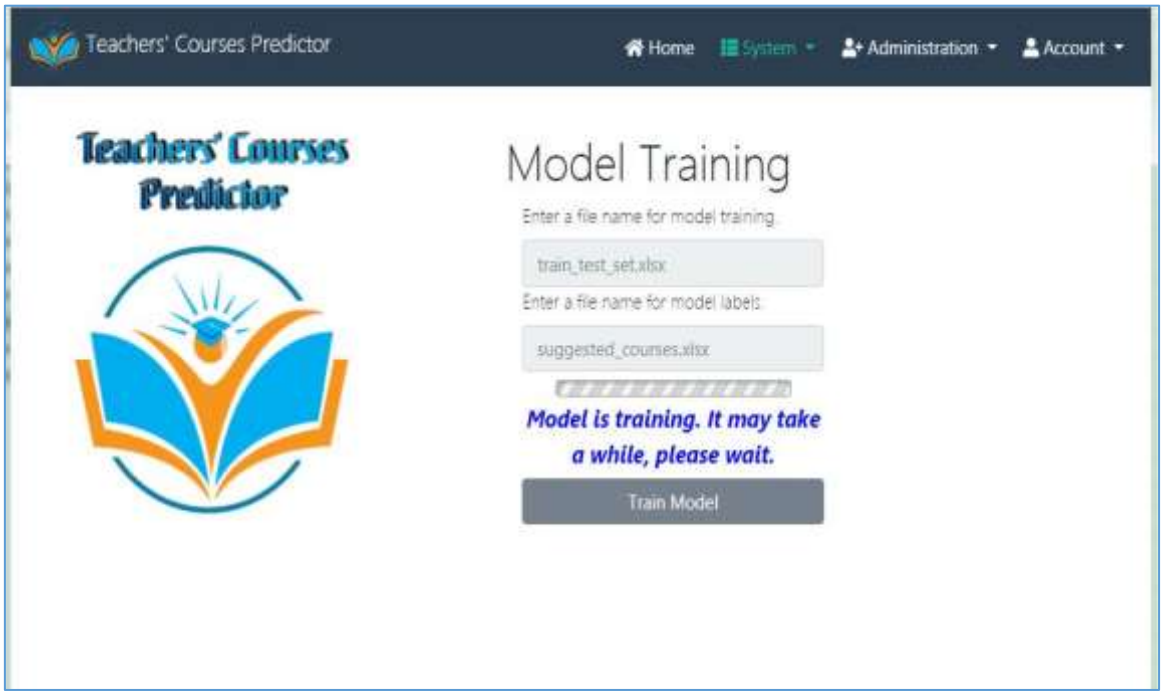


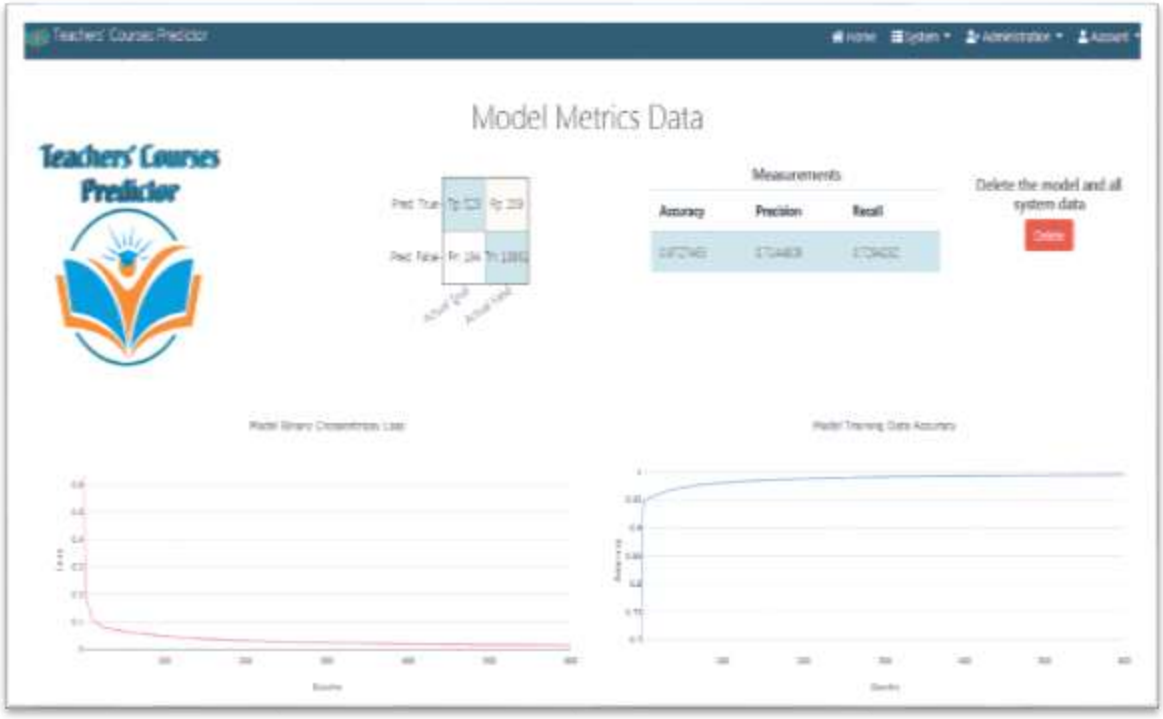


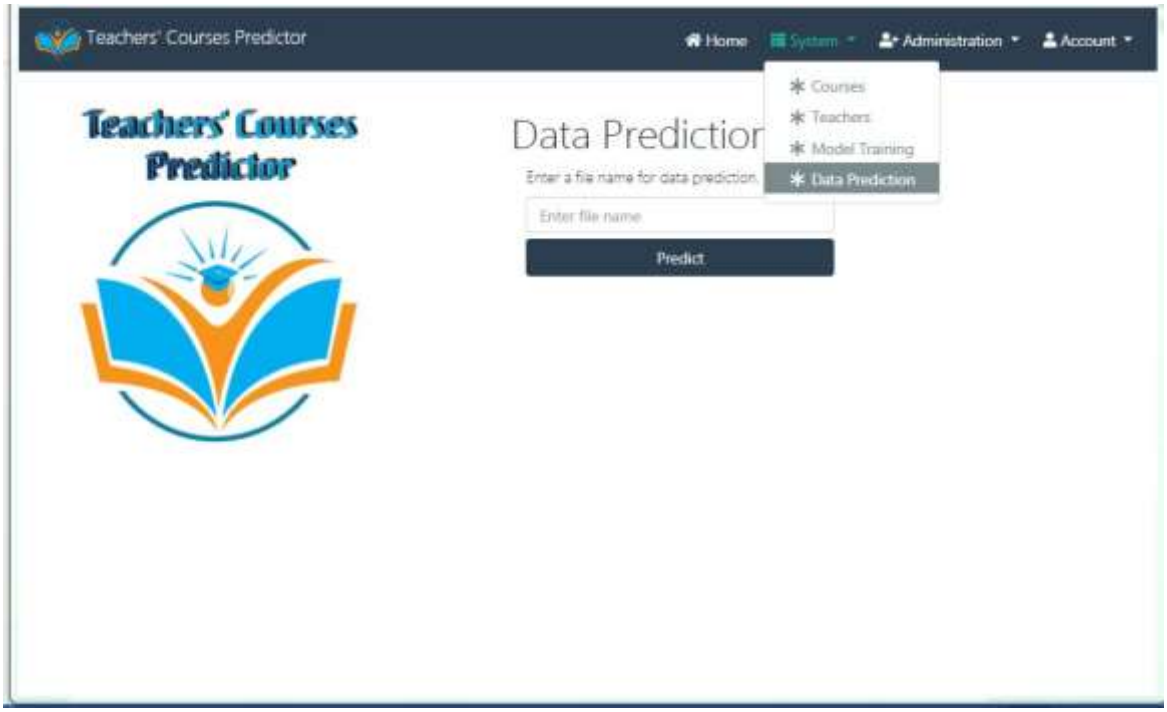

Teachers' Courses Predictor
Home
System
Administration
Account
User management
Users
+ Create a new User
ID
Login
Email
Profiles
Created Date
Last Modified By
Last Modified Date
1 system system@localhost Activated ROLE_USER ROLE_ADMIN system View Edit Delete
3 admin admin@localhost Activated ROLE_USER ROLE_ADMIN system View Edit Delete
4 user user@localhost Activated ROLE_USER system View Edit Delete
5 nidaa nidaafarh@gmail.com Activated ROLE_ADMIN 28/10/20 22:38 admin 28/10/20 22:39 View Edit Delete
Showing 1 - 4 of 4 items
1


Teachers' Courses Predictor
Home
System
Administration
Account
Courses
Teachers
Model Training
Data Prediction
Courses
Code
Specialization
Name
Search by Code
Search by Name
Search by Hours
1 Arabic Achievement_tests C 10
1 English Achievement_tests P 10
1 Math Achievement_tests C 10
2 Arabic Advanced_training_skills P 30
2 Math Advanced_training_skills P 25
3 Arabic intel/1 C 30
3 English Intel P 24


Teachers' Courses Predictor
Home
System
Administration
Account
Courses
Teachers
Model Training
Data Prediction
Teachers
Number
Specialization
Evaluation
Qualification
Stage
Sum Of Hours
Search by Number
Search by Evaluation
Search by Sum of Hours
1 Science 81 Master Secondary 86
2 English 76 Master Primary 70
3 English 82 Diploma Primary 132
4 Science 74 Diploma Secondary 200
5 Science 80 Bachelor Primary 408
6 English 77 Diploma Primary 242


Teachers' Courses Predictor
Home
System
Administration
Account
Teachers' Courses Predictor
Model Training
Enter a file name for model training
train_test_set.xlsx
Enter a file name for model labels
suggested_courses.xlsx
Model is training. It may take a while, please wait.
Train Model


Teachers' Courses Predictor
Home
System
Administration
Account
Model Metrics Data
Teachers' Courses Predictor
Measurements
Accuracy
Precision
Recall
Delete the model and all system data
Delete


Teachers' Courses Predictor
Home
System
Administration
Account
Courses
Teachers
Model Training
Data Prediction
Teachers' Courses Predictor
Data Prediction
Enter a file name for data prediction.
Enter file name
Predict

Appendix C

Facilitating student's mission:

Al-Quds University
College of Science & Technology
Jerusalem

جامعة القدس
كلية العلوم والتكنولوجيا
القدس

Date: 3/2/2020

To: His Excellency the Minister of Education

Subject: Facilitating student's mission

Dear Professor Marwan Awartani;

Greeting from Al-Quds University

The student Nedaa "Mohammad Ghazi" Hasan Farhaneh, ID number (984384099) and registration number (21710109), her email is nidaafarh@gmail.com, and mobile number 0569060450 is a student in Al-Quds university /Computer Science Department, majoring in Software Engineering Management, and she is preparing her thesis entitled:

"An intelligent framework for exploring influential patterns in assessing the performance of Palestinian Teachers and identifying their training needs using Data Mining techniques"

I hope you can agree to facilitate her mission in collecting the necessary data from the Ministry of Education in order to build her model.

Thank you for your help and support

Dr. Rashid Jayousi
Computer Science Department
Al-Quds University

بسم الله الرحمن الرحيم

State Of Palestine
Ministry of Education
Center for Educational Research and Development

دولة فلسطين
وزارة التربية والتعليم
مركز البحث والتطوير التربوي

الرقم : و ت / ٧١/١/١٣ التاريخ : 2 / 3 /2020م

القدس

لمن يهمّه الأمر

"تسهيل مهمّة بحثية (2)"

يهديكم مركز البحث والتطوير التربوي أطيب تحية، ويرجو منكم التكرم بتسهيل مهمة الطالبة:

" نداء "محمد غازي" حسن فرحانه"

من جامعة القدس–كلية العلوم والتكنولوجيا–أبو ديس، للحصول على المعلومات اللازمة لإعداد أطروحتها بعنوان:

" **An intelligent framework for exploring influential patterns in assessing the performance of Palestinian Teachers and identifying their training needs using Data Mining techniques**

ملاحظات:

- ستجمع الباحثة بيانات محوسبة (حيثما توفرت) من وزارة التربية والتعليم عن عينات من المعلمين مثل: المستوى الأكاديمي، التخصص، المراحل التي درسها، الدورات التي اجتازها، معدل تحصيل طلابه، والتقييم السنوي له.
- تتولى الباحثة أنشطة جمع البيانات بما لا يعيق سير العمل.

مع الاحترام،،

د. محمد مطر
/مدير مركز البحث والتطوير التربوي

دولة فلسطين
وزارة التربية والتعليم

نسخة: معالي وزير التربية والتعليم المحترم.
عطوفة وكيل الوزارة المحترم.
السيد مدير عام الشؤون الإدارية المحترم
السيد مدير عام التعليم العام المحترمة
السيد مدير عام تكنولوجيا المعلومات المحترم

Tel (+ 970-562-501092) E-mail (ncerd@moe.edu.ps)

إطار ذكي لتحديد الاحتياجات التدريبية للمعلمين الفلسطينيين باستخدام تقنيات التنقيب عن البيانات

إعداد: نداء فرحانه

إشراف: د. رشيد الجيوسي

الملخص

تهدف هذه الدراسة إلى وضع إطار عمل ذكي يعمل كنظام قائم على البحث عن البيانات المتعلقة بمعلمي اللغة العربية والعلوم والرياضيات ودراسات اللغة الإنجليزية ضمن معايير ومتغيرات محددة، من أجل التنبؤ بالاحتياجات التدريبية وتحديديها باستخدام خوارزمية الذكاء الاصطناعي للتنقيب عن البيانات وهو أمر ضروري لتعزيز أداء المعلم في البيئة الصفية.

أخذت الباحثة عينة من معلمي المدارس الحكومية في العام الدراسي 2015/2014 م، وحددت المتغيرات التالية: المؤهل الأكاديمي، التخصص، المراحل التي يقوم بتدريسها، التقويم السنوي، الدورات التدريبية وعدد ساعاتها في مجالات المحتوى وطرق التدريس.

تم إجراء البحث باستخدام (Angular) كأداة (JavaScript) للواجهة الامامية من جانب المستخدم، و(spring boot) الخادم في النظام والمتصل بقاعدة البيانات (MySQL)، وخادم (Flask python) المسؤول عن عملية التعلم الآلي للشبكة العصبية العميقة (DNN) كخوارزمية الذكاء الآلي، ويتم توصيل جميع الخوادم عبر API في بيئة (jhipster). وقد أظهرت نتائج تطبيق النموذج الأولي فاعلية عالية في تحديد الاحتياجات التدريبية لتحسين أداء المعلمين بطريقة منظمة وتفاعلية.

توصي الباحثة وزارة التربية والتعليم باستخدام أداة التكنولوجيا الذكية هذه لفعاليتها في التنبؤ بالدورات التدريبية اللازمة للمعلم لتحسين أدائه، وكذلك لتنظيم بيانات المعلّم بطريقة موجهة.