International Journal of Applied Engineering Research ISSN 0973-4562 Volume 13, Number 10 (2018) pp. 7621-7630

Prediction of General High School Exam Result Level Using Multilayer Perceptron Neural Networks

Mohammed Awad[1] and Ahmed Ewais[2]

[1]*Computer System Engineering Department, Arab American University, Palestine.*

[2]*Computer Science Department, Arab American University, Palestine.*

Abstract

The general high school exam is considered one of the most important exams for the student. The achievement of this academic qualification enables him to build his future career and determine the course of his life through joining a bachelor program in a university based on the marks obtained in this exam. On another hand, the universities determine the admission rates for each discipline and program depending on the general high school exam results in each year. Some universities with some specialization start the acceptance from high-level grade of high school exam and decrease it gradually to reduce the acceptance grade level depending on the demand for a specific specialization. A smart prediction model of the general high school exam results in next year will help the universities to determine the level from the beginning of the acceptance process. In this paper, we present a prediction model which has the ability to analyze previous patterns of general high school exam result and use them to predict the future results. The time series dataset consists of general high school exam results for the previous eleven years in Palestine, which was used as a target for the Multilayer Perceptron Neural Networks with Backpropagation learning algorithm (MLPBP) to predict the future general high school exam results. The accuracy of the prediction results of the proposed model is significant as it appears in the result of Minimum Mean Square Error (MSE). Moreover, the prediction result of each level in the next year is very accurate regarding the compiled pattern of the historical data.

Keywords: General high school exam levels, Prediction, Artificial Neural Networks, Backpropagation Algorithm

INTRODUCTION

Sometimes, the number of graduate students from general high school exceeds the capacity of universities. As a result, universities will be not able to accept all graduate students, from the general high school, applying for admission in different fields. So they resort to apply criteria for students' admission in the most requested specializations like medicine and its branches. The majority of universities adopted general high school exam rate. This admission procedure increases demand for educational institutions resulting in the increased student competition for seats in the confused requested specializations. Determining the minimum admission limits is better and it goes in the right direction to organize the enrollment process by high school students passed the general high school exam. Moreover, setting the acceptance level from the beginning will positively affect the students so that they will not wait for admission to a particular specialization based on the university decision to increase or decrease the admission or entrance grade level. Among the possible solutions for organizing process of students' admission to universities, prediction models for the future results of high school exam.

During the last decades, the universities in Palestine accepted the students who pass the general high school exam using the process of starting the acceptance in some specializations from high level of the general high school exam mark and then descending the required grade for acceptance till it reaches the available number of seats for each specialization. Such a process deprives and prevent many students from enrolling in their intended disciplines. This can be caused by a lack of information or forecasting results about the expected exam results for the general high school exam.

Time series forecasting got researchers' interest in recent years due to the wide variety of possible applications [1]. Therefore, a large number of techniques have been developed for predicting and modeling of time series. Some prediction strategies seek to approximate a model to a dynamic system by analyzing only the information contained in a series of time by assuming that this is sufficiently detailed to contain all the information required. Many time series possess behaviors that cannot be modeled using a linear model. However, non-linear techniques can be used for modeling and predicting time series such as autoregressive integrates moving average ARIMA [2] artificial neural networks (ANNs) [3]. The general formula of time series can be represented as:

$$\frac{dy(t)}{dt} = F(y(t)), \quad y(t) = [y_1(t), y_2(t), ..., y_d(t)] \tag{1}$$

Where F is a non-linear function, t is the time step, y is the output.

The application of artificial intelligence methods in education is increased [4,5,6] during previous years. This is because of the fact that artificial intelligence can recognize the pattern and useful knowledge to improve the education field. Artificial neural networks ANNs is one of the fields of artificial intelligence which is used as prediction tool. To do so, artificial neural networks (ANNs) consist of processing

units that are used to recognize patterns. One of its fundamental characteristics is that they have the ability to learn and improve their functioning [7]. This paper proposes the use of neural networks as the main method to predict the general high school exam since neural networks are able to recognize past patterns and replicate them to infer behaviors and future results.

In this research work, we will consider the results of general high school exam in both scientific and literary streams during previous years. The main goal of this work is to improve the resolution of outputs related to results of general high school exam in the future (next year). In other words, this applied research will help universities to make the suitable decisions regarding these general high school exam levels of Palestine that depends on the general high school exam level behavior. Therefore, our proposed model uses Multilayer Perceptron Neural Networks with Backpropagation Algorithm as the forecasting tool of the future values of the general high school exam result level. The model uses datasets which includes general high school exam results in the case of Palestine during eleven years ago. The dataset of each year in each stream is divided into 10 levels of the result, each level is categorized by 5 grades. The data of these divided levels will be used as a target for the neural network where the input values will be the time series which have 10 values in each year with a total of 110 values of time. The data is normalized between [0 1]. The neural network will predict the output and compare the result with the target output to find the mean square error (MSE).

This paper is structured as follow: section 2 introduces some related work of using of artificial intelligent in this field of education regardless of this applied research. Section 3 presents a general overview of the ANN algorithm. Section 4 shows the proposed prediction methodology. Both the prediction results and discussions are presented in section 5. Finally, section 6 concludes the paper and future work is presented.

RELATED WORK

The general high school exam level achievement is an important factor in deciding the academic life of any student. In fact, this exam works as a connection point between high school education and the selection of university specializations. The idea of developing a prediction model of the general high school exam level based on artificial neural networks for identifying the percentage of the general high school exam in each level very useful task. Taking into consideration, that this subject has not been investigated in this way previously. This article will present some related work that used artificial intelligence methods in the field of education to draw attention to the capabilities of artificial intelligence in helping to develop education. There is a wide range of studies on data mining applications and techniques which are used to discover knowledge, information, patterns in educational domain. Some of the possible data mining techniques are prediction and classification which are used extract patterns and predict trends from future data in educational context. Therefore, a number of studies compared the effectiveness of using an artificial intelligence method for different purposes and goals such as enhancing student performance, predicting failure, exploring factors that can enhance student's academic performance, etc.

Different data mining techniques have been used for classifications and prediction purposes in an educational context. For instance, Naïve Bayes is used in different models to classify and predict students' performance [8,9]. Another interesting research used decision tree algorithm to identify and predict the failure risk in primary school students [10,11], or to predict students graduation grade using decision tree-ID3 [10,5]

Neural Network is also important data mining technique that has been explored in the context of educational data mining. In particular, it has been used for predicting academic performance for the students. For instance, [12] claimed that Neural Network is capable of finding relationships between dependent and independent variables. As a result, it is considered one of the best prediction techniques. Researchers in [13] proposed a tool which used neural network technique to give insight about student performance and progress and recommend activities such as solving more exercises, reading learning resources, etc. Another data mining technique is the Support Vector Machine which is used when the dataset is small [14] and considered a fast and accurate data mining technique [15].

Other researchers proposed a combination of two or more data mining techniques [16]. For instance, a proposed model used two classification algorithms: decision tree algorithm and fuzzy genetic algorithm [17,18] Another interesting approach was proposed by [19] to use decision trees, random forests, neural networks and support vector machines were used to predict the relationship between the student's previous academic performance and future achievements. Authors in [20] found that the accuracy result was more than 71% by using both algorithms Random Forest and J48.An interesting work [21] compared the results out three algorithms extreme learning machine, neural networks, and support vector machines and the accuracy rate was promising and greater than 93%.

A recent review of the factors that are considered in the prediction process of the academic performance for the students has been presented in [22]. Based on the previous studies, the factors are categorized as follow: Internal Assessment such as quizzes, attendance, assignments, lab exercises, term exams, etc., External Assessment such as final exam of a specific topic, Demographic Data such as gender, family background, and age. Other identified factors, which are used in a limited number of research, are Psychometric Factors such as student interest, family support, engage and study time [20,23]. A number of researchers proposed that there is a major role for students previous results in predicting students grades [24]. For instance, the research [17] showed that depending on the coefficient correlation analysis, the students' cumulative grade point average considered as a significant factor used for predicting students' performance. Other work [18] proposed a prediction model which depends on parameters such as marks of two sessional exams and

assignments, admission marks and attendance marks. Other researchers [24] proposed applying Bayesian classification to forecast students' grades based on their results in a previous year and other possible related factors. The research results showed that students final results can be affected by factors such as mother education level, family income, and students' habits.

Different research work proposed different prediction techniques for different purposes and reasons. This research work used three data mining techniques (decision tree algorithm, Naïve Bayes Algorithm, and ZeroR algorithm) to predict final grade of primary school students. The work considered questionnaire which was distributed to 1000 students. The proposed work managed to predict the grades of the students with 85.53% accuracy rate [24]. The authors in [10] proposed a solution is intended to help instructors in avoiding the possible risk of students' failure in schools. Similarly, [18] proposed a prediction model for both bachelor and master students to identify student academic performance in specific subjects. The results of the proposed model can be used by instructors to improve students' performance. Also, it can be used by companies to find students according to job requirements. Other researchers claimed that the educational data mining can be a better resource for enhancing students performance and courses' developments [25,26], predict students drop out possibility [27] and support administration of higher education institutions [28].

Another interesting research [29] proposed a system using data mining techniques for predicting the chances for a student to get a scholarship based on a number of variables. As a result, students will be able to identify their weak areas and get proper guidance from academic staff to increase students' chance to get scholarships. Researchers in [30,31] proposed the use of educational data mining to identify the courses that can affect students' performance during their study in the bachelor program. The results showed that the second year courses could affect students' academic performance and also the mandatory courses can affect students' performance too. A more specific research work used data mining technique to predict students' performance in programming courses based on student's grades in other courses [32].

To the best of our knowledge, there is no research work conducted to predict general exam results in Palestine. There is some similar work conducted in different contexts such as [19,33,24] who proposed the use of decision tree technique to student final grade in primary and secondary schools. However, the proposed work was conducted in different countries and different examination systems from Palestine. More comprehensive review of proposed approaches and tools for educational data mining are presented in [4, 22].

ARTIFICIAL NEURAL NETWORKS (ANNS)

Artificial neural networks (ANNs) are learning systems inspired by the functionality of the human brain. They can simulate systems of non-linear relationships between the input and output variables. The general structure of the artificial neural network is a network consists of neurons, each neuron can process the information in a parallel and non-linear way. As a result, this makes it suitable for the prediction of non-linear problems [1,7].

ANNs are a powerful tool for the approximation of linear functions; its use is especially useful in modeling complex functions [34]. The application of this technique in the field of time series prediction has provided interesting results [3,34]. The main potential of ANNs is to detect nonlinearities in time series, which make it suitable for the prediction of the future result depending on the historical data of a specified problem. Time series values can be performed in a time step, which is obtained from samples available in time t it generates a value for time $t + 1$ [1,7]. The most used type of ANNs is the Multilayer Perceptron (MLPNN), this type consists of two hidden layers and input layer which receives the input time series data, the outputs layer. Each neuron in one layer is connected to other neuron in the next layer by weights, whose value is different for each of the connections and is determined through the learning process. MLPNN do both forwarding information and backward propagation process to optimize the connection weights. The general structure can be seen in Figure 1.

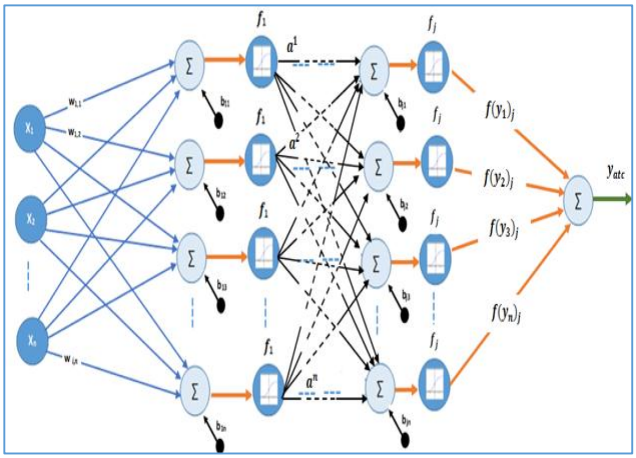

Figure.1: Multi-layer Perceptron NNs Topology

During the training process of the MLP neural network, the tuning of the weights Wji that connect input with hidden layers and hidden layers with output layer are iteratively determined using the learning backpropagation algorithms. The MLPNN produces an output depending on the summation of input data that multiplied by the tuning weight in each iteration. This output is compared with the target one to calculate the prediction error. The training process terminates if the MLPNN reproduces the known outputs for the input parameters. The most activation function used in MLPNN is the sigmoid function. After the training process termination, the second step is applying the testing data in order to evaluate the efficiency of the used learning algorithm.

The training process is a mapping process between input and output of the MLPNN. Input patterns provided to the MLPNN with initial weights usually random. The mathematical expression that describes the output of MLPNN

is illustrated in equation 2 as follow: wji are the connection weights between each layer, X_j is the inputs data variables which present here a time series values, b_i is the NN bias, and f is the activation function of the neuron.

$$y_{ia} = f(\sum_{j=1}^{m} w_{ij} x_j + b_i) \quad (2)$$

To critic the training process, there should be a threshold that depends on the error of the MLPNN, which presents the difference between objective output y_{id} and the current output y_{ia} of the i^{th} element. The following expression is used to calculate the error between output and target:

$$Er = y_{id} - y_{ia} \quad (3)$$

For the used MLPNN model, a general criterion is used as a termination condition to stop the prediction process. In this paper, the mean square error is used as presented in the expression:

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (y_{id} - y_{ia})^2 \leq \theta \quad (4)$$

Where n is the number of the input data (time step), and Θ is the threshold value of the prediction process. The training process continues to tune the connection weights until the error criteria are satisfied regarding Θ, the weight update is performed by the following expression:

$$\Delta\ w_{i+1} = \alpha . Er . x_i \quad (5)$$

Where α is the learning rate, normally between [0 1].

MLPNN has an advantage that it can predict any time series function. MLPNN learning process with backpropagation algorithm changes depending on the application . Any changes in the patterns of training require different training parameters of the MLPNN but the training process remains the same [34].

MODEL CONSTRUCTION

Developing a methodology for predicting general high school exam results levels must be as accurate as possible. Therefore, a number of steps have been conducted. First, a data collections step is performed, followed by initial preprocessing and normalization step. The next step is conducted by entering prepared data to the adopted learning algorithm. As a result of this step, the extracted patterns show a relationship between the values of previous high school exam results. Then, the Neural Networks technique is used to train the produced pattern of the previous high school exam results to predict the high school exam result levels for the next year. A final step shows the prediction results followed by discussion and analysis of the output. The general structure of the proposed model is depicted in Figure 2.

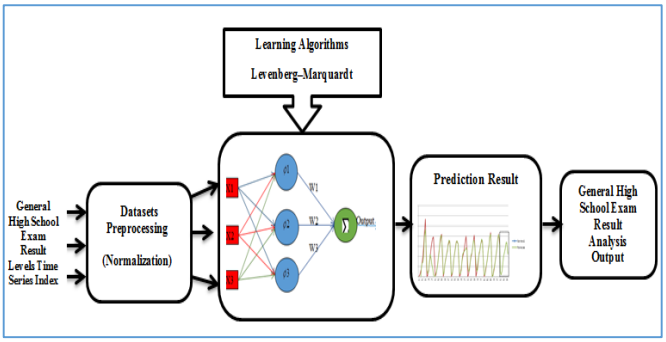

Figure.2: General Structure of the Proposed Model

A. *Datasets and Data Extraction*

The datasets of the two branches scientific branch and the literary branch are collected using historical data of the general high school exam result. Each branch results in each year, from 2006 to 2017, is divided into 10 levels as follow: 95-100, 90-95, 85-90, 80-85, 75-80, 70-75, 65-70, 60-65, 55-60, and 50-55. The dataset of all levels in all years are 110 data (the year 2006 to 2017) will be the target data. Therefore, the time series of 10 levels in each year will be the input to the MLPNN model after normalization step of the dataset. This data will be trained using 70% and tested using 30%. In this way, we used the pattern compiled from the previous 11 years to predict the general high school exam result in the next year.

B. *Model Methodology*

The main goal of time series is to design a model that can predict future unknown data from historical data by recognizing the pattern and minimizing the error function between output and target. To create an MLPNN model for general high school exam result levels, the model will have the ability to produce correct outputs. It is important to mention that the size of input data fed to the model can help to increase the accuracy of model results by providing an excellent level of similarity between predicted and actual output data. The proposed model uses Multilayer Perceptron Feed Forward Neural Networks with Back Propagation which depends on the following steps of training process:

Algorithm: The General Procedure That Was Used in Building the Prediction Model

Input: Datasets consists of time series of 110 points, target data which is the history of general high school exam result levels in the scientific branch, and literary branch, 110 points for each branch.

Output: prediction the general high school exam result levels in the next year;

Data Preprocessing:

Divide the data into the 2 main branches scientific and literary.

Extract the 10 levels of each branch.

Normalize (input time series, level values (target))

Initialize the weights wji of the MLPNN with random values.

Determine input pattern Xt: (Xt_1, Xt_2, ...,X tn).
Determine target output from the collected data for each branch levels.

Forward Phase: Feed Forward

For each iteration in the training process, do

Calculate the prediction output of the first hidden layer L1 of the MLPNN $out_{L1} = f^1(\sum_{i=1}^{n} X_i . w_{iL1})$

Calculate the prediction output of the second layer L2 use output of L1 as input $Y = f^2(\sum_{j=1}^{n} out_{L1} . w_{jL2})$

% Where f^1 and f^2 are the activation functions for output layer and hidden layer calculated using the following expressions:

$$f^1 = \frac{1}{1 + e^{-X}} \quad \text{and} \quad f^2 = X$$

Backsword Phase: Backpropagation

Calculate the error that using the following expression $\nabla E^i = \frac{\partial E}{\partial w^i} = \sum_{l=1}^{N} \text{Ja}\, {}^i_l (y'_l - y_l)$

Where E is error function, **J** is the Hessian matrix of a function f of n variables calculated as: $J^{ij} = \sum_{l=1}^{N} \left[\frac{\partial y'_l}{\partial w^i} \frac{\partial y'_l}{\partial w^j}\right] + \lambda I_M$, I_M is the identity matrix of order M, and λ is a parameter that makes a function similar to learning in the backpropagation algorithm., w_i is i^{th} element of input layer weight, w_j is j^{th} element of output layer weight, y'_l is the derivative output of l^{th} neuron. y_l is output of l^{th} neuron.

Update the weights using the recursive algorithm, starting with the output neurons and backward until reaching the input layer, and tuning the weights in the following way:

$$W(\mu) = W(\mu - 1) - J^{-1}(\mu - 1)\nabla E(\mu - 1)$$

Where μ-ith learning cycle.

Repeat the training process until getting the threshold

Return the output number of model, Result.

End

Test ← dataset.

Select the value of the testing data in each general high school exam result level.

Find the main of each level.

Present the next year prediction result**.**

The backpropagation algorithm calculates the gradient decent between the target and the predicted output using the new weights each backpropagation step. In this research work, one of the good convergence algorithms, which is called Levenberg Marquardt (LM) training algorithms [35], is adopted. It is based on mathematical methods for the optimization of gradient descent method and the Gauss-Newton method. This training algorithm performs prediction results problems methods. This algorithm is based on the idea of using an approximate Newton method in the nearest of the minimum and the method of the gradient in points away from it. The development begins by approaching the calculation of the Hessian [35].

EXPERIMENTAL RESULT

A prediction horizon of one next year results of the general high school exam using the patterns of the previous 11 years. The dataset of each year of the 11 years divided into 2 branches scientific, and literary. Then the data in each branch divided into 10 levels. Which produce 110-time series points in the all the eleven years for each branch. The model is executed in MATLAB 2017 under Windows 10 with processor i7. The result values are presented as; {# of neurons} the set of neurons used in each MLPNN. {# of Iterations} which presents the number of the execution cycle of the MLPNN. MSE_{Train} (train) is the mean squared error of the training. MSE_{Test} (prediction error) is the mean squared error of the testing. {levels} are the divided levels for the pass students in the general high school exam. {Future percentage for each level} is the values of prediction produced by the model for the next year. {Test error Percentage} the value of the error that the result in each level may be changed during ± of its value.

A. Scientific Branch

The scientific branch is the most important branch in the general high school. Usually the students in this branch have good scientific abilities. The success rate in this branch in general and the percentage of students who receive high marks are relatively higher than other branches. The dataset is divided into two set training 70% and testing 30%. Table 1 presents the training and testing MSE, with the number of neurons used in each execution and the number of epoch executed to get the error result.

Table1: Prediction result of the MLPNN for the Scientific Branch

# Of neurons	MSE train	MSE test	# of iteration
5	0.00436	0.0048	8
10	0.00423	0.0053	9
15	0.00157	0.0040	32
20	0.00161	0.0022	15
25	0.00046	0.0021	10
30	0.00052	0.0014	10
35	0.00016	0.0011	58
40	0.00014	0.0007	10
45	0.00004	0.0006	10
50	0.00003	0.0005	10

The training results with the error measurements of the trained and tested MLPNN architectures are shown in Table 1. The time series data used for this process corresponds to the general high school exam levels between 2006-2017 as shown in Figures 3 and 4. The first columns of Table 1 shows that the number of neurons used with the hidden layer in MLPNN. The best result of the testing MSE was obtained in the configuration of 50 neurons. The best result of the testing data obtained with 50 neurons. Depending on this result, Table 2 presents the percentage of the student for each level in the next year with ± the value of the MSE.

Table 2: The percentage of the students in each level for the Scientific Branch.

Levels	**Future Percentage For each level**	**Test error percentage**
50-55	1.17%	± 0.0005
55-60	0.20%	± 0.0005
60-65	3.40%	± 0.0005
65-70	6.56%	± 0.0005
70-75	9.00%	± 0.0005
75-80	13.07%	± 0.0005
80-85	14.26%	± 0.0005
85-90	18.62%	± 0.0005
90-95	19.85%	± 0.0005
95-100	11.04%	± 0.0005

The pattern obtained from MLPNN training and testing showed low errors both with the training data and testing data (prediction). Such results obtained when 50 neurons in the hidden layer are used. The performance of the MLPNN with an input time series data shown in Figures 3 and 4, where a good prediction is found between the real general high school exam levels and the prediction results of the MLPNN. Additionally, the applicability of the proposed model in future periods (next year) as shown in the last 10 result of Figure 4.

As it is already mentioned, the training has been carried out using a sample of 110-time series data to determine the optimal value of the parameters that define the MLPNN. In each iteration in terms of reduction of the mean square error is negligible from the 10 iterations. The Optimal number of neurons in the hidden layer is determined by the application of the incremental method. This means we start with a determined number of neurons and we increase this number in each experimental to get the best value of prediction.

Prediction the future data is not known by the MLPNN for the next year as shown in Table 2 and Figure 4. The proposed model has the ability to predict the future value with minimum MSE. The next year values start from 110 and end with 120 as seen at the right end of the curve of Figure 4.

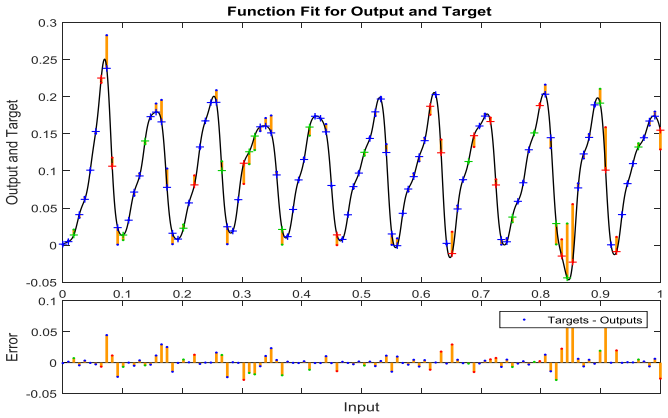

Figure. 3: Prediction result for the training phase using 50 neurons for the Scientific Branch

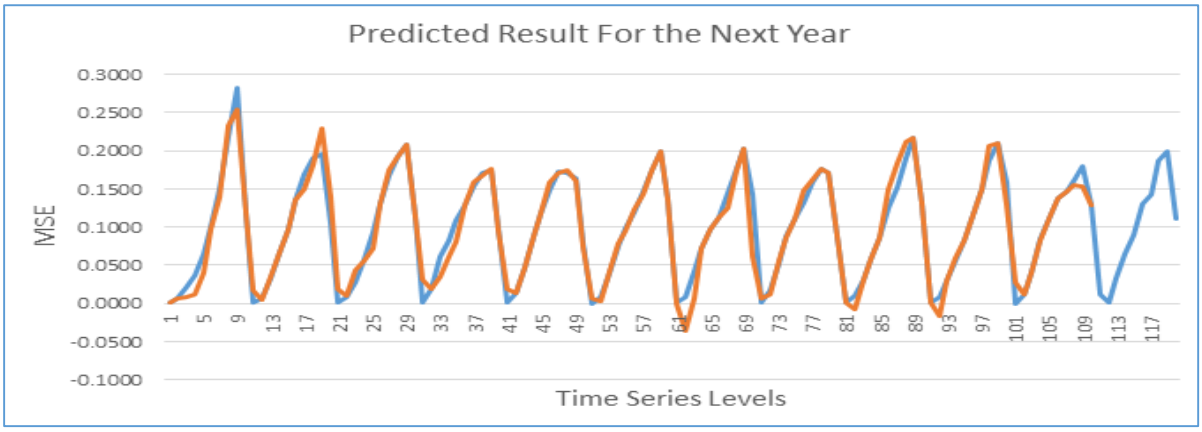

Figure. 4: The prediction result for the testing and the prediction values of the next year for the Scientific Branch

B. Literary Branch

In general, the success rate in this branch and the percentage of students who receive high marks are medium. The succeed students divided into 10 levels for each year from 2006 to 2017, with the difference between the level and the other 5 grades {95-100, 90-95, 85-90, 80-85, 75-80, 70-75, 65-70, 60-65, 55-60, and 50-55}. This process of data divided produced 10-time series point in each year, with 110-time series data for all years. The dataset is divided into two set training 70% and testing 30%. Table 1 present the training and testing MSE, with the number of neurons used in each execution and the number of epoch executed to get the error result.

Table 3: Prediction result of the MLPNN for the Literary Branch

# of neurons	MSE train	MSE test	# of iteration
5	0.00170	0.00177	13
10	0.00132	0.00158	8
15	0.00084	0.00206	17
20	0.00074	0.00163	10
25	0.00072	0.00138	14
30	0.00064	0.00068	8
35	0.00024	0.00051	10
40	0.00018	0.00086	10
45	0.000064	0.00084	9
50	0.000055	0.00038	11

As shown in Table 3, the error results of the training and testing data applied on MLPNN is enough to say that model has the ability to predict the next year of the general high school exam levels. The historical data between 2006-2017 is used to predict the future values of the general high school exam levels. As presented in Table 3, it is clear that with 50 neurons in the hidden layer of MLPNN, the training and testing error are sufficient to predict next values as shown in Figure 6. We can depend on these values of error with 50 neuron to de-normalize the data and get the percentage value of the general high school exam levels as shown in Table 4. Figure 5 shows a comparison between the output of the MLPNN for the series corresponding to Literary Branch and the output of the simulation model with 50 neurons in training phase

Table 4: The percentage of the students in each level for the Literary Branch

Levels	Future percentage for each level	Test error percentage
50-55	8.765%	± 0.00038
55-60	11.726%	± 0.00038
60-65	12.467%	± 0.00038
65-70	12.570%	± 0.00038
70-75	12.294%	± 0.00038
75-80	11.684%	± 0.00038
80-85	10.665%	± 0.00038
85-90	9.339%	± 0.00038
90-95	6.234%	± 0.00038
95-100	4.030%	± 0.00038

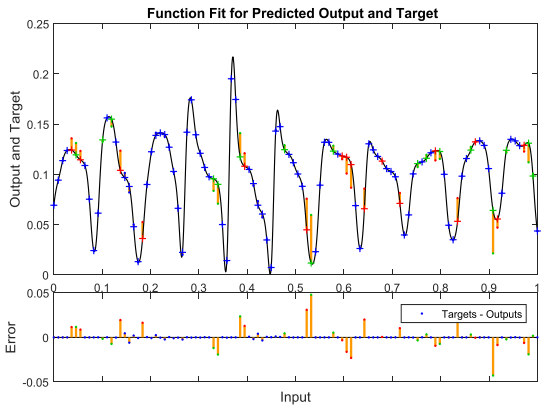

Fig.5: Prediction result for the training phase using 50 neurons for the Literary Branch

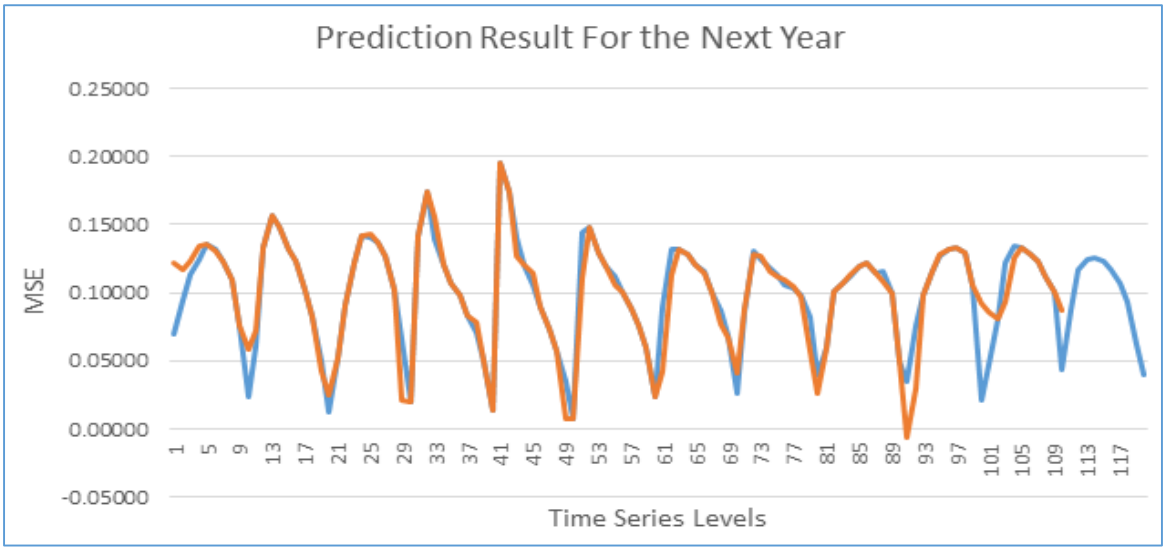

Figure. 6: The prediction result for the testing and the prediction values of the next year for the Literary Branch

Figure 6 shows a comparison between the output of the MLPNN for the series corresponding to Literary Branch and the output of the simulation model with 50 neurons. The predicted data values from 110 to 120 present the next year values of the general high school exam levels.

CONCLUSION

The main aim of this research is to predict the next year of the general high school exam levels in the two main branches depending on the grade level average. In order to organize the acceptance process during the student admission seasons in the universities, two different branches were investigated in this work. The branches are scientific and literary. The previous 11 years comprised the input and target data for the training and testing the prediction process. In general, MLPNN confirms the applicability of the MLPNN in prediction the general high school exam levels. The ability of this type of NN is to model complex systems and obtains small errors in training and testing (prediction) is highlighted. The proposed model becomes a useful tool for the universities; It delivers a percentage of each general high school exam levels that allow the universities to plan the acceptance rate and level with high better precision. The proposed model of neural network obtained excellent results from the testing data. The dataset composed of 110-time series data with 10 data level for each year. The most two selected branches of the general high school exam are applied to be predicted using the proposed model which are scientific, and literary. The Accuracy of the proposed model is very high 99.99% with testing error ± 0.0005.

REFERENCES

[1] Awad, M. (2017) 'Forecasting of Chaotic Time Series Using RBF Neural Networks Optimized By Genetic Algorithms', *International Arab Journal of Information Technology (IAJIT)*, 14(6).

[2] Fleifel, R. T. *et al.* (2017) 'LTE primary user modeling using a hybrid ARIMA/NARX neural network model in CR', *IEEE Wireless Communications and Networking Conference, WCNC*. doi: 10.1109/WCNC.2017.7925756.

[3] Awad, M. and Qasrawi, I. (2016) 'Enhanced RBF neural network model for time series prediction of solar cells panel depending on climate conditions (temperature and irradiance)', *Neural Computing and Applications*, pp. 1–12. doi: 10.1007/s00521-016-2779-5.

[4] Aziz, A. A. *et al.* (2013) 'Mining students' academic performance', *Journal of Theoretical and Applied Information Technology*.

[5] Ogunde and Ajibade (2014) 'A Data Mining System for Predicting University Students' Graduation Grades Using ID3 Decision Tree Algorithm Ogunde A. O 1 . and Ajibade D. A 1 .', *Computer Science and Information Technology*.

[6] Umair, Sajid, and M. M. S. (2018) 'Predicting Students Grades Using Artificial Neural Networks and Support Vector Machine', *Encyclopedia of Information Science and Technology*. IGI Global.

[7] Hamdan, I., Awad, M. and Sabbah, W. (2018) 'Short-Term Forecasting of Weather Conditions in Palestine Using Artificial Neural Networks', *Theoretical and Applied Information Technology*, 96(9).

[8] Osmanbegović, E. and Suljić, M. (2012) 'Data mining approach for predicting student performance', *Journal of Economics and Business*.

[9] Ramesh, V., Parkavi, P. and Ramar, K. (2013) 'Predicting Student Performance : A Statistical and Data Mining Approach', *International Journal of*

Computer Applications. doi: 10.1504/IJTEL.2012.051816.

[10] Márquez-Vera, C. *et al.* (2013) 'Predicting student failure at school using genetic programming and different data mining approaches with high dimensional and imbalanced data', *Applied Intelligence*. doi: 10.1007/s10489-012-0374-8.

[11] Badr, A., Din, E. and Elaraby, I. S. (2014) 'Data Mining : A prediction for Student ' s Performance Using Classification Method', *World Journal of Computer Application and Technology*. doi: 10.13189/wjcat.2014.020203.

[12] Wang, T. and Mitrovic, A. (2002) 'Using neural networks to predict student's performance', *Computers in Education, 2002.* doi: 10.1109/CIE.2002.1186127

[13] Livieris, I., Drakapoulu, K. and Pintelas, P. (2012) 'Predicting students' performance using artificial neural networks', in *Proceedings of the 8th Pan-Hellenic Conference "Information and Communication Technology in Education*.

[14] Hamalainen, W. and Vinni, M. (2006) 'Comparison of machine learning methods for intelligent tutoring systems', *Intelligent Tutoring Systems, Proceedings*.

[15] Sembiring, S. *et al.* (2011) 'Prediction of Student Academic Performance By an Application of Data Mining Techniques.', in *International Conference on Management and Artificial Intelligence (IPEDR)*

[16] Pandey, M. and Taruna, S. (2016) 'Towards the integration of multiple classifier pertaining to the Student's performance prediction', *Perspectives in Science*. doi: 10.1016/j.pisc.2016.04.076.

[17] Ibrahim, Z. and Rusli, D. (2007) 'Predicting students' academic performance: comparing artificial neural network, decision tree and linear regression', *21st Annual SAS Malaysia Forum, 5th September*.

[18] Hamsa, H., Indiradevi, S. and Kizhakkethottam, J. J. (2016) 'Student Academic Performance Prediction Model Using Decision Tree and Fuzzy Genetic Algorithm', *Procedia Technology*. doi: 10.1016/j.protcy.2016.08.114.

[19] Cortez, P. and Silva, A. (2008) 'Using Data Mining To Predict Secondary School Student Performance', *5th Annual Future Business Technology Conference*. doi: 10.13140/RG.2.1.1465.8328.

[20] Osmanbegović, E., Suljić, M. and Agic, H. (2014) 'Determining Dominant Factor For Student Performance Prediction by Using Data Mining Classification Algorithms', *Vitez-Tuzla-Zagreb-Beograd-Bukarest*.

[21] Tekin, A. (2014) 'Early Prediction of Students' Grade Point Averages at Graduation: A Data Mining Approach', *Eurasian Journal of Educational Research*. doi: 10.14689/ejer.2014.54.12.

[22] Shahiri, A. M., Husain, W. and Rashid, N. A. (2015) 'A Review on Predicting Student's Performance Using Data Mining Techniques', in *Procedia Computer Science*. doi: 10.1016/j.procs.2015.12.157.

[23] Abu Saa, A. (2016) 'Educational Data Mining & Students' Performance Prediction', *International Journal Of Advanced Computer Science And Applications*, 7(5), pp. 212–220. doi: 10.14569/IJACSA.2016.070531.

[24] Singh, M., Nagar, H. and Sant, A. (2016) 'Using Data Mining to Predict Primary School Student Performance', *Internation Journal Of Advance Research And Innovative Ideas In Education*, 2(1).

[25] Agarwal, S., Pandey, G. N. and Tiwari, M. D. (2012) 'Data Mining in Education : Data Classification and Decision Tree Approach', *International Journal of e-Education, e-Business, e-Management and e-Learning*.

[26] Al-Twijri, M. I. and Noaman, A. Y. (2015) 'A New Data Mining Model Adopted for Higher Institutions', in *Procedia Computer Science*. doi: 10.1016/j.procs.2015.09.037.

[27] Yadav, S. K., Bharadwaj, B. and Pal, S. (2012) 'Mining Education Data to Predict Student's Retention: A comparative Study', *International Journal of Computer Science and Information Security*. doi: 10.1016/j.aaspro.2015.01.034.

[28] Chalaris, M. *et al.* (2014) 'ScienceDirect ICININFO Improving Quality of Educational Processes Providing New Knowledge using Data Mining Techniques', *Procedia - Social and Behavioral Sciences*. doi: 10.1016/j.sbspro.2014.07.117.

[29] Khan, I. A. and Choi, J. T. (2014) 'An application of educational data mining (EDM) technique for scholarship prediction', *International Journal of Software Engineering and its Applications*. doi: 10.14257/ijseia.2014.8.12.03.

[30] Al-Barrak, M. A. and Al-Razgan, M. (2016) 'Predicting Students Final GPA Using Decision Trees: A Case Study', *International Journal of Information and Education Technology*. doi: 10.7763/IJIET.2016.V6.745.

[31] Altujjar, Y. *et al.* (2016) 'Predicting Critical Courses Affecting Students Performance: A Case Study', in *Procedia Computer Science*. doi: 10.1016/j.procs.2016.04.010.

[32] Badr, G. *et al.* (2016) 'Predicting Students' Performance in University Courses: A Case Study and Tool in KSU Mathematics Department', in *Procedia Computer Science*. doi: 10.1016/j.procs.2016.04.012.

[33] Khan, B., Khiyal, M. S. H. and Khattak, D. M. (2015) 'Final grade prediction of secondary school student using decision tree', *International Journal of Computer Applications*. doi: 10.5120/20278-2712.

[34] Kanchymalay, K. *et al.* (2017) 'Multivariate Time Series Forecasting of Crude Palm Oil Price Using Machine Learning Techniques', in *IOP Conference Series: Materials Science and Engineering*. doi: 10.1088/1757-899X/226/1/012117.

[35] Lourakis, M. (2005) *A brief description of the Levenberg-Marquardt algorithm implemented by levmar*, *Foundation of Research and Technology (FORTH)*. doi: citeulike-article-id:8257861.