Sigma J Eng Nat Sci, Vol. 39, No. 5, Special Issue 22, ICOMAA 2021, pp. 33–40, December, 2021

Sigma Journal of Engineering and Natural Sciences
Web page info: https://sigma.yildiz.edu.tr
DOI: 10.14744/sigma.2021.00039

Research Article

Modeling mathematics achievement with deep learning methods

Ibrahim DEMIR[1], Hasan Aykut KARABOĞA[2,*]

[1]Yıldız Technical University, Department of Statistics, Istanbul, Turkey
[2]Amasya University, Department of Educational Measurement and Evaluation, Amasya, Turkey

ARTICLE INFO

Article history
Received: 31 July 2021
Accepted: 01 November 2021

Key words:
Deep Learning, Elman method, Jordan method, PISA, Mathematics Achievement

ABSTRACT

Deep learning methods are the subfield of the machine learning models that have spread rapidly in the field of engineering in the last decade. But, these methods are a fairly new in educational literature. The aim of this study was modeling and predicting mathematics achievement of successful and unsuccessful students via deep learning methods. For this purpose, Turkey's Programme for International Student Assessment (PISA 2018) survey data was used. Deep learning methods were displayed comparable performance to multi-layer perceptron and logistic regression. Jordan neural network method was found the most successful method among Elman neural network, Logistic regression and multi-layer perceptron methods with 0.826 accuracy and 0.739 area under curve scores. It was understood that deep learning methods can be used in the modelling and predicting of students' mathematics achievement.

Cite this article as: Ibrahim D, Hasan A K. Modeling mathematics achievement with deep learning methods. Sigma J Eng Nat Sci 2021;39(Supp 22):33–40.

INTRODUCTION

International exams that evaluate the educational outcomes of countries by taking into account student and school performances like PISA. This exam has been administered by The Organisation for Economic Co-operation and Development (OECD) to 15-year-old students every 3 years since 2000 [1]. PISA consists of student, school and teacher questionnaires. The main purpose of the PISA survey is to ensure equality in education [1,2]. Student tests are contains questions from 3 key areas: reading, science and math. Its results are combined with various personal factors that provide important information to education authorities.

According to PISA 2009 results, Turkey ranked 32nd among 34 OECD countries in mathematics. In addition, 40% of the students could not reach the basic proficiency level in mathematical literacy [3]. It was ranked 48th out of 72 countries in mathematics in PISA 2015 [4]. According to the results of PISA 2018, Turkey ranked 42 in mathematics with an average of 454 points. But the OECD average was 459 points. Although Turkey has reached the

***Corresponding author.**
*E-mail address: h.aykut.karaboga@amasya.edu.tr

This paper was recommended for publication in revised form by Regional Editor Yusuf Zeren

Published by Yıldız Technical University Press, İstanbul, Turkey

highest average mathematics score level since 2003, this success is still not enough according to the educational investments [5].

As stated in many studies, variables such as the gender, age, family background, disability, self-confidence, attendance to classes, self-motivation, learning preferences, student goals, academic self-efficacy, socioeconomic and cultural status, quality of educational resources, home education resources, parents' education level play a vital role in mathematical performance [6–12]. Various studies show that digital technology is also important to increase the success rate in mathematical education [11].

However, Turkey is below the standards of OECD countries in terms of educational resources due to low financial resources in schools, low teacher salaries, high number of students per class and high number of students per teacher. In addition, most of the Turkish citizens are at a lower level than other OECD countries in terms of socio-cultural status and parental education level [13] and these situations affect students' mathematics achievement negatively. In order to reduce negative effects, it is necessary to determine the variables that affect students' achievement and to predict student achievement or failure. Modeling achievement or failure will allows education authorities to easily identify areas for improvement.

But, the structure of effective variables with complex relationships makes the analysis difficult with classical statistical methods. This encourages the use of artificial intelligence algorithms, mentioned as educational data mining, which includes advanced methods for modeling complex relationships [14]. Zawacki-Richter et al. stated that one of the most important research topics of the next 20 years as educational technology will be artificial intelligence applications [15]. The increase in the number of artificial intelligence studies in education supports this idea. However, most of the studies are related to the estimation and evaluation of student performance in higher education [16–18]. On the other hand, investments in education for primary education yield many times higher returns than investments made in higher education.

Besides, although artificial intelligence is a field that has become widespread especially in the last 30 years, there is no definite information about the pedagogical benefits of this field and its contributions to education. However, it should be known that artificial intelligence algorithms and methods based on artificial intelligence will make great contributions to increasing success in education. For this reason, it is estimated that the techniques mentioned will make great contributions to performance prediction [15,19]. Artificial intelligence techniques and algorithms used in education are grouped under the title of Educational Data Mining (EDM). The most popular up-to-date of EDM techniques is deep learning.

Although there have been many studies on deep neural networks (DNN), especially in the field of engineering, very few studies have been conducted on EDM. Coelho and Silveira addressed this issue in their review [20]. They searched a large database from many sources but, they found only 6 articles using DNN in the field of EDM. Studies with DNN have great potential in order to increase performance by predicting student success and to make development continuous.

DNN can perform prediction, classification or clustering operations with higher success than other EDM techniques. Unlike other EDM methods, DNN can easily model complex relationships and use error propagation more effectively to achieve higher accuracy than classical EDM [21]. Since the deep learning approach produces successful results in educational studies [22], it encourages researchers to use deep learning for EDM in the field of education [23].

The main purpose of this study is to predict mathematics achievement by using key factors of students' mathematics achievement with the DNN approach. We also compared our results with multilayer perceptron and logistic regression, which are the most widely used methods in EDM [24,25].

MATERIAL

In the literature, gender is an effective variable on mathematics achievement in which, it is stated that male students get higher scores than female students [6]. Mathematics study time is also an effective variable. As stated in the literature, as the working time increases, the success also increases [6,8]. Fear of failure is an effective variable on all educational careers of students. Students who are afraid of failure generally fail more. In addition, students' sense of belonging to the school also affects their success.

In addition, socio-economic status in developing countries such as Turkey has an impact on success [5,26]. Since educated and high-status families are more conscious, their children are also more successful. These parents' emotional support, father's education and index highest parental occupational status are effective variables on achievement. Also, having a room at home, having educational resources and computers to study affects achievement. The use of information and communication technology (ICT) resources in lessons also has an impact on achievement. These resources make it easy for students to understand topics that are difficult to understand.

Educational materials in home increase students' mathematics achievement [6,27]. Digital device usage is also important variable for learning mathematics [28]. In addition, the effect of information technology (IT) resources on mathematics achievement increases with the influence of educated parents [29].

In the light of this information obtained from the literature, variables related to mathematics achievement are given in Table 1.

Table 1. Mathematics Achievement Related Features

#	Variable Code	Description of Variables
y	PV1MATH	Mathematical Achievement
x_1	ST004D01T	Student Gender
x_2	MMINS	Mathematics Learning time per week (minutes)
x_3	GFOFAIL	General fear of failure
x_4	BELONG	Subjective well-being: Sense of belonging to school
x_5	FISCED	Father's Education
x_6	HISEI	Index highest parental occupational status
x_7	ESCS	Index of economic, social and cultural status
x_8	HOMEPOS	Home possessions
x_9	EMOSUPS	Parents' emotional support perceived by student
x_{10}	ICTHOME	ICT available at home
x_{11}	IC152Q02HA	Digital device used for learning or teaching during Mathematics lessons within the last month

The first variable in the Table 1 is mathematics achievement. The other variables in the table are gender, learning time per week, fear of failure, belonging to school, father's education, highest parental occupational status, ESCS index, home possessions, parents' emotional support, available ICT devices at home and digital device usage for learning or teaching during mathematics lessons in the last month respectively. These variables are assumed effective variables on mathematics achievement. Also, we used the most successful 30% and the most unsuccessful 30% of the students of PISA 2018 Turkey.

METHOD

Multilayer Perceptron (MLP): MLP has designated as a result of the studies done to solve the XOR Problem. MLP is a class of feed forward neural networks. This model is also called the 'Back Propagation Model' or 'Error Propagation Model' because it propagates the error to the network [30]. MLP works particularly well in classification and generalization situations. The structure of Multi-Layer Networks is as follows.

A MLP consists of at least three layers; input, hidden, and output layers. The input layer receives the data ($x_1, x_2, \ldots, x_n$)

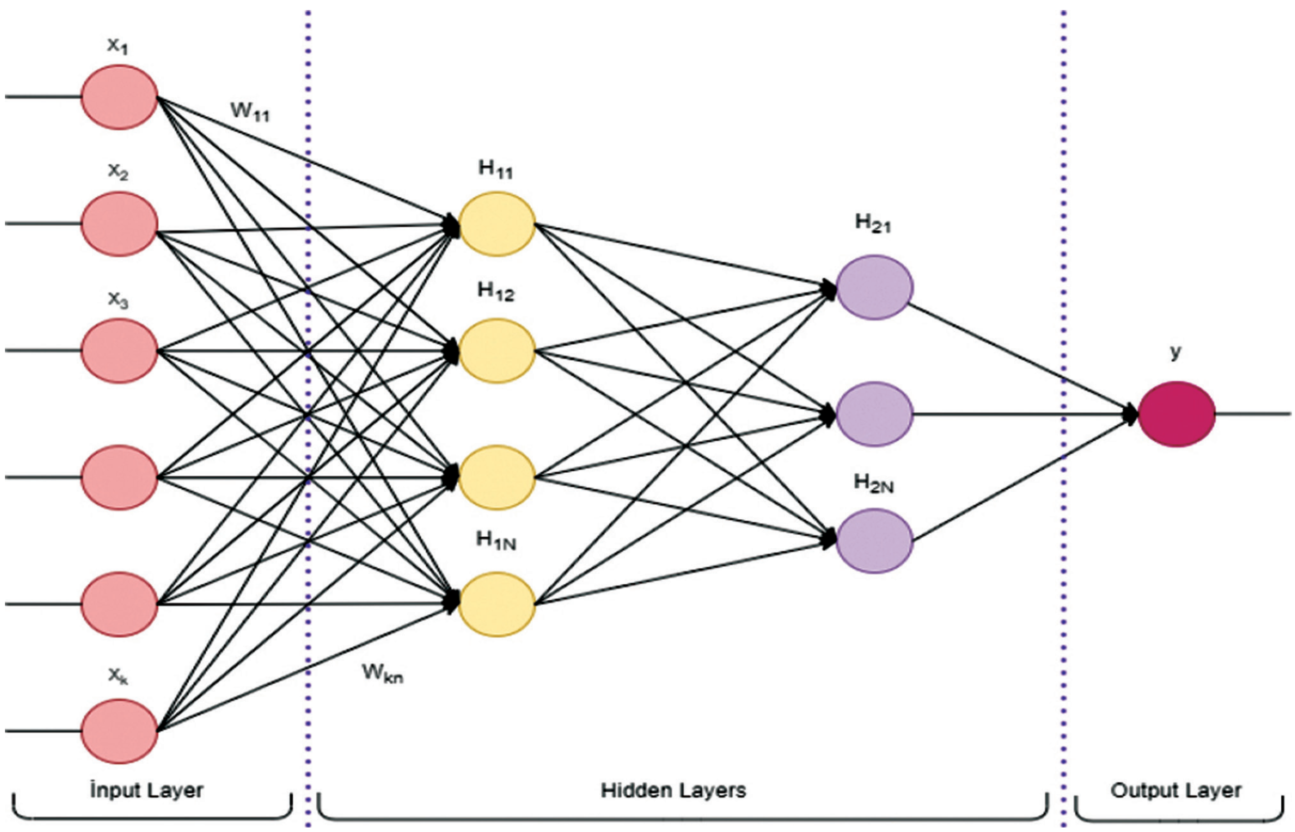

Figure 1. Architectural graph of MLP with two hidden layers.

values and sends it to the middleware. In this process, the input values are multiplied by the weights and combined via a function and transferred to the next layer. The output of the processing element is calculated by passing the value obtained as a result of the sum function through a linear function. Also, it can be calculated by a nonlinear differentiable transfer function. Since the weights represent the importance of the inputs in the model, the accuracy of the model depends on the optimum weights. The net function is usually obtained as the sum of the weighted inputs. The sum function and activation function are given by equations 1 and 2. Each node in a layer use a nonlinear activation function like sigmoid, tang, linear, threshold and hard limiter function [16].

$$net = \sum_{i=1}^{n} w_i x_i + b \tag{1}$$

$$y = f(net) = f\left(\sum_{i=1}^{n} w_i x_i + b\right) \tag{2}$$

$$f(x) = \frac{1}{1+e^{-x}} \tag{3}$$

In the hidden layer, the processed data from the input layer is transmitted to the output layer. The number of hidden layers changes according to the problem, at least one, and is adjusted according to the need. The output of each layer becomes the input of the next layer and there is no certainty about the number of neurons in the hidden layers. This is determined by the researcher. Each node in the hidden layers is connected to each node in the next layer with a certain weight. Since the most used activation function in hidden layers is the sigmoid function, we used sigmoid function in this study. The sigmoid function is given by equation 3.

The output layer processes data from previous layers and determines the output of the network. In the output layer, a linear or soft activation function is generally used. The output number of the system is equal to the number of elements in the output layer. In general, process in MLP takes place in two stages, forward calculation and backward calculation. Therefore, such algorithms are also called feedforward backpropagation artificial neural network algorithms [17].

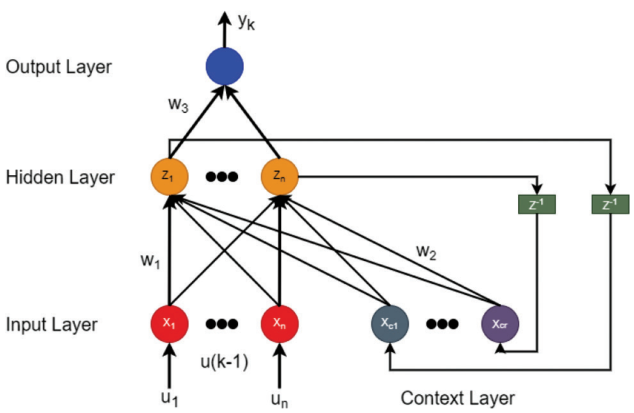

Figure 2. Architectural graph of Elman Neural Network.

Elman Neural Network (ENN): The ENN was developed by Elman in 1990. This method is a kind of back-propagation neural networks (NNs). And, these are constituted by a large number of neurons with certain rules. As a recurrent network type with a context layer as a self-referential layer, ENN is trained in a supervised manner based on the inputs and targets [31,32].

As shown in Fig. 2, ENN has four layers; namely, input, hidden, context and output layers. The input layer, the hidden layer and the output layer are able to be taken into consideration as a feed-forward network, which is similar when compared to the traditional MLP. Besides, there exists another layer named the context layer, which store and use the hidden layer's previous step output values. So, content layer elements carry the activation values from the hidden layer as input to the next iteration.

We suppose r neurons with n inputs produce m outputs in hidden and context layers. Then,

$$h_t = \sigma_h\left(w_2 x_{ct} + w_1 u_t + b_h\right) \tag{4}$$

$$x_{ct} = h_{t-1} \tag{5}$$

$$y_t = \sigma_y\left(w_3 h_t + b_y\right) \tag{6}$$

Where, w_1 is the weight between input and hidden layers, while w_2 is weight from context layer to hidden layer and, w_3 is weight between hidden and output layers. Inputs are represented with u_t and hidden layer outputs represented with h_t. Outputs of the model are represented with y and context layer outputs are represented with x_c.

Jordan Neural Networks (JNN): It is one of the first iterative networks. Like ENN, the JNN is a multi-layered backpropagation (recursive) neural network. In JNN, which have a structure similar to MLP, there are special feedback connection elements called state units in addition to input, output and hidden layers. State unit elements carry the activation values they receive from the output layer as input to the next iteration [33].

Each state unit element has a self-feedback connection. The weights of the connections between the state unit and the output unit elements are fixed to +1. Therefore, learning in JNN occurs in the connections between the input and the hidden layers, and between the hidden and the output layers [33].

Processor elements in other layers are similar to MLP nodes. Similar to MLP, both linear and nonlinear activation functions can be used in the hidden layer, and the learning rules used in MLP can also be used in the training of JNN [30]. The equations used in the JNN are as follows.

$$h_t = \sigma_h \left(w_2 y_{t-1} + w_1 u_t + b_h \right) \tag{7}$$

$$y_t = \sigma_y \left(w_3 h_t + b_y \right) \tag{8}$$

ENN and JNN are similar to each other, but there are two important differences between them. First of all, the activation values used in the context layer in ENN are taken from the hidden layer, not from the output layer. The second is that the context layer elements do not have self-feedback connections. In ENN, the connection weights between hidden and context layer elements are fixed and equal to +1. Generalized delta learning rule is used in ENN as in MLP.

Logistic Regression: Logistic regression, which produces very successful results in comparisons with machine learning methods, is one of the important statistical techniques [34]. This is a special case of regression analysis. Logistic regression method is widely used in educational applications to determine the factors affecting student achievement [35–37]. Regression assumptions are same with classical regression in logistic regression. Only difference from the classical regression is dependent variable. Because in logistic regression, the dependent variable is categorical [38,39].

RESULTS AND DISCUSSION

The comparison matrix given in Table 2 is used to compare model achievements. Comparison criteria are used in order to understand the results obtained in Confusion matrix correctly. Comparison criteria are calculated using the True Positive (TP), False Negative (FN), False Positive (FP) and True Negative (TN) values given in this matrix. In this study, we examined Accuracy, Error Rate, Sensitivity, Specificity and False Positive Rate values.

Accuracy rate refers to the total rate of predictions with correct predictions. The Error Rate value is the complement of the accuracy. Sensitivity is a measure of true positive rate in a model. Specificity measures how exact the assignment to the positive class. False positive rate measures the false negative rate in negative predicted class. The F1 score is an equilibrium measure between positive predictive value and sensitivity [34,38]. The formulations of these criteria are given in equations (9)-(14).

$$Accuracy = \frac{TP + TN}{p + n} \tag{9}$$

$$Error\ Rate = \frac{FP + FN}{p + n} = 1 - Accuracy \tag{10}$$

$$Sensitivity = \frac{TP}{p} \tag{11}$$

$$Specificity = \frac{TN}{n} \tag{12}$$

$$False\ Positive\ Rate = \frac{FP}{n} \tag{13}$$

$$F1 - Score = \frac{2 * TP}{2 * TP + FP + FN} \tag{14}$$

Table 2. Comparison Matrix for Two Classes

	Predicted Class		
Real Class	**Positive**	**Negative**	***Total***
Positive	True Positive (TP)	False Negative (FN)	*p*
Negative	False Positive (FP)	True Negative (TN)	*n*
Total	*p'*	*n'*	

Table 3. Comparison of the Mathematics Achievement Class Prediction Models

	Elman		**Jordan**		**MLP**		**Logistic**	
	Train	*Test*	*Train*	*Test*	*Train*	*Test*	*Train*	*Test*
Accuracy	0.706	0.711	0.716	**0.826**	0.698	0.705	0.684	0.671
Error Rate	0.294	0.289	0.284	**0.174**	0.302	0.295	0.316	0.329
Sensitivity	0.706	0.711	0.716	**0.826**	0.698	0.705	0.684	0.671
Specificity	0.705	0.708	0.717	**0.869**	0.698	0.706	0.684	0.673
False Positive Rate	0.295	0.292	0.283	**0.131**	0.302	0.294	0.316	0.327
F1-Score	0.706	0.711	0.716	**0.826**	0.698	0.705	0.684	0.671

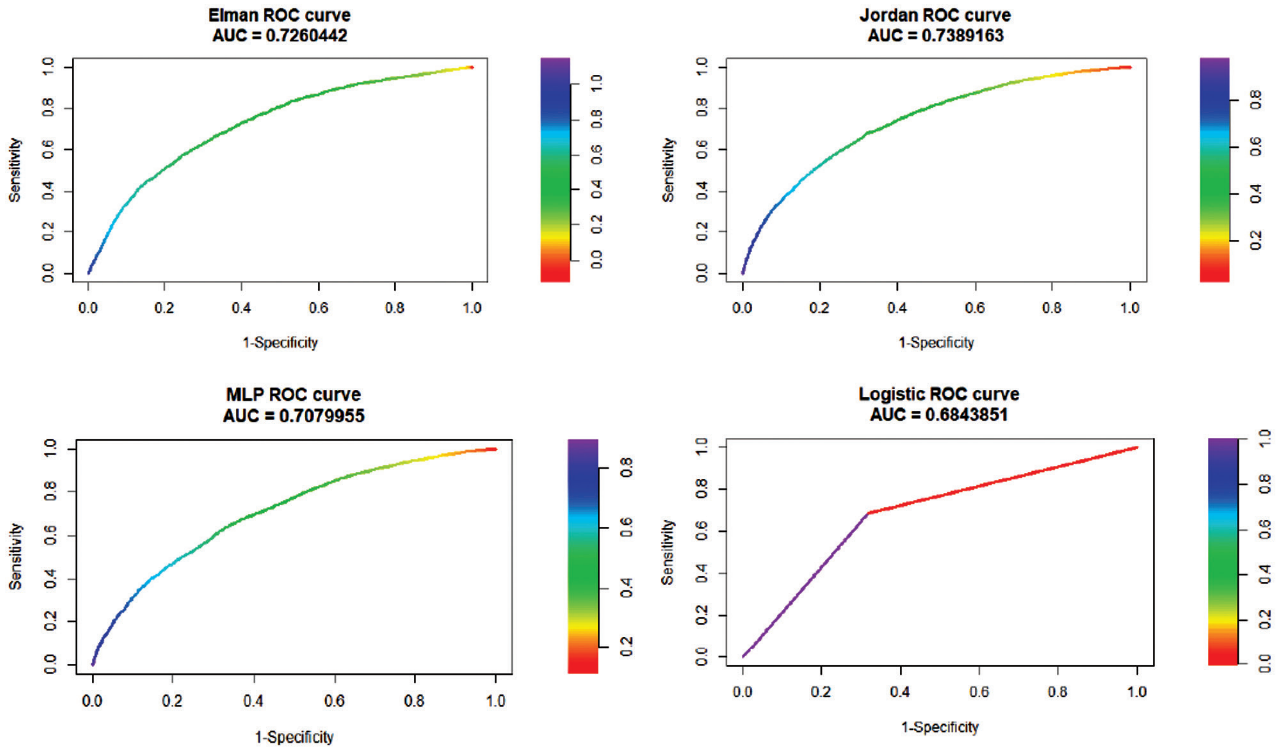

Figure 3. ROC Curves of the Models.

In Table 3, the comparison results of the models according to the training and test sets are given. The result of the most successful method according to the test set is marked in bold. According to the results given, it is understood that the most successful method is JNN.

In addition, the ROC curves of the models are given in Figure 3. ROC curves are a benchmark for reliable comparison of models. The sum of the area under the curve (AUC) is used as the comparison value of the curves. According to the given values, it was seen that the Jordan method produced successful results.

As a result of the analysis, it was observed that the students' gender, sense of belonging to the school, socio-economic status, household items, emotional support of the family, digital devices at home and use of digital devices in lessons were effective on success, respectively. In addition, DNN algorithms produced more successful results in modeling student mathematics achievement than MLP and logistic regression methods.

CONCLUSION

In this study, we compared machine learning methods against deep learning algorithms on PISA 2018 Turkey dataset to explore the DNN classification performance as EDM for predicting mathematics achievement of successful or unsuccessful students. We selected Elman neural networks and Jordan neural networks as DNN, and multi-layer perceptron and logistic regression as classical machine learning methods. According to the analysis results, the Jordan neural networks produced more successful results than other methods with 0.826 predictive accuracy and 0.739 area under curve scores. Multilayer perceptron and logistic regression methods also produced results close to DNN algorithms. According to the results of the analysis, it was understood that deep learning methods can be used in the classification and prediction of mathematics achievement. Also DNN methods displayed comparable performance to multi-layer perceptron and logistic regression.

According to the results, it was seen that the success of the student can be predicted if the value of the factors affecting the mathematics achievement is known. However, the study has several limitations. First of all, it is necessary to correctly determine the variables that affect success. In addition, it is recommended to conduct studies on different data sets in order to examine the success of deep learning methods compared to other machine learning algorithms in more detail. In addition, increasing the size of the data set and the number of variables may also affect the results. The data set should be expanded with new variables and model successes should be examined in larger and smaller data sets.

AUTHORSHIP CONTRIBUTIONS

Authors equally contributed to this work.

DATA AVAILABILITY STATEMENT

The authors confirm that the data that supports the findings of this study are available within the article. Raw data that support the finding of this study are available from the corresponding author, upon reasonable request.

CONFLICT OF INTEREST

The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

ETHICS

There are no ethical issues with the publication of this manuscript.

REFERENCES

[1] OECD. PISA 2018 Assessment and Analytical Framework. OECD; 2019.

[2] OECD. PISA 2018 Results (Volume I): What Students Know and Can Do. Paris: OECD Publishing; 2019.

[3] Blanchy NK, Şaşmaz A. PISA 2009: Where does Turkey stand. Turkish Policy Quarterly 2011;10:126–134.

[4] OECD. PISA 2015 Technical Report. 2017.

[5] MEB. PISA 2018 Turkiye Ön Raporu. Ankara: Milli Eğitim Bakanlığı; 2019.

[6] Demir İ, Ünal H, Kılıç S. The effect of quality of educational resources on mathematics achievement: Turkish Case from PISA-2006. Proced Soc Behav Sci 2010;2:1855–1859. [CrossRef]

[7] Jansen BRJ, Louwerse J, Straatemeier M, Van der Ven SHG, Klinkenberg S, Van der Maas HLJ. The influence of experiencing success in math on math anxiety, perceived math competence, and math performance. Learn Individ Differ 2013;24:190–197. [CrossRef]

[8] Kilic S, Cene E, Demir I. Comparison of learning strategies for mathematics achievement in turkey with eight countries. Educ Sci Theory Pract 2012;12:2594–2598.

[9] Gürsakal S. Pisa 2009 Öğrenci başarı düzeylerini etkileyen faktörlerin değerlendirilmesi. Suleyman Demirel University Journal of Faculty of Economics & Administrative Sciences 2012;17:441–452.

[10] Khanna L, Singh SN, Alam M. Educational data mining and its role in determining factors affecting students academic performance: A systematic review. 2016 1st India International Conference on Information Processing (IICIP), Delhi, India: IEEE; 2016, p. 1–7. [CrossRef]

[11] Pierce R, Stacey K, Barkatsas A. A scale for monitoring students' attitudes to learning mathematics with technology. Comput Educ 2007;48:285–300. [CrossRef]

[12] Koğar H. PISA 2012 matematik okuryazarlığını etkileyen faktörlerin aracılık modeli ile İncelenmesi. Educ Sci 2015;40:45-55.

[13] Aydin A, Erdagf C, Tas N. A Comparative Evaluation of Pisa 2003-2006 Results in Reading Literacy Skills: An Example of Top-Five OECD Countries and Turkey. Educ Sci Theory Pract 2011;11:665–673.

[14] Sun Y, Li Q. The Application of Deep Learning in Mathematical Education. 2018 1st IEEE International Conference on Knowledge Innovation and Invention (ICKII), 2018, p. 130–133. [CrossRef]

[15] Zawacki-Richter O, Marín VI, Bond M, Gouverneur F. Systematic review of research on artificial intelligence applications in higher education – where are the educators? Int J Educ Technol High Educ 2019;16:39. [CrossRef]

[16] Akgün E, Demir M. Modeling course achievements of elementary education teacher candidates with artificial neural networks. Int J Assess Tools Educ 2018:491–509. [CrossRef]

[17] Bahadır E. Using neural network and logistic regression analysis to predict prospective mathematics teachers' academic success upon entering graduate education. Educ Sci-Theor Pract 2016;16:943–964. [CrossRef]

[18] Agaoglu M. Predicting instructor performance using data mining techniques in higher education. IEEE Access 2016;4:2379–2387. [CrossRef]

[19] Rodrigues MW, Isotani S, Zárate LE. Educational Data Mining: A review of evaluation process in the e-learning. Telem Inform 2018;35:1701–1717. [CrossRef]

[20] Coelho OB, Silveira I. Deep Learning applied to Learning Analytics and Educational Data Mining: A Systematic Literature Review. Brasil: Recife, Pernambuco; 2017:143. [CrossRef]

[21] Kreinovich V. From Traditional Neural Networks to Deep Learning: Towards Mathematical Foundations of Empirical Successes. In: Shahbazova SN, Kacprzyk J, Balas VE, Kreinovich V, editors. Recent Developments and the New Direction in Soft-Computing Foundations and Applications, vol. 393, Cham: Springer International Publishing; 2021:387–397. [CrossRef]

[22] Doleck T, Lemay DJ, Basnet RB, Bazelais P. Predictive analytics in education: a comparison of deep learning frameworks. Educ Inf Technol 2020;25:1951–1963. [CrossRef]

[23] Shin D, Shim J. A systematic review on data mining for mathematics and science education. Int J Sci Math Educ 2021;19:639–659. [CrossRef]

[24] İnal H, Turabik T. Matematik başarısını etkileyen bazı faktörlerin yordama gücünün yapay sinir ağları ile belirlenmesi. Uşak Üniversitesi Eğitim Araştırmaları Dergisi 2017;3:23–50. [CrossRef]

[25] Filiz E, Öz E. Educational data mining methods for TIMSS 2015 mathematics success: turkey case. Sigma J Eng Nat Sci 2020;38:963-977.

[26] Anil D, Özkan YÖ, Demir E. PISA 2012 Araştırması Ulusal Nihai Rapor 2012:198.

[27] Azina IN, Halimah A. Student factors and mathematics achievement: evidence from TIMSS 2007. EURASIA J Math Sci Tech Ed 2012;8:249–255. [CrossRef]

[28] Tan CY, Hew KF. The impact of digital divides on student mathematics achievement in confucian heritage cultures: A critical examination using pisa 2012 data. Int J Sci Math Educ 2019;17:1213–1232. [CrossRef]

[29] Tan CY, Hew KF. Information technology, mathematics achievement and educational equity in developed economies. Educ Stud 2017;43:371–390. [CrossRef]

[30] Kröse B, Smagt P van der, Smagt P. An introduction to Neural Networks. 8th ed. Amsterdam: The University of Amsterdam; 1996.

[31] Goodfellow I, Bengio Y, Courville A. Deep learning. Cambridge, Massachusetts: The MIT Press; 2016.

[32] Ren G, Cao Y, Wen S, Huang T, Zeng Z. A modified Elman neural network with a new learning rate scheme. Neurocomputing 2018;286:11–18. [CrossRef]

[33] Pham DT, Karaboga D. Training Elman and Jordan networks for system identification using genetic algorithms. Artif Intell Eng 1999;13:107–117. [CrossRef]

[34] Karaboga HA, Gunel A, Korkut SV, Demir I, Celik R. Bayesian network as a decision tool for predicting ALS disease. Brain Sci 2021;11:150. [CrossRef]

[35] Boulhrir T. Using binary logistic regression to predict long-term effects of early-age home literacy environment on reading motivation. Education 2017;5:858–862. [CrossRef]

[36] Alija S, Snopce H, Aliu A. Logistic regression for determining factors influencing students' perception of course experience. Eurasia Proc Educ Soc Sci 2016;5:99–106.

[37] Peng C-YJ, So T-SH, Stage FK, St. John EP. The use and interpretation of logistic regression in higher education journals: 1988–1999. Res High Educ 2002;43:259–293. [CrossRef]

[38] Demir I. SPSS ile İstatistik Rehberi. 1. Baskı. İstanbul: Efe Akademi; 2020.

[39] Huang FL, Moon TR. What are the odds of that? a primer on understanding logistic regression. Gifted Child Quarterly 2013;57:197–204. [CrossRef]