

Predicting Achievement with Artificial Neural Networks: The Case of Anadolu University Open Education System

Hilal Seda Yıldız Aybek^{1,*}, Muhammet Recep Okur²

¹Center for Open and Distance Learning, Yasar University, Turkey

²Faculty of Open Education, Anadolu University, Turkey

Abstract: This study aims to predict the final exam scores and pass/fail rates of the students taking the Basic Information Technologies – 1 (BIL101U) course in 2014-2015 and 2015-2016 academic years in the Open Education System of Anadolu University, through Artificial Neural Networks (ANN). In this research, data about the demographics, educational background, BIL101U course mid-term, final and success scores of 626,478 students was collected and purged. Data of 195,584 students, obtained after this process was analysed through Multilayer Perception (MLP) and Radial Basis Function (RBF) models. Sixteen different networks attained through the combination of ANN parameters were used to predict the final exam scores and pass/fail rates of the students. As a result of the analyses, it was found out that networks established through MLPs make more exact predictions. In the prediction of the final exam scores, it was determined that there is a low level of correlation between the actual scores and predicted scores. In the analyses for the prediction of pass/fail rates of the students, networks established through MLPs ensured more exact prediction results. Moreover, it was determined that the variables as mid-term exam scores, university entrance scores and secondary school graduation year were of highest importance in explaining the final exam scores and pass/fail rates of the students. It was found out that in the higher institutions serving for Open and Distance Learning, pass/fail state of the students can be predicted through ANN under favour of variables of students which have been found as most the important predictors.

ARTICLE HISTORY

Received: 05 February 2018

Revised: 13 May 2018

Accepted: 29 May 2018

KEYWORDS

Prediction of Student Achievement,
Achievement in the Higher Education,
Open and Distance Learning,
Artificial Neural Networks

1. INTRODUCTION

Considering the great number of students registered in the Open Education System (OES) of Anadolu University, which is one of the mega universities of the world, it is necessary to analyse diverse features of this group of students in detail, with the aim of ensuring various student support services effectively and efficiently (da Silva, de Fátima Nunes, Santos, Queiroz & Leles, 2012; Kose & Arslan, 2017). Particularly, identification and presentation of the variables explaining the achievement levels of the students may provide key information to the

CONTACT: Hilal Seda YILDIZ-AYBEK ✉ hilal.aybek@yasar.edu.tr 📍 Center for Open and Distance Learning, Yasar University

institutions for the planned student support services. Identification of the variables explaining the achievement level of the present students may help to make predictions about the future achievement of the students. Machine learning algorithms such as Artificial Neural Networks (ANN) and various statistical models are used to predict the achievement of the student. It is considered that prediction of student's achievement in advance may be beneficial in various terms for both higher education institutions and students.

It has been observed in the related literature that there is a limited number of studies on the variables explaining the academic achievement of students registered in mega universities such as Anadolu University. Therefore, this study was carried out to develop intelligent systems and applications based on the prediction of student achievement.

1.1. Artificial Neural Networks

Core component of the nervous system of biological organism is neurons, i.e. neural cells (Mangels, 2003). It is estimated that the nervous system of a human has 100 billion neurons in average (Mastin, 2010). These cells do work together. As in each system, in biological systems, source (stimulus), receptor (collecting information), processing system (neural network or brain), effector (that turns brain's signals into movements or responses) and response (feedback) processes interoperate (Siegelbaum & Hudspeth, 2000).

Neurons are generally composed of soma, dendrites, axon and synapses (Finger, & Tansey, 1994; Guillery, 2005; Kandel, Schwartz, & Jessell, 2000). Stimuli (signals) from the external environment are transferred to axon from the dendrite in the nerve cell. During this transfer, nonlinear, complex processes take place. Information (signals) transferred to synapse following these processes are transferred to other neurons by means of synapse (Rojas, 2013).

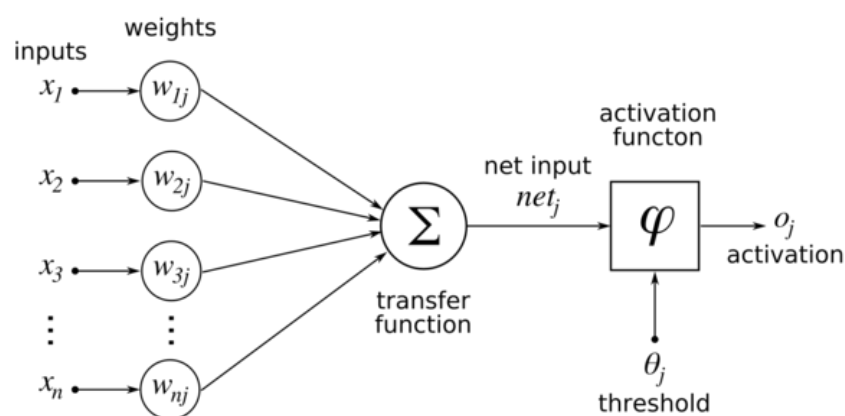


Figure 1. Artificial Neural Network

(Source: https://en.wikibooks.org/wiki/File:ArtificialNeuronModel_english.png)

In engineering sciences, artificial neurons are also called as “processing unit”. In Figure 1, x_n represents input while w_{nj} represents weight coefficients. Dendrite in biological nerve cell acts similar to the combining function of artificial neural network; cell body to transfer function; axons to element output and synapses to weights (Bullinaria, 2015). Functioning of an artificial neural network depends on the threshold value; neuron is activated when the result of multiplying input from external environment or another cell ($w_{nj}x_n$) is higher than the threshold value of the cell (Basheer & Hajmeer, 2000; Kotsiantis, Pierrakeas & Pintelas, 2003).

ANNs are systems learning from their experiences, making inferences based on the prior learning, and taking decisions, in a way similar to people (Öztemel, 2012). They have non-linear fields of application such as image processing, image classification, verification, speech

analysis, optimization problems, robot navigation, processing of incomplete or indefinite data, quality assurance, stock market prediction and simulation (Lippmann, 1987).

As seen in Figure 2, ANNs are basically composed of input layer, hidden layer(s) and output layer. ANNs learning is based on iterative loops. In the first stage, network output is produced from the input in the training set. Input used for training is the data enabling ANNs to learn, i.e. to gain experience (Çayiroğlu, 2013). Later on, weights calculated for the network ties are changed according to the accuracy level of the output. Determination of the network output and changes in the weights are realized variously depending on the learning rule and learning algorithm (Öztemel, 2012).

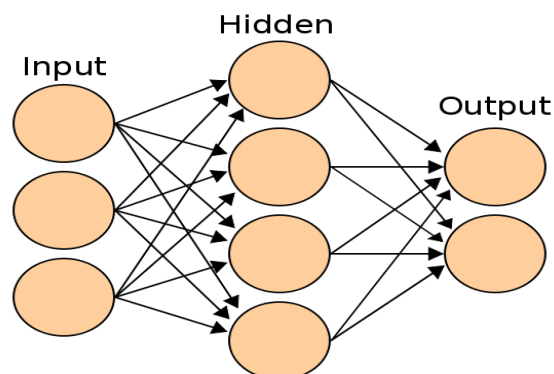


Figure 2. Single Layer Artificial Neural Network Model (Burnett, 2006)

In this study, Multilayer Perception (MLP) and Radial Basis Function (RBF) networks were utilized. MLPs are one of the feed-forward ANN models. They utilize Back Propagation Algorithm (BPA). BPAs are used to minimize the error rate in the network output (Yılmaz, Yavuz & Erkmen, 2013).

RBF operated and popularized by Moody & Darken (1989) are another type of feed-forward BPAs. RBFs can solve non-linear problems but it has been observed that they are poor in determining independent variables explaining dependent variables (Akbiğiç, 2011).

According to Koca (2006), RBFs learn faster than MLPs, and can make classifications and generalisations. In comparison to MLPs, RBFs have simpler architectures (Yu, Xie, Paszczynski & Wilamowski, 2011).

1.2. Achievement and Factors Explaining Achievement

Achievement of an individual in his/her school life is called as academic achievement and today, may represent almost the whole achievement of an individual throughout his/her life. Academic achievement may be defined as an indicator of achieving certain learning objectives (Choi, 2005). Criteria such as grade-point averages, cumulative grade-point averages and course notes are described as academic achievement (Astin, 1991; Snyder, et al., 2002).

A certain part of the studies on academic achievement concerns the examination of the variables explaining the academic achievement of the student. Variables addressed in these studies include *gender* (Amro, Mundy & Kupczynski, 2015; Collins, McLeod & Kenway, 2000; Hajovsky & Kaufman, 2015; Pike, Schroeder & Berry, 1997; Scheiber, Reynolds, Mlambo, 2012); *attitude* (Brown, et al., 2015; Odom & Bell, 2015), *anxiety* (Khalaila, 2015; Macher, Paechter & Papousek, 2015), *socio-economic level* (Jurdak, 2014; Suphi & Yaratana, 2012), *prior learning* (Musso, Kyndt, Cascallar & Dochy, 2013; Power, Robertson & Baker, 1987; Strayhorn, 2006), *self-efficacy* (Valentine, DuBois & Cooper, 2004).

In his study Hattie (2009) examined the factors affecting achievement, by synthesising more than 800 meta-analysis and pointed that gender has a small effect size ($d = .12$), while

socio- economic level has medium ($d = .57$), computer assisted training has medium ($d = .48$), attitude has medium ($d = .36$) and previous achievement has medium ($d = .67$) effect sizes. In addition to these, he addressed 138 more factors affecting achievement, in his study. Considering the studies benefiting from machine learning approaches in the prediction of achievement, findings in Table 1 are determined.

Table 1. Machine learning approaches in the prediction of achievement

Study	Approach
Yukselturk, Ozekes & Türel, 2014	k - Nearest Neighbour, Decision Tree, Naive Bayes Classifier and ANN (3 th rank)
Turhan et al., 2013	ANN (best result) and regression analyses
Lykourantzou et al., 2009	ANN displayed more effective performance, in comparison to linear regression.
Aydın, 2007	C5.0, Logistic Regression, ANN, C&RT, CHAID and QUEST
Rusli et al., 2008	Prediction through ANN is provided more exact results than decision tree and linear regression.
Naik & Ragothaman, 2004	Prediction through ANN is 93.38% exact.

This study generally aims to predict the final exam scores and pass/fail rates of the students taking Basic Information Technologies – 1 (BIL101U) course in 2014-2015 and 2015-2016 academic years in the Open Education System (OES) of Anadolu University, through Artificial Neural Networks (ANN). On the basis of this main aim, below-given questions were tried to be answered:

1. What are the variables that explain the final exam scores of the students taking BIL101U course?
2. At which level do MLP and RBF networks explain final exam scores?
3. Which one of the MLP and RBF type ANN models provide more exact results when pass/fail rates are determined on the basis of the predicted final exam scores?
4. Which one of the MLP and RBF type ANN models can predict pass/fail state of the student more exactly?

In this study, it was examined whether the final exam scores and pass/fail state of the students registered in OES can be predicted by means of ANN models, and accordingly, variables that explain achievement best were determined. In line with the relations among independent variables, dependent variables were explained. Therefore, correlational research model was utilised.

2. MATERIAL METHOD

2.1. Participatiants

Research population of this is composed of students taking common and compulsory BIL101U course in the fall semester of 2014-2015 (SG-1) and 2015-2016 (SG-2) academic years in different faculties (Open Education Faculty, Faculty of Business Administration and Faculty of Economics) and departments in the OES of Anadolu University. In order to prevent difference among the variables in the dataset, it was preferred to utilise data of two academic years (2014-2015 and 2015-2016).

Number of students in SG-1 was 306.633, while the number of students in SG-2 was 319.845. Total number of students taking BIL101U course in two academic years was 626.478. Following the data cleaning process, as explained in detail under the next title, the

total number of students included in the analysis in two years was decreased to 195.584. 93% of the students taking BIL101U course passed.

2.2. Preparing data for the Analysis

Registration procedures of the students in the OES are carried out at the beginning of each fall semester. During registration, TR identity numbers of the students are identified as unique, and data transfer to OES database is realized according to certain variables in the OSYM (Student Selection and Placement Centre) database.

Table 2 displays the data obtained from the OES and Computer Research Centre (CRC) of Anadolu University, under “demographic”, “educational background”, “OES” and “other” columns.

Table 2. Variables in the analysis

Demographic	Educational Background	OES	Other
Year of birth, TR identity no, Nationality, Gender, Province	Year of Graduation from the High School, University entrance score (UES), High school type, High school code, High School GPA, Foreign Language, UES Score type, Quota type, University placement ranking	Mid-term exam, final exam, letter grade	Faculty, department, type of registration

OSYM based data was taken from OES and data regarding the variables in Table 3 was taken from CRC; however, some variables were converted during the preparation phase of the analysis. This conversion process is displayed in Figure 3.

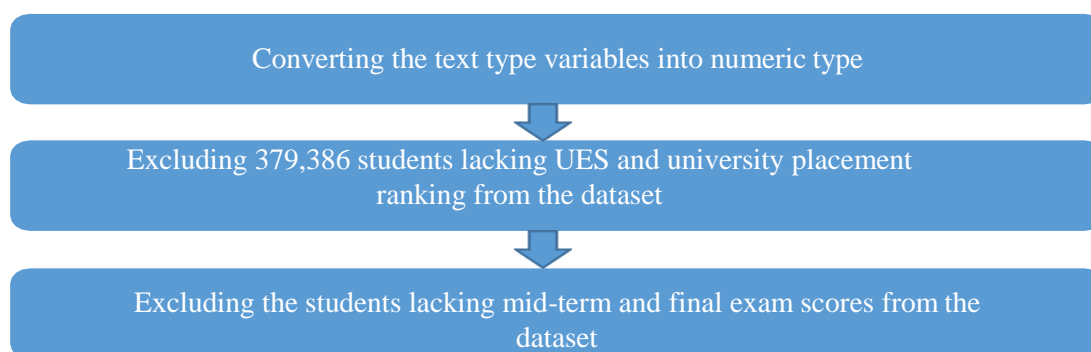


Figure 3. Preparation Process of the Data for the Analysis

Variables of TR identity number, which is used only for identification; and school code, given to each school separately were not used in the data analysis. Variable of university placement ranking was not included in the analysis, as this data was missing for many students. Instead, UES was preferred.

2.3. Data Analysis

In compliance with the aim of the research, ANN was established according to different parameters with MLP and RBF. SG-1 dataset was used as training and test set for the networks; while SG-2 was used as validity set in the determination of the prediction level of the network.

Training set composed 70% of SG-1, while test set composed 30%. In the determination of training and test set, IBM SPSS Statistics v21, 136940 core initial value selected randomly

by the researcher, and $2*rv.bernoulli(0.7)-1$ relation were utilized (IBM Knowledge Center, n.d.).

Accordingly, by benefiting from the training, test and control groups, twelve MLPs consisting of various combinations of six parameters, and four RBFs consisting of various combinations of three parameters were established. Characteristics of these networks are displayed in Table 3 and Table 4.

Table 3. Parameters used for MLP analyses

Network Name	Hidden Layer	Min. Unit	Max. Unit	Training	Activation Function	Re-scaling of continuous variables
MLP-A	1	1	50	Batch	Hyperbolic	Standardized
MLP-B	1	1	50	Batch	Hyperbolic	Normalized
MLP-C	1	1	50	Online	Hyperbolic	Standardized
MLP-D	1	1	50	Online	Hyperbolic	Normalized
MLP-E	2	Auto	Auto	Batch	Hyperbolic	Standardized
MLP-F	2	Auto	Auto	Batch	Hyperbolic	Normalized
MLP-G	2	Auto	Auto	Batch	Sigmoid	Standardized
MLP-H	2	Auto	Auto	Batch	Sigmoid	Normalized
MLP-I	2	Auto	Auto	Online	Hyperbolic	Standardized
MLP-J	2	Auto	Auto	Online	Hyperbolic	Normalized
MLP-K	2	Auto	Auto	Online	Sigmoid	Standardized
MLP-L	2	Auto	Auto	Online	Sigmoid	Normalized

As seen in Table 3, 12 MLP networks were established with different parameters, which are explained below:

- **Hidden Layer:** is a place which is composed of one or more layers and where the network realizes learning process (Nabiyev, 2005). IBM SPSS Statistics v21 software supports one or two layers option for MLP.
- **Min - Max Unit:** is the number of processing units within the hidden layers. The number varies between 1 and 50.
- **Training:** Batch training, which is one of the parameters utilized in this study, minimizes the total error, and changes weight values after all data is known (Shalev- Shwartz, 2011). Online training is preferred in comprehensive datasets and changes weight value of each data during learning.
- **Activation Function:** Under IBM SPSS Statistics v21 software, there is Hyperbolic Tangent which provides output between [-1,1] by subjecting input to tangent function, and Sigmoid function which enables input to be converted into values between [0,1].
- **Re-scaling of Continuous Variables:** In order to develop the training of the network, continuous variables may be re-scaled. Under this study, parameters of Standardized, which converts data to a range between [0, 1] and Normalized, which converts data to a range between [-1, 1] were tested.

Table 4. Parameters used for RBF Analyses

Network	Hidden Layer	Activation Function	Re-scaling of Continuous Variables
A	Auto	Normalized	Standardized
B	Auto	Normalized	Normalized
C	Auto	Ordinary	Standardized
D	Auto	Ordinary	Normalized

In addition to MLP parameters displayed in Table 4 and explained later on, activation function type: “simple” was used for RBF-C and RBF-D among RBF networks. In simple RBF selected for activation function, basically exponential activation function was used to ensure the usual distribution of hidden layers (Matignon, 2005).

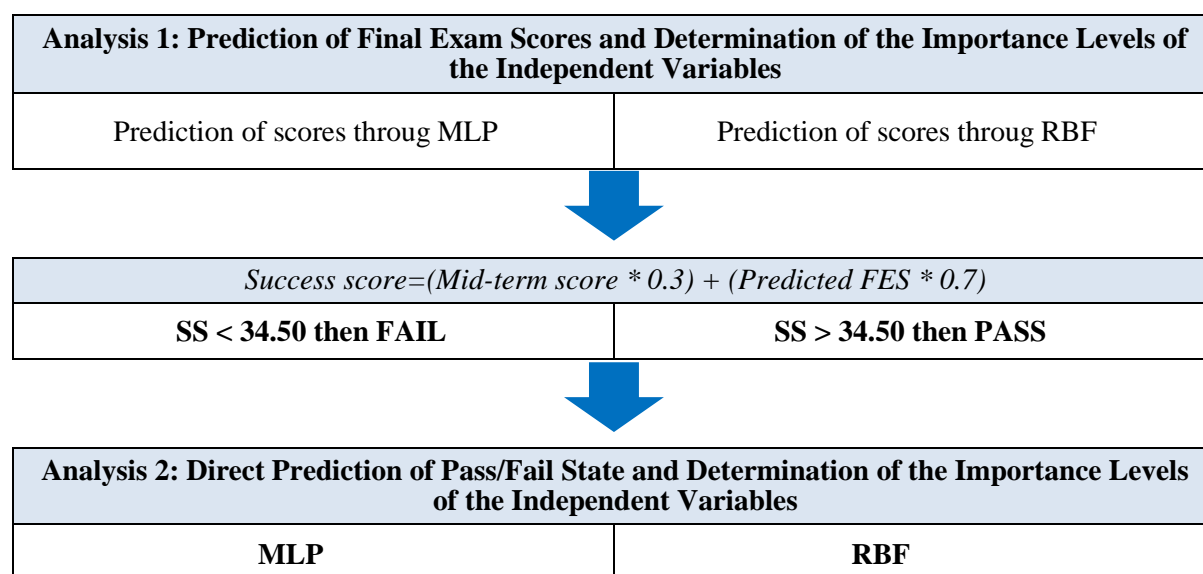


Figure 4. Data analysis process

16 different networks presented in Table 3 and Table 4 was operated twice, one to predict the final exam scores of the students and one to predict the pass/fail state of the students. 32 analyses were carried out in total.

Correlation between the final exam score predicted as a result of the first analysis and the actual score of the student was determined through Pearson Product-Moment Correlation Coefficient. When pass/fail state of the students were determined based on the predicted final exam scores, inconsistencies were examined through crosstabs. During the second analyses, in the networks predicting pass/fail state of the students, consistency between the predicted and actual state was tried to be determined through crosstabs.

3. FINDINGS

3.1. Analysis 1: Prediction of Achievement in Final Exam through MLP and RBF

When the correlation between the results of twelve MLP analyses realized to predict the final exam scores of the students and the actual final exam scores of the students is examined, findings displayed in Table 5 were obtained. Moreover, variables used in each network to explain the final exam scores of the students, and their importance levels were determined.

Table 5. Importance levels, correlation and determination coefficients of the independent variables, obtained as a result of the MLP analyses

Network Name	r_{gk}	r_{gk}^2	Importance Level	Network Name	r_{gk}	r_{gk}^2	Importance Level
A	.482**	.232	mes (100), ues (72.4)	G	.462**	.213	mes (100), ues (55.7)
B	.486**	.236	mes (100), ues (61.6)	H	.445**	.198	mes (100), hst (50.3)
C	.431**	.186	mes (100), ues (71.8)	I	.438**	.192	mes (100), ues (64.2)
D	.311**	.097	mes (100), hst (93.8), ues (79.8)	J	.171**	.029	mes (100), hst (99.6), dept (91.4)
E	.469**	.220	mes (100), ues (71.6)	K	.469**	.220	mes (100), ues (56)
F	.391**	.153	mes (100), prov (67.1), hst (63.0)	L	.369**	.136	mes (100), dept (64.1)

mes: Mid-term exam score, ues: University entrance score, hst: High school type, prov: Province, dept: Department

When Table 5 is examined, it is seen that the network that has highest correlation with the actual final exam scores of the students is MLP-B. MLP-A, MLP-E, MLP-G and MLP-K networks are also among the networks with highest correlation. However, when all the correlation coefficients are examined, medium level correlation coefficients are observed between the predicted and actual scores. Considering the determination coefficients (r^2), it has been found out that the variables in these networks explain the final exam scores of students in a range between 21.3% and 23.6%. Accordingly, it may be asserted that 76.3% and 78.7% of the final exam scores of the students are explained by other variables.

Table 6. Importance levels, correlation and determination coefficients of the independent variables, obtained as a result of the RBF analyses

Network	r_{gk}	r_{gk}^2	Importance Level
A	.224	.05	ues (100), yghs (59.8)
B	.141	.01	tor (100), faculty (87.2)
C	.112	.01	ues (100), faculty (98.8), mes (77.4)
D	.135	.01	tor (100), tor (75.7), fl (68.7)

yghs: Year of Graduation from the High School, tor: Type of registration, fl: Foreign language

Considering the determination coefficients in Table 6, low correlation between the output of RBF-C network and the actual state is remarkable. In this context, it may be claimed that RBF networks do not provide consistent results in explaining final exam scores.

3.2. Benefiting from Final Exam Scores Predicted through MLP and RBF in Pass/Fail Decisions

After the final exam scores of the students were predicted through MLP, their success scores were calculated with mid-term and final exam scores. Following the pass/fail decisions taken on the basis of these scores, consistency between the ANN and actual state was examined through crosstabs. Crosstabs were calculated separately for SG-1 utilised in training and test set; and for SG-2 utilised for control. Accordingly, when pass/fail decisions are taken on the basis of the predicted final exam scores, consistency with the actual state for SG-1 was displayed in percentages in Table 7.

Table 7. Determination of pass/fail state with final exam scores predicted through MLP for SG-1

Network	P _M -P ₀	F _M -F ₀	F _M -P ₀	P _M -F ₀	Network	P _M -P ₀	F _M -F ₀	F _M -P ₀	P _M -F ₀
A	76.2	6.9	3.6	13.4	G	75.7	7.3	4.1	13.0
B	76.1	7.1	3.6	13.2	H	76.0	6.7	3.8	13.5
C	76.2	6.3	3.6	13.9	I	75.0	7.7	4.7	12.5
D	75.5	6.0	4.2	14.3	J	79.5	0.8	0.3	19.5
E	76.2	6.7	3.5	13.6	K	74.1	8.7	5.6	11.5
F	77.1	5.1	2.7	15.2	L	76.5	5.4	3.2	14.8

P_M – P₀: State of passing the course both in actual state and according to MLP

F_M – F₀: State of failing the course both in actual state and according to MLP

F_M – P₀: Fail decision according to MLP, when actually the student passes

P_M – F₀: Pass decision according to MLP, when actually the student fails

In interpreting **Table 7** and similar tables, P_M-F₀ column displaying that ANN decides that the student passes when actually the student fails has been determined as intolerable error by the researchers. Accordingly, it may be asserted that the network “K” leads to lower level of errors. Network “J” provided “pass” decision for 99% of the students.

Table 8. Determination of pass/fail state with final exam scores predicted through MLP for SG-2

Network	P _M -P ₀	F _M -F ₀	F _M -P ₀	P _M -F ₀	Network	P _M -P ₀	F _M -F ₀	F _M -P ₀	P _M -F ₀
A	80.9	4.0	11.6	3.5	G	79.5	4.2	13.0	3.3
B	79.8	4.2	12.8	3.3	H	81.4	3.8	11.2	3.7
C	82.5	3.4	10.0	4.0	I	77.0	4.6	15.5	2.9
D	83.1	2.7	9.4	4.8	J	91.6	0.5	0.9	6.9
E	80.9	4.0	11.6	3.5	K	76.2	4.8	16.4	2.7
F	85.3	2.8	7.2	4.6	L	83.5	3.2	9.1	4.3

In the case when SG-2 is used as validity set, it is remarkable that there is a decrease, in the level of errors in all networks, as seen in **Table 8**. Moreover, it is also seen that network “K” is the network with lowest level of errors.

According to the analysis with RBF, when pass/fail decisions are taken on the basis of predicted final exam scores, consistency with the actual state for 2014-2015 academic year was displayed in percentages in **Table 9**.

Table 9. Determination of pass/fail state with final exam scores predicted through MLP for SG-1 and SG-2

SG-1					SG-2				
Network	P _M -P ₀	F _M -F ₀	F _M -P ₀	P _M -F ₀	Network	P _M -P ₀	F _M -F ₀	F _M -P ₀	P _M -F ₀
A	79.3	0.9	0.5	19.3	A	91.7	0.5	0.8	7.0
B	79.7	0.3	0.0	20.0	B	92.3	0.2	0.2	7.3
C	79.7	0.0	0.0	20.3	C	92.5	0.0	0.0	7.5
D	79.7	0.3	0.0	20.0	D	92.4	0.2	0.2	7.3

When the data about the students, in the validity set is examined, actual state is consistent with RBF networks for nearly 92% of the students as similar to the results of training and test set, but RBF networks lead to decisions different from the actual state for about 7.5% of the students.

3.3. Analysis 2: Direct Prediction of Pass/Fail State and Determination of the Importance Levels of the Independent Variables

Instead of the final exam scores of the students, pass/fail state of the students were tried to be predicted directly through ANN. During this process, parameters in Table 3 were used for MLP networks, while those in Table 4 were used for RBF networks. In predicting the pass/fail state of the students, *passfail* variable set by the researcher was used as the dependent variable. This variable enables binary prediction with regard to the pass/fail state of the students.

Table 10. In MLP networks, Importance Levels of the Independent Variables Utilised in the Prediction of Pass/Fail State

Network	Importance Levels	Network	Importance Levels
A	mes (100), ues (40.0)	G	mes (100), age (33.2)
B	mes (100), ues (35.3)	H	mes (100), hst (83.4)
C	mes (100), hst (44.6)	I	mes (100), age (57.9)
D	mes (100), yghs (29.7)	J	prov (100), hst (83.5)
E	mes (100), ues (34.2)	K	mes (100), ues (23.0)
F	hst (100), prov (93.7), mes (82.2)	L	hst (100), mes (90.8), prov (81.3)

It was observed that F, H and L networks, among MLP networks are different from the other networks in terms of their breakdown of relative importance levels. When the erroneous prediction rates of these networks in Table 11 are examined, it is seen that these networks have more errors in the training set, in comparison to the other networks, but this error level is observed to decrease in validity set.

Table 11. Decisions of pass/fail state for SG-1 and SG-2 taken through MLP

Network	SG-1				Network	SG-2			
	P_M-P_0	F_M-F_0	F_M-P_0	P_M-F_0		P_M-P_0	F_M-F_0	F_M-P_0	P_M-F_0
A	75.2	7.9	4.5	12.3	A	78.0	4.5	14.5	3.0
B	75.1	7.9	4.7	12.4	B	79.2	4.2	13.3	3.3
C	74.9	6.9	4.8	13.4	C	79.3	3.9	13.2	3.5
D	75.0	5.3	4.8	15.0	D	83.7	2.5	8.9	5.0
E	74.7	7.9	5.0	12.3	E	78.5	4.2	14.0	3.3
F	79.7	0.0	0.0	20.2	F	92.5	0.0	0.0	7.5
G	75.3	7.5	4.5	12.8	G	79.4	4.1	13.1	3.4
H	78.4	2.0	1.4	18.3	H	89.6	1.1	2.9	6.4
I	74.4	7.8	5.3	12.4	I	78.6	4.1	13.9	3.4
J	76.1	3.2	3.7	17.1	J	86.9	1.2	5.6	6.3
K	77.5	5.0	2.2	15.3	K	82.5	3.0	7.4	4.4
L	79.7	0.1	0.0	20.2	L	92.5	0.0	0.0	7.5

It is seen in Table 11 that the error ratios of the networks in SG-1 are higher than that of SG-2. Erroneous prediction rates of MLP networks in training, test and validity sets are presented in Table 12. As seen in this table, while the error rate of MLP-F, MLP-H and MLP-L networks is close to 20% in training and test sets, their error rate in validity set is below 10%.

Table 12. Erroneous prediction rates of MLP networks

Network	Training	Test	Validity	Network	Training	Test	Validity
A	16.8	17.0	17.5	G	17.3	17.3	16.5
B	17.1	17.1	16.6	H	19.5	20.0	9.3
C	18.3	18.0	16.7	I	17.7	17.9	17.3
D	19.7	19.7	13.9	J	20.6	21.2	11.9
E	17.4	17.2	17.3	K	17.5	17.5	11.8
F	20.1	20.6	7.5	L	20.1	20.6	7.5

In RBF analyses where pass/fail states instead of the final exam scores of the students are predicted through ANN, pass/fail variables and parameters in Table 13 were used, as in MLP analyses.

Table 13. Pass/fail decisions for SG-1 and SG-2 taken through RBF

Network	SG-1				Network	SG-2			
	P _R -P ₀	F _R -F ₀	F _R -P ₀	P _R -F ₀		P _R -P ₀	F _R -F ₀	F _R -P ₀	P _R -F ₀
A	79.3	0.9	0.5	19.3	A	91.7	0.5	0.8	7.0
B	79.7	0.3	0.0	20.0	B	92.3	0.2	0.2	7.3
C	79.7	0.0	0.0	20.3	C	92.5	0.0	0.0	7.5
D	79.7	0.3	0.0	20.0	D	92.4	0.2	0.2	7.3

In Table 13, it is seen that although 92.5% percentage of passing, which is same for all the networks is the closest result to the actual state as 93%, these networks provided “pass” decision for all students. In this context, it was determined that RBF networks are not successful in classifying pass/fail states or the students.

Table 14 displays the erroneous prediction rates in training, test and validity sets when pass/fail states of the students are predicted through RBF instead of predicting the final exam scores of the students through ANN.

Table 14. Erroneous prediction rates of RBF networks

Network	Training	Testing	Holdout
A	20.1	20.6	7.5
B	20.1	20.6	7.5
C	20.1	20.6	7.5
D	20.1	20.6	7.5

Consistency between the predictions and the actual state was examined by means of crosstabs and the results are presented in Table 15.

Table 15. Pass/Fail Decisions taken through RBF for SG-1 and SG-2

Network	SG-1				Network	SG-2			
	P _R -P ₀	F _R -F ₀	F _R -P ₀	P _R -F ₀		P _R -P ₀	F _R -F ₀	F _R -P ₀	P _R -F ₀
A	79.7	-	-	20.3	A	92.5	-	-	7.5
B	79.7	-	-	20.3	B	92.5	-	-	7.5
C	79.7	-	-	20.3	C	92.5	-	-	7.5
D	79.7	-	-	20.3	D	92.5	-	-	7.5

When Table 15 is examined, it is seen that the erroneous decision rates in validity sets decreased to 7.5%. However, it was found out that RBF networks provided “pass” decision for

all students. These findings are similar to the inconsistencies in the results of RBF networks, through which final exam scores of the students were predicted.

4. DISCUSSION

According to MLP analyses, final exam scores of the students taking BIL101U course and included in SG-2 are mostly explained by the variables of *Mid-term Score, Year of Graduation from the Secondary School, Placement Score, Type of High School and Province of the Address*. According to RBF analyses, the variables that mostly explain final exam scores are the variables of *Registration Type, Placement Score, Type of High School and Faculty*.

When the variables affecting the student achievement are examined, it is seen that the variable of gender is important (Amro, Mundy & Kupczynski, 2015; Mlambo, 2012; Collins, McLeod & Kenway, 2000; Pike, Schroeder & Berry, 1997; Scheiber, Reynolds, Hajovsky & Kaufman, 2015; Zheng, 2002). Considering all the analyses carried out within the framework of this study, it has been found out that the variable of gender has low level of importance within SG-1 and SG-2. This study asserting that the gender is not an important explanatory variable of achievement does not coincide with the previous researches in terms of this variable.

With regard to the courses at the level of bachelor's degree, concerning information technologies, it was found out that gender is not a important explanatory variable of achievement (Fan, 1998; Werth, 1986; Wilson, 2002; Wilson & Shrock, 2001). Considering the literature on the achievement in courses related to information technologies, the results of this study coincide with the studies claiming that gender is not a significant explanatory variable.

Another factor determined to explain student achievement is the educational background (Power, Robertson & Baker, 1987; Strayhorn, 2006; Zheng, 2002). In his study on the examination of explanatory variables of student achievement, Wolfe (1995) found out that high school grade-point average is the first explanatory variable while score in SAT, which is an exam taken by American students for placement in university is the third explanatory variable of achievement. In this study, variables of score (SSPE score) required in registration to OES, year of graduation from the secondary school, type of high school being graduated were addressed under the educational background category. It may be asserted that these variables have high importance for the overall MLP networks, are significant in terms of explaining student achievement, and thus, the results of this study are similar to the results of previous studies.

In general, maths competence of the students acquired during high school years explain their achievement in information technologies courses at the level of bachelor's degree, (Oman, 1986; Wilson & Shrock, 2001; Wilson, 2002). This study did not include a variable of maths competence, but it is observed in the related literature that the SAT score addressed in the studies, which is similar to SSPE score in Turkey, is an explanatory variable with low significance (Ventura, 2005). Considering that SSPE score is an important explanatory variable in this study, it may be said that this study is not in parallel with the previous studies in this regard.

It was found out that predictions through MLP provided more exact results than predictions through RBF, with a difference of nearly 4% in error rate. In their study, Huang and Fang (2013) compared MLP and RBFs, and as in this study, they asserted that MLP network provided more exact predictions than RBF network.

When the studies on the prediction of student achievement are examined, it is seen that in general, different mathematical models such as Regression, Artificial Neural Networks,

Decision support Systems, Decision Trees and Baynes were compared and that ANN displayed a better performance than others (Herzog, 2006; Lykourantzou et al., 2009; Naik and Ragothaman, 2004; Schumacher, Olinsky, Quinn, & Smith, 2010; Şen, Uçar & Delen, 2012; Rusli, Ibrahim & Janor, 2008; Turhan et al., 2013). As distinct from these results, a practice with ANN was ranked the second in the study carried out by Aydın (2007) with an accuracy rate of 77.80%; and ANN analysis was ranked the third in the study carried out by Yükseltürk et al. with a classification rate of 79.7%. Although comparison of ANN models with other mathematical models is not aimed or practiced in this study, the accurate prediction rate ranging between 85% and 87%, obtained through different parameters should be taken into account.

It was observed that certain parameters led to low performance in predictions as each network was established on different parameters. When MLP-D and MLP-J out of MLP networks established to predict the final exam scores of the students are examined, it was found out that they have a low correlation with the actual state. These networks were established with *online* training method, *Hyperbolic Tangent* function and *Normalized* scaling parameters. Therefore, utilisation of the combination of these parameters in the prediction of scores does not ensure a good performance.

When interpreting the findings obtained from the second application set in which predictions are made directly on pass/fail state of the students, error rates resulting from the “failed” decision of MLP although the student passes in the actual state were defined as *negligible errors* in this study. With an error rate below 3%, MLP-K and MLP-I networks were identified as the networks with the fewest error. Both networks were established with online training and Standardized scaling parameters. When the parameters of MLP-A, MLP-B, MLP-C, MLP-E, MLP-G and MLP-H networks, whose error rate is below 4% are examined, it is seen that Hyperbolic Tangent function was used in these networks, in general. Similarly, the study carried out by Özkan & Erbek (2003) pointed that in classification problems, Hyperbolic Tangent function displays better performance than the sigmoid function; and thus, this result complies with the results of this study.

When the findings of MLP and RBF analyses are examined with an overall perspective, it is seen that in comparison to RBF networks, MLP networks provided more exact results in the prediction of both final exam scores and pass/fail states. Moreover, considering the importance levels of the independent variables in RBF analyses, there is no covariance among the networks. According to Akbilgiç (2011), there are problems in determining the independent variables affecting the dependent variable in RBF and hybrid networks.

Final exam score and pass/fail state predictions, where SG-1 was used as training and test set and SG-2 as validity set provided more exact results than the cases where SG-1 was used singly with a division into training and test sets. Accordingly, it may be claimed that when ANN is used in the prediction of student achievement in a course, obtaining training and test sets from the data sets of previous years may ensure more exact results.

Suggestions for the future studies and practices to be carried out within the scope of the findings and results of this research are as follows:

- In this research aiming at the prediction of the student achievement, the student scores only in BIL101U course were utilized. As the content of BIL101U course is related with information technologies, variables such as math competence and educational background play an important role in terms of explaining achievement in this course or similar courses. In this context, data concerning the variables such as right/wrong rates in SSPE, which is stated to be a good explanatory variable of achievement, experience of a computer course, student’s expectations related to the course may also be collected for analysis.

- Considering that each course has different aims and objectives and thus different learning outcomes, explanatory variables of achievement may not be same. Therefore, carrying out similar analyses with similar or different variables for other courses may provide a scientific conclusion in this topic.

- As a result of the analyses with ANN, it has been found out that the variable with the highest effect on explaining achievement is mid-term score. Accordingly, data obtained from the pilot tests and similar practices in the “Anadolum eKampüs” application of Anadolu University may be included in ANN analyses to examine whether these process evaluation variables create a difference in the prediction of achievement.

- In line with the findings of the analyses carried out, pass/fail states of the students may be predicted. Most suitable and goal-oriented model may be selected and integrated into web-based learning environments. This integrated system may be considered as an early warning system for the students. When the student is registered, data obtained from the student himself/herself or from the institution-based systems may be processed through ANN, and predictions regarding the pass/fail state of the students in a course or courses may be made. These predictions should not be communicated to the student through notifications such as “you have passed” or “you have failed” but they may receive suggested topics of study, course materials or pilot test warnings. By this means, “fail” prediction about a student who actually passes – which was considered as a “negligible error” in this study – may be reflected as an additional support to the student, in this process.

- In order to predict the student achievement and to integrate these predictions with the systems used in the learning process, exactness level of these predictions is important. Exact predictions are basically ensured through clear and organized data. Accordingly, this research is important in terms of encouraging educational institutions and organizations to identify the explanatory variables of achievement and to collect data about these or similar variables.

- SSPE score is also one of the most important variables in predicting the student achievement. This variable provides quantitative data about the educational background of the student. Similar studies on the prediction of student achievement (Oman, 1986; Wilson & Shrock, 2001; Wilson, 2002) provide data about the courses taken students during high school education. It is considered that a similar study may be carried out with the OES students to achieve more extensive results, and more exact predictions.

Acknowledgements

This research was produced from master's thesis titled “Predicting Achievement with Artificial Neural Networks: The Case of Anadolu University Open Education System” by Hilal S. Yıldız Aybek in supervising of M. Recep Okur, and supported by Anadolu University Scientific Research Projects Commission.

ORCID

Hilal Seda Yıldız Aybek  <https://orcid.org/0000-0001-9128-734X>

Muhammet Recep Okur  <https://orcid.org/0000-0003-2639-4987>

5. REFERENCES

- Akbilgiç, O. (2011). *Hybrid radial basis function networks and variable selection and estimation: an application for securities investment decisions*. (Doctoral dissertation). Istanbul University. Istanbul.
- Amro, H. J., Mundy, M., & Kupczynski, L. (2015). The effects of age and gender on student achievement in face-to-face and online college algebra classes. *Research in Higher Education Journal*, 27.
- Astin, A. W. (1991). *Assessment for excellence: The philosophy and practice of assessment and*

- evaluation in higher education*. New York: Macmillian.
- Aydın, S. (2007). *Data mining and an application in Anadolu University Open Education System*. (Doctoral dissertation). Anadolu University Graduate School of Social Sciences. Eskisehir.
- Basheer, I. A., & Hajmeer, M. (2000). Artificial neural networks: fundamentals, computing, design, and application. *Journal of microbiological methods*, 43(1), 3-31.
- Bower, B. L., & Hardy, K. P. (2004). From correspondence to cyberspace: Changes and challenges in distance education. *New Directions for Community Colleges*, 2004(128), 5-12.
- Brown, S. J., White, S., Sharma, B., Wakeling, L., Naiker, M., Chandra, S., & Bilimoria, V. (2015). Attitude to the study of chemistry and its relationship with achievement in an introductory undergraduate course. *Journal of the Scholarship of Teaching and Learning*, 15(2), 33-41.
- Bullinaria, J. A. (2015) Biological neurons and neural networks, artificial neurons: neural computation: Lecture 2. Retrieved April 20, 2016, from <http://www.cs.bham.ac.uk/~jxb/INC/l2.pdf>
- Burnett, C. M. (2006, December 27). Artificial neural network. Retrieved April 20, 2016, from https://commons.wikimedia.org/wiki/File:Artificial_neural_network.svg#/media/File
- Collins, C. W., McLeod, J., & Kenway, J. (2000). *Factors influencing the educational performance of males and females in school and their initial destinations after leaving school*. Canberra: Department of Education, Training and Youth Affairs.
- da Silva, E. T., de Fátima Nunes, M., Santos, L. B., Queiroz, M. G., & Leles, C. R. (2012). Identifying student profiles and their impact on academic performance in a Brazilian undergraduate student sample. *European Journal of Dental Education*, 16(1), 27-32.
- Finger, S., & Tansey, T. (1994). Origins of neuroscience: a history of explorations into brain function. *Trends in Neurosciences*, 17(7), 310-310.
- Guillery, R. W. (2005). Observations of synaptic structures: origins of the neuron doctrine and its current status. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 360(1458), 1281-1307.
- Hattie, J.C. (2009). *Visible learning. a synthesis of over 800 meta-anaylses relating to achievement*. New York: Routledge.
- Herzog, S. (2006). Estimating student retention and degree completion time: Decision trees and neural networks vis-à-vis regression. *New Directions for Institutional Research*, 2006(131), 17-33.
- Huang, S., & Fang, N. (2013). Predicting student academic performance in an engineering dynamics course: A comparison of four types of predictive mathematical models. *Computers & Education*, 61, 133-145.
- Ibrahim, Z., & Rusli, D. (2007). Predicting students' academic performance: comparing artificial neural network, decision tree and linear regression. In *21st Annual SAS Malaysia Forum*, 5th September.
- IBM Knowledge Center. (n.d.). Preparing the Data for Analysis. Retrieved April 21, 2016, from https://www.ibm.com/support/knowledgecenter/en/SSLVMB_24.0.0/spss/tutorials/mlp_bankloan_dataprep.html
- Jurdak, M. (2014). Socio-economic and cultural mediators of mathematics achievement and between-school equity in mathematics education at the global level. *Zdm*, (7), 1025.
- Kandel, E. R., Schwartz, J. H., & Jessell, T. M. (Eds.). (2000). *Principles of neural science*, (4), 1227-1246. New York: McGraw-hill.
- Khalaila, R. (2015). The relationship between academic self-concept, intrinsic motivation, test

- anxiety, and academic achievement among nursing students: *Mediating and moderating effects*. *Nurse Education Today*, (3), 432. doi:10.1016/j.nedt.2014.11.001
- Koca, Z. (2006). *Speed control with artificial neural networks with vector control of three phase asynchronous motors* (Master's thesis). Kahramanmaraş Surcu Imam University Graduate School of Natural and Applied Sciences. Kahramanmaraş.
- Kose, U., & Arslan, A. (2017). Optimization of self-learning in Computer Engineering courses: An intelligent software system supported by Artificial Neural Network and Vortex Optimization Algorithm. *Computer Applications in Engineering Education*, 25(1), 142-156.
- Kotsiantis, S., Pierrakeas, C., & Pintelas, P. (2003). Preventing student dropout in distance learning systems using machine learning techniques. In *Knowledge -Based Intelligent Information and Engineering Systems*, (pp. 267–274).
- Lippmann, R. (1987). An introduction to computing with neural nets. *IEEE ASSP Magazine IEEE ASSP Mag.*, 4-22.
- Lykourentzou, I., Giannoukos, I., Mpardis, G., Nikolopoulos, V., & Loumos, V. (2009). Early and dynamic student achievement prediction in e-learning courses using neural networks. *Journal of the American Society for Information Science and Technology*, 60(2), 372-380.
- Macher, D., Paechter, M., Papousek, I., Ruggeri, K., Freudenthaler, H. H., & Arendasy, M. (2013). Statistics anxiety, state anxiety during an examination, and academic achievement. *British Journal of Educational Psychology*, 83(4), 535-549.
- Mangels, J. (2003). Cells of the Nervous System. Retrieved May 3, 2016, from <http://www.columbia.edu/cu/psychology/courses/1010/mangels/neuro/neurocells/neurocells.html>.
- Mastin, L. (2010). Neurons & Synapses. Retrieved from May, 4, 2016, from http://www.human-memory.net/brain_neurons.html.
- Matignon, R. (2005). *Neural network modeling using SAS enterprise miner*. AuthorHouse.
- Mlambo, V. (2012). An analysis of some factors affecting student academic performance in an introductory biochemistry course at the University of the West Indies. *The Caribbean Teaching Scholar*, 1(2).
- Moore, M. G., & Kearsley, G. (2011). *Distance education: A systems view of online learning*. Cengage Learning.
- Musso, M. F., Kyndt, E., Cascallar, E. C., & Dochy, F. (2013). Predicting general academic performance and identifying the differential contribution of participating variables using artificial neural networks. *Frontline Learning Research*, 1(1), 42-71.
- Naik, B., & Ragothaman, S. (2004). Using neural networks to predict MBA student success. *College Student Journal*, 38(1), 143.
- Odom, A. L., & Bell, C. V. (2015). Associations of middle school student science achievement and attitudes about science with student-reported frequency of teacher lecture demonstrations and student-centered learning. *International Journal of Environmental and Science Education*, 10(1), 87-97.
- Özkan, C., & Erbek, F. S. (2003). The comparison of activation functions for multispectral Landsat TM image classification. *Photogrammetric Engineering & Remote Sensing*, 69(11), 1225-1234.
- Öztemel, E. (2012). *Yapay Sinir Ağları [Artificial Neural Networks]*. İstanbul: Papatya Yayıncılık.
- Pike, G. R., Schroeder, C. C., & Berry, T. R. (1997). Enhancing the educational impact of residence halls: The relationship between residential learning communities and first-year college experiences and persistence. *Journal of College Student Development*, 38(6), 609-

621.

- Power, C., Robertson, F., & Baker, M. (1987). *Success in higher education* (No. 94). National Institute of Labour Studies.
- Rojas, R. (2013). *Neural Networks: A Systematic Introduction*. Springer Science & Business Media.
- Rusli, N. M., Ibrahim, Z., & Janor, R. M. (2008, August). Predicting students' academic achievement: Comparison between logistic regression, artificial neural network, and Neuro-fuzzy. In *Information Technology, 2008. ITSIM 2008. International Symposium on (I)*, 1-6. IEEE.
- Scheiber, C., Reynolds, M. R., Hajovsky, D. B., & Kaufman, A. S. (2015). Gender differences in achievement in a large, nationally representative sample of children and adolescents. *Psychology in the Schools*, 52(4), 335-348.
- Schumacher, P., Olinsky, A., Quinn, J., & Smith, R. (2010). A comparison of logistic regression, neural networks, and classification trees predicting success of actuarial students. *Journal of Education for Business*, 85(5), 258-263.
- Shalev-Shwartz, S. (2011). Online learning and online convex optimization. *Foundations and Trends in Machine Learning*, 4(2), 107-194.
- Siegelbaum, S. A., & Hudspeth, A. J. (2000). *Principles of neural science* (Vol. 4, pp. 1227-1246). E. R. Kandel, J. H. Schwartz, & T. M. Jessell (Eds.). New York: McGraw-hill.
- Simonson, M., Smaldino, S., Albright, M. & Zvacek, S. (2003). *Teaching and Learning at a Distance*. Ohio: Columbus.
- Snyder, C. R., Shorey, H. S., Cheavens, J., Pulvers, K. M., Adams, V. H. III, & Wiklund, C. (2002). Hope and academic success in college. *Journal of Educational Psychology*, 94(4), 820-826.
- Strayhorn, T. L. (2006). Factors influencing the academic achievement of first-generation college students. *NASPA Journal*, 43(4), 82-111.
- Suphi, N., & Yaratan, H. (2012). Effects of learning approaches, locus of control, socio-economic status and self-efficacy on academic achievement: a Turkish perspective. *Educational Studies*, 38(4), 419-431.
- Şen, B., Uçar, E., & Delen, D. (2012). Predicting and analyzing secondary education placement-test scores: A data mining approach. *Expert Systems with Applications*, 39(10), 9468-9476.
- Turhan, K., Kurt, B., & Engin, Y. Z. (2013). Estimation of Student Success with Artificial Neural Networks. *Education and Science*, 38(170).
- Valentine, J. C., DuBois, D. L., & Cooper, H. (2004). The relation between self-beliefs and academic achievement: A meta-analytic review. *Educational Psychologist*, 39(2), 111-133.
- Ventura, P. R. (2005). Identifying predictors of success for an objects-first CS1. *Computer Science Education*, 15(3).
- Yu, H., Xie, T., Paszczynski, S., & Wilamowski, B. M. (2011). Advantages of radial basis function networks for dynamic system design. *Industrial Electronics, IEEE Transactions on*, 58(12), 5438-5450.
- Yukselturk, E., Ozekes, S., & Türel, Y. K. (2014). Predicting dropout student: an application of data mining methods in an online education program. *European Journal of Open, Distance and e-Learning*, 17(1), 118-133.
- Zheng, J. L., Saunders, K. P., Shelley, I. I., Mack, C., & Whalen, D. F. (2002). Predictors of academic success for freshmen residence hall students. *Journal of College Student Development*, 43(2), 267.