Hacettepe Üniversitesi Eğitim Fakültesi Dergisi

Hacettepe University Journal of Education

e-ISSN: 2536-4758

A Model Proposal for Predicting Students' Academic Performances Based on Data Mining*

Murat ALTUN, Kemal KAYIKÇI***, Sezgin IRMAK******

Article Information

Received:
06.07.2020

Accepted:
19.03.2021

Online First:
10.04.2021

Published:
31.07.2022

ABSTRACT

The purpose of this research is to propose a model to predict the academic performances of students and prevent their possible academic failures in the future. This research was conducted as an educational data mining application. The academic data of 1570 students who graduated from four departments of Akdeniz University, Faculty of Education between the year 2012 and 2017 were used in the research. All the exam scores, final grades, weighted averages of semester grades, and graduation grades were used in the study. Two main models have been developed for predicting students' academic success using data mining techniques and algorithms. The First Model is the Student Graduation Grade Estimation Model. This model is aimed to predict the future graduation grades of the students. Sub-models were developed using Artificial Neural Networks and Multiple Linear Regression Analysis. It was observed that the developed models predict the graduation grade of the students with an accuracy of 94% to 97% from the 1st semester's data. The second model developed in this research is the Early Warning Model for Students' Possible Academic Failures in the Future. The model predicts whether the general weighted average grades will fall below 2 in the future, according to the students' 1st year's 1st-semester grades. Under this model, the accuracy of the sub-models which were developed using Logistic Regression and Decision Trees was found to be 72% to 87%. As a result of the research, a model was proposed to prevent the academic failures in the future by predicting the student's academic performances. It can be asserted that educational institutions can benefit effectively and efficiently to increase students success by using the proposed model.

Keywords: Prediction of the academic performance, academic warning system, educational data mining, decision trees, artificial neural networks

doi: 10.16986/HUJE.2021068491 Article Type: Research Article

Citation Information: Altun, M., Kayıkçı, K., & Irmak, S. (2022). A model proposal for predicting students' academic performances based on data mining. *Hacettepe University Journal of Education, 37*(3), 1080-1098. doi: 10.16986/HUJE.2021068491

1. INTRODUCTION

This section covers the definition of data mining, data mining process, and data mining activities conducted in the field of education.

1.1. Statement of the Problem

A success indicator of an educational institution is the learning outcomes of students. Therefore, one of the most important duties of educational organizations and administrators is to improve student success (Karpicke & Murphy, 1996). Educational organizations must continuously achieve improvement to raise student success. Ensuring continuous improvement is possible by continuously monitoring and assessing the organization (Öztürk, 2009). For a health monitoring and assessment, information obtained from the processing of data acquired from the educational environment is required.

Process of turning data to information, method, and all approaches are expressed as Data Mining (DM). Various concepts are used for DM in the literature: data mining from databases, information retrieval, data pattern analysis, and data archaeology. The objective in the DM process is information hidden and not visible in "big data" (Fayyad, Piatetsky-Shapiro & Smyth, 1996).

* This study was conducted as a part of the Ph.D. thesis study of the first author and supported by Akdeniz University SRP Unit with project no SDK-2018-4077. The ethic permission was received from Akdeniz University Social Sciences and Humanities Scientific Research and Publication Ethics Committee on 02.05.2018 and issue no 54.
** Dr., Teacher, Antalya-TURKEY. e-mail: emuratatltun@gmail.com (ORCID: 0000-0001-9548-8317)
*** Prof. Dr., Akdeniz University, Faculty of Education, Department of Educational Sciences, Division of Educational Administration, Antalya-TURKEY. e-mail: kemalkayikci@akdeniz.edu.tr (ORCID: 0000-0003-3330-5452)
**** Assoc. Prof. Dr., Akdeniz University, Faculty of Applied Sciences, Department of Management Information Systems, Division of Decision Sciences, Antalya-TURKEY. e-mail: sezgin@akdeniz.edu.tr (ORCID: 0000-0003-2188-7344)

DM is based on the idea of finding connections or patterns and creating computer programs to automatically scan databases estimating the future (Witten & Frank, 2005). In DM applications computer programs are used to search patterns, relations, and rules that would enable estimations in large amounts of data (Alpaydın, 2000). One of the most frequently used terms for DM is knowledge discovery in databases "KDD" (Akgöbek & Kaya, 2011). KDD process is presented in Figure 1. In this approach, DM is discussed as a part of the knowledge discovery process (Fayyad et al, 1996).

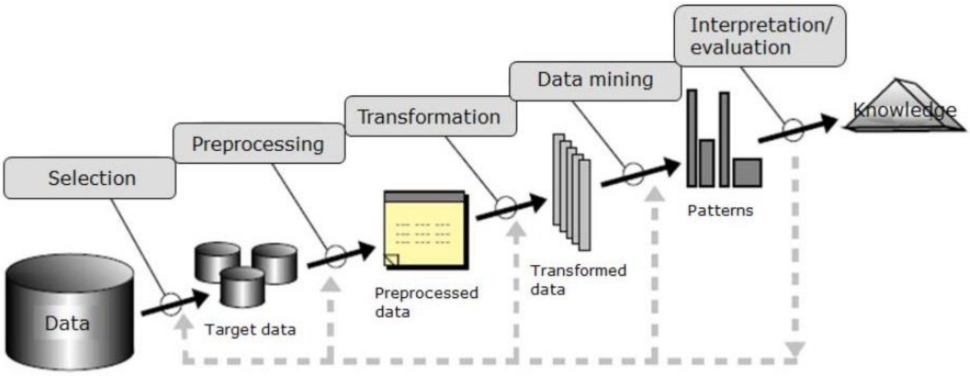

Figure 1.KDD Process
Source: Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). From data mining to knowledge discovery in databases. AI magazine, 17(3), 37-54 (1996).

KDD (Figure 1) consists of selection, preprocessing, transformation, data mining, and evaluation processes (i) Selection: Determination and sorting out of data that shall be used in all data sources. (ii) Preprocessing: This stage is where errors are sorted out in selected data, deficient records are elicited or completed. Improving the quality of data in preprocessing stage also improves the quality of information (iii) Transformation: This is the stage where integration, unification, and necessary transformation procedures are conducted on data received from different data sources (dataset, database). (iv) Data Mining: This is the stage where suitable DM methods are used on the prepared dataset to create models. (v) Evaluation: At this stage, the discovered information is assessed according to simplicity, validity, innovativeness, and usefulness criteria (Fayyad et al, 1996).

DM is used in science, engineering, security-intelligence, health-biomedical, banking finance, customer relations management, and education fields. When studies on data mining in the field of education are scanned at the Turkish Council of Higher Education Thesis Board, it is noted that terms such as educational data mining, instructional data mining, and data mining in education are used. Among such concepts, educational data mining is the most used (CoHE, 2020). Educational Data Mining (EDM) is a new discipline that uses a data mining approach to understand situations of students and learning environments better from large-scale data stacks on learning environments, students, and teachers (IEDMS, 2019). EDM process is a cyclical process whereby data collected from educational environments go through preprocessing, data mining, and final processing stages to turn into information and the acquired information is put into application in educational environments (García et al, 2011). Using EDM, works and processes such as assessment and evaluation, feedback to the student, and student profiling that require more time and labor using classical methods can be conducted in a short time and automatically (Lopez, Luna, Romero & Ventura, 2012). Student data is analyzed with EDM to determine student profiles (learning methods, success levels). Students with similar profiles can be aggregated to present suitable education-learning environments and content. Models created using EDM can be used to predict the academic performance of students. Thus, they can be used to improve a student's academic success and to provide guidance (Bienkowski, Feng & Means, 2012). Data at student affairs can be processed with DM to research subjects such as reasons of success and failure of students, which must be focused on to improve success, the relationship between university entrance scores and school success to improve quality and performance of education (Akgöbek & Çakır, 2009). It is seen that statistics and data-based management mentality that is one of the principles of Total Quality Management has been increasing its impact on the field in Turkey (Kavrakoğlu, 1998; Kayıkçı, 1999). Ministry of National Education 2023 Education Vision Document places a major emphasis on the application of data-based decision making, using large data in education, application of learning analytics, and educational data mining on educational administration (MoNE, 2018).

When EDM studies in the literature are evaluated, those that predict student's exam scores, course grades, graduation grades and classify their exam, class, or general success situations (successful/unsuccessful, pass/fail) are noted. In studies by Ayesha, Mustafa, Sattar, and Khan (2010) and Tsai C-F., Tsai C-T., Hung & Hwang (2011) models are developed to predict students' exam scores. Baradwaj & Pal (2011) conducted a study to predict success at "computer applications" class and determine factors impacting success at class. In a study by Şen, Uçar & Delen (2012) classification method is used to predict the success of students at the transition to higher education. In a study by Şengür (2013) models are developed to predict graduation grades of undergraduate students. Bahadır (2013) used academic grades of students during their undergraduate education to predict their success levels at graduate education as "successful/unsuccessful". Akçapınar (2014) developed a model in his study whereby he used datasets in online learning environments to classify their class successes as "pass/fail". In his study, Alsuwaiket

(2018) used data based on TIMSS 2015 surveys to predict student success in mathematics and determine the important factors affecting success.

Education keeps a large amount of data regarding students, teachers, learning environments, assessment, and evaluation results. Processing of the data at hand with EDM can be used to improve effectiveness, efficiency, and quality of learning at education institutions. In this study which is an applied EDM study models are developed on real education data. It can be argued that when models developed in the study are assessed according to performance criteria, they can be used to predict student success and be useful in improving student success.

1.2. Purpose of the Study

The objective of this study was to conduct an applied EDM study to predict the graduation grades and success levels of students at Akdeniz University Faculty of Education. Different DM methods were used according to this objective and models were developed. It was considered that models developed in the study could be used to improve student success.

1.3. Problem of the Study

Could data of students at Akdeniz University Faculty of Education on Student Information System (SIS) be processed to develop models predicting their academic success levels?

1.3.1. Sub-problems of the study

1. Could exam and class grades of Turkish Teaching, Social Sciences Teaching, Teaching Mathematics in Primary Schools, and Science Teaching students be used to predict their graduation grades?
   a. Freshman 1st Semester Midterm Grades
   b. Freshman 1st Semester Class Grades
   c. Freshman 1st and 2nd Semesters Grade Point Averages
2. Could the possible failure of Turkish Teaching, Social Sciences Teaching, Teaching Mathematics in Primary Schools, and Science Teaching students be predicted using freshman 1st-semester class grades?

2. METHODOLOGY

KDD process was followed in this study in a correlational survey model. Datasets on SIS were selected to perform preprocessing and transforming procedures. Data mining techniques were used to create models and such models were assessed and interpreted according to performance criteria. At the end of the process, the study tried to develop valid models to predict the academic success of students and information based on such models.

2.1. Participants

The count of students that graduated from departments of Akdeniz University Faculty of Education between 2013 and 2017 in the scope of the study was presented in Table 1.

Table 1.
Number of Graduates between 2013 and 2017 According to Departments

SN	Years	2013	2014	2015	2016	2017	Total
1	Turkish Teaching	30	98	117	64	69	378
2	Social Sciences Teaching	32	125	151	80	75	463
3	Teaching Mathematics in Primary Schools	0	97	90	54	81	322
4	Science Teaching	23	103	132	84	65	407
	Number of Faculty of Education Graduates	85	423	490	282	290	1570

The study was conducted using datasets of 1570 students that graduated from Akdeniz University Faculty of Education between 2013 and 2017 (Table 1) on SIS.

2.2. Data Collection

The required permissions were received from the ethics committee to access data on SQL Server relational database on SIS in the scope of the study. Following procedures to receive permission SIS datasets were anonymized and private areas were filtered by the relevant unit of the university to define a user profile for access. Datasets that could be required for the study were accessed through this user profile. From the data collection stage to assessment, the training version of the Rapid Miner program that is a data mining software developed by Dortmund Technology University Artificial Intelligence unit was used (Şeker, 2016). Rapid Miner is an integrated data science platform where data analysis, data preprocessing, data mining methods, algorithms are used, models created and assessed (Rapid Miner, 2019). Datasets used in the research as a result of preprocessing using Rapid Miner were presented in Table 2.

Table 2.
SIS Database Tables Used in the Study

SN	Table Name	Explanation	Number of Registries	Columns *
1	**Academic calendar**	**Start and end dates of academic calendar**	43	3
2	**Class department**	**Classes that are given at departments**	367817	35
3	**Class department exam**	**Exams given at departments**	839439	14
4	**Graduation**	**Graduated students**	49979	4
5	Graduate student transcript	Student class transcript grade	277521	12
6	**Student**	**Table used for student-person mapping**	161227	2
7	Student success grading	Academic success grading of students	3509	10
8	**Student unit**	**Unit of student and organization mapping**	971	5
9	**Student class enrollment**	**Data on student class enrollments and class grades of enrolled classes**	5590847	14
10	Student disciplinary punishment	Student Disciplinary Affairs	925	12
11	Student Registration	Student Registration information Separate from Student Program Registration Table students from other universities that take classes in summer school etc. is also registered.	171465	5
12	Student Registration Additional Information	Where more registration information. Does not include data on all students.	66812	14
13	Student SSPC	Student data from SSPC	106883	175
14	**Student Program**	**University departments**	686	10
15	Student Program Term	Semester grade point average of students	695337	9
16	Student Program Enrollment	Enrollment data of students	156793	12
17	**Student Exam Results**	**Exam results of students**	15147471	8
18	**Student Vocabulary**	**Vocabulary**	472404	3
19	Student Single Course Exam Enrollment	Single Course Exam data	7672	12
20	Student Transfer Information	Data of students transferred to other universities	7818	10
21	**Student and Class**	**All classes opened for students**	58044	43
22	Academic	Data on academics at university	3458	11
23	**Organization**	**Data on all units and organization at the university**	17053	6
24	**Vocabulary**	**Database vocabulary for general purposes includes equivalents of coding used.**	22530	4

* Number of properties in the table

Procedures of data selection, pre-processing and transformation processes are as follows: 1. Creation of datasets including students who graduated from the Faculty of Education, 2. Creation of datasets including classes and exam results in students graduated from Faculty of Education took in their freshman year 1st semester, 3. Creation of datasets including semester grade point averages of students who graduated from the Faculty of Education. At the last stage of data pre-processing, these datasets are integrated. Properties of dataset acquired as a result of data preprocessing and used for DM were presented in Table 3.

Table 3.
Dataset Used for EDM

Property	Explanation	Data type
Student Program Registration Id	Encoded student registration number	Numerical
Department	Department student graduated from	Date
Mid-Term Scores	Mid-term scores of classes taken in freshman year 1st semester	Numerical
SGPA	Grade point average for the semester. Separate for every semester.	Numerical
SG-4N	Class grades of classes taken in freshman year, 1st semester	Numerical
GPA	Student’s graduation grade	Numerical

Datasets are separately created for every department at the Faculty of Education to use at the modeling stage. Registration id, student’s department, midterm exam scores of a student, semester grade point averages, final class grades, and graduation grade as the primary key to access data in the dataset (Table 3) take place in a single dataset.

2.3. Modelling

The general model (Figure 2) developed to predict the academic performances of students in the scope of the study is as follows.

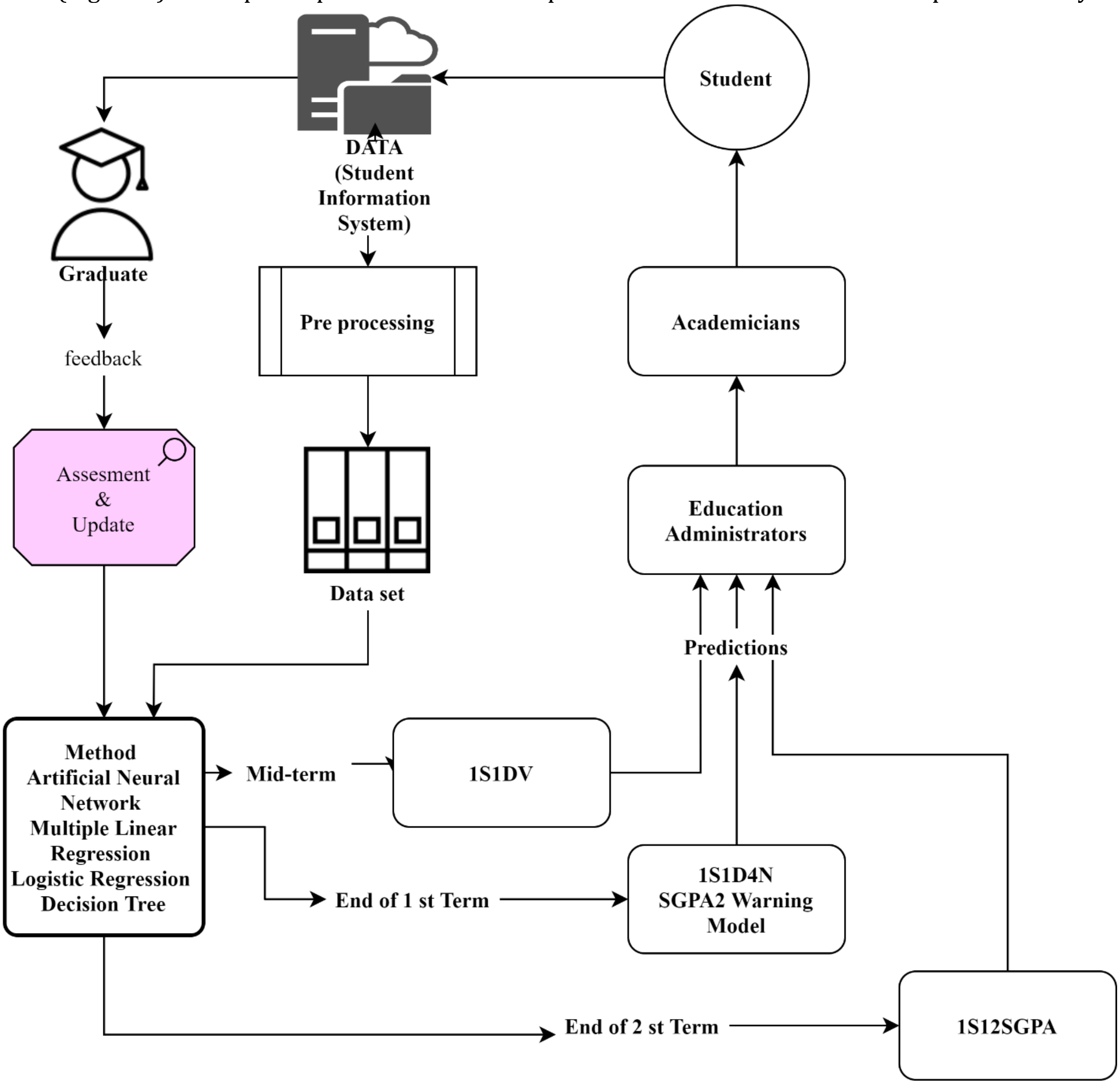

Figure 2. A Suggested Model to Predict Student Academic Performance

Explanation of sub-models in the suggested model to predict student academic performance:

1S1DV: Graduation grade prediction model using freshman year 1st-semester mid-term scores of students. Used to predict academic success in the early period. Thanks to this model, the graduation grade of new students can be predicted as soon as their midterm scores are announced. Thus, initial predictions on student success/failure status could be achieved in the shortest possible time.
1S1D4N: Model for prediction of graduation grade using student's freshman year 1st-semester class grades.
1S12SGPA: Used for prediction of graduation grade using freshman year 1st and 2nd-semester grade point averages of the student.
SGPA2 Warning Model: Used to predict failure (semester grade point average being lower than two) using freshman year 1st-semester class grades of the student.

In the model on predicting student academic performance (Figure 2), when freshman year 1st-semester midterm scores of students are registered to SIS, the 1S1DV model would become functional. Midterm scores of a student were entered in the student's department's 1S1DV model to predict a student's graduation grade. When used for new students this model enables prediction in the early period. Thus, a prediction of success status could be made for students that only took midterms and whose class grades were not given yet. It was targeted to predict success status from as early stage as possible (right after first exam results) to take necessary precautions and actions against students' graduation grades and the possibility of failure. At the end of freshman year 1st semester, students would have taken final exams of the semester and their semester class grades would be calculated. One semester class grades were announced on SIS, 1S1D4N model would be functional. SGPA2 Warning Model that predicts a student's academic failure status acting as a warning system would also be active during this period. SGPA2 Warning Model would determine students whose grade point averages have the possibility of falling under 2 in the following semesters. 1S12SGPA model would become functional when the student's complete freshman year. When class grades would be announced on SIS at the end of the freshman year, students would have a grade point average data for both the 1st and the 2nd semesters and the 1S12SGPA model could predict students' graduation grades. Information received thanks to the model would be conveyed to academic managers to the relevant academic counselors to present guidance and academic counseling to students. The model used for midterms gives an advantage of early prediction of graduation grades. Models used in the following periods (end of semester, freshman year) would be used for providing better prediction of graduation grade and failure. Developing more than one model would present alternatives at various stages between early prediction and late but more precise prediction of student success.

2.3. Data Analysis

In this chapter, there are explanations regarding DM models used at the modeling stage. While the student graduation grade prediction model was being prepared, Multiple Linear Regression Analysis (MLRA) and Artificial Neural Networks (ANN) methods among DM techniques were used.

In data analysis, missing data were coded as zero points (0: did not take the exam). Missing data or points lower than 1 (disciplinary punishment, semester repeat, horizontal transfer, vertical transfer, missing data, etc.) in semester grade point average were not included in the analysis to filter such cases.

Multiple Linear Regression Analysis (MLRA): Used to formulate the relationship of more than one independent variable with the dependent variable. For a model with P number of explanatory variables MLRA equation can be written as $Y = b_0 + b_1X_1 + b_2X_2 + ... + bp\ Xp$. In case Y is the answer of a dependent variable, X's represent p explanatory variables and b_0 are regression coefficients (SAS, 2019). Required assumptions for MLRA were checked with Rapid Miner. Graduation grade, semester grades, and exam scores were continuous variables measured in equal-interval scale (Büyüköztürk, 2018). Whether or not the dataset had normal distribution was controlled with skewness and kurtosis values. The existence of a linear relationship was assessed using the Pearson correlation coefficient. Since the minimum number of observations in the dataset was 259, it was found sufficient for analysis. Variables showing multicollinearity (exam scores, grades) were not included in the analysis.

ANN: Modern neural networks are usually used to model complicated relationships between inputs and outputs or to find hidden patterns in data. ANN improves the network model using the backpropagation algorithm. The backpropagation algorithm is a controlled learning method with two stages: diffusion and weight update. Diffusion and weight update are repeated until the performance of the network are good enough. At backpropagation algorithms output values are compared to the correct answer to calculate values of some predefined effort functions. Error feeds back from the network using various methods. Using this information algorithm sets the weight of every connection to lower the value of the error function to a small extent. After this process is repeated enough times for the training cycle, the network usually turns into a case where the calculation error is smaller. In this case, the network turns into a situation that can model results with fewer errors (Şeker, 2008; SAS, 2019; Rapid Miner, 2019).

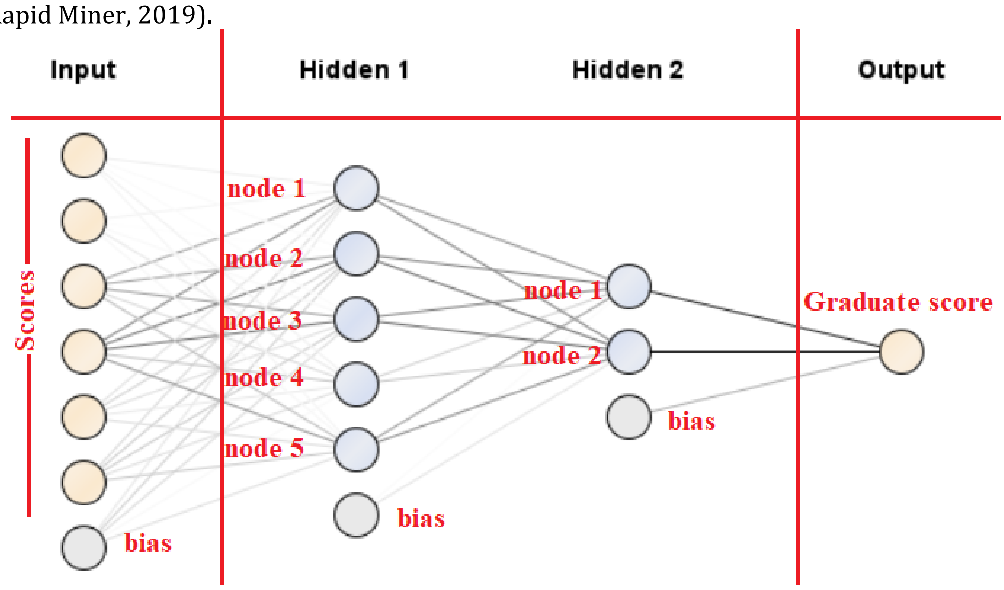

Figure 3. ANN Model

While Student Graduation Score Prediction Models were developed in the study ANN was designed as in Figure 3. ANN was made up of input layers, hidden layers, and output layers. Based on the model, the input layer was the layer when students' midterm, semester class grades, or 1st semester and 2nd-semester grade point averages were entered. Decision layers were layers that had at least one hidden layer and where calculations were made. In the decision layer data from the entry layer was linked to nodes. For every node weight was calculated using a function and was linked to the following layer as input. The output layer was the node where graduation grade was. When the ANN model was trained, inputs and weights at the dataset were constantly updated to develop the model that would calculate graduation grade with minimum errors. In this model, the number of classes and scores of students changed based on departments they graduated from. To acquire the most successful model, the number of layers in decision layers and nodes in every layer was tried using different values.

While Student Graduation Score Prediction Models were developed, Logistic Regression and Decision Tree among DM methods were used (for giving more successful results in practice). They were used to classify students' success status (successful/unsuccessful) in Student Graduation Score Prediction Model. Logistic Regression is a method that classifies on a dual categorical label. Therefore, student success status on the dataset was turned into binominal data type in the study.

Logistic Regression: Used to create a model to classify a series labeled as one of two categories. In Logistic Regression a mathematical model is established and for instance probability of belonging to a class is calculated (Rapid Miner, 2019).

Decision Trees: A decision tree is a tree-like collection of nodes targeting a classification or creation of the prediction of a target value. A decision tree is created by algorithms defining various ways of separating a dataset into sections. Every node represents a division rule for a certain feature. For classification, this rule separates values belonging to different classes and tries to lower error most suitably according to the selected parameter criteria. Exploration of decision rule determining branches is based on a method revealing the relationship between the target (result) and one or more inputs (SAS, 2019; Rapid Miner, 2019).

Assessment of Models: Cross-validation is used for training and test of models created in the study. Cross-validation is composed of testing and training sub-processes. Training sub-process is used for training a created model. The training model is then used at the test sub-process. The performance of the model is measured in the test stage. Dataset input is separated into "k" subsets in equal size. As a test dataset of "k" subsets, one subset is separated (for the test). The remaining "k – 1" subsets are used for training the model. Afterward cross-validation process is repeated for k times, every one of the k subsets is used as test data. To produce one prediction, an average of "k" results from "k" repeats is taken (Rapid Miner, 2019). In the study, the dataset was separated into 4 equal pieces and in training and testing of models 3 subsets were used for training and 1 subset was used for testing, whereby all datasets were used. The shuffle sampling method was used while using the dataset in cross-validation. This method was chosen to avoid possible errors in training and testing of the model with mixed-use of the dataset.

Performance indicators used in Student Graduation Score Prediction Models are as follows: RMSE, AE, MSE, R^2, MAPE.

RMSE: Root Mean Squared Error

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n} e_t^2}$$

MAE: Mean Absolute Error

$$MAE = \frac{1}{n}\sum_{t=1}^{n} |e_t|$$

MSE: Mean Squared Error

$$MSE = \frac{1}{n}\sum_{t=1}^{n} e_t^2$$

MAPE: Mean Absolute Percentage Error

$$\mathrm{MAPE} = \frac{100\%}{n}\sum_{t=1}^{n}\left|\frac{e_t}{y_t}\right|$$

R^2: Least squares

$$R^2 = 1 - \frac{\sum_{t=1}^{n} e_t^2}{\sum_{t=1}^{n}(y-\bar{y})^2}$$

Terms

e: Error term (the difference between the real value and prediction)

y: Actual value

t: Index number

n: Number of observations

y': Prediction

$\bar{y}$: Mean of actual value.

In Student Academic Early Warning Model, the Confusion Matrix was used as a performance indicator. True positive, false positive, true negative, false-negative classification rates, and the model's general correct classification success were assessed.

The class labeled negative are those that had at least one-semester grade point average (SGPA) lower than 2. The class labeled positive are those that had no SGPA lower than 2.

True Positive: Positive in fact and classified positive by the model
False Positive: Negative in fact but classified positive by the model.
True Negative: Negative in fact and classified negative by the model.
False Negative: Positive in fact but classified negative by the model.
General correct classification rate: Sum of the number of students classified as true positive and the number of students classified as true negative divided by the total number of students.

3. FINDINGS

Findings on the First Sub Problem of the Study

Using midterm scores, class grades, and grade point averages of 1st and 2nd semesters were used for 4 departments in the Education Faculty to develop 3 prediction models. In every prediction model, two separate methods (MLRA and ANN) were used to make comparisons.

Turkish Teaching Freshman Year 1st Semester Midterm (1S1DV) Model: Data of 321 students that had no missing information out of 378 students who graduated from Turkish Teaching were used to develop 1S1DV-MLRA and 1S1DV-ANN models. In the 1S1DV-MLRA model Ottoman Turkish class was not included in the analysis because it had multicollinearity (Table 4). Results of Turkish Teaching 1S1DV-MLRA were presented in Table 4.

Table 4.
Turkish Teaching 1S1DV-MLRA

Variables	***B***	***Standard Error***	***ß***	***t***	***p***
Ataturk's Principles and History of Turkish Revolution I	.001	.001	.067	1.434	.153
Knowledge and Theories of Literature I	.003	.001	.177	3.704	.000
Introduction to Education Science	.003	.001	.175	3.382	.001
Oral Narrative I	.004	.001	.188	3.710	.000
Turkish Grammar I: Phonetics	.004	.001	.247	4.928	.000
Turkish I Written Narrative	.004	.001	.150	3.178	.002
Writing Methods	.006	.002	.198	4.152	.000
English I	.003	.001	.168	3.514	.001
RMSE=.207 MAE=.168 MSE=.043, R^2=.347 MAPE= 5.172%					

Classes with significant coefficients on Table 4 ($p< .05$) and regression constant were used to formulate graduation grade as below. Graduation Grade **=** 1.091 + .003 * Knowledge and Theories of Literature I + .003 * Introduction to Education Science + .004 * Oral Narrative I + .004 * Turkish Grammar I: Phonetics + .004 * Turkish I Written Narrative + .006 * Writing Methods + .003 * English I

Turkish Teaching 1S1DV second model was developed using ANN. Turkish Teaching 1S1DV-ANN model had a structure that used midterm scores of 9 classes input, two layers, 2 nodes in the 1st layer, 5 nodes in the 2nd layer, and used graduation grade as output. As a result of modeling, the best parameters in terms of performance indicators were determined as – activation function: Sigmoid, learning cycle: 2000, learning ratio .01, and moment .1. It was understood that the model with 1st layer 2 nodes and 2nd layer 5 nodes were more successful compared to other models. Performance indicators of the model that was developed were: RMSE=.201 AE=.164 MSE=.041 R^2=.387 MAPE= 5.040%. The model that was developed could explain 36.5% of the variance between midterm scores of 9 classes in freshman year 1st semester (classes on Table 4 and Ottoman Turkish) and graduation grades (R^2 =.365). Based on MAPE value it could be argued that the model predicts graduation grades with a 5.040% error.

Turkish Teaching Freshman Year 1st Semester Grade (1S1D4N) Model: Data of 377 students that had no missing class grade information out of 378 students graduated from Turkish Teaching were used to develop 1S1D4N-MLRA and 1S1D4N-ANN models. Results of Turkish Teaching 1S1D4N-MLRA were presented in Table 5.

Table 5.
Turkish Teaching 1S1D4N MLRA

Variables	B	Standard Error	ß	t	p
Ataturk's Principles and History of Turkish Revolution I	.035	.013	.106	2.584	.010
Knowledge and Theories of Literature I	.068	.014	.197	4.724	.000
Introduction to Education Science	.042	.012	.138	3.416	.001
Ottoman Turkish I	.034	.010	.139	3.267	.001
Oral Narrative I	.030	.013	.100	2.340	.020
Turkish Grammar I: Phonetics	.057	.013	.193	4.280	.000
Turkish I Written Narrative	.046	.013	.140	3.427	.001
Writing Methods	.081	.017	.188	4.779	.000
English I	.060	.010	.225	5.790	.000
RMSE=.193 MAE=.156 MSE=.037 R^2 =.446 MAPE= 4.812%					

Classes with significant coefficients on Table 5 ($p< .05$) and regression constant were used to formulate graduation grade as below. Graduation Grade **=** .035 * Ataturk's Principles and History of Turkish Revolution I .068 * Knowledge and Theories of Literature I + .042 * Introduction to Education Science + .034 * Ottoman Turkish I + .030 * Oral Narrative I + .057 * Turkish Grammar I: Phonetics + .046 * Turkish I Written Narrative + .081 * Writing Methods + .060 * English I + 1.949

Turkish Teaching 1S1D4N second model was developed using ANN. Turkish Teaching 1S1D4N-ANN model had a structure that used midterm scores of 9 classes input, two layers (Hidden 1, Hidden 2), 2 nodes in the 1st layer, 5 nodes in the 2nd layer, and used graduation grade as output. As a result of modeling, the best parameters in terms of performance indicators were determined as – activation function: Sigmoid, learning cycle: 3000, learning ratio .01, and moment .1. It was understood that the model with 1st layer 2 nodes and 2nd layer 5 nodes were more successful compared to other models. Performance indicators of the model that was developed were: RMSE=.189 AE=.152 MSE=.036 R^2=.466 MAPE= 4.700%. The model that was developed could explain 46.6% of the variance between midterm scores of 9 classes in freshman year 1st semester and graduation grades (R^2 =.466). Based on MAPE value it could be argued that the model predicts graduation grades with 4.700% error on average.

Turkish Teaching Freshman Year 1st and 2nd Semester Grade Point Averages (1S12SGPA) Model: Data of 330 students that had no missing information out of 378 students graduated from Turkish Teaching were used to develop 1S12SGPA-MLRA and 1S12SGPA-ANN models. Results of Turkish Teaching 1S12SGPA-MLRA were presented in Table 6.

Table 6.
Turkish Teaching 1S12SGPA-MLRA

Variables	B	Standard Error	ß	t	p
1st Semester	.193	.023	.339	8.320	.000
2st Semester	.312	.023	.547	13.437	.000
RMSE=.162 MAE=.128 MSE=.026 R^2 =.606 MAPE= 3.960%					

Classes with significant coefficients on Table 6 ($p< .05$) and regression constant were used to formulate graduation grade as below. = Graduation Grade= .193 * 1st Semester + .312* 2nd Semester GGPA + 1.743

Turkish Teaching second 1S12SGPA model was developed using ANN. Turkish Teaching 1S12SGPA-ANN model had a structure that used grade point averages of the 1st and 2nd semesters of freshman year as 2 input, single layer (Hidden 1), 3 nodes, and used graduation grade as output. As a result of modeling, the best parameters in terms of performance indicators were determined as – activation function: Sigmoid, learning cycle: 2000, learning ratio .01, and moment .1. It was understood that the model with a single layer and single node was more successful compared to other models. Performance indicators of the model that was developed were: RMSE=.156 AE=.121 MSE=.025 R^2=.634 MAPE = 3.7774%. The model that was developed could explain 63.4% of the variance between grade point averages of 1st and 2nd semesters of freshman year and graduation grades (R^2 =.634). Based on MAPE value it could be argued that the model predicts graduation grades with a 3.774% error on average.

Social Sciences Teaching Freshman Year 1st Semester Midterm (1S1DV) Model: Data of 404 students that had no missing information out of 463 students graduated from Social Sciences Teaching were used to develop 1S1DV-MLRA and 1S1DV-ANN models. Ataturk's Principles and History of Turkish Revolution class was not included in the analysis because it had multicollinearity (Table 7). Results of Social Sciences Teaching 1S1DV-MLRA were presented in Table 7.

Table 7.
Social Sciences Teaching 1S1DV-MLRA

Variables	B	Standard Error	ß	t	p
Archaeology	.002	.001	.115	2.653	.008
Computer I	.001	.001	.078	1.780	.076
Introduction to Education Science	.002	.001	.143	3.297	.001
Basics of Social Sciences	.002	.001	.167	3.601	.000
Social Psychology	.004	.001	.214	4.835	.000
Sociology	.002	.001	.159	3.443	.001
Turkish I Written Narrative	.005	.001	.267	5.872	.000
English I	.002	.001	.174	4.064	.000
RMSE=.188 MAE=.151 MSE=.035 R^2 =.343 MAPE(%)= 4.769					

Classes with significant coefficients on Table 7 (p< .05) and regression constant were used to formulate graduation grade as below.

Graduation Grade = 1.708 + .002 * Archaeology + .002 * Introduction to Education Science + .002 * Basics of Social Sciences + .004 * Social Psychology + .002 * Sociology + .005 * Turkish I Written Narrative + .002 * English I

Social Sciences Teaching second 1S1DV model was developed using ANN. Social Sciences Teaching 1S1DV-ANN model had a structure that used midterm scores of 9 classes input, single layer (Hidden 1), single node, and used graduation grade as output. As a result of modeling, the best parameters in terms of performance indicators were determined as – activation function: Sigmoid, learning cycle: 1000, learning ratio .01, and moment .1. It was understood that the model with a single layer and single node was more successful compared to other models. Performance indicators of the model that was developed were: RMSE=.185 AE=.148 MSE=.034 R^2=.365 MAPE= 4.673%. The model that was developed could explain 36.5% of the variance between midterm scores of 9 classes in freshman year 1st semester (classes on Table 7 and Ataturk's Principles and History of Turkish Revolution) and graduation grades (R^2 =.365). Based on MAPE value it could be argued that the model predicts graduation grades with a 4.673% error on average.

Social Sciences Teaching Freshman Year 1st Semester Class Grade Model (1S1D4N): Data of 463 students who graduated from Social Sciences Teaching were used to develop 1S1D4N MLRA and 1S1D4N ANN models. Results of Social Sciences Teaching 1S1D4N-MLRA were presented in Table 8.

Table 8.
Social Sciences Teaching 1S1D4N-MLRA

Variables	B	Standard Error	ß	t	p
Archaeology	.025	.010	.086	2.370	.018
Ataturk's Principles and History of Turkish Revolution	.019	.009	.077	2.201	.028
Computer I	.032	.007	.148	4.215	.000
Introduction to Education Science	.057	.011	.180	5.010	.000
Basics of Social Sciences	.025	.008	.116	3.018	.003
Social Psychology	.059	.011	.193	5.314	.000
Sociology	.052	.010	.198	5.227	.000
Turkish I Written Narrative	.072	.012	.206	5.788	.000
English I	.049	.008	.213	6.022	.000
RMSE=.163 MAE=.131 MSE=.027 R^2 =.517 MAPE= 4.130%					

Social Sciences Teaching second 1S1D4N model was developed using ANN. Social Sciences Teaching 1S1D4N-ANN model had a structure that used midterm scores of 9 classes input, two layers (Hidden 1, Hidden 2), 5 nodes in the 1st layer, 2 nodes in the 2nd layer, and used graduation grade as output. As a result of modeling, the best parameters in terms of performance indicators were determined as – activation function: Sigmoid, learning cycle: 1000, learning ratio .01, and moment .1. It was understood that the model with 1st layer 5 nodes and 2nd layer 2 nodes were more successful compared to other models. Performance indicators of the model that was developed were: RMSE=.159 AE=.128 MSE=.026 R^2=.541 MAPE= 4.018%. The model that was developed could explain 54.1% of the variance between semester grades of 9 classes in freshman year 1st semester and graduation grades (R^2 =.541). Based on MAPE value it could be argued that the model predicts graduation grades with a 4.018% error on average.

Social Sciences Teaching Freshman Year 1st and 2nd Semester Grade Point Averages (1S12SGPA) Model: Data of 414 students that had no missing information out of 463 students graduated from Social Sciences Teaching were used to develop 1S12SGPA-Multiple Regression Analysis Model and 1S12SGPA-ANN model. Results of Social Sciences Teaching 1S12SGPA-Multiple Regression Analysis Model were presented in Table 9. As it can be seen in Table 9, classes with significant coefficients (p<.05) and regression coefficient are used to formulate graduation grade as below. Graduation Grade = .232 * 1st Semester GGPA + .1178* 2nd Semester GGPA + 2.113

Table 9.
Social Sciences Teaching 1S12SGPA-MLRA

Variables	B	Standard Error	ß	t	p
1st Semester	.232	.021	.446	10.886	.000
2st Semester	.178	.018	.401	9.792	.000
RMSE=.147 MAE=.120 MSE=.022 R^2 =.581 MAPE(%)= 3.791					

It could be argued that in the calculation of student's graduation grade in 1S12SGPA-MLRA developed for Social Sciences Teaching, 1st-semester grade point average and 2nd-semester grade point average has significant contributions. The developed model could explain 58.1% of the variance between the grade point average of the 2nd semester and graduation grade (R^2 =.581). When MAPE value is considered it could be argued that the model predicted graduation grades with 3.791% error on average. Social Sciences Teaching second 1S12SGPA model was developed using ANN. Social Sciences Teaching 1S12SGPA-ANN model had a structure that used grade point averages of the 1st and 2nd semesters of freshman year as 2 input, single layer (Hidden 1), single node, and used graduation grade as output. As a result of modeling, the best parameters in terms of performance indicators were determined as – activation function: Sigmoid, learning cycle: 1000, learning ratio .01, and moment .1. It was understood that the model with a single layer and single node was more successful compared to other models. Performance indicators of the model that was developed were: RMSE=.145 AE=.117 SE=.021 R^2=.591 MAPE = 3.700%. The model that was developed could explain 59.1% of the variance between grade point averages of 1st and 2nd semesters of freshman year and graduation grades (R^2 =.591). Based on MAPE value it could be argued that the model predicts graduation grades with a 3.700% error on average.

Mathematics Teaching Freshman Year 1st Semester Midterm (1S1DV) Model: Data of 259 students that had no missing information out of 322 students who graduated from Mathematics Teaching were used to develop 1S1DV-MLRA and 1S1DV-ANN models. The Computer I class was not included in the analysis because it had multicollinearity (Table 10). Results of Mathematics Teaching 1S1DV-MLRA were presented in Table 10.

Table 10.
Mathematics Teaching 1S1DV-MLRA Analysis

Variables	B	Standard Error	ß	t	p
Ataturk's Principles and History of Turkish Revolution	.001	.001	.034	.547	.585
Introduction to Education Science	.006	.001	.270	4.892	.000
General Mathematics	.007	.001	.390	7.134	.000
Turkish I Written Narrative	.002	.002	.075	1.133	.258
English I	.005	.002	.132	2.340	.020
RMSE=.281 MAE=.222 MSE=.079 R^2 =.226 MAPE(%)= 7.225					

As it can be seen on Table 10, classes with significant coefficients (p< .05) and regression constant were used to formulate graduation grade as below. Graduation Grade **=** 1.700 + .006 * Introduction to Education Science + .007 * General Mathematics + .005 * English I

Mathematics Teaching second 1S1DV model was developed using ANN. Mathematics Teaching 1S1DV-ANN model had a structure that used midterm scores of 6 classes (Classes on Table 10 and Computer I) input, single layer (hidden 1), two nodes, and used graduation grade as output. As a result of modeling, the best parameters in terms of performance indicators were determined as – activation function: Sigmoid, learning cycle: 1000, learning ratio .01, and moment .1. It was understood that the model with a single layer and two nodes was more successful compared to other models. Performance indicators of the model that was developed were: RMSE=.275 AE=.216 MSE=.076 R^2=.255 MAPE = 7.037%. The model that was developed could explain 25.5% of the variance between midterm scores of 6 classes in freshman year 1st semester and graduation grades (R^2 =.255). Based on MAPE value it could be argued that the model predicts graduation grades with a 7.037% error on average.

Mathematics Teaching Freshman Year 1st Semester Grade Model (1S1D4N): Data of 322 students who graduated from Mathematics Teaching were used to develop 1S1D4N MLRA and 1S1D4N-ANN models. The computer I class was not included in the analysis because it had multicollinearity (Table 11). Results of Mathematics Teaching 1S1D4N-MLRA were presented in Table 11.

Table 11.
Mathematics Teaching 1S1D4N-MLRA

Variables	B	Standard Error	ß	t	P
Ataturk's Principles and History of Turkish Revolution I	.067	.017	.181	3.959	.000
Computer I	.032	.019	.077	1.677	.095
Introduction to Education Science	.079	.018	.204	4.375	.000
General Mathematics	.119	.014	.395	8.599	.000
Turkish I Written Narrative	.087	.021	.018	4.073	.000
RMSE=.245 MAE=.197 MSE=.060 R^2 =.386 MAPE(%)= 6.338					

Classes with significant coefficients on Table 11 (p< .05) and regression constant were used to formulate graduation grade as below. Graduation Grade = 2.053 + .067 * Ataturk's Principles and History of Turkish Revolution I + .079 * Introduction to Education Science + .119 * General Mathematics + .087 * Turkish I Written Narrative

Mathematics Teaching second 1S1D4N model was developed using ANN. Social Sciences Teaching 1S1D4N-ANN model had a structure that used semester grades of 6 classes (Classes on Table 11 and English I) input, two layers (Hidden 1, Hidden 2), 3 nodes in the 1st layer, 5 nodes in the 2nd layer, and used graduation grade as output. Various layers, numbers of nodes, and parameters were tested in the artificial neural network model. As a result of modeling, the best parameters in terms of performance indicators were determined as – activation function: Sigmoid, learning cycle: 2000, learning ratio .01, and moment .1. It was understood that the model with 1st layer 3 nodes and 2nd layer 5 nodes were more successful compared to other models. Performance indicators of the model that was developed were: RMSE=.239 AE=.193 MSE=.057 R^2=.422 MAPE = 6.173%. The model that was developed could explain 42.2% of the variance between semester grades of 6 classes in freshman year 1st semester and graduation grades (R^2 =.422). Based on MAPE value it could be argued that the model predicts graduation grades with a 6.173 % error on average.

Mathematics Teaching Freshman Year 1st and 2nd Semester Grade Point Averages (1S12SGPA) Model: Data of 260 students that had no missing information out of 322 students graduated from Mathematics Teaching were used to develop 1S12SGPA-MLRA and 1S12SGPA-ANN models. Results of Social Sciences Teaching 1S12SGPA-Multiple Regression Analysis Model were presented in Table 12.

Table 12.
Mathematics Teaching 1S12SGPA-MLRA

Variables	***B***	***Standard Error***	***ß***	***t***	***p***
1st Semester	.152	.030	.263	5.036	.000
2st Semester	.375	.033	.593	11.366	.000
RMSE=.195 MAE=.155 MSE=.038 R^2 =.637 MAPE= 5.016%					

As it can be viewed in Table 12, classes with significant coefficients on Table 6 (p< .05) and regression constant were used to formulate graduation grades as below. Graduation Grade= 1.560 + .152 * 1st Semester GGPA + .375* 2nd Semester GGPA Mathematics Teaching second 1S12SGPA model was developed using ANN. Mathematics Teaching 1S12SGPA-ANN model had a structure that used grade point averages of the 1st and 2nd semesters of freshman year as 2 input, single layer (Hidden 1), single node, and used graduation grade as output. Various layers, numbers of nodes, and parameters were tested in the artificial neural network model. As a result of modeling, the best parameters in terms of performance indicators were determined as – activation function: Sigmoid, learning cycle: 2000, learning ratio .01, and moment .1. It was understood that the model with a single layer and single node was more successful compared to other models. Performance indicators of the model that was developed were: RMSE=.190 AE=.150 SE=.037 R^2=.648 MAPE= 4.845%. The model that was developed could explain 64.8% of the variance between grade point averages of 1st and 2nd semesters of freshman year and graduation grades (R^2 =.648). Based on MAPE value it could be argued that the model predicts graduation grades with a 4.845% error on average.

Science Teaching Freshman Year 1st Semester Midterm (1S1DV) Model: Data of 357 students that had no missing information out of 407 students who graduated from Science Teaching were used to develop 1S1DV-MLRA and 1S1DV-ANN models. General Lab I class was not included in the analysis because it had multicollinearity (Table 13). Results of the Science Teaching 1S1DV Multi Linear Regression Analysis model were presented in Table 13.

Table 13.
Science Teaching 1S1DV-MLRA

Variables	***B***	***Standard Error***	***ß***	***t***	***p***
Ataturk's Principles and History of Turkish Revolution I	.001	.001	.043	.875	.382
Introduction to Education Science	.004	.001	.198	4.156	.000
General Physics I	.001	.001	.086	1.832	.068
General Physics Lab. I	.003	.001	.148	3.064	.002
General Chemistry I	.002	.001	.110	2.199	.029
General Mathematics I	.003	.001	.265	5.365	.000
Turkish I Written Narrative	.003	.001	.168	3.443	.001
RMSE=.234 MAE=.187 MSE=.055 R^2 =.291 MAPE (%) = 6.076					

Classes with significant coefficients on Table 13 (p< .05) and regression constant were used to formulate graduation grade as below. Graduation Grade = .004 * Introduction to Education Science + .003 * General Physics Lab. I + .002 * General Chemistry I + .003 * General Mathematics I + .003 * Turkish I Written Narrative + 1.959

Science Teaching second 1S1DV model was developed using ANN. Science Teaching 1S1DV-ANN model had a structure that used midterm scores of 8 classes (Classes on Table 13 and General Chemistry I) input, single layer (Hidden 1), two nodes, and used graduation grade as output. As a result of modeling, the best parameters in terms of performance indicators were determined as – activation function: Sigmoid, learning cycle: 1000, learning ratio .01, and moment .1. It was understood that the model with a single layer and two nodes was more successful compared to other models. Performance indicators of the model that was developed were: RMSE=.231 AE=.186 MSE=.054 R^2=.296 MAPE= 6.046%. The model that was developed could explain 29.6% of the variance between midterm scores of 8 classes in freshman year 1st semester and graduation grades (R^2 =.296). Based on MAPE value it could be argued that the model predicts graduation grades with a 6.046% error on average.

Science Teaching Freshman Year 1st Semester Grade Model (1S1D4N) Model: Data of 406 students that had no missing information out of 407 students who graduated from Science Teaching were used to develop 1S1D4N-MLRA and 1S1D4N-ANN models. Results of Science Teaching 1S1D4N-MLRA were presented in Table 14.

Table 14.
Science Teaching 1S1D4N-MLRA

Variables	*B*	*Standard Error*	*ß*	*t*	*p*
Ataturk's Principles and History of Turkish Revolution I	.030	.010	.121	3.045	.002
Introduction to Education Science	.079	.014	.219	5.739	.000
General Physics I	.047	.010	.177	4.510	.000
General Physics Lab. I	.030	.014	.084	2.111	.035
General Chemistry I	.044	.012	.152	3.646	.000
General Chemistry Lab. I	.047	.011	.171	4.163	.000
General Mathematics I	.067	.011	.249	6.227	.000
Turkish I Written Narrative	.037	.012	.124	3.091	.002
RMSE=.205 MAE=.165 MSE=.042 R^2=.447 MAPE (%) = 5.401					

Classes with significant coefficients on Table 14 (p< .05) and regression constant were used to formulate graduation grade as below.

Graduation Grade = 2.090 + .030 * Ataturk's Principles and History of Turkish Revolution I + .079 * Introduction to Education Science + .047 * General Physics I + .030 * General Physics Laboratory I + .044 * General Chemistry I + .047 * General Chemistry Laboratory I + .067 * General Mathematics I + .037 * Turkish I Written Narrative

Science Teaching second 1S1D4N model was developed using ANN. Science Teaching 1S1D4N-ANN model had a structure that used semester grades of 8 classes input, two layers (Hidden 1, Hidden 2), 1 node in the 1st layer, 5 nodes in the 2nd layer, and used graduation grade as output. As a result of modeling, the best parameters in terms of performance indicators were determined as – activation function: Sigmoid, learning cycle: 2000, learning ratio .01, and moment .1. It was understood that the model with 1st layer 1 node and 2nd layer 5 nodes were more successful compared to other models. Performance indicators of the model that was developed were: RMSE=.201 AE=.161 MSE=.040 R^2=.476 MAPE (%)= 5.274. The model that was developed could explain 47.6% of the variance between semester grades of 8 classes in freshman year 1st semester and graduation grades (R^2 =.476). Based on MAPE value it could be argued that the model predicts graduation grades with a 5.274% error on average.

Science Teaching Freshman Year 1st and 2nd Semester Grade Point Averages (1S12SGPA) Model: Data of 735 students that had no missing information out of 940 students graduated from Science Teaching were used to develop 1S12SGPA-MLRA and 1S12SGPA-ANN models. Results of Preschool Teaching 1S12SGPA-MLRA were presented in Table 15.

Table 15.
Science Teaching 1S12SGPA-MLRA

Variables	*B*	*Standard Error*	*ß*	*t*	*p*
1st Semester	.174	.024	.358	7.294	.000
2st Semester	.227	.026	.443	8.822	.000
RMSE=.190 MAE=.152 SE=.036 R^2 =.533 MAPE= 4.935%					

As it can be viewed in Table 15, classes with significant coefficients on Table 6 (p< .05) and regression constant were used to formulate graduation grade as below. Graduation Grade= .174 * 1st Semester GGPA + .227 * 2nd Semester GGPA + 2.007 Science Teaching second 1S12SGPA model was developed using ANN. Science Teaching 1S12SGPA-ANN model had a structure that used grade point averages of the 1st and 2nd semesters of freshman year as 2 input, single layer (Hidden 1), single node, and used graduation grade as output. As a result of modeling, the best parameters in terms of performance indicators were determined as, activation function: Sigmoid, learning cycle: 2000, learning ratio .01, and moment .1. It was understood that the model with a single layer and single node was more successful compared to other models. Performance indicators of the model that was developed were: RMSE=.187 AE=.151 MSE=.035 R^2=.546 MAPE = 4.902 %. The model that was developed could explain 61% of the variance between grade point averages of 1st and 2nd semesters of freshman year and graduation grades (R^2 =.546). Based on MAPE value it could be argued that the model predicts graduation grades with a 4.902% error on average.

Findings on Second Sub-Problem of the Study

In the SGPA2 Model developed for the second sub-problem of the study, the prediction was made on students' semester grade point averages remaining lower than 2. In every prediction model, two different methods (Logistic Regression and Decision Trees) were used to make predictions.

Turkish Teaching SGPA2 Warning Model (Logistic Regression): Results of the logistic regression model developed to detect students whose semester grade point averages would remain under 2 for at least one semester using freshman year 1st-semester grade point averages of Turkish Teaching graduates were presented in Table 16.

Table 16.
Turkish Teaching SGPA2 Logistic Regression Classification Results

Predicted/Actual	Successful	Unsuccessful	Sensitivity-Specificity (%)
Successful	317	41	88.55
Unsuccessful	8	11	57.89
Accuracy (%)	97.54	21.15	

There were data of 377 students who graduated from Turkish Teaching used in the models. 41 of 377 students had semester grade point averages lower than 2 (unsuccessful) for at least one semester. 325 students had semester grade point averages lower than 2 in no semesters (successful). When Table 14 is examined it could be seen that 317 of the 325 students who were, in fact, successful (true positive: 97.54%) were predicted as successful. The number of students who were predicted as unsuccessful despite being successful (false positive) was 8. 11 of 52 students who were, in fact, unsuccessful were classified as unsuccessful (true negative: 31.18%), 41 were classified as successful (false positive). The total correct classification rate of the model was calculated as 87.00%.

Turkish Teaching SGPA2 Warning Model (Decision Trees): It was noted that 320 of the 325 students who were, in fact, successful (true positive: 98.46%) were predicted as successful. The number of students who were predicted as unsuccessful despite being successful (false negative) was 5. Three of the 52 unsuccessful students were classified as unsuccessful (true negative: 5.77%) and 49 were classified as successful (false positive). The total correct classification rate of the model was calculated as 85.68%.

Social Sciences Teaching SGPA2 Warning Model (Decision Trees): It was noted that 359 of the 370 successful students (true positive: 97.03%) were predicted as successful. The number of students who were predicted as unsuccessful despite being successful (false negative) was 11. 17 of the 93 unsuccessful students were classified as unsuccessful (true negative: 18.28%) and 76 were classified as successful (false positive). The total correct classification rate of the model was calculated as 81.21%.

Social Sciences Teaching SGPA2 Warning Model (Logistic Regression): Results of logistic regression model that was developed to predict students whose semester grade point averages would remain under 2 for at least one semester using freshman year 1st-semester grade point averages of Social Sciences Teaching graduates were presented on Table 17.

Table 17.
Social Sciences Teaching SGPA2 Logistic Regression Classification Results

Predicted/Actual	Successful	Unsuccessful	Sensitivity-Specificity (%)
Successful	352	64	84.62
Unsuccessful	18	29	61.70
Accuracy (%)	95.14	31.18	

There were data of 463 students who graduated from Preschool Teaching used in the models. 93 of 463 students had semester grade point averages lower than 2 (unsuccessful) for at least one semester. 370 students had semester grade point averages lower than 2 in no semesters (successful). When Table 17 is examined it could be seen that 352 of the 370 successful students (true positive: 95.14%) were predicted as successful. The number of students who were predicted as unsuccessful despite being successful (false positive) was 1. 29 of 93 unsuccessful students were classified as unsuccessful (true negative: 31.18%), 64 were classified as successful (false positive). The total correct classification rate of the model was calculated as 82.29%.

Mathematics Teaching SGPA2 Warning Model (Decision Trees): It was noted that 217 of the 228 successful students (true positive: 95.18%) were predicted as successful. The number of students who were predicted as unsuccessful despite being successful (false negative) was 11. 26 of the 94 unsuccessful students were classified as unsuccessful (true negative: 27.66%) and 68 were classified as successful (false positive). The total correct classification rate of the model was calculated as 75.47%.

Mathematics Teaching SGPA2 Warning Model (Logistic Regression): Results of logistic regression model that was developed to predict students whose semester grade point averages would remain under 2 for at least one semester using freshman year 1st-semester grade point averages of Mathematics Teaching graduates were presented on Table 18.

Table 18.
Mathematics Teaching SGPA2 Logistic Regression Classification Results

Predicted/Actual	Successful	Unsuccessful	Sensitivity-Specificity (%)
Successful	207	52	79.92
Unsuccessful	21	42	66.67
Accuracy (%)	90.79	44.68	

There were data of 322 students who graduated from Preschool Teaching used in the models. 94 of 322 students had semester grade point averages lower than 2 (unsuccessful) for at least one semester. 228 students had semester grade point averages lower than 2 in no semesters (successful). When Table 18 is examined it could be seen that 207 of the 228 successful students (true positive: 90.79%) were predicted as successful. The number of students who were predicted as unsuccessful despite being successful (false positive) was 21. 42 of 94 unsuccessful students were classified as unsuccessful (true negative: 44.68%), 52 were classified as successful (false positive). The total correct classification rate of the model was calculated as 77.33%.

Science Teaching SGPA2 Warning Model (Decision Trees): It was noted that 229 of the 252 successful students (true positive: 90.87%) were predicted as successful. The number of students who were predicted as unsuccessful despite being successful (false negative) was 23. 53 of the 152 unsuccessful students were classified as unsuccessful (true negative: 23.01%) and 99 were classified as successful (false positive). The total correct classification rate of the model was calculated as 72.60%.

Science Teaching SGPA2 Warning Model (Logistic Regression): Results of logistic regression model that was developed to predict students whose semester grade point averages would remain under 2 for at least one semester using freshman year 1st-semester grade point averages of Science Teaching graduates were presented in Table 19.

Table 19.
Science Teaching SGPA2 Logistic Regression Classification Results

Predicted/Actual	Successful	Unsuccessful	Sensitivity-Specificity (%)
Successful	225	71	76.01
Unsuccessful	27	42	60.87
Accuracy (%)	89.29	37.17	

There were data of 365 students who graduated from Science Teaching used in the models. 13 of 365 students had semester grade point averages lower than 2 (unsuccessful) for at least one semester. 252 students had semester grade point averages lower than 2 in no semesters (successful). When Table 20 is examined it could be seen that 225 of the 252 successful students (true positive: 89.29%) were predicted as successful. The number of students who were predicted as unsuccessful despite being successful (false positive) was 27. 42 of 113 unsuccessful students were classified as unsuccessful (true negative: 37.17%), 71 were classified as successful (false positive). The total correct classification rate of the model was calculated as 73.15%.

4. RESULTS, DISCUSSION, AND RECOMMENDATIONS

In this section, the findings were analyzed, and the results of the study were presented. The results of the study were compared to other studies and discussed.

Results Regarding First Sub-Problem of the Study

3 types of models, namely the 1S1DV Model, 1S12SGPA Model, and 1S1D4N Model were developed to predict graduation grades of Turkish Teaching, Social Sciences Teaching, Teaching Mathematics and Science in Primary Schools students.

Regarding prediction of graduation grade, it could be seen that models developed for Social Sciences Teaching were the models with the highest success rates. Similarly, models developed for Teaching Mathematics in Primary Schools were models with the lowest success rates. All models developed in the study could be used to predict graduation grades.

A study conducted by Bahadır (2013) using data of 139 undergraduate students used ANN and Logistic Regression analysis. It was noted that models developed using ANN gave better results in the prediction of academic successes of students at the graduate level. In this study, it was seen that MLRA and ANN models developed to predict graduation grades gave approximate results. It could be argued that models developed with MLRA and ANN could be used to predict the graduation grades of students.

When models were compared, it was noted that 1S1D4N models had better performances than 1S12SGPA models while 1S12SGPA models had better performances than 1S1DV. Considering graduation grade was calculated cumulatively, the fact that models developed using semester class grades and semester grade point averages had better performances than models developed using only mid-term scores could be understood. A study conducted by Şengür (2013) used class grades in the first 3 years of education of 127 students which similarly gave more successful results compared to the model created using class grades in freshman and sophomore years.

Results Regarding Second Sub-Problem of the Study

In the scope of the study, SGPA2 Warning Models were developed to predict student failure. Models developed in the study could be used in the prediction of students with a likelihood of failure.

In this study, models developed using Logistic Regression and Decision Trees gave approximate results in the prediction of student success in terms of total classifying success and that both models could be used in the prediction of student failure. Among developed models, the Turkish Teaching SGPA2 Model had the highest success rate (87%) of the general classification. Science Teaching SGPA2 Model had the lowest success rate (72.60%) of the general classification.

When discussed in terms of classes, it could be seen that Logistic Regression Analysis was more successful than Decision Trees in terms of classifying unsuccessful students. While the success rate of the Turkish Teaching SGPA2 Model in predicting unsuccessful students was 5.77% at Decision Trees, it was 21.15% at Logistic Regression. Similarly, rates were in favor of logistic regression models as 27.66%-44.68% for Mathematics Teaching in Primary Schools while they were 18.28% - 31.18% for Social Sciences Teaching and 31.86%-37.17% for Science Teaching. In a Study by Aydın (2007) success status of students was correctly predicted with a 74% rate. In Aydın's study (2007) Decision Trees gave more successful results compared to Logistic regression. Among late models, decision trees had 82.10% success while the logistic regression model had 78.53% success.

A study by Akçapınar (2014) using online data of 76 students, predicted students' passing status of classes with 86% success. Bahadır (2013) conducted a study where the author used bachelor's grades of 139 students to accurately classify graduate success statuses with 92% success. ANN model created in the same study using grades students took from bachelor's classes accurately predicted academic graduate exam scores with a 78.78% rate. A study conducted by Aksoy (2014) using data of 113 students used models that could identify gifted children with a 71% accuracy rate. Compared to such studies, it could be argued that the SGPA2 Warning models developed had good performances.

It could be argued that graduation grade prediction and failure prediction models developed in the study were successful. With the developed failure prediction models, it could be possible to conduct efforts to raise the success levels of detected students. Prediction of graduation grade could be used to determine students who would like to have a graduate education, those that receive scholarships, those that would like to receive scholarships, those faced with the risk of losing scholarship based on academic success. In case students have an early prediction of their graduation grades, opportunities, and risks associated with this graduation grade, this could have a positive academic effect on them. Because data used in the system belong to graduates, updating models (re-training) could be necessary. To do this, real graduation grades of students in their senior years and semester grade point averages could be used to assess the success of models. Conditions that change over time (newly added classes, changes in higher education policies) could require recreation and update of models. The general model suggested in the study could be used by Akdeniz University to predict the academic performances of students and to determine students that would fail. In the scope of the suggested model, the software could be developed that is integrated into SIS and that would report results to relevant units online to conduct academic and administrative efforts to improve the academic successes of students. Based on results of student success prediction models faculty administrations, university guidance services, and social units could cooperate to improve the success levels of students

This study described almost half of the process from acquisition of data to its preparation for modeling. When student data on SIS was examined, it was seen that there were features that were added to the database later and there were missing data on such areas. Areas, where there were many missing data, were not included in the dataset prepared for modeling. The fact that processes of choosing data on SIS, cleaning of data, it's structuring, and integration took a long time demonstrated that institutions did not have a database infrastructure that would enable DM efforts. This study demonstrated that there were missing data on tables where student records were kept and different data types (library usage, health services, social services, socio-economic information) were not integrated. This situation prevents the determination of the role such factors play in affecting student success. Institutions could ensure data integrity and integration to ensure studies based on data analysis are more efficient.

There are different dynamics affecting student success other than research models and datasets used in research. In the study factors that could affect student success directly such as academics that give classes, socio-economic statuses of students, changing class content based on years could not be analyzed. The study is limited to models developed according to data acquired from the dataset to predict academic performance.

The study revealed some findings especially with MLRA on the effect of classes at the Faculty of Education on the graduation grades of students. Detailed studies in this area could lead to an assessment of classes at the level of departments, class levels to update class catalogs. Authors could add socio-economic and demographic datasets to academic datasets of students in educational institutions to study the effect of these factors on the prediction of student success. A similar study could be conducted at schools under the Ministry of National Education to predict and prevent possible academic failure.

Research and Publication Ethics Statement

The ethic permission was received from Akdeniz University Social Sciences and Humanities Scientific Research and Publication Ethics Committee on 02.05.2018 and issue no 54.

Contribution Rates of Authors to the Article

The article is derived from the Murat Altun's doctoral dissertation supported by Akdeniz University SCR Unit. The authors in the article are also members of the research project.

Dr. Murat Altun: The article is based on the author's work within the scope of his doctoral dissertation.
Prof. Dr. Kemal Kayıkçı: Doctoral thesis advisor of the first author and research project director. He contributed to the acquisition of data, the evaluation of the analysis results, and their use in education.
Assoc. Prof. Dr. Sezgin Irmak: Research project member and thesis monitoring jury member. He contributed to the analysis of the database, the analysis, and the evaluation of the DM models.

Support Statement

This study was conducted as a part of the Ph.D. thesis study of the first author and supported by Akdeniz University SRP Unit with project no SDK-2018-4077.

Acknowledgment

We would like to thank to the Rectorate of Akdeniz University, the Computer Science Research, and Application Center, and the Scientific Research Projects Unit for their support for the study within the framework of ethical rules.

Statement of Interest

There is no conflict of interest between the authors.

5. REFERENCES

Akçapınar, G. (2014). *Çevrimiçi öğrenme ortamındaki eğitim verinine göre öğrenclerin akademik performanslarını veri madenciliği yaklaşımı ile modellenmesi* (Unpublished doctoral dissertation). Hacettepe Üniversitesi Eğitim Bilimleri Enstitüsü, Ankara. Retrieved from Netional Thesis Center of the Council of Higher Education. https://tez.yok.gov.tr/UlusalTezMerkezi/TezGoster?key=CwVIqqBuz1VkysVpueogASpo2VXm10JbNpABPw6T79VxuAwtThPlofETv7T4quMr

Akgöbek, Ö. & Çakır, F. (2009). Veri madenciliğinde uzman bir sistem tasarımı. *Akademik bilişim konferansları.* Şanlıurfa: İnternet Teknolojileri Derneği.

Akgöbek, Ö. & Kaya, S. (2011). Veri madencliği teknikleri ile veri kümlerinden bilgi keşfi: medikal veri madenciliği uygulaması. *e-Journal of New World Sciences Academy , 6* (1), 237-245.

Aksoy, E. (2014) *Matematik alanında üstün yetenekli ve zekalı öğrencilerin bazı değişkenler açısından veri madenciliği ile belirlenmesi* (Unpublished master's thesis, Thesis No. 368273). Dokuz Eylül Üniversitesi Eğitim Bilimleri Enstitüsü, İzmir. Retrieved from Netional Thesis Center of the Council of Higher Education.
https://tez.yok.gov.tr/UlusalTezMerkezi/TezGoster?key=gyLHMouPes-CvnhRcjQsKTy8nL6mSNtwsXf5MZA6M8LueCZzWsverimadenciliği4xNxTCDF7GSO

Alpaydın, E. (2000). Zeki veri madenciliği: ham bilgiden altın bilgiye ulaşma yöntemleri. *Bilişim 2000 eğitmi semineri.* 01 20, 2019. Retrieved from https://www.cmpe.boun.edu.tr/~ethem/files/papers/veri-maden_2k-notlar.doc
Alsuwaiket, M. (2018). *Measuring the academic performance of students in higher education using data mining techniques* (Unpublished doctoral dissertation). Retrieved from ProQuest Dissertations & Theses Global. (2204780401). Loughborough University, Loughborough.

Aydın, S. (2007). *Veri madenciliği ve Anadolu Üniversitesi uzaktan eğitim sisteminde bir uygulama* (Unpublished master's thesis, Thesis No: 220873). Anadolu Üniversitesi Sosyal Bilimler Enstitüsü, Eskişehir. Retrieved from Netional Thesis Center of the Council of Higher Education. https://tez.yok.gov.tr/UlusalTezMerkezi/TezGoster?key=-Z0vbSUgrhM9fXoGkRe6Qxsck9sPUR5ObbGWZ7WOup0UYUtvr3aoRm0czfWR5exU

Ayesha, S., Mustafa, T., Sattar, A. R., & Khan, M. I. (2010). Data mining model for the higher education system*. Europen Journal of Scientific Research*, 43 (1), 24-29.

Bahadır, E. (2013). *Yapay sinir ağları ve lojistik regresyon analizi yaklaşımları ile öğretmen adaylarının akademik başarılarının tahmini* (Unpublished master's thesis, Thesis No: 349939). Marmara Üniversitesi Eğitim Bilimleri Enstitüsü, İstanbul Retrieved from Netional Thesis Center of the Council of Higher Education https://tez.yok.gov.tr/UlusalTezMerkezi/TezGoster?key=iTkOhwevEenJZ3onUvs52rtmytHX_zk1oFoRubryMJD86xeUawYC0RBnNE_xezeP

Baradwaj, B. K. & Pal, S. (2011). Mining educational data to analyze students' performance. *International Journal of Advanced Computer Science and Applications*, 2 (6), 63-69.

Bienkowski, M., Feng, M., & Means, B. (2012). *Enhancing teaching and learning through educational data mining and learning analytics: An issue brief.* Washington, D.C

Büyüköztürk, Ş. (2018). *Sosyal bilimler için veri analizi el kitabı*. Ankara: Pegem Akademi Yayınları.

Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). From data mining to knowledge discovery in databases. *AI magazine*, 17(3), 37-37.

García, E., Romero, C., Ventura, S., & De Castro, C. (2011). A collaborative educational association rule mining tool*. The Internet and Higher Education*, 14(2), 77-88.

Han, J., Pei, J., & Kamber, M. (2011). *Data mining: concepts and techniques.* Waltham: Elsevier.

IDC. (2016). *IDC Global Research.* 12 25, 2016. Retrieved from http://www.idc.com

IEDMS. (2019). *International Educational Data Mining Society*. 07 20, 2019. Retrieved from http://www.educationaldatamining.org/

Kavrakoğlı. İ. (1998). *Toplam kalite yönetimi*, İstanbul: Kal-Der yayını.

Kayıkçı, K. (1999). Toplam kalite yönetiminde liderlik, *Kuram ve uygulamada eğitim yönetimi*, Pegem yayınları.

Tekin, A. (2014). Early prediction of students' grade point averages at graduation: A data mining approach. *Eurasian Journal of Educational Research* 54, 207-226.

Karpicke, H. & Murphy, M. E. (1996). Productive school culture: Principals working from the inside, *NASSP Bulletin*, 80: 26-34

Lopez, M. I., Luna, J. M., Romero, C., & Ventura, S. (2012). Classification via clustering for predicting final marks based on student participation in forums. *5th International Conference on Educational Data Mining*. Chania, Greece.

MoNE, (2018). *2023 Eğitim vizyonu.* 11 20, 2019. Retrieved from http://2023vizyonu.meb.gov.tr/doc/2023_EGITIM_VIZYONU.pdf

Öztürk, A. (2009). *Toplam Kalite Yönetimi ve Planlamas*ı. Bursa: Ekin

Pektaş, A. O. (2013). *SPSS ile veri madenciliği.* İstanbul: Dikeyeksen

Rapid Miner, (2019). *Rapid miner documentation*. 07 20, 2019 Retrieved from https://docs.rapidminer.com/

SAS. (2019). *SAS*. 07 20, 2019 Retrieved from e http://documentation.sas.com/

Şeker, Ş. E. (2008). *Bilgisayar kavramları*. 06 23, 2019. Retrieved from http://bilgisayarkavramlari.sadievrenseker.com/2008/10/02/yapay-sinir-aglari-artificial-neural-networks/

Şeker, Ş. E. (2016). *Bilgisayar kavramları*. 02 14, 2021. Retrieved from http://ybsansiklopedi.com/wp-content/uploads/2017/04/veri-madencili%C4%9Fi_araclari.pdf

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.

Şengür, D. (2013). *Öğrencilerin akademik başarılarının veri madenciliği metotları ile tahmini* (Unpublished master's thesis, Thesis No: 333873) Fırat Üniversitesi Eğitim Bilimleri Enstitüsü, Elazığ. Retrieved from Netional Thesis Center of the Council of

Higher Education. https://tez.yok.gov.tr/UlusalTezMerkezi/TezGoster?key=vVNzTGHHhjH-u3WMToxQ-hMSW7wzafvB8MKpauiON91WQBGBsu6OcVNWVoJDe6cF

Tsai, C. F., Tsai, C. T., Hung, C. S. & Hwang, P. S. (2011). Data mining techniques for identifying students at risk of failing a computer proficiency test required for graduation. *Australasian Journal of Educational Technology*, 27 (3), 481-498.

Witten, I. H., Frank, E. (2005). Data Mining: *Practical machine learning tools and techniques*. Morgan Kaufmann.

CoHE. (2020). National Thesis Center of the Council of Higher Education. 19, 05 2020. Retrieved from https://tez.yok.gov.tr/UlusalTezMerkezi/