

Comparison of Classification Performances of Mathematics Achievement at PISA 2012 with the Artificial Neural Network, Decision Trees and Discriminant Analysis

Emre Toprak ^{1,*}, Selahattin Gelbal ²

¹Erciyes University, Faculty of Education, Department of Educational Sciences, Kayseri, Turkey

²Hacettepe University, Faculty of Education, Department of Educational Sciences, Ankara, Turkey

ARTICLE HISTORY

Received: Aug. 10, 2020

Revised: Sep. 29, 2020

Accepted: Dec. 04, 2020

KEYWORDS

Artificial Neural Network,
Decision Trees,
Discriminant Analysis,
PISA,
Student Achievement

Abstract: This study aims to compare the performances of the artificial neural network, decision trees and discriminant analysis methods to classify student achievement. The study uses multilayer perceptron model to form the artificial neural network model, chi-square automatic interaction detection (CHAID) algorithm to apply the decision trees method and linear discriminant analysis. The performance of each method has been investigated in different sample sizes when classifying into different numbered subgroups. The study has revealed that the artificial neural network has the best performance in large, medium and small sample sizes when classifying into six, three and two subgroups. In the very small sample size, which has homogeneous variance-covariance matrices, the discriminant analysis performs the best, while in the very small sample size, which does not have homogeneous variance-covariance matrices, it is the discriminant analysis which performs the best when classifying into six subgroups and the artificial neural network performs the best when classifying into two and three subgroups. Considering the performances of the methods with respect to sample size, it can be concluded that as the sample size gets smaller, the performance of the decision trees method gets worse, whereas the performance of the discriminant analysis method improves. No correlation of this kind has been found with regard to the artificial network method.

1. INTRODUCTION

One of the aims of the studies in the field of educational sciences is to determine the current characteristics of students and draw a road map for their development based on these characteristics. To this end, curriculum specialists try to prepare qualified curricula, education managers try to maintain order and control in the implementation of these, and educational psychologists or guidance experts try to provide guidance and counseling services where students need them. This system, which targets the development of the students, works implicitly. In order for the system to operate efficiently and for the outputs obtained at the end of the process to be interpreted reliably, the most necessary component is measurement and evaluation. After the curricula are developed, implemented and provided with guidance, it is of

CONTACT: Emre Toprak ✉ emrettoprak@gmail.com 📍 Erciyes University, Faculty of Education, Department of Educational Sciences, Kayseri, Turkey

great importance to measure and evaluate the achievement of individuals accurately for the improvement the individual, society and the future of the country.

Following a valid and reliable assessment and evaluation of students, it is ensured that individuals who are qualified and who have the required characteristics are selected and placed in the areas required by the country with an appropriate classification. From this point of view, it can be said that the main purpose of measurement and evaluation studies conducted in education is to analyze whether students have the characteristics related to the subject being measured and to determine their achievement accurately. The answers given by students to the questions prepared for the purpose of measurement lead them to fall into one of the classification groups such as “passed/failed, adequate/inadequate, incomplete/complete, low/medium/high, bad/medium/good/very good”. These classifications, which form the basis of the positions which students will have in the future, are seen in all the examinations at national and international level. Today, the High-School Entrance Examination (LGS) classifies students by their performance in placing them in different types of schools in the transition from basic education to secondary education. Similarly, the Higher Education Foundations Examination (YKS) classifies students by their performance in placing them in different education programs in the transition from secondary education to higher education. Also, in examinations such as Language Proficiency Test administered in Turkey (YDS) at the national level and Test of English as a Foreign Language (TOEFL) at the international level conducted for the purpose of assessing language proficiency, student achievements are classified by the performance at certain levels.

There are different variables that affect classification in the studies conducted on the achievement classification of students. These variables, which affect the achievement of students, can be made up of very different data that can be further increased such as students' anxiety about the lessons, attitudes towards the school, feelings of belonging to the school, motivation towards the course, interest in the course, self-efficacy of the course, the influence of the teacher, classroom atmosphere, educational opportunities at home, socioeconomic values, income levels, ages, the number of siblings, time management skills, and academic self-confidence. These data about students are important predictors of their achievement classification (Arslantaş, Özkan, & Külekçi, 2012; İbrahim & Rusli, 2007; Keser & Sarıbay, 2007; Tosun, 2007).

In recent years, predictive studies conducted for the purpose of determining student achievement and the factors that affect this achievement have increased considerably (Altun & Yazıcı, 2013; Anıl, 2008, 2009, 2011; Arslantaş et al., 2012; Çiftçi & Çağlar, 2014; Doğan, 2009; Erdil, 2010; Gelbal, 2008; Özer & Anıl, 2011; Sadi, Uyar, & Yalçın, 2014; Şahin, 2011). Determining the effect size of the factors that are related to the shaping of achievement will also guide which points should be focused on the development and higher achievement. This requires careful selection of the methods used in studies conducted for this purpose and thus ensuring the most accurate prediction and classification. However, it is difficult to classify in groups whose characteristics are similar.

Different applications have been developed in the literature for the purpose of predicting student achievement and classifying accordingly. Each model used in classification applications has its own unique algorithm. Determining the performance of these algorithms in changing conditions will make studies more efficient and increase classification achievement in studies. Comparative evaluation of the algorithms is of great importance in terms of revealing which algorithm is successful in which situations and increasing classification performances (Kuyucu, 2012). When the literature is examined, researchers generally use regression-based methods such as different regression analysis and DA and structural equation modeling in predicting and classifying student achievement and the factors affecting this achievement (Altun & Yazıcı,

2013; Ercan, Işık, & Çakır 2005; Okioga, 2013; Özdemir & Koruklu, 2011; Yıldırım, 2000). Apart from these methods, there are artificial neural networks (ANNs) that form the experience of their information processes with the information they receive from the samples given. ANNs have the ability to produce solutions to many problems by making the same decisions in the face of similar issues (Haykin, 1994; Öztemel, 2012). Compared to traditional approaches, ANNs offer more exciting alternatives (Jain, Mao, & Mohiuddin 1996). Although numerous studies have been conducted with ANNs applications in different fields such as business, statistics, mathematics, biostatistics, economics, medicine, banking, engineering, tourism, agriculture and insurance (Bayru, 2007; Burmaoğlu, 2009; Çuhadar, 2006; Kayıkçı, 2014; Kibar, 2015; Kocadağlı, 2012; Köktürk, 2012; Sabancı, 2013; Şirvan, 2010; Tolon, 2007; Torun, 2007; Yüksek, 2007), studies conducted in the field of education are limited. ANNs applications do not require any statistical assumption and they also achieve successful results with incomplete data. They even produce successful results in situations when the data is defective or multidimensional, nonlinear or has a high probability of error (Çırak, 2012; Öztemel, 2012; Tepehan, 2011). These features make the application superior to regression-based methods and stand out as important advantages that can be preferred in research.

Another classification model, which is frequently used especially in science, has easy installation and interpretation, adapts easily to databases, does not require the assumptions of parametric regression techniques to be met, has high reliability and is therefore commonly preferred, is the decision trees (DTs) method (Chang & Wang, 2006; Pehlivan, 2006). In DTs, the aim is the creation of homogeneous sub-sets of data about the dependent variable as much as possible (Kuyucu, 2012). The DTs method is a classification technique “which does not require the meeting of the assumptions of parametric regression techniques and which can establish the relationships between the dependent variable(s) and independent variables in its space without interfering in the values in the data set” (Chang & Wang, 2006; Yamauchi et al., 2001 as cited in Kayri & Günüş, 2010, p. 2472). In addition, the DTs method’s visualization of the independent variables affecting the predicted variable in the form of trees by their levels of significance makes DTs more impressive than traditional regression methods (Hebert et al., 2006 as cited in Kayri & Boysan, 2008).

One of the traditional statistical methods used in classification and prediction studies is the discriminant analysis (DA) method. The DA, which serves the purpose of building models to predict membership to any group in studies, is a statistical method with strong foundations if their assumptions are met (Çokluk, Şekercioğlu, & Büyüköztürk, 2012; Johnson & Wichern, 1992; Kachigan, 1991). This decision or classification rule is commonly based on the maximum likelihood principle, where each observation is assigned to the group in which it has the greatest likelihood of occurrence (Huberty, 1994). The DA, which is one of the multivariate analysis methods, has been used in many studies in the literature (Atar, 2012; Avcılar & Yakut, 2015; Bektaş, 2012; Ceylan, 2009; Çakmak & Kara, 2011; Çankaya et al., 2009; Demircioğlu et al., 2004; Güzeller & Kelecioğlu, 2006; Oğuzlar, 2006; Öztürk, Coşkun, & Dirsehan, 2012; Serinkan & Bardakcı, 2007).

In the literature, there are different studies on binary comparison of the ANNs and DTs methods with the regression-based methods (Burmaoğlu, 2009; Çölkesen, 2009; Köktürk, 2012; Torun, 2007; Tosun, 2007). Studies conducted with ANNs applications can be exemplified as the comparison of the logistic regression analysis with the ANNs (Benli, 2005; Burmaoğlu, 2009; Güneri & Apaydın, 2004; Kurt & Türe, 2005; Naik & Ragothaman, 2004; Ocakoğlu, 2006), the comparison of the DA with the ANNs (Burmaoğlu, 2009) and the comparison of the multivariate regression analysis with the ANNs (Baş, 2006; Brown, 2007; Thigpen, 2000; Yüksek, 2007).

Studies conducted with the DTs method can be exemplified as the comparison of the logistic regression analysis with the DTs (Kuyucu, 2012; Zurada & Lonial, 2005) and the comparison of the ANNs applications with the DTs (Kuyucu, 2012; Tosun, 2007).

However, among the methods of data mining, these methods, which do not have the problem of assumption, which can be easily interpreted and which can easily integrate into database systems, are not used much in the field of educational sciences and a few studies conducted are in the form of binary method comparison. It is considered that the use of these methods in different applications in the field of educational sciences with traditional methods will be useful for researchers. Also, each of the classification methods developed can give different results under different conditions. From this point of view, the main aim of the study is the comparative examination of the classification of performances of the DA method, which is one of the traditional statistical methods, with the methods of ANNs and DTs. In addition to this main aim, when it is considered that both examinations which are conducted for the purpose of measuring cognitive characteristics and groups in which scales used for the purpose of measuring affective characteristics are administered have different sample sizes and that studies of classification are conducted in different subgroup numbers, it is believed that the examination of the performances all the three methods will produce in different sample sizes and in classifying different subgroup numbers will also be useful. Therefore, different sample sizes and different subgroup numbers have been added to the study and the main problem of the study have been determined as “Do the ANNs, the DTs and the DA performances differ in large (126,126), medium (6,186), small (603) and very small samples (102) in classifying student achievements into 6, 3 and 2 subgroups?”

It is thought that comparing these three methods with each other in the research will expand the results of existing studies and contribute to rich discussions on the subject. Also, results to be obtained from the study will form a significant model in that the three methods can be used in different applications in the field of educational sciences. The classifications to be made will be in different sample sizes and on 6, 3 and 2 subgroups separately. This is also significant in terms of revealing which method gives successful results as the number of groups and sample size of the study change.

In this section of the study, the methods used in the comparison are briefly mentioned.

1.1. The Artificial Neural Networks (ANNs)

The ANNs, which were developed based on the characteristics of the working system of the human brain, is a simulation of the biological nervous system. Considering the known structure of the biological nervous system, the ANNs, which consist of intense connections of simple computational elements in order to achieve high performance, can be defined as the generalization of a mathematical model of human perception and biological nerves (Akpınar, 2014, p. 239; Fausett, 1993 as cited in Yakut, 2012, p. 52). Considering that there are approximately ten billion nerve cells and sixty trillion connections in the human brain (Garson, 1998, p. 25), the brain can be regarded as a flawless computer running very fast (Munakata, 2008, p. 7). While Haykin (1999) describes the ANNs as “a parallel distributed processor consisting of simple units with a natural tendency to store information”, Zurada (1992) describes the ANNs as systems with physical cells that receive, store and use information” (Sağiroğlu, Beşdok, & Erler, 2003, p. 25).

The ANNs have many features such as “non-linearity, learning, parallelism, adaptability, generalization, working with missing data, tolerance of error, retention of information, pattern recognition” which are effective in starting to use them widely in different fields (Bayru, 2007; Burmaoğlu, 2009; Dikmen, 2001; Haykin, 1994; Kayıkçı, 2014; Köktürk, 2012; Öztemel, 2012; Seven, 1993; Simpson, 1990; Tosun, 2007; Yüksek, 2007).

The ANNs, formed by artificial neural cells coming together, is generally composed of three layers (Demiryürek, 2009; Elmas, 2003; Öztemel, 2012; Tolon, 2007, Tosun, 2007; Yurtoğlu, 2005):

- The input layer has at least one artificial nerve cell and transmits information from the outside world to the intermediate layer.
- The intermediate layers, process the information from the input layer and send it to the output layer. There can be more than one intermediate layer in a network. The hidden layer of the ANNs is called a black box due to the fact that what is happening in this layer can not be fully explained.
- The output layer has at least one artificial nerve cell and processes the information from the intermediate layer, creates the output that the network must produce for the input set presented from the input layer. The output produced is sent to the outside world.

Learning in ANNs is the development of problem-solving ability by assimilating past experiences (Tosun, 2007, p. 41). The basis of learning, defined as the process of improving behavior through the discovery of new knowledge over time, is the process of experience (Simon, 1983 as cited in Öztemel, 2012). This process is accomplished by learning rules that modify or adjust the connection weights of the network depending on the input examples and preferably the outputs of these inputs (Durmuş, 2008, p. 44). In the learning process, it is accepted that the relationship between the inputs and outputs of each example represents the general aspect of the event from different perspectives and thus it is thought that the event is learned from different perspectives with different examples. Only examples are shown to the computer in the learning process and no other prior information is given (Öztemel, 2012).

1.2. The Decision Trees (DTs)

The DTs, which are used to classify data with the values it has by different characteristics, is a method used to divide a data set into small groups over certain decision steps and make the elements that come together in groups more similar to each other after each division process (Berry & Linoff, 2004; Sun & Hui, 2008).

The DTs consist of roots, branches and leaves in the form of a natural tree. In the DTs with a structure similar to a flowchart, each of the attributes is represented by a node. The last structure in the tree is called “leaf”, the top structure is called “root” and the structures between them are called “branches” (Akpınar, 2014; Quinlan, 1993 as cited in Özkan, 2013, p. 53). In the DTs application, many questions are asked about the data and the results are trying to be reached based on the answers received. DTs are derived from top-down, general-specific data, starting from the root node (Oğuzlar, 2004). The basic logic in this process is based on the division of the relevant group into two more homogeneous subgroups at each stage. Decision rules are formed based on the answers received during the question and answer process and the classification process starting from the root node continues until nodes or leaves without branches are found, in other words, until a statistically significant difference is reached (Köktürk, 2012; Thomas, 2000).

There are “decision nodes” in the DTs indicating the test used for classification. The decision nodes perform the tests and branch successfully. Each branch of the tree is a candidate for completing the classification process and a decision node is formed when the classification process does not take place. If a class has been reached, at the end of that branch there is a leaf, one of the classes to be determined on the data. These processes start from the top root node and continue until the lowest leaves are reached (Özekes, 2003).

The DTs application operates in an order with a two-tiered structure, the first step is “learning” and the second step is “classifying”. A previously known training data in the learning step is analyzed by the classification algorithm in order to form a model. The model learned is shown

as classification rules or DTs. In the classification step, the test data is used to determine the accuracy of the classification rules or the DTs. If the accuracy obtained as a result of the analysis is at an acceptable rate, the rules are used to classify the new data. This acceptability ratio is obtained by comparing the known class and the estimated class (Argüden & Erşahin, 2008; Kıran, 2010; Köktürk, 2012; Özekes, 2003; Silahtaroglu, 2013).

When the DTs algorithms are used for classification, they are generally called classification trees and when they are used for regression purposes, they are called regression trees (Rokach & Maimon, 2008). In order for the DTs algorithms to be applied to a problem, events and objects must be expressed with certain property values and there must be distinctive properties that affect the determination of classes (Altıntaş, 2010).

1.3. The Discriminant Analysis (DA)

The concept of discriminant was first used by James Joseph Sylvester in 1851 (Akpınar, 2014, p. 189). As a statistical method, while the DA, which was first introduced by Fisher (1936), could initially ensure the division of only two groups, today with its increasing computational power, it ensures the division of data series into more categories (Albayrak, 2006; Akpınar, 2014, p. 189).

The DA, which is one of the multivariate analysis methods, is a method that ensures the division of the variables in the X data set into two or more real groups and that derives discrimination functions that enable the optimal assignment of the units to the real groups in the natural environment by considering the p feature to be examined (Özdamar, 2010). The DA (Çokluk et al., 2012), which is used in the models created to predict membership to any group in studies, is used when it is desired to subdivide the p units with known features into subcategories by their characteristics.

What is aimed at the DA is to determine one or more functions consisting of a linear combination of variables that maximize the differences between individuals in groups (Çakmak, 1992). By means of the DA, which has basically two purposes, first, functions that separate the groups from each other are found, and then, through the functions found, it is ensured that a newly observed unit is assigned to one of the groups in such a way that the classification error is minimum (Güzeller & Kelecioğlu, 2006). Due to these two functions of the DA, it has been considered appropriate by some authors to be named differently. For example, if the DA is applied to determine a discrimination function, it refers to as the Descriptive DA, and if it is applied for classification purposes, it refers to the Predictive DA (Özdamar, 2010). The descriptive DA identifies discrimination functions and through these functions, it enables us to determine the discriminating variables which reveal the difference between the groups the most. The predictive DA allows predicting which group a unit whose group is unknown will be included in (Grimm & Yarnold, 1995; Tatlıdil, 1996; Özdamar, 2002 as cited in Güzeller & Kelecioğlu, 2006). The DA, which is the process of revealing the differences between two or more groups by means of discrimination variables, is a broad concept that covers several closely related statistical approaches (Klecka, 1980). The DA has many assumptions such as normal distribution, sample size, variable selection, homogeneity of variance-covariance matrices, extreme values and multiple connections that affect the performance of the discrimination performance.

2. METHOD

This section includes the type of research, the study group, the data collection instruments, data collection process and data analysis.

2.1. The Type of the Research

This research, in which the data obtained from PISA 2012 application is used, is a field research

in terms of taking it from real life and daily life and not being artificial due to any intervention in the research (Kaptan, 1995); it is a cross-sectional study in terms of examining the event, which is the subject of the research for a specific time unit only (Büyüköztürk et al., 2008) and it is descriptive research because an existing event or situation is defined as it is (Karasar, 2014).

2.2. Study Group

The number of the subgroups to be classified and the sample sizes were determined in a way that they were interconnected by taking into account the main problem and sub-problems of the research. The achievement levels determined by PISA were utilized in the formation of classification groups and the samples were selected in a way that they fit the classification into 6, 3 and 2 subgroups;

- The classification for the 6 subgroups was made as Level 1 / Level 2 / Level 3 / Level 4 / Level 5 / Level 6. Each achievement level included in the PISA classification constituted a group.
- The classification for the 3 subgroups was made as Lower Level / Medium Level / Upper Level. The Sub-Level Group consists of the students in Level 1 and Level 2 in PISA. The Intermediate Group consists of the students in Level 3 and Level 4 in PISA. The Top-Level Group consists of the students in Level 5 and Level 6 of PISA.
- The classification for 2 subgroups was made as Lower Level / Upper Level. The Sub-Level Group consists of the students in Level 1, Level 2 and Level 3 in PISA. The Top-Level Group consists of the students in Level 4, Level 5 and Level 6 in PISA.

The population of the study consists of 15-year-old students studying in all the countries participating in PISA 2012 Mathematics practice. There is a sample of 485,490 students from all the countries participating in this practice. The distribution of the students in the sample by their achievement levels is given in [Table 1](#).

Table 1. *The distribution of all the students participating in PISA 2012 by their achievement levels*

<i>Level</i>	<i>f</i>	<i>%</i>
Less than Level 1	69,691	14.4
Level 1	91,369	18.8
Level 2	109,383	22.5
Level 3	99,016	20.4
Level 4	67,520	13.9
Level 5	34,652	7.1
Level 6	13,859	2.9
Total	485,490	100.0

However, by taking into account the independent/predictive variables used for classification purposes, the students with missing data were excluded from the study and the number of the students constituting the largest sample was determined as 126,126. In addition, in order to form 6, 3 and 2 subgroups, the students below the first level were not included in the research. The distribution of the students representing the large sample by their achievement level is given in [Table 2](#).

Table 2. *The distribution of 126,126 people by their achievement levels*

6 Subgroups	f	%
Level 1	26,292	20.8
Level 2	32,893	26.1
Level 3	30,424	24.1
Level 4	20,917	16.6
Level 5	11,105	8.8
Level 6	4,495	3.6
Total	126,126	100.0

As can be seen in [Table 2](#), the largest sample was selected as 126,126 people for comparison on different sample sizes. In addition, 6,186 medium-sized, 603 small-sized and very small samples of 102 for each where homogeneity was and was not ensured in variance-covariance matrices were prepared.

The systematic sampling method was utilized in the preparation of large, medium and small samples. All the students were ranked from the lowest achievement level to the highest achievement level before making the selection. Following the ranking, a study was conducted in a way that there were 1,000 students at each achievement level and a sample of 6,186 people was first formed. This sample is given in [Table 3](#).

Table 3. *The creation of 6,186 sample and distribution by achievement levels*

6 Subgroups	f	%	Systematic sampling
Level 1	1,012	16.4	Starting from the 1st person, one person in every 26th person
Level 2	1,028	16.6	Starting from the 1st person, one person in every 32th person
Level 3	1,015	16.4	Starting from the 1st person, one person in every 30th person
Level 4	997	16.1	Starting from the 1st person, one person in every 21th person
Level 5	1,010	16.3	Starting from the 1st person, one person in every 11th person
Level 6	1,124	18.2	Starting from the 1st person, one person in every 4th person
Total	6,186	100.0	

Following the ranking, a study was conducted in a way that there were 100 students at each achievement level and a small sample of 603 people was formed. This sample is given in [Table 4](#).

Table 4. *The creation of 603 sample and distribution by achievement levels*

6 Subgroups	f	%	Systematic sampling
Level 1	100	16.6	Starting from the 1st person, one person in every 262th person
Level 2	100	16.6	Starting from the 1st person, one person in every 328th person
Level 3	101	16.7	Starting from the 1st person, one person in every 304th person
Level 4	101	16.7	Starting from the 1st person, one person in every 209th person
Level 5	101	16.7	Starting from the 1st person, one person in every 111th person
Level 6	100	16.6	Starting from the 1st person, one person in every 45th person
Total	603	100.0	

Apart from these samples, in order to see the effect of homogeneity on the performances of the methods in variance-covariance matrices, two very small samples of 102 people where homogeneity was ensured and was not ensured in variance-covariance matrices were prepared. These samples are also given in [Table 5](#) and [Table 6](#).

Table 5. The formation of very small and homogeneous sample (102 persons) in variance-covariance matrices and distribution by achievement levels

6 Subgroups	f	%	Purposive sampling
Level 1	100	16.7	17 people
Level 2	100	16.7	17 people
Level 3	101	16.7	17 people
Level 4	101	16.7	17 people
Level 5	101	16.7	17 people
Level 6	100	16.7	17 people
Total	603	100.0	

Table 6. The formation of very small and homogeneous sample (102 persons) in variance-covariance matrices and distribution by achievement levels

6 Subgroups	f	%	Purposive sampling
Level 1	100	16.7	17 people
Level 2	100	16.7	17 people
Level 3	101	16.7	17 people
Level 4	101	16.7	17 people
Level 5	101	16.7	17 people
Level 6	100	16.7	17 people
Total	603	100.0	

The sizes and distribution of the samples formed by the classification groups after the selections are given in Table 7.

Table 7. All the sample sizes and distribution by classification groups

	6 Subgroups			3 Subgroups	2 Subgroups				
	f	%	f		%	f	%		
126126	Level 1	26,292	20.8	Low level	59,185	46.9	Low level	89,609	71.0
	Level 2	32,893	26.1						
	Level 3	30,424	24.1	Medium level	51,341	40.7	Top level	36,517	29.0
	Level 4	20,917	16.6						
	Level 5	11,105	8.8	Top level	15,600	12.4			
	Level 6	4,495	3.6						
6186	Level 1	1,012	16.4	Low level	2,040	33.0	Low level	3,055	49.4
	Level 2	1,028	16.6						
	Level 3	1,015	16.4	Medium level	2,012	32.5	Top level	3,131	50.6
	Level 4	997	16.1						
	Level 5	1,010	16.3	Top level	2,134	34.5			
	Level 6	1,124	18.2						
603	Level 1	100	16.6	Low level	200	33.2	Low level	301	49.9
	Level 2	100	16.6						
	Level 3	101	16.7	Medium level	202	33.5	Top level	302	50.1
	Level 4	101	16.7						
	Level 5	101	16.7	Top level	201	33.3			
	Level 6	100	16.6						
10	6 Subgroups	f	%	3 Subgroups	f	%	2 Subgroups	f	%

	Level 1	17	16.7	Low level	34	33.3	Low level	51	50.0
	Level 2	17	16.7						
	Level 3	17	16.7						
	Level 4	17	16.7	Medium level	34	33.3			
	Level 5	17	16.7						
	Level 6	17	16.7	Top level	34	33.3	Top level	51	50.0
102 (Unhomogenized)	6 Subgroups	f	%	3 Subgroups	f	%	2 Subgroups	f	%
	Level 1	17	16.7	Low level	34	33.3	Low level	51	50.0
	Level 2	17	16.7						
	Level 3	17	16.7						
	Level 4	17	16.7	Medium level	34	33.3			
	Level 5	17	16.7						
	Level 6	17	16.7	Top level	34	33.3	Top level	51	50.0

2.3. The processing and analysis of the data

In the study, the multilayer perceptron model was used in the creation of the ANNs model, the CHAID algorithm and the linear DA, which is one of the DA types, were used in the application of the DTs method. The dependent variable included in the models consists of 6, 3 and 2 subgroups made up of the students' achievement levels and the independent variables consist of 17 variables used in the classification.

In the selection of an activation function in the ANNs analyses, the hyperbolic tangent function was applied to the cells in the hidden layer and the Softmax function was applied to the cells in the output layer. 70% of the data set in the analyses was chosen as the sample used in education and 30% was selected as the sample used in the test. Since the ANNs determined a different 70% of the data set as the training set and a different 30% as the test set, 50 different experimental analyses were performed. As a result of the 50 different analyses, the performance of the method was revealed by selecting the highest classification percentage obtained by the method.

The SPSS program was used in all the analyses. Before proceeding with the analysis of the data used within the scope of the research, some arrangements were made for the sub-problems of the research. With the arrangements made, it was tried to meet these assumptions of the DA, which has different assumptions. The variables which do not have missing data within the data, which do not have a distribution closest to normal, which do not have extreme values, whose variance-covariance matrices are homogeneous and which do not have multiple connections were tried to be selected. The examinations are presented respectively.

- **Missing data**

Firstly, independent/predictive variables that are thought to be used in the study were examined and the data of the students who did not have missing data in these variables were included in the study. Since the comparison was made according to the sample sizes in the research, the number of data was requested to be large enough. At this point, 17 variables with the best data were determined. In the PISA sample of 485,490 students, the data of the students with missing data in the variables included in the study were cleared and a total of 126,126 students remained.

- **Sample size**

The sample size assumption was met in all the subgroups to be classified. Each subgroup included students at least as many as the number of the independent variables. The size of the subgroups in the samples created and their distribution by the classification groups are given in [Table 7](#).

• **Normal distribution**

When the kurtosis and skewness values of the variables in the study are examined, it can be said that the distributions are close to normal. However, the significance of skewness and kurtosis values is evaluated by dividing by their standard errors. When the values were divided by their standard errors, it was observed that the results obtained were significant and some of the variables did not have a normal distribution in different samples. Considering the state of the variables in all the sample sizes, it was observed that as the sample size decreased, the skewness and kurtosis of the distributions approached the normal distribution; however, the normal distribution assumption was not fully met.

• **Variable Selection**

The selection of excessive and unnecessary variables was avoided and 17 variables to be used in classification were determined. In the selection of the variables, the variables which do not have missing data, which are closest to the normal distribution, which do not have extreme values and which do not have multiple connections were selected. The dependent and independent variables used in the research are given in [Table 8](#).

Table 8. *The dependent and independent variables used in the research*

<i>The dependent variables</i>	<i>The independent variables</i>
	Math anxiety
	Attitudes towards school: Learning outcomes
	Attitude towards school: Learning activities
	Sense of belonging to school
	Math teacher's classroom management
	Cognitive activities in mathematics
6 subgroups	Class climate
3 subgroups	Education at home
2 subgroups	Math Motivation
	Mathematics interest
	Math behavior
	Mathematics self-efficacy
	Mathematical intention
	Math teacher support
	Openness to problem solving
	Self-perception of mathematics
	Teacher Student Relations

• **The homogeneity of variance-covariance matrices**

The homogeneity of variance-covariance matrices was tested for each sample size. Comments were made depending on whether the samples met this assumption.

• **Extreme values**

Whether there are extreme values in the data set was examined and all the values belonging to the variables were converted to standard values. When the sample size is 100 or less, if the z-score of any observation is not in the range of (-3, +3) and when the sample size is more than 100, if the z-score of any observation is not in the range of (-4, +4), the observation is the extreme value (Mertler & Vannatta, 2005 as cited in Çokluk et al., 2012). As a result of the examinations carried out, it was determined that there were no extreme values. This assumption was met.

• **Multicollinearity**

Multicollinearity is a problem arising in case that there are very high correlations between variables. The inclusion of variables with multicollinearity in analyses increases errors and weakens the analysis. A correlation of 0.90 and over between two variables indicates that there is multicollinearity (Çokluk et al., 2012). When correlations between variables were examined in different sample sizes, while it was determined that there was no multicollinearity in a large, medium and small sample and in a very small sample which was homogenized, multicollinearity between variables was found in a very small sample which was unhomogenized. An examination of correlation coefficients with two variables is not enough to determine the problem of multicollinearity because the problem is not just that the correlation between two variables is high, but that an independent variable has a high degree of correlation with all the other independent variables. For this reason, different examinations were made to determine the multicollinearity status, and tolerance, VIF and CI statistics (Tabachnick & Fidell, 2007).

The variables do not have multicollinearity in a large, medium and small sample and in a very small sample where homogeneity was ensured in variance-covariance matrices. In addition, while it was established that in a very small sample where homogeneity was not ensured in variance-covariance matrices, the tolerance values belonging to the variables “sense of belonging to school, mathematics motivation, mathematics interest, mathematics teacher’s support and teacher-student relationship” were smaller than 0.10 and that the VIF values were higher than 10, it was observed that the CI values were within normal limits.

As a result of the examinations conducted, the independent/predictive variables which are in [Table 8](#) were used. The reasons for the selection of these variables for classification and prediction in the research can be summarized as follows:

1. There is no missing data in variables.
2. They are the variables that have the closest level to the normal distribution even though they can not meet the assumption of normal distribution.
3. The variables do not have extreme values.
4. There is no problem with multiple connections between variables except for a very small sample where homogeneity was not ensured in variance-covariance matrices.
5. The comparisons have been made by creating separate samples in which the assumption of the homogeneity of variance-covariance matrices has been met and has not been met.
6. In addition, these variables are the ones which are frequently used in studies aimed at predicting student achievement in the literature (Altun & Yazıcı, 2013; Anıl, 2008, 2009, 2011; Arslantaş et al., 2012; Çiftçi & Çağlar, 2014; Erayman, 2004; Ercan et al., 2005; Erdil, 2010; Gelbal, 2008; Kaysılı, 2008; Keser & Sarıbay, 2007; Özabacı & Acat, 2005; Özer & Anıl, 2011; Sadi et al., 2014; Şahin, 2011).

3. RESULTS

The results of the methods’ classification performance of student achievements into 6, 3 and 2 subgroups in the large, medium, small and very small sample sizes are given in [Figure 1](#), [Figure 2](#) and [Figure 3](#).

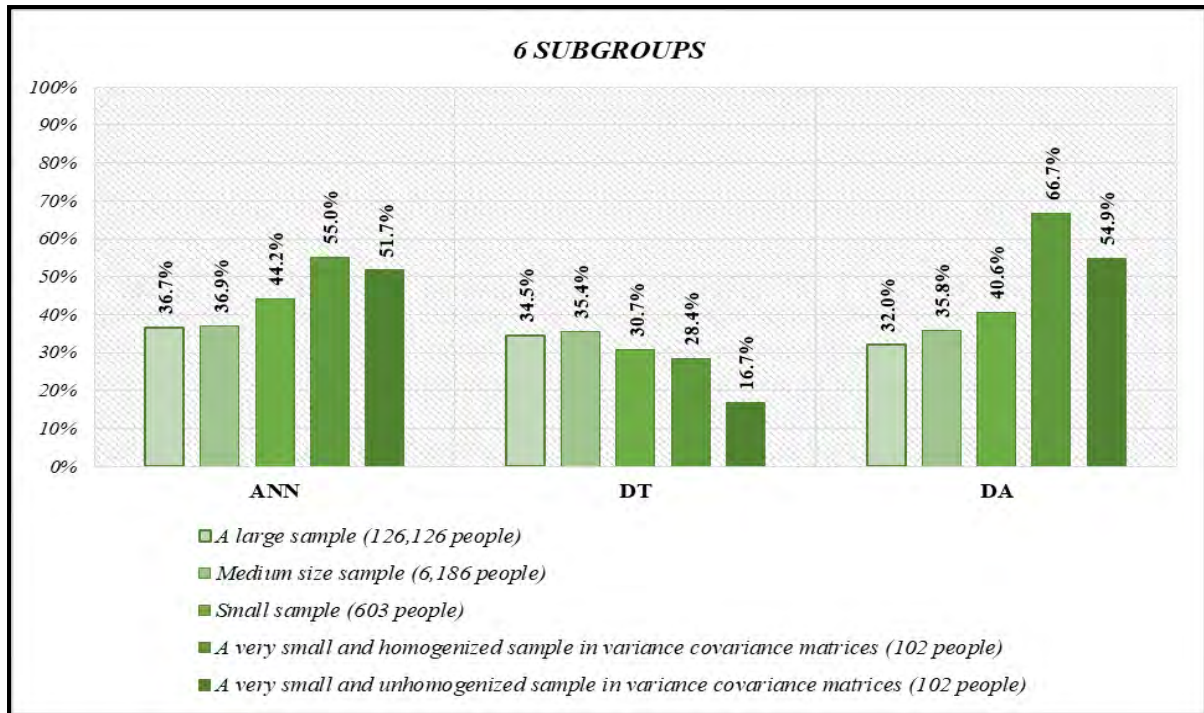


Figure 1. The performance of the methods to classify student achievement into 6 subgroups in different sample sizes

When the findings obtained are examined, the most successful method for making the classification into 6 subgroups is the ANNs in large, medium and small sample sizes. While the second-best method is the DTs in the large sample, the DA is the second-best method in the medium and small samples. While the DA is the most successful method for classifying 6 subgroups in the very small sample size where homogeneity was ensured and was not ensured in variance-covariance matrices; the second-best method is the ANNs.

As the sample size changes in classification into 6 subgroups, different findings regarding the performance of the ANNs have been obtained. While the performance averages of 50 different trial analyses conducted for the purpose of determining the best performance of the ANNs indicate that the best performance average is 39.9% and in 50 different trial analyses, the highest classification performance was achieved with 55.0% in the very small sample size where homogeneity was ensured in variance-covariance matrices. While this performance was followed by the small sample with an average performance of 38.4%; the second-highest performance was obtained in the very small sample size which homogeneity was not ensured in variance-covariance matrices. While the third-best average performance was obtained in the large sample, the third-highest performance was observed in the small sample. While the fourth-best average performance was obtained in the medium-sized sample, the fourth-highest performance was again observed in the medium-sized sample. It was observed that the lowest average performance was in the very small sample size where homogeneity was not ensured in variance-covariance matrices and that the lowest performance was in the large sample. These findings suggest that increasing the number of trials for the purpose of achieving the best performance in applications to be made with the ANNs in the classification of 6 subgroups will increase the opportunity to achieve the best performance. At the same time, ensuring the homogeneity of variance-covariance matrices increases the performance of the ANNs classification into 6 subgroups.

The DTs showed the highest performance in the classification of 6 subgroups in the medium-sized sample, while the performance of the DTs in the large sample was approximately 0.9%

lower than that in the medium-sized sample, this loss was further increased in small and very small samples. While the classification performance was 35.4% in the medium-sized sample where the method achieved the highest performance, the performance of the method decreased to 16.7%, especially in the very small sample size where homogeneity was not ensured in variance-covariance matrices. In addition, although the DTs method lost performance when the sample size decreased, if homogeneity was ensured in variance-covariance matrices, it showed higher performance in the very small sample than the very small sample where homogeneity was not ensured in variance-covariance matrices.

The DA is very sensitive to the homogeneity of variance-covariance matrices. The high performance of the DA in the very small sample and in the sample where homogeneity was ensured in variance-covariance matrices reveals this situation. The classification performance which increased to 32.0% in the large sample, 35.8% in the medium-sized sample and 40.6% in the small sample for classifying into 6 subgroups, approximately doubled and was 66.7% in the very small sample size where homogeneity was ensured in variance-covariance matrices. Also, the classification performance in the very small sample where homogeneity was not ensured in variance-covariance matrices is 54.9% for classifying into 6 subgroups. This supports the fact that the performance of the method increases as the sample size decreases.

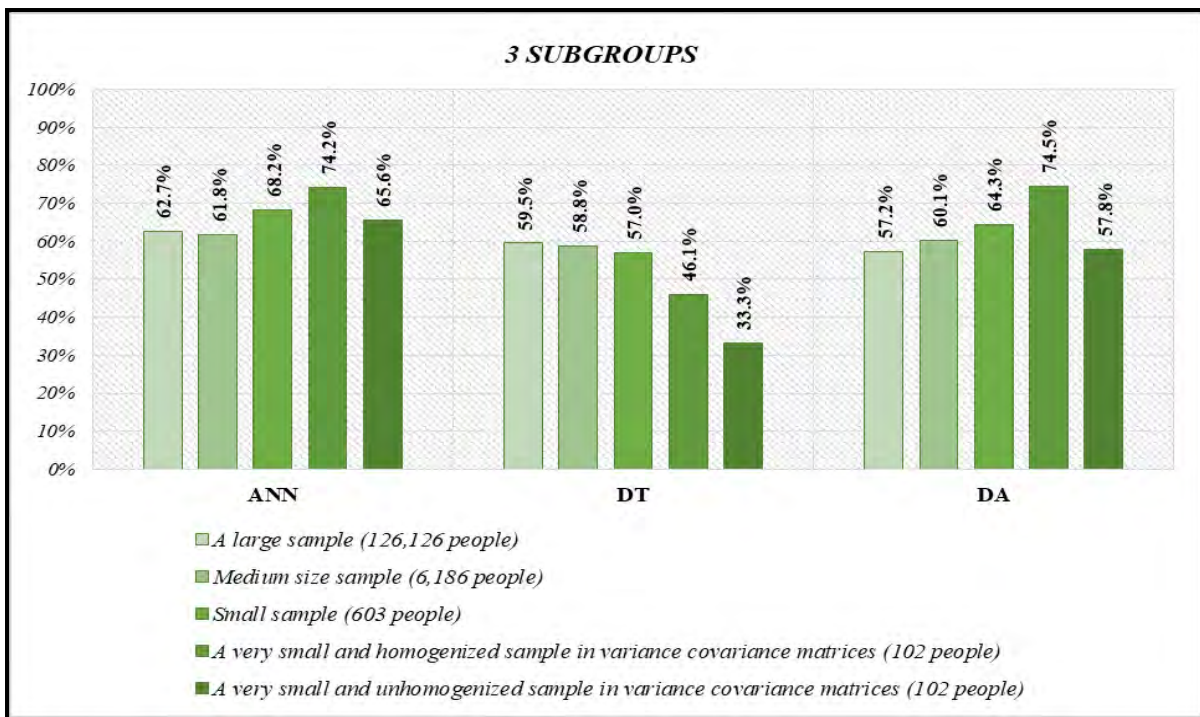


Figure 2. The performance of the methods to classify student achievement into 3 subgroups in different sample sizes

The ANNs are also the most successful method in the classification of 3 subgroups in large, medium and small sample sizes. While the second-best method is the DTs in the large sample, it is the DA method in medium and small samples. While the DA is the most successful method in classifying into 3 subgroups in the very small sample size where homogeneity was ensured in variance-covariance matrices, the second-best method is the ANNs. While the ANNs are the most successful method in classifying into 3 subgroups in a very small sample size where homogeneity was not ensured in variance-covariance matrices; the second-best method is the DA.

As the sample size changes in classifying into 3 subgroups, different findings regarding the performance of the ANNs have been obtained. In order to determine the best performance of the ANNs, the performance averages of 50 different trial analyses conducted in each sample size showed that the best performance average was in the small sample with 63.1%. In 50 different trial analyses, the highest classification performance was achieved in the very small sample size where homogeneity was ensured in variance-covariance matrices with 74.2%. While this performance was followed by a very small sample size where homogeneity was ensured in variance-covariance matrices where the average performance was 62.3%; the second-highest performance was obtained in the small sample. While the third-best average performance was obtained in the large sample, the third-highest performance was observed in the very small sample size where homogeneity was not ensured in variance-covariance matrices. While the fourth-best average performance was obtained in the medium sample, the fourth-highest performance was observed in the large sample. It was observed that the lowest average performance was in the very small sample size where homogeneity was not ensured in variance-covariance matrices and that the lowest performance was in the medium size sample. These findings suggest that increasing the number of trials in order to achieve the best performance in applications to be made with the ANNs in classifying into 3 subgroups will increase the opportunity to achieve the best performance. Also, ensuring the homogeneity of variance-covariance matrices seems to increase the performance of the ANNs.

The classification performance of the DTs method decreased in classifying into 3 subgroups as the sample size decreased. The DTs showed the highest performance in the large sample. While the performance of the DTs was lost approximately 0.7% in the medium-sized sample compared to the larger sample, this loss was further increased in small and very small samples. While the classification performance of the DTs method was 59.5% in the medium-sized sample where it achieved the highest performance, the performance of the method decreased to 33.3%, especially in the very small sample size where homogeneity was not ensured in variance-covariance matrices. In addition, although the DTs method lost performance when the sample size decreased, if homogeneity was ensured in variance-covariance matrices, it showed higher performance in the very small sample than the very small sample where homogeneity was not ensured in variance-covariance matrices.

The DA is very sensitive to the homogeneity of variance-covariance matrices. The highest performance shown by the DA in the very small sample and in the sample where homogeneity was ensured in variance-covariance matrices reveals this. The classification performance in classifying into 3 subgroups which increased to 57.2% in the large sample, 60.1% in the medium sample and 64.3% in the small sample, was 74.5% in the very small sample size where homogeneity was ensured in variance-covariance matrices. The fact that the classification performance is 57.8% in the very small sample size where homogeneity was not ensured in variance-covariance matrices suggests that the method's classification performance into 3 subgroups is low in cases where homogeneity was not ensured in variance-covariance matrices. When this assumption is not fulfilled, the sample where the performance is the highest is determined as a small sample.

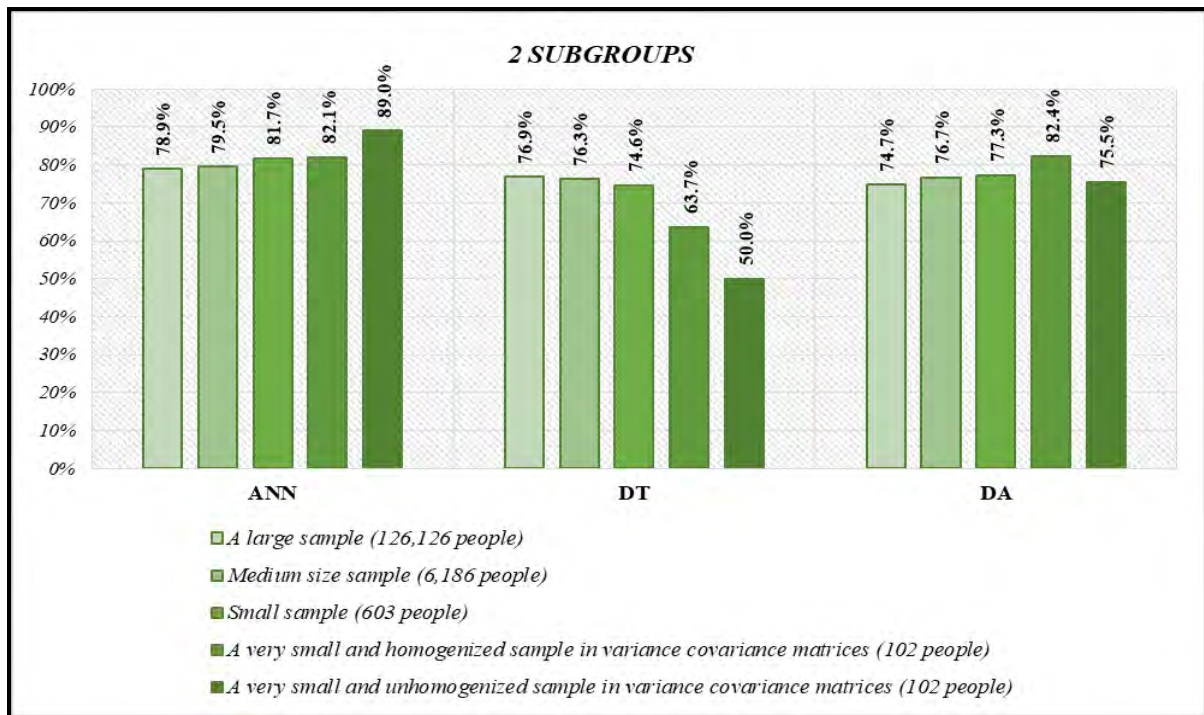


Figure 3. The performance of the methods to classify student achievement into 2 subgroups in different sample sizes

The ANNs are the most successful method in classifying into 2 subgroups in large, medium and small sample sizes. While the second-best method is the DTs in the large sample, the DA is the second-best method in the medium and small samples. While the most successful method for making 2 subgroupings in a very small sample size where homogeneous was ensured in variance-covariance matrices is the DA; the second-best method is the ANNs. While the ANNs are the most successful method for classifying 2 subgroups in a very small sample size where homogeneity was not ensured in variance-covariance matrixes; the second-best method is the DA.

As the sample size changes in classifying into 2 subgroups, different findings regarding the performance of the ANNs have been obtained. While the performance averages of 50 different trial analyses in each sample size for the purpose of determining the best performance of the ANNs, showed that the best performance average was in the large sample with 78.2%; in 50 different trial analyses, the highest classification performance was achieved with 89.0% in the very small sample size where homogeneity was not ensured in variance-covariance matrices. While this performance was followed by a small sample where the average performance was 78.0%, the second-highest performance was obtained in the very small sample size where homogeneity was ensured in variance-covariance matrices. While the third-best average performance was obtained in the medium sample, the third-highest performance was observed in the small sample. While the fourth-best average performance was obtained in the very small sample size where homogeneity was ensured in variance-covariance matrices, the fourth-highest performance was observed in the medium-sized sample. It was observed that the lowest average performance was in the very small sample size where homogeneity was not ensured in variance-covariance matrices and that the lowest performance was in the large sample. These findings suggest that increasing the number of trials in order to achieve the best performance in applications to be performed with the ANNs in classifying into 2 subgroups will also increase the opportunity to achieve the best performance. Also, ensuring the homogeneity of variance-covariance matrices increases the performance of the ANNs.

The classification performance of the DTs method decreased in classifying into 2 subgroups as the sample size decreased. The DTs showed the highest performance in the large sample. While the performance of DTs was lost approximately 0.6% in the medium-sized sample compared to the large sample, this loss was further increased in small and very small samples. While the classification performance of the DTs method was 76.9% in the medium-sized sample where it achieved the highest performance, especially in a very small sample size where homogeneity was not ensured in variance-covariance matrices, the performance of the method decreased to by 50.0%. In addition, although the DTs method lost performance when the sample size decreased, if homogeneity was ensured in variance-covariance matrices, the method showed higher performance in the very small sample than the very small sample size where homogeneity was not ensured in variance-covariance matrices.

The DA is very sensitive to the homogeneity of variance-covariance matrices. The high performance that the DA showed in a very small sample and in a sample where homogeneity was ensured in variance-covariance matrices reveals this. The classification performance which increased to 74.7% in the large sample, 76.7% in the medium sample and 77.3% in the small sample in classifying into 2 subgroups was 82.4% in the small sample where homogeneity was ensured in variance-covariance matrices. The fact that the classification performance into 2 subgroups is 75.5% in the very small sample where homogeneity was not ensured in variance-covariance matrices reveals that the performance of the method is low when homogeneity was not ensured in variance-covariance matrices. If this assumption is not fulfilled, the sample with the highest performance is determined as a small sample.

4. DISCUSSION and CONCLUSION

In this study, the classification performances of different classification methods were evaluated under varying conditions. Apart from comparing the performances of the methods used in the study in classifying different sample sizes and subgroup numbers with each other, results regarding the conditions under which each method performed the best were obtained. A result that can be said in general for each of the methods; when the number of subgroups classified was less, the methods showed higher performance. When the number of subgroups is high, classification becomes difficult as expected.

As a result of 50 different trial analyses performed to achieve the highest performance with ANNs in each sample size, the highest performance was reached with 89.0% in a very small sample where there was no homogeneity in variance-covariance matrices. However, when the average of 50 different trial analyses was examined, it was determined that the highest performance average was obtained with a large sample with 78.2%. Again, the second-highest performance with 82.1% was observed in a very small sample where homogeneity was achieved in variance-covariance matrices; the fact that the average of 50 different trial analyses in this sample was lower than the performance average obtained in the large sample showed a similar situation.

ANNs, use a different part of the research data as a training set and another part as a test set in each analysis. For this reason, the findings obtained in each analysis differ. These results show that increasing the number of trials analyzes increases the chance of achieving the highest performance. As a result, it is possible to say that ANNs show the highest performance in a very small sample where homogeneity is not provided in variance-covariance matrices. This situation also revealed that there is no need to meet the assumption of ensuring homogeneity in the variance-covariance matrices in the classification studies to be performed with ANNs.

The highest classification performance in terms of DTs method was obtained in classifying 2 subgroups in the large sample. However, it is possible to say that the performance of the method decreases linearly as the sample size decreases, except that the performance of classifying into

6 subgroups in the large sample is approximately 0.9% lower than the performance of classifying into 6 subgroups in the medium sample. In addition, the fact that the classification performance for all subgroups is higher in case of homogeneity in variance-covariance matrices in a very small sample, compared to a very small sample where homogeneity is not provided, shows that the performance of the method will increase if this assumption is met.

The DTs method does not give good results in estimating the values of continuous variables. In addition, it does not give good results in a/the model building when the number of classes is very high and learning set examples are low (Akpınar, 2014; Büyükkışıklar, 2014; Köktürk, 2012). The findings obtained in this study are similar to those obtained in the literature. At the same time, it was determined that ANNs showed higher classification performance in the studies conducted to compare the classification performances of ANNs and DTs, and it was observed that this result coincided with the research findings.

Tosun (2007) used the ANNs and the DTs in classifying student achievements into 2 subgroups in a small sample of 424 people. As a result of the analysis, he found that the performance of the ANNs was higher. İbrahim and Rusli (2007) compared the performance of the DTs, the regression analysis and the ANNs in classifying students' academic achievement in a small sample of 206 students. As a result of the analysis, they determined that the ANNs yielded more achievementful results than the other two methods. Çölkesen (2009) compared the achievement of the classification of satellite images into 6 subgroups in medium-sized samples of 6000 and 3750 people with the ANNs, the DTs and the k-star algorithms. As a result of the analysis, he determined that advanced classification techniques were a better and more effective alternative than conventional classifiers in the classification of remotely sensed images. Köktürk (2012) compared the achievements of the classification of the K-nearest neighborhood, the ANNs and the DTs in a small sample of 240 patients who applied to the gynecology and obstetrics clinic. She compared the data obtained from pregnant women in classifying into two subgroups. As a result of the analysis, he determined that the achievement of the classification of the ANNs technique was better than the other two methods. Kuyucu (2012) compared the classification performance of the logistic regression, the ANNs and the DTs in classifying into two subgroups in a small sample of 236 people. As a result of the analysis, he determined that the method which showed the highest performance was the ANNs.

The performance of DA was significantly affected by the sample size and homogeneity in variance-covariance matrices, which is one of the basic assumptions of the method. As the sample size decreased, the performance of the DA increased linearly. In addition, ensuring homogeneity in variance-covariance matrices made the performance of the method even stronger. DA achieved the highest performance in classifying 2 subgroups in the sample where homogeneity was achieved in variance-covariance matrices. On the other hand, even if the classification performance into 2 and 3 subgroups in very small samples was higher than in the large sample when homogeneity could not be achieved, it remained behind the performances obtained in medium and small samples. These results show that if the homogeneity is achieved in variance-covariance matrices, DA reaches a very high performance. DA showed higher performance than ANNs in classifying into 6, 3 and 2 subgroups in a very small sample where homogeneity was achieved in variance-covariance matrices. On the other hand, in the classification of large, medium and small samples into 6, 3 and 2 subgroups, ANNs have performed better than DA, and these results are similar to the literature.

Türe et al. (2005) compared the ANNs, logistic regression analysis and flexible DA methods in the prediction of primary hypertension in classifying into two subgroups in a small sample of 276. As a result of the analysis, they determined that the performance of the neural networks was higher than that of the DA. Burmaoğlu (2009) compared the achievement of the DA, the logistic regression analysis and the classification of the ANNs into two subgroups in a small

sample of 120 where homogeneity was not ensured in variance-covariance matrices using United Nations Development Program Human Development Index Data. As a result of the analysis, he determined that the multilayer perceptron model made better classification than the DA. Avcılar and Yakut (2015) compared the classification performance of the ANNs, the logistic regression and the DA methods in classifying into three subgroups in a small sample of 500 in determining voter preferences in local elections. As a result of the analysis, they found that the performance of the ANNs was higher than the DA. Wheeler (1993), compared the methods of ANNs and DA in classifying the achievement of the law school students' attorney exams into 2 subgroups in a small sample of 460 people. As a result of the analysis, he determined that the ANNs showed higher performance.

The results obtained from this study were similar to the results of the studies in the literature. In addition, the performance results of the methods used in the study in different sample sizes and in classifying into different subgroup numbers added new findings to the literature.

The suggestions developed based on the results obtained from the research can be listed as suggestions based on the results of this research and suggestions for the researchers in new studies.

As a result of the research, the highest performance of ANNs in the classification of student achievement into 6, 3 and 2 subgroups in large, medium-sized and small samples revealed that the method can be used reliably in these sample sizes and in these subgroup numbers. For this reason, it is recommended that ANNs should be preferred to classify 6, 3 and 2 subgroups in these sample sizes. The highest performance of ANNs in classifying into 2 and 3 subgroups in very small sample sizes where homogeneity is not provided in variance-covariance matrices is an important reason for the method to be preferred in classifying these subgroup numbers.

DTs lost their classification performance significantly as the sample size decreased. In addition to the decrease in the sample size, the performance loss of the method increased even more when the variance-covariance matrices were not homogenous. For this reason, reaching large samples in classification studies to be carried out with DTs is important and necessary for the high performance of the method.

The performance of the DA method increased as the sample size decreased. In addition, DA showed a much higher performance than ANNs and DTs in the condition that homogeneity is provided in variance-covariance matrices. According to the results obtained from the research, it can be said that it would be appropriate to prefer DA in case of homogeneity in variance-covariance matrices in classification studies to be made in very small samples.

The ANNs findings obtained in the research are limited to the multilayer perceptron model. It may be useful to examine different network models in other studies in order to obtain more information about the performance of ANNs in different sample sizes and classification into subgroup numbers.

In the analysis made with ANNs in the research, 70% of the data was separated as a training set and 30% as a test set. In new researches, the performance of the method can be handled under different conditions by changing the ratios of training and test sets.

DTs findings obtained in the research are limited by the CHAID algorithm. It may also be useful to examine different algorithms in other studies in order to obtain more information about DTs' performance in classifying different sample sizes and subgroup numbers.

DA findings obtained in the research are also limited to linear DA. It may be useful to use different types of DA such as quadratic DA and flexible DA in other studies in order to obtain more information about DA's performance in classifying different sample sizes and subgroup numbers.

In the study, variables with no missing data, the closest to normal distribution, no extreme values, no multicollinearity problems among variables as much as possible and the most frequently used variables in the literature were used in the classification of student achievement. New research with different variables other than these variables can contribute to the generalizability of the performance of the methods.

It is possible to work with data sets with missing data in researches in the field of education. Therefore, examining the performance of the methods with a different study in case of missing data in the data sets may contribute to the literature.

PISA 2012 mathematics test and survey results were used in the study. Working with data sets from different fields other than mathematics can gain new perspectives to researchers by revealing the performance of the methods in other fields.

In PISA applications, there are a total of 7 levels, which are lower than level one, level one, level two, level three, level four, level five and level six in order to determine student achievement. In this study, in order to create 6, 3 and 2 subgroups, students with a level lower than one were excluded from the study and the studies were carried out on 6, 3 and 2 subgroups. In different studies, 4, 5, 7 etc. comparisons can be made in samples with different subgroup numbers.

In this study, the performances of the methods in large, medium, small and very small sample sizes were discussed. New research plans can be prepared by creating different sample sizes to compare methods.

Acknowledgements

This paper was extracted from the first author's Ph.D. thesis.

Declaration of Conflicting Interests and Ethics

The authors declare no conflict of interest. This research study complies with research publishing ethics. The scientific and legal responsibility for manuscripts published in IJATE belongs to the author(s).

ORCID

Emre TOPRAK  <https://orcid.org/0000-0002-4131-4888>

Selahattin GELBAL  <https://orcid.org/0000-0001-5181-7262>

5. REFERENCES

- Akpınar, H. (2014). *DATA, veri madenciliği veri analizi [DATA, data mining data analysis]*. İstanbul: Papatya Publishing.
- Albayrak, A. S. (2006). *Uygulamalı çok değişkenli istatistik teknikleri [Applied multivariate statistical techniques]*. Ankara: Asil Publishing.
- Altıntaş, Y. (2010). *Veri madenciliğinin tıpta kullanımı ve bir uygulama: Hemodiyaliz hastaları için risk seviyelerine göre risk faktörlerinin etkileşimlerinin incelemesi [The usage of data mining in medicine and an application: Analysis of risk factors' interactions according to risk levels for hemodialysis patients]* (Unpublished Master Thesis). Gazi University, Institute of Science, Ankara.
- Altun, F., & Yazıcı, H. (2013). Ergenlerin benlik algılarının yordayıcıları olarak: Akademik öz-yeterlik inancı ve akademik başarı [As predictors of self-perception of adolescents: Academic self-efficacy belief and academic success]. *Kastamonu Education Journal*, 21(1), 145-156.
- Anıl, D. (2008). The analysis of factors affecting the mathematical achievement of Turkish students in the PISA 2006 evaluation program with structural equation modeling. *American-Eurasian Journal of Scientific Research*, 3(2), 222-227.

- Anıl, D. (2009). Uluslararası öğrenci başarılarını değerlendirme programı (PISA)'nda Türkiye'deki öğrencilerin fen bilimleri başarılarını etkileyen faktörler [Factors effecting science achievement of science students in programme for international students' achievement (PISA) in Turkey]. *Education ve Science*, 34(152), 87-100.
- Anıl, D. (2011). Türkiye'nin PISA 2006 fen bilimleri başarısını etkileyen faktörlerin yapısal eşitlik modeli ile incelenmesi [Investigation of factors affecting PISA 2006 science success of Turkey with structural equation model]. *Educational Sciences: Theory & Practice*, 11(3), 1253-1266.
- Argüden, Y., & Erşahin, B. (2008). *Veri madenciliği: Veriden bilgiye, masraftan değere [Data mining: From data to information, from cost to value]*. İstanbul: ARGE Consultancy.
- Arslandaş, İ. H., Özkan, M., & Külekçi, E. (2012). Eğitim fakültesi öğrencilerinin akademik başarı düzeylerinin bazı demografik değişkenler açısından incelenmesi [The analysis of academic achievement for some of demographic variables in education faculty of students]. *Electronic Journal of Social Sciences*, 11(39), 395-407.
- Atar, H. Y. (2012). Resim-iş öğretmenliği özel yetenek sınavlarının sınıflama doğruluğu üzerine bir çalışma [A study on the classification accuracy of art teaching special aptitude exams]. *Education ve Science*, 37(163), 283-296.
- Avcılar, M. Y., & Yakut, E. (2015). Yapay sinir ağları çoklu lojistik regresyon ve çoklu diskriminant analiz yöntemlerinden yararlanarak yerel seçimlerde seçmen tercihlerinin belirlenmesi: Osmaniye ili uygulaması [Determination of the voter preferences by using ANNs, multiple logistic regression and multiple discriminant analysis techniques: An investigation local elections in Osmaniye province]. *International Journal of Alanya Faculty of Business Administration*, 7(2), 207-224.
- Baş, N. (2006). *Yapay sinir ağları yaklaşımı ve bir uygulama [Artificial neural networks approach and an application]* (Unpublished Master Thesis). Mimar Sinan Fine Arts University, Institute of Science, İstanbul.
- Bayru, P. (2007). *Elektronik basında tüketici tercihleri analizi: yapay sinir ağları ile lojit modelin performans değerlendirilmesi [Electronic media consumer choice analysis: Artificial neural networks to evaluate the performance of the model with lojit]* (Unpublished Doctoral Dissertation). İstanbul University, Institute of Social Sciences, İstanbul.
- Bektaş, S. (2012). Çok şeritli bölünmüş karayollarında kaza tahmin modeli [A crash prediction model for multilane divided highway]. *Journal of Advanced Technology Sciences*, 1(1), 27-34.
- Benli, Y. K. (2005). Bankalarda mali başarısızlığın öngörülmesi lojistik regreyon ve yapay sinir ağı karşılaştırması [Prediction of financial failure in banks, comparison of logistic regression and artificial neural network]. *The Journal of the Industrial Arts Education Faculty of Gazi University*, 16, 31-46.
- Berry, M. J., & Linoff, G. S. (2004). *Data mining techniques: For marketing, sales, and customer relationship management* (2nd ed). USA: Wiley.
- Brown, J. D. (2007). *Neural network prediction of math and reading proficiency as reported in the educational longitudinal study 2002 based on non-curricular variables* (Unpublished Doctoral Dissertation). Duquesne University, Pennsylvania, ABD.
- Burmaoğlu, S. (2009). *Birleşmiş milletler kalkınma programı beşeri kalkınma endeksi verilerini kullanarak diskriminant analizi, lojistik regresyon analizi ve yapay sinir ağlarının sınıflandırma başarılarının değerlendirilmesi [Evaluating classification success of discriminant analysis, logistic regression analysis and neural network models using UN Developing Programme's Human Development Index]* (Unpublished Doctoral Dissertation). Atatürk University, Institute of Social Sciences, Erzurum.

- Büyükışıklar, A. (2014). *Karar ağaçları sınıflandırma algoritması ile toprak özgül direnci tespitinde jeolojik veri kullanımı [Use of geological data to determine soil resistance with decision tree classification algorithm]* (Unpublished Master Thesis). Bilecik Şeyh Edebali University, Institute of Science, Bilecik.
- Büyüköztürk, Ş., Çakmak, E. K., Akgün, Ö. E., Karadeniz, Ş., & Demirel, F. (2008). *Bilimsel araştırma yöntemleri [Scientific research methods]*. Ankara: Pegem Academy.
- Ceylan, E. (2009). PISA 2006 sonuçlarına göre Türkiye’de fen okuryazarlığında düşük ve yüksek performans gösteren okullar arasındaki farklar [Differences between low-and high-performing schools in scientific literacy based on PISA 2006 results in Turkey]. *Van Yüzyüncü Yil University Journal of Education*, 6(2), 55-75.
- Chang, L. Y., & Wang, H. W. (2006). Analysis of traffic injury: an application of non-parametric classification tree techniques. *Accident Analysis Prevention*, 38, 1019-1027.
- Çakmak, Z. (1992). *Çoklu ayırma ve sınıflandırma analizi: Eğitimde öğrencilerin meslek seçimine uygulanması [Multiple discriminant and classification analysis: Application to students' choice of profession in education]*. No: 658, Eskişehir: Anadolu University Publications.
- Çakmak, Z., & Kara, H. (2011). Yöneticilerde benlik algılamalarının belirlenmesi: Sanayi örgütlerinde bir araştırma [Determination of self-perception of managers: A research in industrial organizations]. *Dumlupınar University Journal of Social Sciences*, 30, 301-310.
- Çankaya, A. B., Taşdemir, G., Taşdemir, S., & Zilelioğlu, O. (2009). Delici göz yaralanması olgularımızın uzun dönem sonuçları ve görsel prognozu etkileyen faktörlerin analizi [Long term results of our penetrating eye injury cases and factors influencing final visual outcome]. *Turkish Journal of Ophthalmology*, 39, 220-226.
- Çırak, G. (2012). *Yükseköğretimde öğrenci başarılarının sınıflandırılmasında yapay sinir ağları ve lojistik regresyon yöntemlerinin kullanılması [The usage of artificial neural network and logistic regression methods in the classification of student achievement at higher education]* (Unpublished Master Thesis). Ankara University, Institute of Educational Sciences, Ankara.
- Çiftçi, C., & Çağlar, A. (2014). Ailelerin sosyo-ekonomik özelliklerinin öğrenci başarısı üzerindeki etkisi: Fakirlik kader midir? [The effect of socio-economic characteristics of parents on student achievement: Is poverty destiny?]. *International Journal of Human Sciences*, 11(2), 155-175.
- Çokluk, Ö., Şekercioğlu, G., & Büyüköztürk, Ş. (2012). *Sosyal bilimler için çok değişkenli istatistik SPSS ve LISREL uygulamaları [Multivariate statistics for social sciences SPSS and LISREL applications]* (2nd ed). Ankara: Pegem Academy.
- Çölkesen, İ. (2009). *Uzaktan algılamada ileri sınıflandırma tekniklerinin karşılaştırılması ve analizi [Comparing and analyzing of advanced classifier techniques in remote sensing]* (Unpublished Master Thesis). Gebze Technical University, Institute of Engineering and Science, Kocaeli.
- Çuhadar, M. (2006). *Turizm sektöründe talep tahmini için yapay sinir ağları kullanımı ve diğer yöntemlerle karşılaştırmalı analizi (Antalya ilinin dış turizm talebinde uygulama) [Forecasting tourism demand by artificial neural networks and time series methods (A comparative analysis in inbound tourism demand to Antalya)]* (Unpublished Doctoral Dissertation). Süleyman Demirel University, Institute of Social Sciences, Isparta.
- Demircioğlu, N., Ayan, S., Avanoğlu, B., & Sıvacıoğlu, A. (2004). Kastamonu-Taşköprü orman fidanlığında üretilen 2+0 yaşlı sarıçam (Pinus sylvestris L.) fidanlarının TSE normlarına göre değerlendirilmesi [Evaluation of 2+0 aged nursery of the scotch pine (Pinus sylvestrisl.) raised in Kastamonu-Taşköprü forest nursery as to TSE quality classification]. *Pamukkale University Journal of Engineering Sciences*, 10(2), 243-251.

- Demiryürek, O. (2009). *Polyester/viskon karışumlu open-end rotor iplik özelliklerinin yapay sinir ağları ve istatistiksel modeller kurularak tahmin edilmesi [Predicting the properties of polyester/viscose blended open-end rotor spun yarns by establishing artificial neural networks and statistical models]* (Unpublished Doctoral Dissertation). Çukurova University, Institute of Science, Adana.
- Dikmen, İ. (2001). *Strategic decision making in construction companies: An artificial neural network based decision support system for international market selection* (Unpublished Doctoral Dissertation). Middle East Technical University, Institute of Science, Ankara.
- Doğan, N. (2009). Bilgisayar destekli istatistik öğretiminin başarıya ve istatistiğe karşı tutuma etkisi [The effect of computer-assisted statistics instruction on achievement and attitudes toward statistics]. *Education and Science*, 34(154), 3-16.
- Durmuş, G. (2008). *Çimentolu harç özelliklerine yüksek sıcaklık etkisinin belirlenmesi ve yapay sinir ağı ile modellenmesi [Examining of the effect of high temperature on cement mortar properties and modelling by artificial neural network]* (Unpublished Doctoral Dissertation). Gazi University, Institute of Science, Ankara.
- Elmas, Ç. (2003). *Yapay sinir ağları [Artificial neural networks]*. Ankara: Seçkin Bookstore.
- Erayman, Y. (2004). *KSÜ öğrencilerinin sosyo-ekonomik yapılarının başarıları üzerine etkisi [The effect of socio-economical structure of the students in KSU on their success]* (Unpublished Master Thesis). Kahramanmaraş Sütçü İmam University, Institute of Science, Kahramanmaraş.
- Ercan, S., Işık, O., & Çakır, V. (2005). *HHO öğrencilerinin akademik başarılarına etki eden faktörlerin çoklu regresyon yöntemiyle incelenmesi [Investigation of factors affecting academic success of HHO students with multiple regression method]* V. National Production Research Symposium (25-27 November 2005), İstanbul Commerce University, İstanbul.
- Erdil, Z. (2010). Sosyoekonomik olarak risk altında bulunan çocuklara yönelik erken müdahale programları ve akademik başarı ilişkisi [Relationship of academic achievement and early intervention programs for children who are at socio-economical risk]. *Hacettepe University Faculty of Health Sciences Nursing Journal*, 17(1), 72-78.
- Fisher, R. A. (1936). The use of multiple measurements in taxonomic problems. *Annals of Human Genetics*, 7(2), 179-188.
- Garson, D. G. (1998). *Neural networks, an introductory guide for social scientists*. London: Sage Publications, 25.
- Gelbal, S. (2008). Sekizinci sınıf öğrencilerinin sosyoekonomik özelliklerinin Türkçe başarıları üzerindeki etkisi [The effect of socio-economic status of eighth grade students on their achievement in Turkish]. *Education and Science*, 33(150), 1-13.
- Grimm, L. G., & Yarnold, P. R. (1995). *Reading and understanding multivariate statistics*. Washington D. C.: American Psychological Association.
- Güneri, N., & Apaydın, A. (2004). Öğrenci başarılarının sınıflandırılmasında lojistik regresyon analizi ve sinir ağları yaklaşımı [Logistic regression analysis and neural networks approach in the classification of students' achievement]. *Gazi University Journal of Commerce and Tourism Education Faculty*, 1, 170-188.
- Güzeller, C., & Kelecioğlu, H. (2006). Ortaöğretim kurumları öğrenci seçme sınavının sınıflama geçerliği üzerine bir çalışma [The study on classification validity of secondary education student selection & placement exam]. *H. U. Journal of Education*, 30, 140-148.
- Haykin, S. (1994). *Neural networks: a comprehensive foundation*. New York: Mcmillan Press.
- Huberty, C. J. (1994). *Applied discrimination analysis*. New York: John Wiley and Sons.
- İbrahim, Z., & Rusli, D. (2007). *Predicting students' academic performance: Comparing neural network, decision tree and linear regression*. 21. Annual SAS Malasia Forum (5 September 2007), Kuala Lumpur, Malezya.

- Jain, A. K., Mao, J., & Mohiuddin, K. (1996). Artificial neural networks: A tutorial. *IEEE Computer Society*, 29(3), 31-44.
- Johnson, R. A., & Wichern, D. W. (1992). *Applied multivariate statistical analysis* (3rd ed.). New Jersey: Prentice-Hall, Englewood Cliffs.
- Kachigan, S. K. (1991). *Multivariate statistical analysis: A conceptual introduction* (2nd ed.). New York: Radius Press.
- Kaptan, S. (1995). *Bilimsel araştırma ve istatistik teknikleri [Scientific research and statistical techniques]*. Ankara: Tekışık Web Offset.
- Karasar, N. (2014). *Bilimsel araştırma yöntemi [Scientific research method]* (27th ed). Ankara: Nobel Academic Publishing.
- Kayıkçı, Ş. (2014). *Web sayfalarının yapay sinir ağları ile sınıflandırılması [Classification of web pages using neural networks]* (Unpublished Doctoral Dissertation). Marmara University, Institute of Social Sciences, İstanbul.
- Kayri, M., & Boysan, M. (2008). Bilişsel yatkınlık ile depresyon düzeyleri ilişkisinin sınıflandırma ve regresyon ağacı analizi ile incelenmesi [Assesment of relation between cognitive vulnerability and depression's level by using classification and regression tree analysis]. *H. U. Journal of Education*, 34, 168-177.
- Kayri, M., & Günüş, S. (2010). Türkiye'deki ortaöğretim öğrencilerinin internet bağımlılık düzeyini etkileyen bazı faktörlerin karar ağaçları yöntemleri ile incelenmesi [An analysis of some variables affecting the internet dependency level of Turkish adolescents by using decision tree methods]. *Educational Sciences: Theory & Practice*, 10(4), 2465-2500.
- Kaysılı, B. (2008). Akademik başarının artırılmasında aile katılımı [Parent involvement to improve academic achievement]. *Ankara University Faculty of Educational Sciences Journal of Special Education*, 9(1), 69-83.
- Keser, İ., & Sarıbay, E. (2007). İzmir'deki özel ve devlet üniversitelerindeki öğrencilerin başarılarını etkileyen faktörlerin belirlenmesi ve karşılaştırılması [Determining and comparing the factors effecting the performance of the students at the state and private universities in İzmir]. *Muğla University Journal of Institute of Social Sciences*, 18. 39-48.
- Kıran, Z. (2010). *Lojistik regresyon ve C&RT analizi yöntemleriyle sosyal güvenlik kurumu ilaç provizyon sistemi üzerinde bir uygulama [An application on Pharmacy Provision System data of Social Security Institution by logistic regression and CART analysis technics]* (Unpublished Master Thesis). Gazi University, Institute of Science, Ankara.
- Kibar, F. (2015). *Türkiye'de kamyon kazaları ile trafik ve karayolu geometrik özellikleri arasındaki ilişkinin istatistiksel ve yapay sinir ağları yöntemleri ile modellenmesi [Modeling the relationship between truck accidents and traffic and highway geometric characteristics in turkey with statistical and artificial neural networks methods]* (Unpublished Doctoral Dissertation). Karadeniz Technical University, Institute of Science, Trabzon.
- Klecka, W. (1980). *Discriminant analysis*. London: Sage Publications.
- Kocadağlı, O. (2012). *Genetik algoritmalar ve bulanık üyelik fonksiyonlarıyla hibrit bayes yapay sinir ağları [Hybrid bayesian neural networks with genetic algorithms and fuzzy membership functions]* (Unpublished Doctoral Dissertation). Mimar Sinan Fine Arts University, Institute of Science, İstanbul.
- Köktürk, F. (2012). *K-en yakın komşuluk, yapay sinir ağları ve karar ağaçları yöntemlerinin sınıflandırma başarılarının karşılaştırılması [Comparing classification success of k-nearest neighbor, artificial neural network and decision trees]* (Unpublished Doctoral Dissertation). Bülent Ecevit University, Institute of Health Sciences, Zonguldak.
- Kurt, İ., & Türe, M. (2005). Tıp öğrencilerinde alkol kullanımını etkileyen faktörlerin belirlenmesinde yapay sinir ağları ile lojistik regresyon analizinin karşılaştırılması

- [Comparison of artificial neural networks and logistic regression analysis in determining factors affecting alcohol consumption among medicine students]. *Medical Journal of Trakya University*, 22(3), 142-153.
- Kuyucu, Y. E. (2012). *Lojistik regresyon analizi (LRA), yapay sinir ağıları (YSA) ve sınıflandırma ve regresyon ağaçları (C&RT) yöntemlerinin karşılaştırılması ve tıp alanında bir uygulama [Comparison of logistic regression analysis (LRA), artificial neural networks (ANNs) and classification and regression trees (C&RT) methods and an application in medicine]* (Unpublished Master Thesis). Gaziosmanpaşa University, Institute of Health Sciences, Tokat.
- Munakata, T. (2008). *Fundamentals of the new artificial intelligence, neural, evolutionary, fuzzy and more*. Springer-Verlag London Limited.
- Naik, B., & Ragothaman, S. (2004). Using neural networks to predict MBA student achievement. *College Student Journal*, 38(1), 143-149.
- Ocakoğlu, G. (2006). *Lojistik regresyon analizi ve yapay sinir ağı tekniklerinin sınıflama karşılaştırması ve bir uygulama [Logistic regression analysis and comparison of classification characteristics of artificial neural network techniques and an application]* (Unpublished Master Thesis). Uludağ University Institute of Health Sciences, Bursa.
- Oğuzlar, A. (2004). CART analizi ile hanehalkı işgücü anketi sonuçlarının özetlenmesi [Summary of results of household labor force with CART analysis]. *Atatürk University Journal of Economics and Administrative Sciences*, 18(3-4), (79-90).
- Oğuzlar, A. (2006). Hanehalkı tipi ve kır-kent ayırımının diskriminant analizi ile incelenmesi [Assessment of household type and rural-urban area distinctions by means of discriminant analysis]. *Akdeniz University Faculty of Economics & Administrative Sciences Faculty*, 11, 70-84.
- Okioga, C. K. (2013). The impact of students' socio-economic background on academic performance in universities, a case of students in Kisii University College. *American International Journal of Science*, 2(2), 38-46.
- Özabacı, N., & Acat, M. B. (2005). Sosyo ekonomik çevreye göre ilköğretim öğrencilerinin başarısızlık nedenleri [Causes of academic underachievement among socio-economic level for secondary school student]. *Eskişehir Osmangazi University Journal of Social Sciences*, 6(1), 145-170.
- Özdamar, K. (2010). *Paket programlar ile istatistiksel veri analizi (Çok değişkenli analizler) [Statistical data analysis with package programs (Multivariate analysis)]* (7th ed), Eskişehir: Kaan Bookstore.
- Özdemir, Y., & Koruklu, N. (2011). Üniversite öğrencilerinde değerler ve mutluluk arasındaki ilişkinin incelenmesi [Investigating relationship between values and happiness among university students]. *Van Yuzuncu Yil University Journal of Education*, 8(1), 190-210.
- Özekes, S. (2003). Veri madenciliği modelleri ve uygulama alanları [Data mining models and application areas]. *Journal of İstanbul Commerce University*, 3, 65-82.
- Özer, Y., & Anıl, D. (2011). Öğrencilerin fen ve matematik başarılarını etkileyen faktörlerin yapısal eşitlik modeli ile incelenmesi [Examining the factors affecting students' science and mathematics achievement with structural equation modeling]. *H. U. Journal of Education*, 41, 313-324.
- Özkan, Y. (2013). *Veri madenciliği yöntemleri [Data mining methods]* (2nd ed). İstanbul: Papatya Publishing.
- Öztemel, E. (2012). *Yapay sinir ağları [Artificial neural networks]* (1st ed). İstanbul: Papatya Publishing.
- Öztürk, S., Coşkun, A., & Dirsehan, T. (2012). Fırsat sitelerine yönelik e-sadakati belirleyen boyutların incelenmesi [Analyzing dimensions determining e-loyalty towards daily deal

- sites]. *Eskişehir Osmangazi University Journal of Economics and Administrative Sciences*, 7(2), 217-239.
- Pehlivan, G. (2006). *CHAID analizi ve bir uygulama [CHAID analysis and an application]* (Unpublished Master Thesis). Yıldız Technical University, Institute of Science, İstanbul.
- Rokach, L., & Maimon, O. (2008). *Data mining with decision trees: Theory and applications. Series in Machine Perception and Artificial Intelligence*. Vol. 69, USA: World Scientific Publishing Co. Pte. Ltd.
- Sabancı, K. (2013). *Şeker pancarı tarımında yabancı ot mücadelesi için değişken düzeyli herbisit uygulama parametrelerinin yapay sinir ağlarıyla belirlenmesi [Determination of variable rate herbisit application parameters with artificial neural networks for weed contention in agriculture of sugar beet]* (Unpublished Doctoral Dissertation). Selçuk University, Institute of Science, Konya.
- Sadi, Ö., Uyar, M., & Yalçın, H. (2014). Lise öğrencilerinin biyoloji dersi başarılarında, cinsiyet, sınıf düzeyi ve aile yapısının rolü [The role of gender, grade level and family environment in high school students' biology achievement]. *Journal of Research in Education and Teaching*, 3(2), 138-151.
- Sağiroğlu, Ş., Beşdok, E., & Erler, M. (2003). *Mühendislikte yapay zeka uygulamaları-I, yapay sinir ağları [Artificial intelligence applications in engineering-I, artificial neural networks]* (1st ed). Kayseri: Ufuk Publishing.
- Serinkan, C., & Bardakçı, A. (2007). Pamukkale Üniversitesinde çalışan öğretim elemanlarının iş tatminlerine ilişkin bir araştırma [Job satisfaction: An empirical research towards academicians working in Pamukkale University]. *Selçuk University Karaman Journal of FEAS*, 12, 152-163.
- Seven, A. (1993). *Yapay sinir ağları ile doku sınıflandırma [Tissue classification using artificial neural networks]* (Unpublished Master Thesis). İstanbul Technical University, Institute of Science, İstanbul.
- Silahtaroglu, G. (2013). *Veri madenciliği kavram ve algoritmaları [Data mining concepts and algorithms]* (2nd ed). İstanbul: Papatya Publishing.
- Simpson, P. K. (1990). *Artificial neural systems foundations, paradigms, application and implementation*. Elmsford NY: Pergamon Press.
- Sun, J., & Hui L. (2008). Data mining method for listed companies' financial distress prediction. *Knowledge-Based Systems*, 21(1), 1-5.
- Şahin, A. (2011). İlköğretim 6. sınıf öğrencilerinin dinleme becerisi farkındalıklarının sosyo-ekonomik düzeye göre incelenmesi [A study on 6th grade students' self-awareness on listening skills according to their socio-economic level]. *Çankırı Karatekin University Journal of Institute of Social Sciences*, 2(1), 178-188.
- Şirvan, O. (2010). *Yapay sinir ağları kullanılarak retina görüntülerinden hastalık tanılama sistemi tasarımı ve gerçekleştirimi [Design and implementation of disease recognition system in retinal images using artificial neural networks]* (Unpublished Doctoral Dissertation). Ege University, Institute of Science, İzmir.
- Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Allyn & Bacon/Pearson Education.
- Tatlıdil, H. (1996). *Uygulamalı çok değişkenli istatistiksel analiz [Applied multivariate statistical analysis]*. Ankara: Cem Web Offset.
- Tepehan, T. (2011). *Türk öğrencilerinin PISA başarılarının yordanmasında yapay sinir ağı ve lojistik regresyon modeli performanslarının karşılaştırılması [Performance comparison of artificial neural network and logistic regression model in predicting Turkish students' PISA success]* (Unpublished Doctoral Dissertation). Hacettepe University, Institute of Social Sciences, Ankara.

- Thomas, L. C. (2000). A survey of credit and behavioral scoring: Forecasting financial risk of lending to consumer. *International Journal of Forecasting*, 16(2), 149-172.
- Thigpen, M. K. (2000). *Data mining techniques in education: a comparison of conventional statistical linear regression and neural network based tools* (Unpublished Doctoral Dissertation). University of Alabama, Alabama, ABD.
- Tolon, M. (2007). *Tüketici tatmininin yapay sinir ağları yöntemiyle ölçülmesi ve Ankara'daki perakendeci mağazaların müşterileri üzerinde bir uygulama [Measuring customer satisfaction with artificial neural networks and an application of retail consumers in Ankara]* (Unpublished Doctoral Dissertation). Gazi University, Institute of Social Sciences, Ankara.
- Torun, T. (2007). *Finansal başarısızlık tahmininde geleneksel istatistikî yöntemlerle yapay sinir ağlarının karşılaştırılması ve sanayi işletmeleri üzerinde uygulama [Comparison of traditional statistical techniques with artificial neural networks in financial failure prediction and an application on industry firms]* (Unpublished Doctoral Dissertation). Erciyes University, Institute of Social Sciences, Kayseri.
- Tosun, S. (2007). *Sınıflandırmada yapay sinir ağları ve karar ağaçları karşılaştırması: Öğrenci başarıları üzerine bir uygulama [Artificial neural networks and decision tree comparison in classification analysis: An application on students' success]* (Unpublished Master Thesis). İstanbul Technical University, Institute of Science, İstanbul.
- Türe, M., Kurt, İ., Yavuz, E., & Kürüm, T. (2005). Hipertansiyonun tahmini için çoklu tahmin modellerinin karşılaştırılması (Sinir ağları, lojistik regresyon ve esnek ayırma analizleri) [Comparison of multiple prediction models for hypertension (Neural network, logistic regression and flexible discriminant analyses)]. *The Anatolian Journal of Cardiology*, 5(1), 24-28.
- Wheeler, M. C. (1993). *A comparative case study of neural network analysis and statistical discriminant function analysis for predicting law students passing the bar examination* (Unpublished Doctoral Dissertation). Gonzaga University Spokane, WA, USA.
- Yakut, E. (2012). *Veri madenciliği tekniklerinden c5.0 algoritması ve destek vektör makineleri ile yapay sinir ağlarının sınıflandırma başarılarının karşılaştırılması: İmalat sektöründe bir uygulama [The comparison of the classification successes of the artificial neural networks through data mining techniques of C5.0 algorithm and supporting vector machines: An application in manufacturing sector]* (Unpublished Doctoral Dissertation). Atatürk University, Institute of Social Sciences, Erzurum.
- Yıldırım, İ. (2000). Akademik başarının yordayıcısı olarak yalnızlık, sınav kaygısı ve sosyal destek [Loneliness, exam anxiety and social support as predictors of academic success)]. *H. U. Journal of Education*, 18, 167-176.
- Yurtoğlu, H. (2005). *Yapay sinir ağları metodolojisi ile öngörü modellemesi: Bazı makroekonomik değişkenler için Türkiye örneği [Predictive modeling with artificial neural network methodology: Turkey examples for some macroeconomic variables]* (Unpublished Master Thesis). SPO General Directorate of Economic Models and Strategic Research, Ankara.
- Yüksek, A. G. (2007). *Hava kirliliği tahmininde çoklu regresyon analizi ve yapay sinir ağları yönteminin karşılaştırılması [Comparison of multiple regression analysis and neural network methods for predicting air pollution]* (Unpublished Doctoral Dissertation). Cumhuriyet University, Institute of Social Sciences, Sivas.
- Zurada, J., & Lonial, S. (2005). Comparison of the performance of several data mining methods for bad debt recovery in the healthcare industry. *The Journal of Applied Business Research*, 21(2), 37-54.