Education and Science

Vol 47 (2022) No 211 195-222

Original Article

Comparison of Classification Performances of MARS and BRT Data Mining Methods: ABİDE- 2016 Case *

Hikmet Şevgin [1], Emine Önen [2]

Abstract

This research examined the relationships between student, teacher, school and instructional qualifications and 8th grade students' science achievement, based on the conceptual framework created by Nilsen and Gustafsson (2016), using data mining methods MARS and BRT. Research data (n=10407 students, n=941 teachers and n=865 school administrators) were obtained from the ABİDE study conducted at the national level by the Ministry of National Education in 2016. MARS and BRT analyzes were performed in the SPM 8.2 program. The science achievement classification performances of these methods were compared by considering the correct classification rate, sensitivity and specificity rates, F1 statistical value and the area under the ROC curve. It was found that the BRT method was more successful than the MARS method in terms of all these criteria, and the most important predictors of science achievement were similar compared to these two methods. The results revealed that the most important predictors of science success are the student's perception of science self-efficacy, the father's occupation, the family's monthly income, the instructional activities of the teacher, the teacher's experience and preparation for the lesson, and the school administrators' perception of school climate. It is thought that the reason why BRT outperforms the MARS method in terms of the criteria considered in this study is that BRT learns from errors with the additive combination of various regression trees and provides a stronger classification performance by minimizing the errors that may occur in classification. This study revealed the benefits of using these two data mining methods in the field of Educational Sciences and discussed the contribution of the related methods in this field.

Keywords

Data mining

Multivariate adaptive regression splines

Boosted regression trees

ABİDE

Science achievement

Article Info

Received: 02.09.2021

Accepted: 07.06.2022

Online Published: 07.29.2022

DOI: 10.15390/EB.2022.10575

* This article is derived from Hikmet Şevgin's PhD dissertation entitled "Predicting the ABIDE 2016 science achievement: The comparison of MARS and BRT data mining methods", conducted under the supervision of Emine Önen.

[1] Van Yüzüncü Yıl University, Faculty of Education, Department of Educational Sciences, Turkey, hikmetsevgin@gmail.com

[2] Gazi University, Gazi Faculty of Education, Department of Educational Sciences, Turkey, emine_onen@yahoo.com

Introduction

Academic Skills Monitoring and Evaluation (ABIDE), financed and administered by the Ministry of National Education (MoNE), is a type of large-scale exam designed to be administered biennially on 8th grade students enrolled in formal education (ABİDE 8. Sınıflar Raporu, 2017). It was first applied in schools selected in Ankara as a pilot application in 2015, and then in schools in the sample selected to represent all schools in Turkey in 2016 and 2018. ABIDE is used to measure 8th grade students' math and science literacy, reading skills and social knowledge. In addition, it is a large-scale examination application at the national level, in which various information about students (such as students' attitudes and motivations towards the lessons, their interests and concerns about the lessons, the peer bullying they encounter in the social and school environment, and belonging to the school), branch teachers (such as professional development, professional satisfaction, instructional activities) and school administrators are obtained.

Large-scale examinations serve many purposes including monitoring standards at the national level, providing feedback to students and their families, guiding teachers' instructional activities, and directing education policies (EACEA, 2009). It is also used in international level to provide information about education systems at different levels and cultures, to compare various subjects in education between countries, to examine the factors that affect education at the micro and macro level at the level of countries, to encourage international cooperation and to follow the development processes related to education (Torney-Purta & Amadeo, 2013). The results obtained from these applications allow to examine the relations between the factors related to the student, teacher and school, the possible changes that may occur in these relations and the developments that can be observed over time (Nilsen, Gustafsson, & Blömeke 2016). In addition, these exam applications allow the selection of representative samples from more than one education system and the application of multivariate analyzes (Nilsen et al., 2016). Ultimately, these results reveal the relationships between student achievement and student, teacher, school qualifications and instructional qualifications.

Characteristics related to students, teachers, and schools as components of the education system can be extensively studied through studies such as PISA (Program for International Student Assessment) and TIMSS (Trends in International Mathematics and Science Study). ABIDE, similar to PISA and TIMSS, is an application that aims to obtain information at student, teacher and school level about the academic achievement of students and the variables that are thought to be related to this success, because ABIDE is similar to PISA in terms of focusing on the measurement of skills, and TIMSS in terms of being based on achievements (ABİDE 8. Sınıflar Raporu, 2017; Taş, Arıcı, Ozarkan, & Özgürlük, 2016; Yıldırım, Özgürlük, Parlak, Gönen, & Polat, 2016). Therefore, the implementation of such a practice in the form of large-scale exams at the national level was considered valuable in terms of measuring student skills and achievements in terms of student, teacher, school and instructional qualifications. The need for detailed analysis of the data obtained has made the ABIDE application the basis of this study.

It is seen in the literature (Blömeke, Olsen, & Suhl, 2016; Gustafsson & Nilsen 2016; Nilsen vd., 2016; Nortvedt, Gustafsson, & Lehre, 2016; Özçınar, 2006; Rutkowski & Rutkowski, 2016; Scherer & Nilsen 2016) that conceptual frameworks have been developed to explain the relationships between student, teacher, school and instructional qualifications based on the researches carried out on the basis of data obtained from large-scale exam applications. One of them is the “Conceptual framework for determinants of student outcomes” developed by Nilsen and Gustafsson (2016) on the basis of studies based on TIMSS. This framework was created based on the "Dynamic Model of Educational Effectiveness" proposed by Creemers and Kyriakides (2007). Nilsen and Gustafsson's (2016) conceptual framework for the determinants of student outcomes based on the Dynamical Model of Educational Effectiveness consists of national level, school level, class level and student level, as in Figure 1.

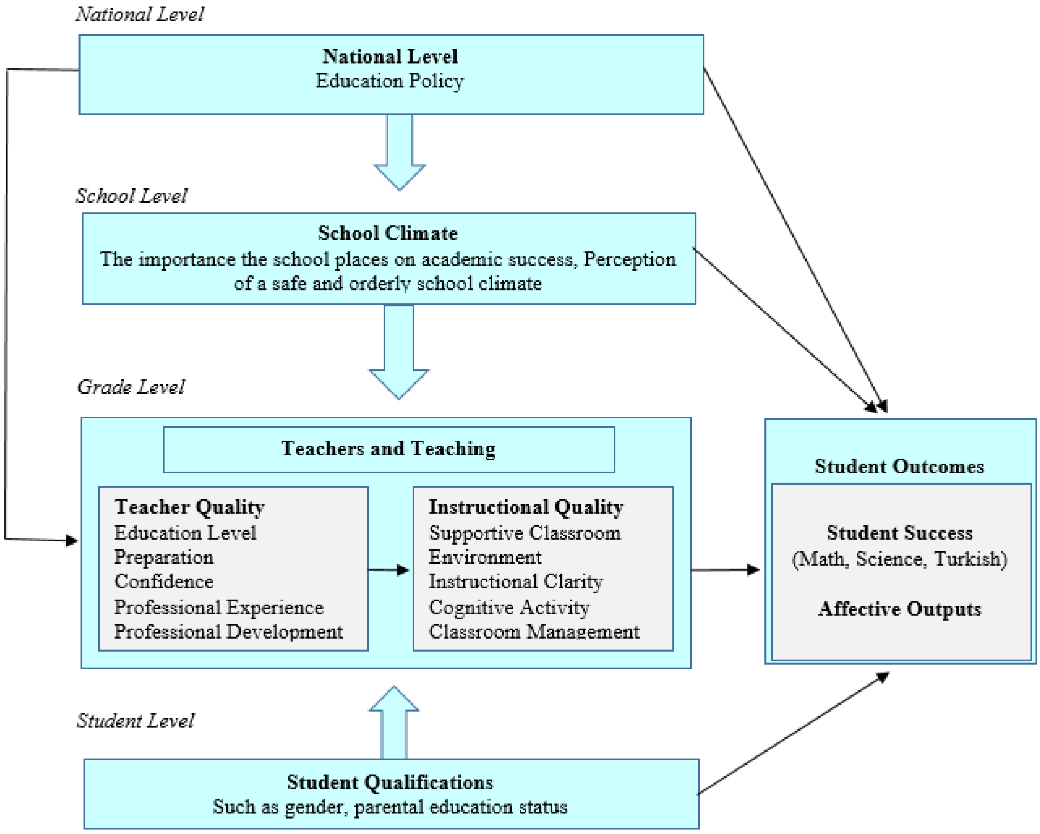

Figure 1. Conceptual framework for determinants of student outcomes (Nilsen et al., 2016)

This model takes into account the educational system's complexity and categorizes the predictors of student outcomes at the national, school, grade, and student levels (Nilsen et al., 2016). This model was not directly tested for the data obtained with the ABIDE application in the current study, but instead, the qualities that could be related to the students' science achievement were determined based on this conceptual framework at the student, teacher and school level. In other words, this model was used to determine possible factors that may affect the science achievement of students participating in the ABIDE application, student, teacher, school qualifications and instructional qualifications by MARS and BRT analysis methods. As a matter of fact, Dobert and Sroka (2004) and Azigwe (2016) made cross-country comparisons based on the PISA application, one of the international large-scale tests, within the framework of the dynamic model of educational effectiveness.

The number of studies using the data of the ABIDE application is low in the relevant literature (Akıncı, 2020; Çalık, 2020; Doğru, 2019; Elkonca, 2020; Kılıç, 2019; Özgürlük, 2019; Uysal, 2019; Ülkü, 2019) and there is no study that uses both analysis methods together in education. The current study is important because it is the first study in which MARS from the family of nonlinear regression methods under the umbrella of data mining and BRT analysis methods from the family of Decision Trees are used together in education. Again, since both analysis methods are among the non-parametric methods that do not need to provide various assumptions in the data sets, it can be said that they provide flexibility in terms of use.

A certain part of the data obtained with ABIDE at the national level, and PISA and TIMSS at the international level regarding the characteristics measured in the field of education tend not to provide the assumptions of parametric methods such as normality, linearity and homogeneity of variances. The use of various analysis methods on data sets where assumptions cannot be provided is a limitation. There is no such limitation in non-parametric methods. Therefore, in case the assumptions cannot be met, it is considered important to introduce and apply various non-parametric approaches and methods so that the aforementioned data sets will yield more accurate and valid results. Here, it is believed that data mining methods will play a significant role in education, particularly in studies using enormous data sets generated from large-scale assessments, when parametric methods' assumptions cannot be satisfied. Parallel to this, it is observed in the literature (Akçapınar, 2014; Aksu, 2018; Al-Saleem, Al-Kathiry, Al-Osimi, & Badr, 2015; Amrieh, Hamtini, & Aljarah, 2016; Baradwaj & Pal, 2011; Bilen, Hotaman, Aşkın, & Büyüklü, 2014; Mazman, 2013; Saa, 2016; Tepehan, 2011; Yu, Kaprolet, Jannasch-Pennell, & DiGangi, 2012) that data mining methods have been used for such cases. However, no studies in which MARS and BRT data mining methods were used have been encountered. In this study, MARS and BRT analysis methods, which are thought to contribute to the modeling of complex relationships based on classification and estimation, between the existing variables in the datasets of researches in the field of education were preferred. The MARS analysis method tries to flatten the variables as piecewise linear in order to analyze the variables that do not provide linearity with the dependent variable (Friedman, 1991). Thus, in contrast to traditional regression models, which establish the regression equation on a single basis function and obtain a single regression line, the MARS analysis method was chosen for this study because it provides more detailed information about the relationships between the variables by establishing the regression equation on multiple basis functions. Similarly, the BRT analysis method combines multiple decision trees on the basis of the Boosting algorithm, one of the recent ensemble learning algorithms (Elith, Leathwick, & Hastie, 2008; Friedman, 2001, 2002; Friedman & Meulman, 2003; Hastie, Tibshirani, & Friedman, 2009). Therefore, the Boosting algorithm's creation of a series of models in the form of a single predictive model by bringing together weak classifiers (such as decision trees, such as C&RT) has been effective in the preference of the BRT analysis method. In addition, the fact that both analysis methods are little affected by missing data and extreme values of the data set (Elith et al., 2008; Salford Systeam, 2018) has been another reason for preference.

It is possible with MARS and BRT analysis methods to examine in depth the multivariate and complex relationships in large volumes of data that are obtained from applications such as large-scale exams and that do not provide the assumptions of parametric methods. In addition, both methods provide the opportunity to rank the variables that contribute the most to the prediction of the dependent variable, and the variables that contribute the least or even no contribution, according to the level of importance. In this respect, it can be thought that these methods will guide the policies of policy makers (such as the distribution of income transferred, time spent) and increase the efficiency in education. Furthermore, MARS and BRT analysis methods can provide strong predictions for examining the information about students, teachers and schools obtained in large-scale exam applications applied in the field of education and can show successful classification performance (pass-fail about students). Thus, it is essential to use these methods for selection and classification purposes in educational assessment and evaluation in order to achieve high classification/decision validity results for both student achievement and psychological structures (interest, attitude, motivation).

MARS analysis method can reveal the relationship of many more variables with each other in a more robust way, since it can define nonlinear relationships between variables as piecewise linear relationships and include them in the analysis when examining the relationship between psychological structures and other structures in education. Additionally, it can provide measurement and evaluation professionals and associated researchers with observing the interactions (the interaction of the independent variables with each other and the effect of this interaction on the dependent variable) of independent factors. On the other hand, the BRT analysis method may improve classification performance by learning from mistakes via the additive combination of numerous regression trees and by minimizing classification errors in this way (Elith et al., 2008). Finally, it should be known that the less the probability of misclassifying successful and unsuccessful students by the existing statistical methods, the less they will contribute to the classification validity (Erkuş, 2003). Therefore, it is thought that MARS and BRT analysis methods can be used as well as other statistical methods used for selection and classification functions based on measurement results in education.

Within the scope of this research, the MARS and BRT analysis methods were used to examine the relationships between various school, teacher, and student qualifications of eighth-grade students in the field of science and science achievement, and these two data mining methods were compared in terms of their classification performance regarding science achievement. Rather than just modeling the existing situation, the primary goal here is to identify the most critical factors associated with science achievement and to give guidance for educational activities and policies. The relationships between various factors determined in accordance with the theoretical framework discussed in this study and science achievement were examined using MARS and BRT analysis methods. These factors were ranked according to their contribution to predicting science achievement, and the classification performances of these two methods were compared. The research intended to address the following questions in this direction:

1. Do the classification performances of MARS analysis method and BRT analysis method differ in predicting science achievement according to ABIDE 2016 application?
2. What are the importance levels of the most important predictors of science achievement in ABIDE 2016 application based on MARS and BRT analysis methods?

Method

This research, which used data from the ABIDE 2016 study to examine the possible relationships between various factors affecting 8th grade students' science achievement and success, and to compare the effectiveness of the MARS and BRT methods for classifying students' science achievement, is a relational survey model study, one of the general survey methods. Survey models are research techniques that are used to accurately depict a previous or current situation. The general survey model is a collection of survey arrangements done on the whole universe or a subset of the universe in order to form a general judgment about the universe in a multi-element universe. The problem under investigation is attempted to be characterized in its own terms and in its current state (Karasar, 2009).

Data Set

The data for this research were collected from randomly chosen 33590 8th grade students, 1420 teachers, and 1280 school administrators regardless of school type in 2016 using the Ministry of National Education's ABIDE program. Separate questionnaires for each teacher and school administrator were matched, and the results were merged on the basis of students. In other words, the students who fall under the body of each scientific teacher are merged with the science teachers who are under the body of each school administrator. To assign missing data before starting the analysis process, deletion for missing data in demographic variables, value assignment method with regression was performed since missing data for data obtained based on Likert-type rating scales contain less than 5% MCAR missing data (totally unbiased loss) in total for each scale (Tabachnick & Fidell, 2015). Deleting the data of each school administrator due to lost data means that the teachers and students in that school are also removed from the study group. Again, deleting the data of each teacher due to missing data means that his students are also removed from the study group. Following the deletion and assignment of lost data, the data merging procedure was carried out by randomly assigning teachers and school administrators to 14868 students, with the entire data set included in the same group, on the basis of the teachers and administrators on the basis of the teachers. Students are classified on the basis of their achievement scores in 5 levels in ABIDE, as “below basic, basic, intermediate, upper-intermediate and advanced”. At the beginning of the study, a two-step cluster analysis was applied on the data set to determine the standard cut-off score, and it was observed that the data were collected under two clusters with a cut-off point of 468.96, which were homogeneous and heterogeneous with the other cluster. It has been observed that the point of divergence between the two clusters is in the "medium" category among the proficiency levels given with threshold values in the ABIDE preliminary report. Following that, it was determined that the analysis conducted with this data set and the findings of the study conducted with the data set in which the lower-intermediate level was categorized as unsuccessful and the upper-intermediate level as successful. As a result, the data set contains 10407 students, 941 teachers classified as below-intermediate (below basic, basic) unsuccessful, above-intermediate (higher-intermediate, advanced) successful, and 865 school administrators classified as below-intermediate (below basic, basic) unsuccessful. Again, on the basis of the conceptual framework developed based on the "Dynamic Model of Educational Effectiveness," the characteristics related with students' science achievement were determined at the student, teacher, and school levels.

Measurement Tools

This study consists of questionnaires for teachers, students and schools belonging to the ABIDE application and the data collected with the Science achievement test. Science achievement test consists of 20 items, half of which are multiple choice and the other half is open-ended. Multiple-choice items were scored as 1-0, and open-ended items were scored with more than one planner. The result for the inter-rater consistency for each open-ended item was calculated between 0.83 and 0.99 (Cramer's V values) (ABİDE 8. Sınıflar Raporu, 2017). These items were developed with the participation of academicians and field experts, taking into account the skills such as remembering-finding, understanding, interpreting-inferring and evaluation by making use of the curriculum (ABİDE 8. Sınıflar Raporu, 2017).

The school questionnaire was applied to the school principals of the participating schools, the teacher questionnaire was applied to the science teachers and the student questionnaire was applied to the 8th grade students. The school questionnaire includes items such as the type of school, its administrative form, geographical and economic location, security, and the demographic information of the administrators. In the teacher questionnaire, there are items related to demographic information, lesson preparation, self-efficacy, professional development and professional satisfaction for teachers. In the student questionnaire, there are items on many subjects such as the students' home, school and social environment life, demographic information, educational resources, attitudes towards the school, peer bullying, parental approach and the value students give to the lessons. The above-mentioned characteristics were measured using 4- and 5-point Likert type scales in student, teacher, and school questionnaire applications. To match the total scores obtained in different Likert scales, the standard score was transformed to the Z score. Because the factor load values obtained for the scale items were not comparable, the reliability coefficient for the Likert type scales was computed using McDonald's reliability coefficient (Yurdugül, 2006). According to the relevant literature, a computed reliability value of $0.00 \leq \alpha \leq 0.40$ indicates unreliability; $0.40 \leq \alpha \leq 0.60$ indicates low reliability; $0.60 \leq \alpha \leq 0.80$ indicates strong confidence (Özdamar, 2013). McDonald's used values ranging from 0.68 to 0.95 for the measurements. The calculated McDonald's ω reliability coefficients were obtained between the highly reliable and highly reliable categories and were evaluated as evidence for the reliability of the said measurements. In addition to the Likert type scales used, some demographic variables of students, teachers and administrators were also used within the scope of the research (Table 1).

Table 1. Descriptive Statistics of Predictive Variables

Predictive Variables	Variable Type	Sub-categories	%
Student's Gender	Categorical	1: Female	48,2
		2: Male	51,8
Father's Profession	Categorical	1: Not alive	2,4
		2: Public Personnel	17,5
		3: Private Sector Employee	17,4
		4: Self Employed-Tradesman-Business Owner	47,7
		5: Retired, Working	4,8
		6: Retired, Not working	4,6
		7: Not working-Unemployed	5,6
Mother's Profession	Categorical	1: Not alive	0,6
		2: Public Personnel	6,7
		3: Private Sector Employee	6,2
		4: Self Employed-Tradesman-Business Owner	5,0
		5: Retired, Working	0,4
		6: Retired, Not working	2,0
		7: Housewife	76,4
		8: Not working-Unemployed	2,7
Monthly Salary	Categorical	1: Between 0-1500 TL	31,4
		2: Between 1501-2500 TL	23,6
		3: Between 2501-4000 TL	16,6
		4: Between 4001-6000 TL	5,8
		5: 6001 TL and over	2,9
		6: not knowing	19,7
House	Categorical	1: Rental	24,1
		2: It belongs to us	70,8
		3: It belongs to one of our relatives	3,3
		4: Lodging	1,7
Teacher's Gender	Categorical	1: Female	37,4
		2: Male	62,6
Teacher's Education Level	Categorical	1: Two-year degree	3,2
		2. Undergraduate	93,6
		3: Master	3,2
		4: Doctoral	0,0
Type of School the Teacher Graduated From	Categorical	1: Education Institute/Higher Teacher Training School	0,7
		2: Faculty of Education/Faculty of Educational Sciences	79,7
		3: Faculty of Arts and Sciences/ Faculty of Language	19,6
		History and Geography	0,0
		4: Open Education Faculty/Undergraduate Completion	0,0
		5: Other	

As seen in Table 1, categorical variables do not need to be coded as dummy variables when using the MARS analysis method since the type of variables is specified separately during the model design phase.

Data Analysis

This research comparatively examined the possible relationships between the science achievement of various factors related to 8th grade students, teachers and schools in the field of science in ABİDE 2016 application, and MARS and BRT analysis methods, which are data mining methods. The study compared both methods in terms of "accuracy, specificity, sensitivity, precision, F1-statistics" ratios and classification performances as "AUC value (area under the ROC curve)".

The model setup step of the MARS and BRT analysis methods was created using predefined (default) values. Only because the BRT analysis method's default values treat the data set as 80% training and 20% test data, the data set was included in the analysis as 80% training and 20% test data for the MARS analysis method to ensure comparable performance. Thus, the study employs 80% of the whole dataset as training data for modeling and 20% as test data for validity.

MARS-Multivariate Adaptive Regression Splines

MARS is critical for both classification and regression; it has been effectively employed in scientific fields where complex relationships involving a large number of variables must be represented. MARS uses appropriate transformation strategies to turn nonlinear interactions between dependent and independent variables into linear structures (Deichmann, Eshghi, Haughton, Sayek ve Teebagy, 2002; Friedman 1991). In this way, MARS determines the relationships between the dependent variable and independent variables by flattening the small linear particles. In addition, the missing data in the data set is considered unbiased due to its ability to effectively process the missing data in the model, divide the nonlinear models into linear particles and make parameter estimations for each particle separately (Kayri, 2009).

MARS analysis is a two-step process that aims to generate the best suitable model. It begins by calculating the sum of the principal functions, which are the transformations of the independent variables, taking into account non-linear deviations and model interactions. In the second step, it uses the basic functions in estimation on behalf of the independent variables by applying the removal of the basic functions with the least effect with the least squares method (Deichmann et al., 2002). The least squares method is used to find the constants in these basic functions (Friedman, 1991; Hastie et al., 2009). Thus, the regression lines tend to pass through the points closest to the values, and by merging these lines at the nodes, the regression spline function is obtained (Friedman, 1991; Hastie et al., 2009; Oğuz, 2014; Özfalcı, 2008; Statsoft, 2018). This function is as in Equation 1:

$$Y = f(x) = \beta_0 + \sum_{k=1}^{K} a_k \, \beta_k \, (X_t) + \varepsilon_i \qquad (1)$$

As can be seen, the regression equation consists of the constant term (β_0) and the weighted sum of one or more fundamental functions. K is the number of basic functions, k is the number of nodes, X is the independent variable, a_k k. is basic function coefficient and $\beta k \ (Xt)$ is k. basic function for t independent variable in this equation (Hastie et al., 2009).

The basic functions are probably nonlinear versions of X_t. But Y is a linear function of fundamental functions (Friedman, 1991; Özfalcı, 2008). The basic function is defined as in Equation 2;

$$B_m(x) = \prod_{t=1}^{L_m} [s_{1,m}(x_{v(1,m)} - k_{1,m})]_+ \qquad k=1,2,\ldots,K \qquad (2)$$

Here ;

L_m : the degree of interaction,

$S_{1,m}$: $\in [\pm 1]$

$k_{1,m}$: node value and

$x_{v(1,m)}$: the value of independent variable

MARS uses fundamental functions to model linear regression lines made up of particles. Equations 3 and 4 contain equations of piecewise linear basic functions in the form of (x - t)+ and (t - x)+. Here, "+" indicates the positive side and indicates that only the positive results of the related equation are taken into account, and if the desired condition is not met, it will take the value 0 (Deconinck et al., 2005; StatSoft, 2018). So the piecewise linear basic function is as below (Hastie et al., 2009):

$$(x-t)_+ = \begin{cases} (x-t), eğer\ x > t, \\ 0, \quad diğer, \end{cases} \tag{3}$$

$$(t-x)_+ = \begin{cases} (t-x), eğer\ x < t, \\ 0, \quad diğer \end{cases} \tag{4}$$

Again an alternative representation of the basic functions is as $(x - t)_+ = \max(x - t, 0)$ and $(t - x)_+ = \max(t - x, 0)$ (Ferreruela, 2008). Figure 2 shows a graphical representation of the basic functions of $(x - t)_+$ and $(t - x)_+$ for the value t=0.5.

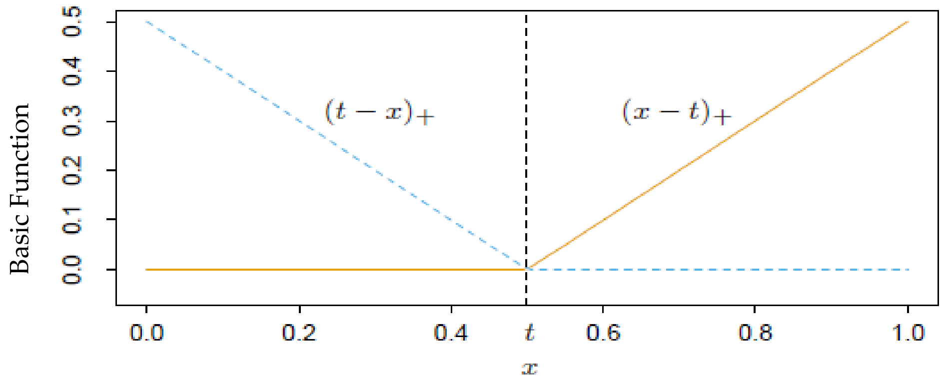

Figure 2. Mirror variable versus basic functions (Hastie et al., 2009)

In figure 2 for the $(x - 0.5)_+$ and $(0.5 - x)_+$ functions, each function value is piecewise linear with a node at t, and these two functions are called reflection pairs (Hastie et al., 2009).

The forward step algorithm, which is the initial phase, identifies all potential nodes in the data set in order to reach at an optimum model. However, if too many basic functions are included at the end of this procedure, the MARS model may suffer overfitting. This problem is reduced using the second step, the backward step algorithm. Using the Generalized Cross Validation (GCV) criterion, redundant principal functions that contribute the least or do not contribute to the model are extracted in this phase from the basic functions utilized in the forward step algorithm to create the best final model. This criterion value is calculated with the help of the following equation, taking into account both the error term and the model complexity:

$$GCV(M) = \frac{1}{N}\frac{\sum_{i=1}^{N}[y_i - f_M(x_i)]^2}{[1 - \frac{C(M)}{N}]^2} \tag{5}$$

In the equation, N refers to the number of samples in the data set, C(M) refers to the number of effective parameters in the model (Friedman, 1991).

The relative contributions of each independent variable entering the final model and the interactions between the variables are reached as a result of ANOVA decomposition (Salford Systeam, 2018).

$$f(x) = \beta_0 + \sum_{k_m=1} \beta_m \beta_m(xi) + \sum_{k_m=2} \beta_m \beta_m(xi, xj) + \sum_{k_m=3} \beta_m \beta_m(xi, xj, xk) + \cdots \quad (6)$$

The first sum expresses all basic functions including a single variable, the second sum includes two variables and their interactions, if any, and the third sum describes all basic functions containing three variables and their interactions, if any (Friedman, 1991). Thus, the MARS analysis method's estimation variables can be clearly interpreted as a result of ANOVA decomposition.

BRT-Boosted Regression Trees

Boosted Regression Trees (BRT) is a non-parametric regression technique that combines classification and regression trees from decision trees with the Gradient Boosting algorithm from Boosting algorithms (Colin, Clifford, Wu, Rathmanner, & Mengersen, 2017). This analysis method, abbreviated as BRT and referred to as such, has been successfully applied in scientific fields where complex relationships among various variables are modeled by adding classification trees when the dependent variable is categorical and regression trees when the dependent variable is continuous. BRT is a technique that uses the classification and regression tree (C&R T) as a weak learner and the gradient boosting algorithm as a model for adding weak learners to each other to increase the prediction performance (Elith et al., 2008). In other words, BRT combines the classification and regression trees into the gradient boosting algorithm while accounting for their errors. It then learns from the errors in each tree structure and trains the next tree addition attempting to reduce the errors in the previous tree. It is a sequential and iterative method in this respect. The graphical representation of the BRT analysis method with C&RT is as in figure 3;

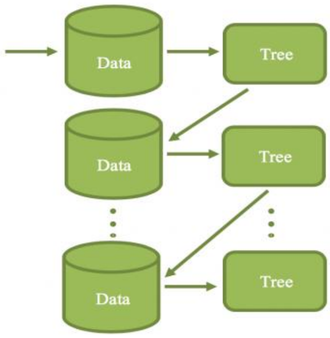

Figure 3. Figurative representation of BRT with the addition of C&RTs

With a single C&RT tree, let the K (x) model have an accuracy rate of 70% and the equation for this model be "Y = K(x) + Error". A second C&RT tree to be added later, taking into account the error of the previous formation, the new equation to be created will be “Error = L(x) + Error2”. Following that, the third C&RT tree will be constructed using the error from the previous formation, using the equation "Error2 = M(x) + Error3". The combination of the three becomes “Y = K (x) + L (x) + M (x) + Error3”. Since the errors are taken into account with each iteration, the errors will decrease and therefore the model obtained will have an accuracy higher than the 70% accuracy rate obtained at the beginning.

Performance Metrics

The confusion matrix is used to analyze the classifier's ability to recognize a pattern across many classes. The confusion matrix of the areas corresponding to the actual and estimated classes of clay is shown in Table 2.

Table 2. Confusion Matrix

		Estimated Class		
		Unsuccessful	**Successful**	**Total**
Real Class	**Unsuccessful**	TN	FP	TN+FP
	Successful	FN	TP	FN+TP
	Total	TN+FN	FP+TP	TP+FP+TN+FN

(**TN:** True Negative, **TP:** True Positive
FN: False Negative, **FP:** False Positive)

Since the confusion matrix was used to determine the classification performance in this study, accuracy rate, specificity rate, sensitivity rate, precision rate, F1-statistic value and AUC value were chosen as performance criteria. The equations for these criteria are given below.

Correct classification rate: A measure of how often the classifier produces an accurate prediction. It indicates how successful the actually successful are anticipated to be, and how unsuccessful the actually unsuccessful are predicted to be.

$$Correct\ classification\ rate = \frac{(\mathrm{TP} + \mathrm{TN})}{(\mathrm{TP} + \mathrm{FP} + \mathrm{TN} + \mathrm{FN})} \quad (7)$$

As seen in the equation, this ratio is obtained by dividing the number of correct predictions by the total number of samples. This value is a value between 0 and 1 and is interpreted as a percentage.

Specificity rate: Also referred to as the true-negative rate, it demonstrates the effectiveness of estimating the degree to which the classifier correctly classifies the rejected one.

$$Specificity\ rate = \frac{(TN)}{(TN + FP)} \quad (8)$$

As seen in the equation, this ratio is obtained by dividing correctly rejected by the number of correctly rejected and incorrectly approved. This value is a value between 0 and 1 and is interpreted as a percentage.

Sensitivity rate: Also known as true positive rate, this metric indicates the classifier's effectiveness at estimating the extent to which it accurately classified the one that confirmed it correctly.

$$Sensitivity\ rate = \frac{(TP)}{(TP + FN)} \quad (9)$$

As seen in the equation, this ratio is obtained by dividing the number of correctly approved and incorrectly rejected. This value is a value between 0 and 1 and is interpreted as a percentage.

Precision ratio: This value is calculated by dividing the number of correctly predicted correct positive predictions by the number of samples of all positive predictions predicted. This value is a value between 0 and 1 and is interpreted as a percentage.

$$Precision\ rate = \frac{(TP)}{(TP + FP)} \qquad (10)$$

F1-Statistics: It is a measure obtained as a result of the harmonic average of the sensitivity and precision measures and it provides information about the classification success.

$$F1 - Statistics = 2x \frac{sensitivity\ x\ precision}{sensitivity + precision} \qquad (11)$$

AUC value (area under the ROC curve): It is a robust method for visualizing, organizing, selecting, and ranking classifiers based on their performance (Olson & Delen, 2008; Provost & Fawcett, 2001). The AUC value is a measure that indicates the model's accuracy. The AUC is the likelihood that a randomly selected positive sample will be rated higher than a randomly picked negative sample. The AUC value for a given model indicates the balance between the true positive rate and the false positive rate. (Han, Pei, & Kamber, 2011). The AUC value refers to the area under the ROC curve. The larger this area, the higher the classification success rate of the model.

The fact that the independent variables are highly correlated with one another creates a multicollinearity problem. If the tolerance value is less than 0.20 and the VIF value is more than 10, then there is a multi-connection problem (Büyüköztürk, 2011; Kalaycı, 2010; Özdamar, 2013). VIF and Tolerance values for continuous and categorical variables in the ABİDE 2016 data set were calculated. The tolerance values were found between the lowest 0.494 and the highest 0.991, and the VIF values between the lowest 1.009 and the highest 2.024. Based on these values, it was understood that there was no multicollinearity problem in the data set.

Findings

Comparison of Classification Performances of MARS and BRT Analysis Methods

A total of 10407 study groups, 8303 training data and 2104 test data, were analyzed. When students are classified as "successful / unsuccessful" in science achievement as a result of MARS analysis method, the number of students falling into these groups is given in Table 3.

Table 3. Confusion Matrix Obtained at Certain Levels as a Result of MARS Analysis

	Training data (%80)			Test data (%20)		
	Unsuccessful	**Successful**	**Total**	**Unsuccessful**	**Successful**	**Total**
Unsuccessful	2150	980	3130	559	237	796
Successful	1441	3732	5173	434	874	1308
Total Number of Students	3591	4712	8303	993	1111	2104

As seen in Table 3, the MARS analysis method used 80% of the data set as training data for modeling and 20% as test data for validity. MARS analysis method classified 3732 students in successful category and 1441 students in unsuccessful category out of 5173 students in the successful category for the training dataset. Out of 3130 students in the unsuccessful category, 2150 were classified in the unsuccessful category and 980 in the successful category. For the MARS analysis method test data set, out of 1308 students in the successful category, 874 were classified in the successful category and 434 in the unsuccessful category. Again, out of 796 students in the unsuccessful category, 559 students were classified in the unsuccessful category and 237 students in the successful category.

As a result of the BRT analysis method, the students were classified as "successful/unsuccessful" in the Science lesson, and the number of students in the relevant groups is given in Table 4.

Table 4. Confusion Matrix Obtained at Certain Levels as a Result of BRT Analysis

	Training data (%80)			Test data (%20)		
	Unsuccessful	**Successful**	**Total**	**Unsuccessful**	**Successful**	**Total**
Unsuccessful	2644	486	3130	556	240	796
Successful	922	4251	5173	379	929	1308
Total Number of Students	3566	4737	8303	935	1169	2104

Table 4 shows that 80% of the BRT analysis method data set is used as training data for modeling and 20% as test data for validity. The BRT analysis method classified 4251 of the 5173 students in the successful category for the training data set, and 922 in the unsuccessful category. Again, out of 3130 students in the unsuccessful category, 2644 were classified in the unsuccessful category and 486 in the successful category. For the BRT analysis method test data set, 929 of 1308 students in the successful category were classified in the successful category and 379 in the unsuccessful category. Again, out of 796 students in the unsuccessful category, 556 students were classified in the unsuccessful category and 240 students in the successful category.

The classification performances of MARS and BRT analysis methods, obtained by comparing the training and test data set with each other in the form of real class spacing and estimated class spacing, are given in Table 5.

Table 5. Classification Performance Rates at Certain Levels as a Result of MARS and BRT Analysis

	Training Data		Test Data	
Criteria	**MARS**	**BRT**	**MARS**	**BRT**
Correct Classification Rate	%70,84	%83,04	%68,11	%70,58
Specificity Rate	%68,69	%84,47	%70,23	%69,85
Sensitivity Rate	%72,14	%82,18	%66,82	%71,02
Precision Rate	%79,20	%89,74	%78,67	%79,47
F1 Statistics	%75,51	%85,79	%72,26	%75,01
AUC Value	%77,81	%91,17	%74,91	%78,20

The classification performances of both analytical approaches were compared over the test data values, since this provides evidence for the established models' validity. As shown in Table 5, the MARS analysis method achieves a correct classification rate of 68.11 %, while the BRT analysis method achieves a correct classification rate of 70.58 %. It is observed that the BRT analysis method performs better than the MARS analysis method in terms of proper classification. In other words, the BRT analysis method more accurately classified a successful student as successful and an unsuccessful student as unsuccessful.

As for Specificity Ratio, the MARS analysis method has a specificity rate of 70.23 %, while the BRT analysis method has a specificity rate of 69.85 %. As can be observed, the MARS method has a higher specificity rate than the BRT method. In other words, 70.23 % predicted to be unsuccessful by the MARS analysis method are actually unsuccessful. As a result, the MARS analysis method has a higher rate of true negative prediction.

In terms of sensitivity, the MARS analysis method achieves a sensitivity rate of 66.82 %, while the BRT analysis method reaches a sensitivity rate of 71.02 %. The BRT analysis method has a higher sensitivity rate than the MARS analysis method. In other words, 71.02 % predicted to be successful by the BRT analysis method are really successful. As a consequence, the BRT analysis method has a better rate of predicting the correct positives.

In terms of precision, the MARS analysis method achieves a precision rate of 78.67 %, while the BRT analysis method reaches a precision rate of 79.47 %. The BRT analysis method is shown to be more precise than the MARS analysis method. In other words, 79.47 percent of students classified as successful using the BRT analysis method were classified as such. Therefore, this value obtained by the BRT analysis method, which is higher than the MARS analysis method, gives the ratio of students who were predicted correctly among all those who were predicted successfully.

In terms of F1-statistics, the F1-statistic obtained with the MARS analysis method is 72.26%, while the F1-statistic obtained by the BRT analysis method is 75.01%. The BRT analysis method showed a higher classification success in terms of the F-1 statistical result, which is the harmonic mean of sensitivity and precision. In other words, the BRT analysis method achieved a higher rate in detecting successful students and distinguishing unsuccessful students compared to the MARS analysis method.

While the AUC value obtained with the MARS analysis method is 74.91%, the AUC value obtained with the BRT analysis method is 78.20%. In other words, the BRT analysis method minimized the false positive rate compared to the MARS analysis method, while increasing the true positive rate to a higher level. In other words, the BRT analysis method classified successful students in the successful category and unsuccessful students in the unsuccessful category with fewer errors compared to the MARS analysis method.

The findings above demonstrated that the MARS analysis method achieved a higher percentage in the specificity ratio criterion, and the BRT analysis method in all other criteria, and showed a more successful performance there.

The most important predictors of science achievement based on the MARS analysis method

MARS analysis was conducted in two stages. The first stage is to identify the maximum number of basic functions (BF). This step included determining the lowest Test MSE value to be given in the analysis using the basic function numbers entered in various trials, which was at least twice the number of independent variables (Statsoft, 2018). In cases where the data set is divided into training data and test data, the lowest point of the Test MSE value represents the maximum number of fundamental functions (Salford Systeam, 2018). The lowest Test MSE value obtained reached its maximum with 49 basic functions. The second is the stage of determining the number of basic functions that make up the most appropriate model. In other words, in the model that starts with the maximum number of basic functions, it is the stage of determining the number corresponding to the lowest Test MSE value. In other words, in the model that started with 49 basic functions, 39 basic functions corresponding to the lowest Test MSE value were used for the formation of the most suitable model. The regression equation of the MARS analysis method is formed as a result of multiplying the model coefficients of each of the 39 basic functions used while creating the most appropriate model. As can be seen in Table 6, the multiplication of each fundamental function by its coefficient gives its contribution to the model.

Table 6. Regression Equation for Optimal Model

Y = 0,706716	- 0,28298 * BF1 – 2,30291 * BF3 – 0,12088 * BF4
	+ 0,0137402 * BF5 + 0,547644 * BF6 + 0,187965 * BF7
	- 0,144839 * BF9 – 0,437399 * BF10 + 0,00353516 * BF11
	+ 0,124828 * BF12 – 0,0526214 * BF13 – 0,196745 * BF14
	+ 0,052752 * BF15 + 0,0182318 * BF17 + 0,312279 * BF18
	+ 0,0519219 * BF19 – 0,00717812 * BF22 – 9,30736 * BF23
	+ 0,0354042 * BF24 + 0,00337613 * BF25 + 0,00840481 * BF26
	- 0,0164451 * BF27 + 0,0821411 * BF29 – 0,438221 * BF31
	- 0,00454103 * BF32 + 0,0190105 * BF33 – 0,169167 * BF35
	+ 0,00907354 * BF36 + 0,0985396 * BF37 – 0,0628089 * BF39
	- 0,044178 * BF40 + 0,128554 * BF42 – 0,0497339 * BF43
	- 0,0182725 * BF44 + 0,0359945 * BF45 + 0,0665624 * BF46
	+ 0,0706328 * BF47 + 0,00407991 * BF48 + 0,0233424 * BF49

The F value (60.02392) and p value ($p < 0.01$) obtained for 3 degrees of freedom in this established equation show that the established model is significant. As a result of the analysis of the data obtained with the ABIDE application with the MARS method, the predictive variables included in the analysis and the importance levels of these variables on science achievement in the established model are listed in Table 7, starting with 100 points.

Table 7. MARS Analysis Method Table of Significance Levels of Variables

Predictive Variables	Score
Self-Efficacy Perception	100,00
Father's Profession	87,50
Monthly Income	87,28
Student-Oriented School Climate Perception of Administrators	82,73
Parent Approach	73,72
Peer Bullying	69,99
Teacher's Professional Year	56,12
Teacher's Instructional Activities	51,63
Mother’s Profession	48,50
Value Given to the Course	38,62
Enjoyment	35,09
Sense of Belonging to school	31,04
Teacher's Preparation for Lesson	22,37
Teacher-Oriented School Climate Perception of Administrators	0,00
General Instructional Activities	0,00
Self-Efficacy	0,00
Residence	0,00
Teacher's Gender	0,00
Level of education	0,00
Professional development	0,00
Professional Satisfaction	0,00
Type of School Graduated	0,00
Comprehensive Instructional Activities	0,00

As seen in Table 7, the most important predictors of science achievement were obtained as self-efficacy perception, father's profession, monthly income, administrators' perception of student-oriented school climate, parental approach, bullying, teacher's professional year, teacher's instructional activities, mother's profession, value given to the course, enjoyment, sense of belonging to school and teacher's preparation for the lesson, respectively. The MARS analysis method assigned a significance level to each variable associated with the dependent variable, starting at 100 points. It was discovered that the variable that most correctly predicted science achievement was self-efficacy perception, whereas the variable that predicted science achievement least accurately was teacher preparation for the session. The variables of administrators' perception of teacher-oriented school climate, general instructional activities, self-efficacy, residence, gender of the teacher, education level, professional development, professional satisfaction, type of school from which he graduated, and comprehensive instructional activities either did not contribute or contributed very little to the prediction of students' science achievement.

The most important predictors of science achievement based on the BRT analysis method

The first important point to consider in order to establish the most appropriate model in the BRT analysis method is to determine the maximum number of trees. In this direction, the sufficient number of trees was tried to be calculated by entering the tree numbers of 200, 500, 750, 1000, 2500 and 5000 as seen in Table 8.

Table 8. BRT Analysis Results to Determine the Maximum Number of Trees

	By	Neg. AvgLL	ROC	Misclass	Lift
200	Measurement	0,54437	0,77791	0,27186	1,50917
	Number of Trees	200	198	113	63
500	Measurement	0,54096	0,78061	0,26331	1,50917
	Number of Trees	344	499	363	63
750	Measurement	0,54015	0,78278	0,26331	1,50917
	Number of Trees	541	702	363	63
1000	Measurement	0,54015	0,78278	0,26331	1,50917
	Number of Trees	541	702	363	63
2500	Measurement	0,54015	0,78278	0,26331	1,50917
	Number of Trees	541	702	363	63
5000	Measurement	0,54015	0,78278	0,26331	1,51682
	Number of Trees	541	702	363	3.438

As shown in Table 8, the findings remain consistent across all analyses conducted once the number of trees in the formation of the most suitable model reaches 702. As a result, the needed number of trees was determined to be 702. Table 9 lists the predictor variables included in the study and their significance levels for science achievement in the established model, beginning with 100 points.

Table 9. BRT Analysis Method Table of Significance Level of Variables

Variables	Score
Self-Efficacy Perception	100,00
Monthly Income	89,31
Father's Profession	88,41
Teacher's Instructional Activities	88,21
Parent Approach	86,09
Student-Oriented School Climate Perception of Administrators	85,76
Peer Bullying	81,68
Enjoyment	76,80
Sense of Belonging to school	73,61
Teacher-Oriented School Climate Perception of Administrators	71,47
Teacher's Professional Year	68,44
General Instructional Activities	66,14
Professional development	63,82
Value Given to the Course	58,81
Mother Profession	58,03
Teacher's Preparation for Lesson	56,93
Comprehensive Instructional Activities	53,22
Self-Efficacy	49,87
Professional Satisfaction	48,34
Residence	30,98
Type of School Graduated	23,36
Teacher's Gender	21,57
Level of education	18,85

The most significant predictors and their associated significance levels as determined by the BRT analysis method are shown in Table 9. Accordingly, the most important predictors of science achievement were determined as self-efficacy perception, monthly income, father's occupation, teacher's instructional activities, parental approach, administrators' perception of student-oriented school climate, bullying, enjoyment, sense of belonging to school, teacher-oriented school climate perception of administrators, and teacher's professional year, respectively. Similar to the BRT analysis method, it was observed that the variable that contributed the most in predicting science achievement in the MARS analysis method was the perception of self-efficacy.

Discussion, Conclusion and Suggestions

The relationships between various school, teacher and student qualities, which are thought to be related to the science success of 8th grade students in science in ABIDE 2016, and science achievement were examined with MARS and BRT analysis methods in line with the theoretical framework based on the research, and these two data mining methods were compared in terms of classifying students in terms of science achievement.

There are few studies in the literature (Akıncı, 2020; Çalık, 2020; Doğru, 2019; Elkonca, 2020; Kılıç, 2019; Özgürlük, 2019; Uysal, 2019; Ülkü, 2019) using educational data such as the ABIDE application. Again, few studies were found in education where the MARS analysis method was used (Gocheva-Ilieva, Kulina, & Ivanov, 2021; Kayri, 2010; Oğuz, 2014; Yu, Digangi, Jannasch-Pennell, & Kaprolet, 2008). Just a few studies were found in which the other method, the BRT analysis method, was used (Mazman, 2013; Stearns et al. 2017; Stone & Tang, 2013; Sinharay, 2016).

Gocheva-Ilieva et al. (2021) examined the factors associated with students' mathematics achievement with MARS, CART and CART-EB analysis methods and stated that the MARS analysis method outperformed them. Kayri (2010) examined students' internet addictions with CART and MARS analysis methods and stated that MARS obtained different findings from CART in estimating addiction level and was more efficient in model estimation. Oğuz (2014) examined the factors related to the academic success of university students with the MARS analysis method. Yu et al. (2008) examined the interest of young and adult students taking online education in terms of various variables with the MARS analysis method and stated that young students prefer online education more, unlike adults who suffer from factors such as workload and marriage and children. Mazman (2013) tried to predict the factors related to the programming performance of CEIT students with BRT and RF analysis methods and stated that the variables found to be significant were the same in both analysis methods, but the importance levels of the factors found were different because the algorithms used by both methods were different. In this study, it was seen that the BRT analysis method gave better results in terms of classification performance and the MARS analysis method was more successful in terms of estimation ability. Both analysis methods are discussed in two dimensions in terms of the results obtained.

In the first dimension, both analysis methods were compared one by one in terms of their classification performance. The results of the research, in terms of specificity rate, according to the BRT analysis method of the MARS analysis method; showed that the BRT analysis method had a more successful classification performance than the MARS analysis method in terms of correct classification rate, sensitivity rate, precision rate, F1 statistical value and AUC value. Instead of producing only a single model as in standard regression analyzes, the BRT analysis method increased its predictive performance by combining multiple models (Hill & Lewicki, 2006). Sevimli-Saitoğlu (2015) compared the MARS method with the C&RT method in terms of classification performance and reported that the MARS method was more successful. In contrast to the C&RT method, which produces a single model, the BRT analysis method, which is based on the combination of multiple C&RT models, was more successful against the MARS analysis method in this study. Again, Elish and Elish (2009) and Mukkamala, Xu, and Sung (2006) compared the BRT analysis method with MARS and other data mining methods and stated that the BRT analysis method was more successful in terms of classification performance.

Both of the existing data mining methods have shown different levels of performance due to the difference in the algorithms they use in the background. While the BRT analysis method uses the boosting technique on the tree-based classification and regression tree (C&RT) method, the MARS analysis method uses the smoothing technique to create partial linearities. The fact that the BRT method uses the combination of many interconnected regression trees to learn from errors has provided a stronger classification performance by minimizing the errors made in the classification. This showed that the BRT analysis method had a higher classification validity in this study.

The second dimension examined variables affecting scientific accomplishment at the student, teacher-class, and school levels within the scope of the dynamic model of educational effectiveness. Self-efficacy perceptions, father's occupation, monthly income, parental approach, peer bullying, mother's occupation, value placed on science lessons, enjoyment and belonging to school variables; teacher's instructional activities, teacher's professional year, and teacher's lesson preparation variables; the administrators' student-centered school climate perception variables were obtained as important predictors at student, grade and school level, respectively in in both analysis methods.

The perception of self-efficacy for the science course was revealed to be the variable with the highest score in both analysis methods based on the predictive significance level. Juan, Hannan, and Namome (2018) found a significant relationship between students' science achievement and their self-efficacy beliefs using TIMSS-2015 data. According to Bandura (1995), students' self-belief, patience and not giving up in the face of negativities are important for their self-efficacy perceptions in order to be successful. Pajares (1996) stated that individuals with high self-efficacy perceptions make a great effort to be successful, do not take a step back in the face of negativities and act patiently. Again, other investigations (Acar & Öğretmen, 2012; Doğan & Barış, 2010; Juan et al., 2018; Sarı, Arıkan, & Yıldızlı 2017; Yazıcı, Seyis, & Altun, 2011) revealed similar findings to those in the present study. In light of this information, the finding that students' self-efficacy perceptions of science courses were the most predictive of their science achievement in ABIDE 2016 implies that students do indeed have a realistic perception of their science success.

After the self-efficacy perception, the qualities that contributed the most to predict science achievement at the student level revealed to be monthly income, father's occupation, parental approach, peer bullying, value given to science lesson, enjoyment and sense of belonging to school, respectively. It has been observed that similar results have been reported in the related literature (Ainley & Ainley, 2011; Andreou, 2000; Austin & Joseph, 1996; Ferguson, 2006; Jeynes, 2005; Juan et al., 2018; Okutan, 2017; Önen, 2018; Pajares, 1996; Young, 1998). In this respect, it is understood that economic status such as monthly income and father's occupation, social status such as parental approach and peer bullying, and affective characteristics such as the value given to science lesson, enjoyment of science lesson and sense of belonging to school contribute significantly to the prediction of students' science achievement. Consistent with these findings, Okatan and Tomul (2020) examined the effects of various economic, social and affective variables on students' mathematics achievement on PISA data. It was concluded that ESCS (economic, social and cultural status) index variable economically, mother education level variable socially, and mathematics self-efficacy variable affectively are more effective than other variables.

The analyses revealed that the variables of monthly income and father's profession in the current study contributed significantly to predicting science achievement in terms of economic status. Abacı (2015) and Okutan (2017) stated in their research that there is a positive and significant relationship between the monthly income variable of the family and the science achievement of 8th grade students who participated in the TEOG exam. Karar (2011) and Yolagiden and Bektaş (2018) stated that there is a significant relationship between the academic achievement of 8th grade students and their father's profession. İpek (2011) stated in his study on SBS data that students' SBS scores differ depending on their fathers' profession. Finally, Young (1998) reported in his study that the father's profession contributed significantly to the explanation of student achievement. Parents with a good job and high income in terms of economic situation can offer better opportunities for their children within their means, so the amount of expenditure per student can increase. It is thought that this will indirectly increase the success in education.

In terms of social status, it was observed that the variables of parental approach and peer bullying contributed significantly to explaining the interindividual differences in science achievement. Jeynes (2005), in his study, stated that the parent approach contributed significantly to the academic achievement of primary school students. Similarly, Ferguson (2006) and Ma, Shek, Cheung, and Lam (2000) stated that students who have good relationships with their peers and parents have high academic achievement. Likewise, many studies have reported that students who have bad relationships with their peers and who are exposed to peer bullying have lower academic achievement than students who are not exposed to such bullying (Kartal & Bilgin, 2009; Kochenderfer & Ladd, 1996; Juvonen, Nishina, & Graham, 2000; Winnaar, Arends, & Beku, 2018). Similarly, the academic achievement of students who bullied their peers was lower than those who did not show bullying (Andreou, 2000; Austin & Joseph, 1996). On the other hand, Önen (2018) stated in his study on the TIMSS 2015 application that students with low achievement levels are exposed to peer bullying more than students with high achievement levels. It can be said that the children of families who can understand their child, communicate with him, trust him and most importantly express this at every opportunity and are aware of the necessary responsibility for their child are more successful. Again, in terms of social status, the student's exposure to peer bullying causes him to face many negative situations such as developing a negative attitude towards school, not going to school, experiencing anxiety and loneliness, getting sick and losing his motivation. As a result, the academic success of the student may be affected directly or indirectly.

It was observed that the value given to science lesson in terms of affective characteristics, the enjoyment of science lesson and the importance of belonging to school variables significantly contribute to predicting science achievement. In his study on the Turkish sample of the PISA 2006 application, Çalışkan (2008) stated that the value students give to science lesson is related to science achievement. Similarly, Ceylan and Berberoğlu (2007) stated in their study that there is a positive relationship between the value given to the science lesson and the science achievement of the students. Anıl (2009), in his study on the PISA 2006 application, stated that students who enjoy science lessons develop a positive attitude towards the lesson. Ainley and Ainley (2011) stated that students' enjoyment of science and their interest in learning science come from previous science learning experiences. Abdollahi and Noltemeyer (2018) stated that there is a positive relationship between belonging to school and academic achievement. Goodenow (1993) with Winnaar et al. (2018) stated that belonging to the school is important in the development of academic success. From the affective point of view, the existence of the value attributed to the science lesson can be observed by exhibiting many behaviors. Such supporting behaviors can increase the value given to the science lesson and thus contribute positively to the increase in success. Again, interest in and enjoyment of the science lesson can create an opportunity for students with low science achievement to experience success. Therefore, teachers, guidance specialists and administrators in schools should pay attention to ensure that the social environment of the school is in a way that will increase the students' sense of belonging to the school and strengthen their emotional bond with the school, and that it is designed as environments where students are more valued and accepted. Moreover, they should take the necessary measures, directly or indirectly, to help students feel safe, accepted, and respect themselves and their environment.

In terms of the teacher qualifications considered at the grade level, it is recognized that the teacher's instructional activities, professional year, and lesson preparation all contribute greatly to predicting students' science success. According to Bloom (2012), the teacher's instructional activities in the learning-teaching environment should inform students about the course's objectives and content, give essential feedback and assistance, and encourage student participation in the course via on-site reinforcement. Such teacher practices also have an effect on student achievement. Ceylan and Berberoğlu (2007) discovered a positive and significant relationship between science success and teacher instructional activities in their research using data obtained from Turkish students in the TIMSS-1999 application. Again, Akyüz (2006) claimed in his research investigating the relationship between teacher and classroom qualities and student achievement in Turkey and European Union nations using TIMSS data that teachers' instructional activities greatly contributed to student success in a number of countries. Indeed, the variables expressing the teacher's teaching activities and the years spent in the profession are the variables indicating the teacher's teaching experience. As a result of the examination of the relations between the variables discussed in the research, it is understood that the students of experienced teachers are more successful than those of the teachers who are in the first years of the profession. Martin, Mullis, Foy, and Stanco (2012), who prepared the TIMSS 2011 international scientific report, stated that students with more experienced and confident teachers in the fourth and eighth grades have higher science achievement. As a result of all these inferences, it is understood that pupils of experienced instructors achieve more success than those of inexperienced ones. Martin et al. (2012), authors of the TIMSS 2011 international scientific report, claimed that children in the fourth and eighth grades who have more experienced and confident teachers achieve higher science achievement. As a result of these inferences, it is possible to conclude that the science achievement of students with inexperienced teachers is low due to both the process of adapting to the profession and the wide range of problems encountered during the early stages of the profession, whereas the science achievement of students with experienced teachers is high due to the time and experience gained. Again, regardless of grade level, the teacher's emotional, intellectual, and technical preparedness prior to the lesson will lead to increased student success and achievement of the lesson's intended objectives. Teachers should allot sufficient time prior to the lesson to prepare for it, carefully plan the learning-teaching process, design materials that are appropriate for the students' intellectual and affective readiness levels for the subject being covered, and inform their students in advance about any materials they will need to bring for the lesson. Teachers must recognize that their pupils cannot comprehend a topic for which they have not been appropriately prepared intellectually, emotionally, and technically prior to the session.

As for the significant predictors of science achievement at school level, based on both MARS and BRT analysis methods, it is seen that the student-centered school climate perception of the administrators is the variable with high predictive importance. On the TIMSS 2003 application, Chen, Lin, Wang, Lin, and Kao (2012) found a significant relationship between 4th and 8th grade students' school climate perceptions and their science achievement. Again, Bahçetepe (2013), in his study on secondary school 8th grade students, concluded that with the increase in students' positive perceptions about the school climate, their success also increased. It should be focused on creating a school environment that encourages self-confidence, excitement, mentoring, belonging and success through programs aimed at improving students' positive perceptions of the school and security measures to be taken (Plucker, 2010). As a matter of fact, simple but positive actions such as acting together with students and building mutual trust with students produce more permanent results, rather than employing officers with detectors and installing security cameras in order to create a safe school climate perception in schools (Bracey, 2011). Therefore, it can be concluded that administrators who create safe environments, support success and are open to communication contribute significantly to students' science success in particular and academic success in general.

The findings of this study discuss the applicability of MARS and BRT analysis methods, both of which are data mining methods, on data gathered from national and international large-scale tests applied in Turkish education. Data mining methods, which are extensively utilized in a variety of fields like economics, health, engineering, and banking, are critical in education, particularly when it comes to assessing the data obtained on students, teachers, and schools in large-scale exams employed in this field. Data mining methods that identify patterns from raw data and generate predictions are considered critical in education because they enable the processing and interpretation of thousands of data gathered from hundreds of students. Magdin and Turcani (2015) claimed that data mining can be used to identify factors that influence learning and academic achievement in education, to gain a better knowledge of the learning process, and to provide teachers with more objective feedback. As indicated before, it is believed that using data mining methods would aid in the study of big data sets and numerous complicated patterns discovered via large-scale testing and other educational applications. While MARS and BRT data mining methods are not as widely used or accessible as other data mining methods, they may be analyzed utilizing paid package programs, instructional applications, and a variety of package programs that allow free R-based open access.

This research is limited to the Science Achievement Test items administered to students participating in the ABIDE 2016 study and data gathered through student, teacher, and school administrator questionnaires. Again, data mining methods are limited to MARS and BRT analysis methods.

The study indicated that the most critical variable in both analysis methods is self-efficacy perception. The student's perception of themselves as successful or unsuccessful effects his actual success or failure. The student experiences a kind of learned helplessness. Within the Ministry of National Education, initiatives should be established to boost students' self-efficacy perceptions, and any activity that contributes to this goal should be promoted. It has also been observed that variables related to the family economy, such as father's occupation and monthly income, significantly contribute to the level of importance in predicting science achievement. In economic terms, the increase in the amount of expenditure per student is a factor affecting success. The study indicated that the most critical variable in both analysis methods is self-efficacy perception. The student's perception of themselves as successful or unsuccessful effects his actual success or failure. The student experiences a kind of learned helplessness. Within the Ministry of National Education, initiatives should be established to boost students' self-efficacy perceptions, and any activity that contributes to this goal should be promoted. Additionally, characteristics relating to the family economics, such as the father's work and monthly income, have been shown to contribute considerably to the degree of relevance in predicting scientific success. Economically speaking, the rise in spending per student is a factor influencing success. The studies that will be conducted in order to regulate education policy based on this data and to ensure equal opportunity will contribute positively to student achievement.

It is thought that the research findings provide important information (regarding student, teacher and school dimension) for educators and politicians in terms of revealing how much student, teacher and school qualifications are related to students' science achievement. Taking into account all of the findings, it can be concluded that the MARS and BRT analysis methods perform admirably in terms of elucidating the strongest predictors of science achievement and their usefulness in classifying students' achievement. The results of the research are guiding that MARS and BRT analysis methods, which are data mining methods, can be used for researchers who want to reveal the relationships between variables and make classification based on a huge data set.

References

Abacı, Ç. Ç. (2015). *Merkezi sistem ortak sınavlarının farklı değişkenler açısından irdelenmesi* (Unpublished master's thesis). Gazi University, Ankara.

Abdollahi, A., & Noltemeyer, A. (2018). Academic hardiness: Mediator between sense of belonging to school and academic achievement?. *The Journal of Educational Research, 111*(3), 345-351. doi:10.1080/00220671.2016.1261075

ABİDE 8. Sınıflar Raporu. (2017). *Akademik becerilerin izlenmesi ve değerlendirilmesi 8. sınıflar raporu.* Ankara: Milli Eğitim Bakanlığı Ölçme Değerlendirme ve Sınav Hizmetleri Genel Müdürlüğü. Retrieved from https://odsgm.meb.gov.tr/meb_iys_dosyalar/2017_11/30114819_iY-web-v6.pdf

Acar, T., & Öğretmen, T. (2012). Çok düzeyli istatistiksel yöntemler ile 2006 PISA fen bilimleri performansının incelenmesi. *Eğitim ve Bilim, 37*(163). Retrieved from http://egitimvebilim.ted.org.tr/index.php/EB/article/view/1040

Ainley, M., & Ainley, J. (2011). Student engagement with science in early adolescence: The contribution of enjoyment to students' continuing interest in learning about science. *Contemporary Educational Psychology, 36*(1), 4-12. doi:10.1016/j.cedpsych.2010.08.001

Akçapınar, G. (2014). *Çevrimiçi öğrenme ortamındaki etkileşim verilerine göre öğrencilerin akademik performanslarının veri madenciliği yaklaşımı ile modellenmesi* (Unpublished doctoral dissertation). Hacettepe University, Ankara.

Akıncı, B. (2020). *Fen bilimleri dersi öğretim programı ve ölçme değerlendirme araçlarının akademik becerilerin izlenmesi ve değerlendirilmesine (ABİDE) göre incelenmesi* (Unpublished master's thesis). Ankara University, Ankara.

Aksu, G. (2018). *PISA başarısını tahmin etmede kullanılan veri madenciliği yöntemlerinin incelenmesi* (Unpublished doctoral dissertation). Hacettepe University, Ankara.

Akyüz, G. (2006). Türkiye ve Avrupa Birliği ülkelerinde öğretmen ve sınıf niteliklerinin matematik başarısına etkisinin incelenmesi. *İlköğretim Online, 5*(2), 75-86. Retrieved from http://ilkogretim-online.org.tr

Al-Saleem, M., Al-Kathiry, N., Al-Osimi, S., & Badr, G. (2015). Mining educational data to predict students' academic performance. In *International workshop on machine learning and data mining in pattern recognition* (pp. 403-414). Cham: Springer.

Amrieh, E. A., Hamtini, T., & Aljarah, I. (2016). Mining educational data to predict student's academic performance using ensemble methods. *International Journal of Database Theory and Application, 9*(8), 119-136.

Andreou, E. (2000). Bully/victim problems and their association with psychological constructs in 8-to 12-year-old Greek schoolchildren. *Aggressive Behavior: Official Journal of the International Society for Research on Aggression, 26*(1), 49-56. doi:10.1002/(SICI)1098-2337(2000)26:1<49::AID-AB4>3.0.CO;2-M

Anıl, D. (2009). Uluslararası öğrenci başarılarını değerlendirme programı (PISA)'nda Türkiye'deki öğrencilerin fen bilimleri başarılarını etkileyen faktörler. *Eğitim ve Bilim, 34*(152), 87-100. Retrieved from http://egitimvebilim.ted.org.tr/index.php/EB/article/view/594/74

Austin, S., & Joseph, S. (1996). Assessment of bully/victim problems in 8 to 11 year-olds. *British Journal of Educational Psychology, 66*(4), 447-456. doi:10.1111/j.2044-8279.1996.tb01211.x

Azigwe, J. B. (2016). Using comparative international studies for modeling educational effectiveness: A secondary analysis of PISA-2009 study. *Journal of Education and Practice, 7*(18), 199-209. Retrieved from https://eric.ed.gov/?id=EJ1105875

Bahçetepe, Ü. (2013). *İlköğretim sekizinci sınıf öğrencilerinin akademik başarıları ile algıladıkları okul iklimi arasındaki ilişki* (Unpublished master's thesis). İstanbul University, İstanbul.

Bandura, A. (Ed.). (1995). *Self-efficacy in changing societies.* Cambridge: Cambridge University Press.

Baradwaj, B. K., & Pal, S. (2011). Mining educational data to analyze students' performance. *International Journal of Advanced Computer Science and Applications, 2*(16), 63-69.

Bilen, Ö., Hotaman, D., Aşkın, Ö. E., & Büyüklü, A. H. (2014). LYS başarılarına göre okul performanslarının eğitsel veri madenciliği teknikleriyle incelenmesi: 2011 İstanbul örneği. *Eğitim ve Bilim, 39*(172), 78-94. Retrieved from http://egitimvebilim.ted.org.tr/index.php/EB/article/view/2733/627

Bloom, B. S. (2012). *İnsan nitelikleri ve okulda öğrenme* (2nd ed., D. A. Özçelik, Trans.). Ankara: Pegem Akademi.

Blömeke, S., Olsen, R. V., & Suhl, U. (2016). Relation of student achievement to the quality of their teachers and instructional quality. In T. Nilsen & J. E. Gustafsson (Eds.), *Teacher quality, instructional quality and student outcomes* (Vol. 2, pp. 21-50). Cham: Springer. doi:10.1007/978-3-319-41252-8_2

Bracey, N. (2011). Student perceptions of high-security environments. *Youth & Society, 43*(1), 365-395. doi:10.1177/0044118X10365082

Büyüköztürk, Ş. (2011). *Sosyal bilimler için veri analizi el kitabı* (14th ed.). Ankara: Pegem Akademi.

Ceylan, E., & Berberoğlu, G. (2007). Öğrencilerin fen başarısını açıklayan etmenler: Bir modelleme çalışması. *Eğitim ve Bilim, 32*(144), 36-48. Retrieved from http://egitimvebilim.ted.org.tr/index.php/EB/article/view/821

Chen, S. F., Lin, C. Y., Wang, J. R., Lin, S. W., & Kao, H. L. (2012). A cross-grade comparison to examine the context effect on the relationships among family resources, school climate, learning participation, science attitude, and science achievement based on TIMSS 2003 in Taiwan. *International Journal of Science Education, 34*(14), 2089-2106. doi:10.1080/09500693.2012.701352

Colin, B., Clifford, S., Wu, P. P., Rathmanner, S., & Mengersen, K. (2017). Using boosted regression trees and remotely sensed data to drive decision-making. *Open Journal of Statistics, 7*(5), 859-875. doi:10.4236/ojs.2017.75061

Creemers, B., & Kyriakides, L. (2007). *The dynamics of educational effectiveness: A contribution to policy, practice and theory in contemporary schools.* London: Routledge.

Çalık, G. (2020). *Investigation of 8th grade students' science achievement in Turkey: Results from monitoring and evaluating academic skills study (ABIDE) 2016* (Unpublished master's thesis). Middle East Technical University, Ankara.

Çalışkan, M. (2008). *The impact of school and student related factors on scientific literacy skills in the programme for international student assessment PISA 2006* (Unpublished doctoral dissertation). Middle East Technical University, Ankara.

Deconinck, E., Xu, Q. S., Put, R., Coomans, D., Massart, D. L., & Vander Heyden, Y. (2005). Prediction of gastro-intestinal absorption using multivariate adaptive regression splines. *Journal of Pharmaceutical and Biomedical Analysis, 39*(5), 1021-1030. doi:10.1016/j.jpba.2005.05.034

Deichmann, J., Eshghi, A., Haughton, D., Sayek, S., & Teebagy, N. (2002). Application of multiple adaptive regression splines (MARS) in direct response modeling. *Journal of Interactive Marketing, 16*(4), 15-27. doi:10.1002/dir.10040

Doğan, N., & Barış, F. (2010). Tutum, değer ve özyeterlik değişkenlerinin TIMSS-1999 ve TIMSS-2007 sınavlarında öğrencilerin matematik başarılarını yordama düzeyleri. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi, 1*(1), 44-50. Retrieved from https://dergipark.org.tr/tr/download/article-file/65998

Doğru, Ş. C. (2019). *Karma testlerin psikometrik özelliklerini belirlemede klasik test kuramı ve Rasch modelinin karşılaştırılması* (Unpublished doctoral dissertation). Hacettepe University, Ankara.

Dobert, H., & Sroka, W. (Eds.). (2004). *Features of successful school systems: A comparison of schooling in six countries* (Vol 4.). Almanya: Waxmann Verlag.

EACEA. (2009). *Avrupa'da öğrencilerin ulusal ölçümü: Hedefler organizasyon ve sonuçların kullanılması*. Retrieved from http://sgb.meb.gov.tr/eurydice/kitaplar/Avrupada_ogrencilerin_Ulusal_olcumu/Avrupada_ogrencilerin_Ulusal_olcumu.pdf

Elish, M. O., & Elish, K. O. (2009, March). Application of treenet in predicting object-oriented software maintainability: A comparative study. In *2009 13th European Conference on Software Maintenance and Reengineering* (pp. 69-78). New Jersey, ABD: IEEE. doi:10.1109/CSMR.2009.57

Elith, J., Leathwick, J. R., & Hastie, T. (2008). A working guide to boosted regression trees. *Journal of Animal Ecology, 77*(4), 802-813. doi:10.1111/j.1365-2656.2008.01390.x

Elkonca, F. (2020). *ABİDE özyeterlik ölçeği DMF kaynaklarının gizil sınıf yaklaşımıyla incelenmesi* (Unpublished doctoral dissertation). Gazi University, Ankara.

Erkuş, A. (2003). *Psikometri üzerine yazılar.* Ankara: Türk Psikologlar Dernegi Yayınları.

Ferguson, K. M. (2006). Social capital predictors of children's school status in Mexico. *International Journal of Social Welfare, 15*(4), 321-331. doi:10.1111/j.1468-2397.2006.00422.x

Ferreruela, I. C. (2008). Explaining patterns of broadband development in OECD countries. In *Handbook of research on global diffusion of broadband data transmission* (pp. 756-775). Pensilvanya, ABD: IGI Global.

Friedman, J. H. (1991). Invited paper multivariate adaptive regression splines. *The Annals of Statistics, 19*(1), 1-67.

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics, 29*(5), 1189-1232. Retrieved from https://www.jstor.org/stable/2699986

Friedman, J. H. (2002). Stochastic gradient boosting. *Computational Statistics & Data Analysis, 38*(4), 367-378. doi:10.1016/S0167-9473(01)00065-2

Friedman, J. H., & Meulman, J. J. (2003). Multiple additive regression trees with application in epidemiology. *Statistics in Medicine, 22*(9), 1365-1381. doi:10.1002/sim.1501

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

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction.* Berlin, Almanya: Springer Science & Business Media.

Hill, T., & Lewicki, P. (2006). *Statistics: Methods and applications: A comprehensive reference for science, industry, and data mining.* Tulsa OK, Oklahoma: StatSoft, Inc.

Gocheva-Ilieva, S., Kulina, H., & Ivanov, A. (2021). Assessment of students' achievements and competencies in mathematics using CART and CART ensembles and bagging with combined model improvement by MARS. *Mathematics, 9*(1), 62. doi:10.3390/math9010062

Goodenow, C. (1993). Classroom belonging among early adolescent students: Relationships to motivation and achievement. *The Journal of Early Adolescence, 13*(1), 21-43. doi:10.1177/0272431693013001002

Gustafsson, J. E., & Nilsen, T. (2016). The impact of school climate and teacher quality on mathematics achievement: A difference-in-differences approach. In T. Nilsen & J. E. Gustafsson (Eds.), *Teacher quality, instructional quality and student outcomes* (pp. 81-95). Cham: Springer.

İpek, C. (2011). Velilerin okul tutumu ve eğitime katılım düzeyleri ile aileye bağlı bazı faktörlerin ilköğretim öğrencilerinin Seviye Belirleme Sınavları (SBS) üzerindeki etkisi. *Pegem Eğitim ve Öğretim Dergisi, 1*(2), 69-79.

Jeynes, W. H. (2005). A meta-analysis of the relation of parental involvement to urban elementary school student academic achievement. *Urban Education, 40*(3), 237-269. doi:10.1177/0042085905274540

Juan, A., Hannan, S., & Namome, C. (2018). I believe I can do science: Self-efficacy and science achievement of Grade 9 students in South Africa. *South African Journal of Science, 114*(7-8), 48-54. doi:10.17159/sajs.2018/20170269

Juvonen, J., Nishina, A., & Graham, S. (2000). Peer harassment, psychological adjustment, and school functioning in early adolescence. *Journal of Educational Psychology, 92*(2), 349-359. doi:10.1037/0022-0663.92.2.349

Kalaycı, Ş. (2010). *SPSS uygulamalı çok değişkenli istatistik teknikleri* (5th ed.). Ankara: Asil Yayın Dağıtım.

Karar, E. E. (2011). *İlköğretim 8. sınıf öğrencilerinin bilimsel süreç becerilerinin bazı değişkenler açısından incelenmesi* (Unpublished master's thesis). Adnan Menderes University, Aydın.

Karasar, N. (2009). *Bilimsel araştırma yöntemi* (23th ed.). Ankara: Nobel Yayınları.

Kartal, H., & Bilgin, A. (2009). Bullying and school climate from the aspects of the students and teachers. *Eurasian Journal of Educational Research (EJER), 36*(36), 209-226.

Kayri, M. (2009). Çok değişkenli uyarlanabilir regresyon uzanımları (multivariate adaptive regression splines) yönteminin yansız ve sapmasız ölçme işlemlerindeki etkililiği: Bir uygulama örneği. In *XVIII. Ulusal Eğitim Bilimleri Kurultayı* (pp. 123-132). İzmir.

Kayri, M. (2010). The analysis of internet addiction scale using multivariate adaptive regression splines. *Iranian Journal of Public Health, 39*(4), 51-63. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3481689/

Kılıç, A. F. (2019). *Karma testlerde doğrulayıcı faktör analizi kestirim yöntemlerinin karşılaştırılması* (Unpublished doctoral dissertation). Hacettepe University, Ankara.

Kochenderfer, B. J., & Ladd, G. W. (1996). Peer victimization: Cause or consequence of school maladjustment?. *Child Development, 67*(4), 1305-1317. doi:10.1111/j.1467-8624.1996.tb01797.x

Ma, H. K., Shek, D. T., Cheung, P. C., & Lam, C. O. B. (2000). Parental, peer, and teacher influences on the social behavior of Hong Kong Chinese adolescents. *The Journal of Genetic Psychology, 161*(1), 65-78. doi:10.1080/00221320009596695

Magdin, M., & Turcáni, M. (2015). Personalization of student in course management systems on the basis using method of data mining. *Turkish Online Journal of Educational Technology-TOJET, 14*(1), 58-67. Retrieved from https://eric.ed.gov/?id=EJ1057349

Martin, M. O., Mullis, I. V., Foy, P., & Stanco, G. M. (2012). *TIMSS 2011 international results in science.* Chestnut Hill: TIMSS & PIRLS International Study Center.

Mazman, S. G. (2013). P*rogramlama performansını etkileyen faktörlerin bilişsel tabanlı bireysel farklılıklar temelinde modellenmesi* (Unpublished doctoral dissertation). Hacettepe University, Ankara.

Mukkamala, S., Xu, D., & Sung, A. H. (2006). Intrusion detection based on behavior mining and machine learning techniques. In *International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems* (pp. 619-628). Berlin, Heidelberg: Springer.

Nilsen, T., & Gustafsson, J. E. (2016). *Teacher quality, instructional quality and student outcomes: Relationships across countries, cohorts and time*. Cham: Springer.

Nilsen, T., Gustafsson, J. E., & Blömeke, S. (2016). Conceptual framework and methodology of this report. In T. Nilsen & J. E. Gustafsson (Eds.), *Teacher quality, instructional quality and student outcomes* (pp. 1-19). Cham: Springer.

Nortvedt, G. A., Gustafsson, J. E., & Lehre, A. C. W. (2016). The importance of instructional quality for the relation between achievement in reading and mathematics. In T. Nilsen & J. E. Gustafsson (Eds.), *Teacher quality, instructional quality and student outcomes* (pp. 97-113). Cham: Springer.

Oğuz, A. (2014). *Çok değişkenli uyarlanabilir regresyon zincirlerinin irdelenmesi ve bir uygulama* (Unpublished master's thesis). Erzincan University, Erzincan.

Okatan, Ö., & Tomul E. (2020). Uluslararası öğrenci başarılarını değerlendirme programı'na (PISA) göre Türkiye'deki öğrencilerin matematik başarıları ile ilişkili değişkenlerin incelenmesi. *Mehmet Akif Ersoy Üniversitesi Eğitim Fakültesi Dergisi,* 57, 98-125. doi:10.21764/maeuefd.663150

Okutan, S. (2017). *Ortaokul öğrencilerinin fen bilimleri başarılarının farklı değişkenler açısından incelenmesi* (Unpublished doctoral dissertation). Ordu University, Ordu.

Olson, D. L., & Delen, D. (2008). *Advanced data mining techniques*. Berlin, Almanya: Springer Science & Business Media.

Önen, E. (2018). Öğrenci, öğretmen ve öğretimsel nitelikler açısından TIMSS-2015'e dayalı olarak öğrencilerin sınıflandırılması. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi, 9*(1) 64-84. doi:10.21031/epod.364319

Özçınar, H. (2006). *KPSS sonuçlarının veri madenciliği yöntemleriyle tahmin edilmesi* (Unpublished master's thesis). Pamukkale University, Denizli.

Özdamar, K. (2013). *Paket programlar ile istatiksel veri analizi: MINITAB 16-IBM SPSS 21*. Eskişehir: Nisan Kitabevi.

Özfalcı, Y. (2008). *Çok değişkenli uyarlanabilir regresyon kesitleri: MARS* (Unpublished master's thesis). Gazi University, Ankara.

Özgürlük, B. (2019). *Örneklem büyüklüğünün ve madde formatının sekizinci sınıf ABİDE testlerinin eşitlenmesine etkisi* (Unpublished doctoral dissertation). Hacettepe University, Ankara.

Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research Winter, 66*(4), 543-578. doi:10.3102/00346543066004543

Plucker, J. A. (2010). The relationship between school climate conditions and student aspirations. *The Journal of Educational Research, 91*(4), 240-246. doi:10.1080/00220679809597549

Provost, F., & Fawcett, T. (2001). Robust classification for imprecise environments. *Machine Learning, 42*(3), 203-231. Retrieved from https://link.springer.com/article/10.1023/A:1007601015854

Rutkowski, L., & Rutkowski, D. (2016). The relation between students' perceptions of instructional quality and bullying victimization. In T. Nilsen & J. E. Gustafsson (Eds.), *Teacher quality, instructional quality and student outcomes* (pp. 115). Cham: Springer.

Saa, A. A. (2016). Educational data mining & students' performance prediction. *International Journal of Advanced Computer Science and Applications, 7*(5), 212-220.

Salford Systeam. (2018). Salford predictive modeller: Introducing MARS. Retrieved from https://www.minitab.com/content/dam/www/en/uploadedfiles/content/products/spm/IntroMARS.pdf

Sarı, M. H., Arıkan, S., & Yıldızlı, H. (2017). Factors predicting mathematics achievement of 8th graders in TIMSS 2015. *Journal of Measurement and Evaluation in Education and Psychology, 8*(3), 246-265. doi:10.21031/epod.303689

Scherer, R., & Nilsen, T. (2016). The relations among school climate, instructional quality, and achievement motivation in mathematics. In T. Nilsen & J. E. Gustafsson (Eds.), *Teacher quality, instructional quality and student outcomes* (pp. 51-80). Cham: Springer.

Sevimli-Saitoğlu, Y. (2015). *Sınıflama ve regresyon ağaçları* (Unpublished doctoral dissertation). Marmara University, İstanbul.

Sinharay, S. (2016). An NCME instructional module on data mining methods for classification and regression. *Educational Measurement: Issues and Practice, 35*(3), 38-54.

StatSoft. (2018). Multivariate adaptive regression splines (MARSplines). Retrieved from http://www.statsoft.com/textbook/multivariate-adaptive-regression-splines

Stearns, B., Rangel, F., Rangel, F., Faria, F. F., & Oliveira, J. (2017). Scholar performance prediction using boosted regression trees techniques. Paper presented at the European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, Bruges, Belgium. Retrieved from https://www.esann.org/proceedings/2017

Stone, C. A., & Tang, Y. (2013). Comparing propensity score methods in balancing covariates and recovering impact in small sample educational program evaluations. *Practical Assessment, Research, and Evaluation, 18*(1), 1-12.

Tabachnick, B. G., & Fidell, L. S. (2015). *Using multivariate statistics* (6th ed., M. Baloğlu, Trans.). Ankara: Nobel Yayınları.

Taş, U. E., Arıcı, Ö., Ozarkan, H. B., & Özgürlük, B. (2016). *PISA 2015 ulusal raporu*. Ankara: Milli Eğitim Bakanlığı.

Tepehan, T. (2011). *Türk öğrencilerinin PISA başarılarının yordanmasında yapay sinir ağı ve lojistik regresyon modeli performanslarının karşılaştırılması* (Unpublished doctoral dissertation). Hacettepe University, Ankara.

Torney-Purta, J., & Amadeo, J. A. (2013). International large-scale assessments: Challenges in reporting and potentials for secondary analysis. *Research in Comparative and International Education, 8*(3), 248-258. doi:10.2304/rcie.2013.8.3.248

Uysal, İ. (2019). Açık *uçlu maddelerde otomatik puanlamanın güvenirliği ve test eşitleme hatalarına etkisi* (Unpublished doctoral dissertation). Hacettepe University, Ankara.

Ülkü, S. (2019). *ABİDE 2016 Türkçe ve Fen bilimleri alt-testlerinin öğretmen özelliklerine göre ölçme değişmezliğinin incelenmesi* (Unpublished master's thesis). Hacettepe University, Ankara.

Winnaar, L., Arends, F., & Beku, U. (2018). Reducing bullying in schools by focusing on school climate and school socio-economic status. *South African Journal of Education, 38*(1). doi:10.15700/saje.v38ns1a1596

Yazıcı, H., Seyis, S., & Altun, F. (2011). Emotional intelligence and self-efficacy beliefs as predictors of academic achievement among high school students. *Procedia-Social and Behavioral Sciences, 15*, 2319-2323. doi:10.1016/j.sbspro.2011.04.100

Yıldırım, A., Özgürlük, B., Parlak, B., Gönen, E., & Polat, M. (2016). *TIMSS 2015 ulusal matematik ve fen bilimleri ön raporu 4. ve 8. sınıflar*. Ankara: Milli Eğitim Bakanlığı Ölçme, Değerlendirme ve Sınav Hizmetleri Genel Müdürlüğü.

Yolagiden, C., & Bektaş, O. (2018). Sekizinci sınıf öğrencilerinin fen bilimleri öğrenme kaygıları ile fen bilimleri öğrenme yönelimleri arasındaki ilişkinin incelenmesi. *Maarif Mektepleri Uluslararası Eğitim Bilimleri Dergisi, 2*(2), 18-41. Retrieved from https://dergipark.org.tr/en/pub/mamulebd/issue/41955/496279

Young, D. J. (1998). Ambition, self-concept, and achievement: A structural equation model for comparing rural and urban students. *Journal of Research in Rural Education, 14*(1), 34-44. Retrieved from https://eric.ed.gov/?id=EJ574238%20

Yu, C. H., Digangi, S., Jannasch-Pennell, A. K., & Kaprolet, C. (2008). Profiling students who take online courses using data mining methods. *Online Journal of Distance Learning Administration, 11*(2), 1-14.

Yu, C. H., Kaprolet, C., Jannasch-Pennell, A., & DiGangi, S. (2012). A data mining approach to comparing American and Canadian grade 10 students' PISA science test performance. *Journal of Data Science, 10*(24), 441-464.

Yurdugül, H. (2006). The comparison of reliability coefficients in parallel, tau-equivalent, and congeneric measurements. *Ankara University, Journal of Faculty of Educational Sciences, 39*(1), 15-37. Retrieved from https://www.researchgate.net/publication/237341588_The_Comparison_of_Reliability_Coefficients_in_Parallel_Tau-Equivalent_and_Congeneric_Measurements