Research Article

Psycho-Educational Research Reviews
12(*2*), 2023, 505-521
www.perrjournal.com

Influential Factors on Mathematical Literacy of Turkish Students: An Educational Data Mining Study Using PISA 2015 Data

Serap Büyükkıdık, Assist. Prof. Dr., Sinop University, Türkiye, sbuyukkidik@gmail.com
0000-0003-4335-2949

Keywords

Educational data mining
J48
Multilayer perceptron
Naïve bayes
Support vector machine

Article Info:

Received : 15-07-2022
Accepted : 17-07-2023
Published : 10-08-2023

DOI: 10.52963/PERR_Biruni_V12.N2.10

Abstract

This study aims to classify students as successful and unsuccessful regarding mathematical literacy on Programme for International Student Assessment (PISA) 2015 database through data mining methods. The sample consists of all Turkish students who participated in PISA 2015. While data mining methods such as Support Vector Machine, Multi-Layer Perceptron, and J48 were used in data analysis, the data set was evaluated with 10-fold Cross-validation. The evaluation criteria included F-measure, Precision, Recall, Matthews Correlation Coefficient, and Receiver Operating Characteristic (ROC). In the classification of successful and unsuccessful students, analyses were conducted with 13 statistically significant variables according to Chi-SquareAttributedEval, GainRatioAttributeEval, and InfoGainAttributeEval methods. The results showed that the most important variables for classifying successful and unsuccessful students were learning time per week in total, and father's education level. The highest ROC value was 0.720. When comparing the precision values, the lowest classification value for the Multilayer Perceptron method was 0.645. There was no single method that performed best for all criteria. Researchers should use at least two methods to obtain more accurate results.

To cite this article: Büyükkıdık, S. (2023). Influential factors on mathematical literacy of Turkish students: An educational data mining study using PISA 2015 data. *Psycho-Educational Research Reviews, 12*(2), 505-521. doi: 10.52963/PERR_Biruni_V12.N2.10

INTRODUCTION

In today's rapidly changing and competitive business environment, prompt decisions must be made by the companies to ensure high productivity. Dealing with big data comprises of one of the problems of industry 4.0 and society 5.0. Data is collected to suggest predictive production in the world of the future with machines. It requires the use of advanced forecasting tools to systematically explain uncertainties and thereby transform data into information capable of making "informed" decisions (Lee, Kao, & Yang, 2014). Tremendous amounts of data are collected by modern computer systems from various sources, such as government statistics, credit card transactions, cash withdrawal machines at the bank, and Earth observation satellites with an unceasingly increasing volume of information available from the internet (Bramer, 2013; Hand, Manila, & Smyth, 2001). The amount of data is big, but it is not likely to make sense of these data entirely these days. "The world is becoming data rich but knowledge poor'' has deservedly been noted (Bramer, 2013, pp.1-2).

In science, engineering, and many disciplines were basically changed by the quick progress of knowledge and computer technology for twenty years. Currently, data-poor era has now been changed by the data-rich era, and new methods to conduct research is needed for the development of the data-intensive era. It is an essential task to develop robust data mining tools to analyze such data (Han & Gao, 2008). One of the changing disciplines in the data-rich period is education. Today, a lot of data about education is collected. These data are analyzed by methods such as machine learning or data mining. Machine learning or data mining is a technology that determines which factors are taken into account in successful decisions based on experience. It is a flourishing new way for mining knowledge from data (Witten & Frank, 2005). Data mining is the method of exploring patterns in data (Witten & Frank, 2005). It is an interdisciplinary analysis. Database technology, statistics, artificial intelligence, pattern recognition, and visualization, machine learning, all play a role. People at the scientific, business and the physical world, or at some other conceptual domain aim at discovering the relationships that exist in the real world (Hand et al., 2001).

People have become increasingly interested in big data in education (Siemens & Baker, 2012). Consistent with this interest, the number of educational data mining (EDM) studies has grown outstandingly in the literature over the last few years (Bousbia & Belamri, 2014). EDM is the sphere of using data mining techniques in educational environments to address important educational questions (Bakhshinategh, Zaiane, El Atia, & Ipperciel, 2018; Bousbia & Belamri, 2014; Romero & Ventura, 2010, 2013). EDM research has been preferred by researchers in recent years as a more effective alternative to classical inferential and multivariate statistics (Martínez-Abad, 2019). Many studies on educational data mining have been conducted so far for different purposes (Aldowah, Al-Samarraie, & Fauzy, 2019; Baker & Yacef, 2009; Bakhshinategh et al., 2018; Romero & Ventura, 2007, 2010). For instance, Bakhshinategh, Zaiane, El Atia, and Ipperciel (2018, p. 541) defined five goals of EDM research: "(1) student modeling (predicting student performance, achievement of learning outcomes or characteristics, detecting undesirable student behaviors, profiling and grouping students, social network analysis), (2) decision support systems, (3) adaptive systems, (4) evaluation, and (5) the scientific inquiry". Baker and Yacef (2009, pp. 6-7) addressed the four key applications of EDM method as (1) "improvement of student models", (2) "discovering or improving models of a domain's knowledge structure", (3) "studying pedagogical support" and (4) "looking for empirical evidence to refine and extend educational theories, and well-known educational phenomena". Romero and Ventura (2007, p. 135) stated that EDM can be applied within the scope of statistics and visualization, text mining, classification and outlier detection, clustering, association rule mining and pattern mining to improve the quality of education. Romero and Ventura (2010, p. 602) stated that the stakeholders of EDM are "students" (to personalize e-learning, etc.), "educators" (for purposes such as obtaining objective feedback about education and predicting student performance etc.), "course developers" (to maintain and evaluate course, etc.), "organizations" (to improve decision processes, etc.) and "administrators" (to organize organizational resources in the best way, etc.). When the literature was examined, it has been seen that data mining provides feedback to diverse stakeholders for various purposes in education.

Data mining has facilitated educational research by providing information about learning time, better understanding learning, and improving education. Such researches have been conducted in various fields (including education, statistics, computer science, etc.) in the last few years (Romero, Ventura, Pechenizkiy, & Baker, 2010). Since EDM is an interdisciplinary field, this method is used in the areas such as education, machine learning, statistics, psychometry, information retrieval, recommendation systems, cognitive psychology, psycho-pedagogy, etc. (Bousbia & Belamri, 2014). We can use educational data mining for social science research that improves our understanding of the learning process. It is also valid for the applied research that improves the quality of learning (Bakhshinategh et al., 2018).

In the field of education, data mining is often used for the purpose of classification, one of the aims of data mining (Aldowah et al., 2019; Bramer, 2013). In this method, an object is put into categories or classes (Hämäläinen & Vinni, 2010). That is, it is the process of supervised learning that separates data into dissimilar predefined classes (Aldowah et al., 2019). There are a lot of examples of classification studies in education (e.g.; Aksu, 2018; Aksu & Güzeller, 2016; Bezek Güre, Kayri, & Erdoğan, 2020; Bresfelean, Bresfelean, Ghisoiu, & Comes, 2008; Bunkar, Singh, Pandya, & Bunkar, 2012; Büyükkıdık, Bakırarar, & Bulut, 2018; Costa, Fonseca, Santana, de Araújo, & Rego, 2017; Bezek Güre et al., 2020; Kaur, Singh, & Josan, 2015; Kılıç Depren, Aşkın, & Öz, 2017; Koyuncu & Gelbal, 2020; Liu & Whitford, 2011; Martínez-Abad, 2019; Martínez-Abad et al., 2020; Ramesh, Parkavi, & Ramar, 2013; Saarela, Yener, Zaki, & Kärkkäinen, 2016; Yukselturk, Ozekes, & Kılıç Türel, 2014). For instance, classifying a student project can be desired as pass or fail in education (Bramer, 2013). Estimating student achievement, performance, knowledge, detecting undesirable student behaviors in online courses/e-learning, estimating/preventing student dropout can be realized with the concepts of classification (Aldowah et al., 2019). Students are always classified by their instructors and teachers on their behavior, motivation, and knowledge in education. Assessing exam answers is also a classification task, where a score is determined by clear evaluation criteria (Hämäläinen & Vinni, 2010). Programme for International Student Assessment (PISA) is one of the first data sets that comes to mind when big data is mentioned in education, it is important to categorize or classify students in terms of their success in PISA. In addition, examining the variables that affect the success of students will contribute to the literature. At this point, data mining methods can be used in PISA dataset.

VARIABLES RELATED TO MATHEMATICAL LITERACY

For accountability in education, countries participate in international assessments like PISA, Progress in International Reading Literacy Study (PIRLS), Trends in International Mathematics and Science Study (TIMSS), Teaching and Learning International Survey (TALIS), etc. as well as national examination practices. "The Organisation for Economic Co-operation and Development's" (OECD) "Program for International Student Assessment (PISA)" is one of the sources that provide researchers with big data in education through collecting international data. PISA 2015 covered science, mathematics, reading, financial literacy, and collaborative problem solving with a primary focus on science, which was conducted in 72 countries and economies (OECD, 2017a). Mathematical literacy was assessed as one of the domains in PISA 2015 (OECD, 2017b).

OECD (2017b, p. 67) defined mathematical literacy as follows:

> Mathematical literacy is an individual's capacity to formulate, employ and interpret mathematics in a variety of contexts. It includes reasoning mathematically and using mathematical concepts, procedures, facts and tools to describe, explain and predict phenomena. It assists individuals to recognize the role that mathematics plays in the world and to make the well-founded judgements and decisions needed by constructive, engaged and reflective citizens.

Approximately 540 000 students from 72 countries participated in the PISA 2015 and provided some feedback to researchers and officials about the quality of education. It is possible to make inferences from the data collected through national, and international exams for accountability in education. In the selected sample representing 29 million 15-year-old students, PISA 2015 provides some data about mathematics, science and reading literacy with a lot of questionnaires and various cognitive indicators (OECD, 2016). It is necessary to analyze this big data and make inferences for accountability.

Although mathematical literacy and mathematics achievement are not exactly the same two concepts, since both concepts are related to student performance, it was seen that they were discussed together in the literature (e.g., D'Agostino et al., 2022; Keller, Preckel, Eccles, & Brunner, 2022). The studies have demonstrated that there is a relationship between mathematical literacy or achievement and variables like students' grade level (Ammermüller, 2004; Bratti, Checchi, & Filippin 2011; Fuchs & Wößmann, 2008; Gamazo & Martínez-Abad, 2020; Gilleece, Cosgrove, & Sofroniou, 2010), gender (Bratti et al., 2011; Else-Quest, Hyde, & Linn, 2010; Fuchs & Wößmann, 2008; Gamazo & Martínez-Abad, 2020; Gilleece et al., 2010; Hyde et al., 1990; Liu & Wilson, 2009; Keller et al., 2022; Reilly, Neumann, & Andrews, 2017), parental support (Hertel & Jude, 2016), test anxiety (Culler & Holohan, 1980; Cassady & Johnson, 2002; D'Agostino, Schirripa Spagnolo, & Salvati, 2022), achievement motivation (Gamazo & Martínez-Abad, 2020; Gunderson, Park, Maloney, Beilock, & Levine, 2018; Keller et al., 2022; Singh, Granville, & Dika, 2002), enjoy cooperation (Bratti et al., 2011; Slavin, 1983), sense of belonging school (Linnakylä & Malin, 2008; Wilms, 2003), learning time in mathematics, class periods, out of school study time in mathematics (Lee & Stankov, 2018; Singh et al., 2002), mother's education, father's education, highest education of parents (Anıl, 2009; Bezek Güre et al., 2020; Bratti et al., 2011; Chevalier & Lanot, 2002; Fuchs & Wößmann, 2008; Kılıç Depren et al., 2017; Liu & Whitford, 2011; Põder, Lauri, Ivaniushina, & Alexandrov, 2016; Ramesh et al., 2013; Yayan & Berberoglu, 2004).

Specifically, the studies reveal that as the grade level increases, mathematics literacy increases (Ammermüller, 2004; Fuchs & Wößmann, 2008; Gamazo & Martínez-Abad, 2020; Gilleece et al., 2010). Exploring gender differences in achievement using international measures such as PISA is a critical starting point for psychological and educational research, policies, and practices in Science, Technology, Engineering, Mathematics (STEM) that target women's underrepresentation (Keller et al., 2022). It can be claimed that gender differences in mathematics achievement continue today. While several studies indicate that females outperform their male peers in mathematics (e.g., Hyde, Fennema, & Lamon, 1990), there are also studies that show that males perform higher than females in mathematics (e.g., Liu & Wilson, 2009). Some studies found no difference by gender (e.g., Hyde, 2005) or statistically insignificant differences (e.g., Else-Quest et al., 2010). Ma (1999) examined the influence of anxiety on mathematics achievement and noted a negative significant relationship between mathematics achievement and anxiety at the elementary and secondary school level, which did not change for gender, ethnic group, scale used for measuring anxiety, grade level, and time of publication) in the meta-analysis study. This finding is supported by a lot of studies in the literature (e.g., D'Agostino et al., 2022; Gunderson et al., 2018; Hembree, 1990; Sherman & Wither, 2003; Tocci & Engelhard, 1991; Wu, Willcutt, Escovar, & Menon, 2014; Zhang, Zhao, & Kong, 2019). Studies have reported that individuals with high test anxiety perform less than individuals with low test anxiety in mathematics (e.g., Culler & Holohan, 1980; Cassady & Johnson, 2002; D'Agostino et al., 2022; Wine, 1971). Gunderson et al. (2018) revealed the relationships between motivational frameworks, math anxiety, and math achievement at the early elementary school level. As the motivation for success increases, the level of mathematics literacy also increases (Gunderson et al., 2018; Singh et al., 2002). In elementary and secondary schools, rewards and collaborative learning are required to increase student achievement (Slavin, 1983). Apart from this, variables of learning time in mathematics class periods and out-of-school study time in mathematics also affect mathematics literacy success. As the education level of the parents increases, the mathematical literacy level of the children increases (Anıl, 2009; Bratti et al., 2011; Chevalier & Lanot, 2002; Fuchs & Wößmann, 2008; Kılıç Depren et al., 2017; Liu & Whitford, 2011; Põder et al., 2016; Ramesh et al., 2013; Yayan & Berberoglu, 2004).

CURRENT STUDY

There are many studies in the field of mathematics literacy or achievement. Similarly, examples of the use of educational data mining for classification are prevalent in the literature (e.g. Aksu, 2018; Aksu & Güzeller, 2016; Bezek Güre et al., 2020; Bresfelean et al., 2008; Bunkar et al., 2012; Büyükkıdık et al., 2018; Costa et al., 2017; Bezek Güre et al., 2020; Kaur et al., 2015; Kılıç Depren et al., 2017; Koyuncu & Gelbal, 2020; Liu & Whitford, 2011; Martínez-Abad, 2019; Martínez-Abad, Gamazo, & Rodríguez-Conde, 2020; Ramesh et al., 2013; Saarela et al., 2016; Yukselturk et al., 2014). Nonetheless, there are limited number of studies that compares the results of PISA dataset via different data mining methods (e.g., Aksu, 2018; Aksu & Güzeller, 2016; Bezek Güre, Kayri, & Erdoğan, 2020; Büyükkıdık et al., 2018; Liu & Whitford, 2011; Martínez-Abad, 2019; Martínez-Abad et al., 2020; Koyuncu & Gelbal, 2020; Saarela et al., 2016). When all these studies were examined, no completely similar study was found comparing the Multilayer Perceptron, J48, Support Vector Machine, and Naïve Bayes methods. Additionally, studies in the literature didn't discuss the significance of the variables affected mathematical literacy in detail using the PISA 2015 dataset. This is another reason for us to conduct the research. In this study, both the importance of variable and the performance of data mining methods in the classification of mathematics literacy were compared under different criteria. In this respect, the current research is thought to contribute to the literature.

For this aim, the following research questions were asked in the study:

(1) Which variables related to mathematical literacy are important for classifying students?

(2) How are the descriptive statistics on the important variables for successful and unsuccessful students?

(3) Based on the five evaluation criteria (F-measure, Precision, Recall, Matthews Correlation Coefficient (MCC) and ROC), which data mining method (Multilayer Perceptron, J48, Support Vector Machine, and Naïve Bayes) performs best in classifying students successful or not in mathematical literacy?

METHOD

RESEARCH TYPE

Since the research includes the classification of PISA mathematical literacy using different data mining methods, it was descriptive research (Fraenkel, Wallen, & Hyun, 2012).

SAMPLE

This study is based on PISA 2015 data for Turkey. A total of 5895 students from 61 provinces and 187 schools were involved in this exam to assess their proficiency in applying skills and knowledge to authentic problems and noncognitive responses. PISA is an age-based assessment, measuring 15-year-old students who are mostly at the end of compulsory education in grade 7 or higher.

The 15-year-old student population in PISA 2015 is defined as 1324089 in Turkey, while the reached universe in Turkey is defined as 925366 students eligible to participate in the application. School in PISA research sampling is determined by stratified random sampling method (Taş, Arıcı, Ozarkan, & Özgürlük, 2016, p. 5). A total of 5895 students' data in Turkey from 61 provinces and 187 schools were analyzed in this research. The sample of the study consists of all Turkish students who participated in PISA 2015.

DATA COLLECTION INSTRUMENTS

The independent variables used in data analysis in the present study were demonstrated at Table 1 and Table 2. The dependent variable in this study was mathematical literacy. When the educational data mining studies in the literature were examined, it was seen that the analyzes were carried out by considering the only plausible values as the dependent variable (Aksu, 2018; Aksu & Güzeller, 2016; Bezek Güre et al., 2020; Büyükkıdık et al., 2018; Gamazo & Martínez-Abad, 2020; Kılıç Depren et al., 2017; Koyuncu & Gelbal, 2020). Students were seen as successful and unsuccessful taken into account in research classify according to the average score of Turkey like similar studies (Aksu, 2008; Aksu & Güzeller, 2016; Büyükkıdık et al., 2018; Koyuncu & Gelbal, 2020). Plausible values were used for indicators of mathematical literacy. In the study, the average mathematical literacy score was obtained by taking the average of PVMATH coded math scores from PISA 2015 Turkey data. Then, the average of the achievement scores ($\bar{x}$ = 420) was taken and this value was determined as cut-off score. Although the mathematics literacy chosen as a dependent variable was a continuous variable as a score (PVMATH), this variable was converted into a categorical variable by comparing the PISA 2015 Turkey average score with 420 points.

In data mining research using WEKA software, several effective feature selection methods were used to achieve the best classification and prediction for performance (Kılıç Depren et al., 2017). Three of these methods are InfoGainAttributeEval, GainRatioAttributeEval, and Chi-SquaredAttributedEval. Since there were too many variables related to mathematics literacy in the PISA 2015 dataset, the importance of variables was examined by using the methods of InfoGainAttributeEval, GainRatioAttributeEval, and Chi-SquaredAttributedEval, and the variables which were determined insignificant or less important by the three methods were excluded from the dataset. A total of 14 variables (13 independent variables and 1 dependent variable) remained in the data set. These variables were students' grade level (ST001D01T, grade), gender (ST004D01T, gender), parental support (ST123, EMOSUPS), test anxiety (ST118, ANXTEST), achievement motivation (ST119, MOTIVAT), enjoy cooperation (ST082, COOPERATE), sense of belonging school (ST034, BELONG), learning time in mathematics (ST059Q02TA, MMINS), learning time per week in total (ST060Q01NA, TMINS), out of school study time in mathematics (ST071Q02NA, OUTHOURS), mother's education (MISCED), father's education (FISCED), highest education of parents (HISCED) and mathematical literacy (PVMATH, MATH). Detailed information about the percentages of the significance of variables according to the dependent variable mathematical literacy (Math) was given (see Figure 1).

All measurements of PISA 2015 questionnaires in this research were reliable since all McDonald's ω coefficients were above 0.70. The lowest coefficient was 0.704 for the enjoy cooperation questionnaire and the highest coefficient was 0.860 for the parental support questionnaire for the Turkey sample.

In addition, "Ethics Committee Permission" was also obtained for the research. Ethics Committee Permission was granted by Sinop University Ethics Committee with the decision number 2022/026 at the Ethics Committee meeting on 24/03/2022.

DATA ANALYSIS

Before starting data analysis, missing data were imputed. The data were analyzed by using WEKA 3.8.6 and SPSS programs. The mean ± standard deviation and median (minimum-maximum) for quantitative variables and the number of students (percentage) were used for qualitative variables. Multilayer Perceptron, J48, Support Vector Machine, and Naïve Bayes which are the classification methods in the WEKA program were used. The model evaluation criteria used in this research were explained in the next section.

MODEL EVALUATION CRITERIA

The data set was evaluated by using the 10-fold Cross Validation test option. Recall, Precision, F-measure, and Matthews Correlation Coefficient (MCC) were used as evaluation criteria. Computations of these measures were calculated with 2 x 2 contingency table including all possible outcomes (false positive (FP), and false negative (FN), true positive (TP), true negative (TN) classifications). All measures are computed as follows:

$$\text{Recall}=\frac{\text{TP}}{\text{TP+FN}} \quad (1)$$

$$\text{Precision}=\frac{\text{TP}}{\text{TP+FP}} \quad (2)$$

$$\text{F-measure}=\frac{\text{2xRecallxPrecision}}{\text{Recall+Precision}} \quad (3)$$

$$\text{Matthews Correlation Coefficient}= \frac{\text{TP x TN}-\text{FP x FN}}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \quad (4)$$

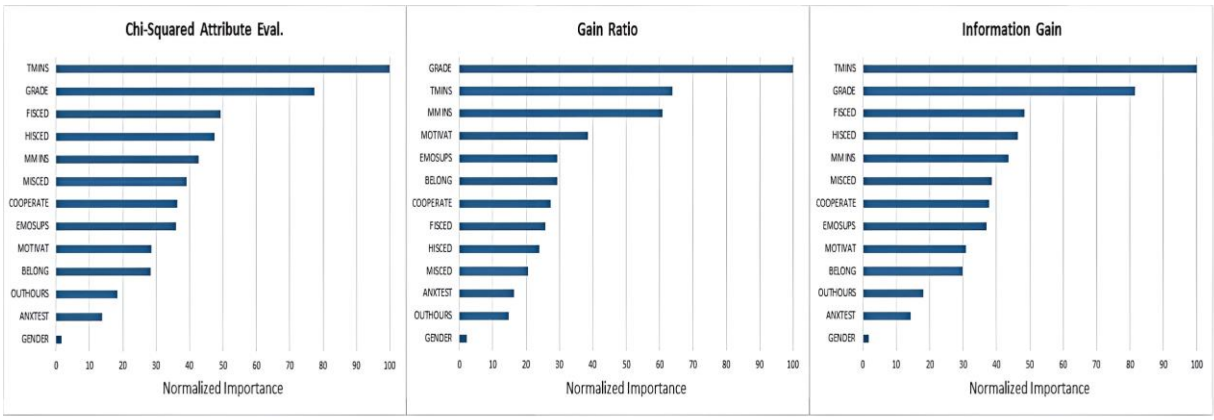
F- measure, Precision, Recall measures range from 0 to 1; the higher the value of these criteria have the better fit. Using the elements in the confusion matrix, the Matthews correlation coefficient (MCC) takes values between -1 and 1. Positive MCCs are indicative of correct predictions, MCC being 1 is an indication of perfect prediction (Kılıç Depren et al., 2017).

FINDINGS

VARIABLE IMPORTANCE

Significant Variables Related to Mathematical Literacy in PISA 2015 In this section, findings related to the first sub-problem were found. The percentages of the significance of variables according to the dependent variable mathematical literacy (Math) were given (see Figure 1).

Figure 1. *Variable Importance for Classifying Mathematical Literacy in Data Set*



The importance of the independent variables selected in the data set in the classification of mathematical literacy was examined with InfoGainAttributeEval, GainRatioAttributeEval, and Chi-SquaredAttributedEval methods (see Figure 1). When the bar graph was examined, values related to the significance of the independent variables were seen. Considering all three criteria, TMINS (minutes of overall school instruction) can be considered as the most important variable. When the InfoGainAttributeEval, and Chi-SquaredAttributedEval criteria were used, the same variable significance order was found; whereas in the GainRatioAttributeEval criterion, it was partially differentiated. Based on all three criteria, gender was a significant variable in classifying mathematics literacy, but it was the least important variable when compared to other independent variables. The number of independent variables in the data set was reduced to thirteen based on InfoGainAttributeEval, GainRatioAttributeEval, and Chi-SquaredAttributedEval criteria.

DESCRIPTIVE STATISTICS

Table 1 shows the descriptive statistics of continuous variables in the classification of mathematical literacy.

Table 1. *Descriptive Statistics for Continuous Variables*

Variables	Mathematical Literacy			
	Unsuccessful		Successful	
	Mean±SD	Median (Min.-Max.)	Mean±SD	Median (Min.-Max.)
EMOSUPS	12.45±3.13	13.00 (4.00-16.00)	13.34±2.47	14.00 (4.00-16.00)
ANXTEST	14.04±3.77	14.00 (5.00-20.00)	13.68±3.47	14.00 (2.00-20.00)
MOTIVAT	16.31±3.61	17.00 (5.00-20.00)	17.25±2.69	18.00 (5.00-20.00)
COOPERATE	22.92±4.68	23.00 (8.00-32.00)	21.14±3.49	24.00 (6.00-32.00)
BELONG	14.86±2.94	15.00 (6.00-24.00)	15.22±2.13	15.00 (6.00-24.00)
MMINS	5.24±1.81	6.00 (0.00-15.00)	5.79±1.39	6.00 (0.00-15.00)
TMINS	38.24±8.25	40.00 (10.00-60.00)	39.51±5.57	40.00 (10.00-60.00)
OUTHOURS	6.48±5.32	6.00 (0.00-30.00)	6.28±4.83	6.00 (0.00-30.00)

Mean ± standard deviation, median and maximum and minimum values obtained from questionnaires used in the study for the groups with successful and unsuccessful mathematical literacy performance were displayed at Table 1. Regarding parental support, the mean ± standard deviation value for students classified as unsuccessful was 12.45 ± 3.13, while it was $\bar{X}$ = 13.34 ± 2.47 for students classified as successful. The median (min-max) value of the group classified as unsuccessful for the same scale was 13.00 (1.00-16.00), while the median (min-max) value for the group classified as successful was 14.00 (4.00-16.00). The maximum score that can be obtained from the four-category and four-item scale is 16 and the minimum score is 4.

Descriptive statistics of categorical variables in the classification of mathematics literacy were presented at Table 2.

Table 2. *Descriptive Statistics for Qualitative Variables*

Variables		Mathematical Literacy			
		Unsuccessful		Successful	
		Frequency	%	Frequency	%
Grade	Grade 7	16	0.5	0	0.0
	Grade 8	94	2.9	11	0.4
	Grade 9	951	29.3	322	12.2
	Grade 10	2077	64.0	2231	84.2
	Grade 11	103	3.2	83	3.1
	Grade 12	4	0.1	3	0.1
Gender	Female	1670	51.5	1268	47.8
	Male	1575	48.5	1382	52.2
MISCED	None	535	16.5	259	9.8
	ISCED 1	1149	35.4	1002	37.8
	ISCED 2	579	17.8	444	16.8
	ISCED 3A, ISCED 4	149	4.6	280	10.6
	ISCED 3B, C	298	9.2	155	5.8
	ISCED 5A, 6	205	6.3	277	10.5
	ISCED 5B	330	10.2	233	8.7
FISCED	None	205	6.3	134	5.1
	ISCED 1	1115	34.4	680	25.6
	ISCED 2	895	27.6	566	21.4
	ISCED 3A, ISCED 4	132	4.1	222	8.4
	ISCED 3B, C	223	6.9	233	8.7
	ISCED 5A, 6	276	8.4	484	18.3
	ISCED 5B	399	12.3	331	12.5
HISCED	None	114	3.5	85	3.2
	ISCED 1	903	27.7	578	21.8
	ISCED 2	834	25.7	505	19.1
	ISCED 3A, ISCED 4	203	6.3	323	12.2
	ISCED 3B, C	362	11.2	249	9.4
	ISCED 5A, 6	362	11.2	564	21.2
	ISCED 5B	467	14.4	346	13.1

When Table 2 was examined, the percentages and frequencies of the students classified as successful and unsuccessful were found for the categorical variables used in the research. For example, 1670 of the female students included in the study were classified as unsuccessful and 1268 of them were classified as successful. In addition, while 51.5% of unsuccessful students are female students, 47.8% of successful students are female students.

PERFORMANCES OF DATA MINING METHODS

In this section, findings related to the second sub-problem were found. F-measure, Precision, Recall, MCC, and ROC were used as evaluation criteria and the results of the dependent variable categories were given in the Table 3.

Table 3. *Results of The Classification Criteria for Multilayer Perceptron, J48, Support Vector Machine, and Naïve Bayes Methods Based on Dependent Variable Categories for the Data Set*

Methods	Math	Performance Criteria				
		Precision	Recall	F-measure	MCC	ROC
Multilayer Perceptron	Unsuccessful	0,676	0,685	0,680	0,283	0,701
	Successful	0,608	0,597	0,602	0,283	0,701
J48	Unsuccessful	0,689	0,714	0,701	0,321	0,661
	Successful	0,634	0,605	0,619	0,321	0,661
Support Vector Machine	Unsuccessful	0,686	0,724	0,704	0,320	0,659
	Successful	0,637	0,594	0,614	0,320	0,659
Naïve Bayes	Unsuccessful	0,744	0,543	0,628	0,318	0,720
	Successful	0,579	0,771	0,661	0,318	0,720

As shown in Table 3, the Support Vector Machine method yields the best results according to Recall and F-measure performance criteria for classifying unsuccessful students. Multilayer Perceptron method gave the worst results in the classification of successful students according to most of performance criteria. When Table 3 was examined, the values obtained according to F-measure, Precision and, Recall criteria were found between 0.543 and 0.771.

Similarly, the average of the general classification results independent of category are given (see Table 4) by taking Precision, Recall, F-measure, MCC and ROC as evaluation criteria. Table 4 shows the results of classification criteria for the data set related to the Multilayer Perceptron, J48, Support Vector Machine, and Naïve Bayes methods.

Table 4. *Results of Classification Criteria for Multilayer Perceptron, J48 and Support Vector Machine Methods*

Methods	Performance Criteria				
	Precision	Recall	F-measure	MCC	ROC
Multilayer Perceptron	0,645	0,646	0,645	0,283	0,701
J48	0,664	0,665	0,664	0,321	0,661
Support Vector Machine	0,664	0,665	0,664	0,320	0,659
Naïve Bayes	0,670	0,645	0,643	0,318	0,720

As shown in Table 4, the J48 method gave the best results according to the F-measure, Recall and MCC criteria. Multilayer Perceptron method performed worst according to Precision, F-measure, Recall and MCC criteria in classifying students. J48 and Support Vector Machine gave same results according to Precision, Recall and F-measure. Naïve Bayes had best performance according to Precision and ROC criteria. When Table 4 was examined, the values obtained according to F-measure, Precision and Recall criteria were found between 0.643 and 0.720. All these values indicate that the correct classification rate of mathematics literacy of the selected independent variables was acceptable.

DISCUSSION, CONCLUSION AND IMPLICATIONS

There are many variables that affect mathematics achievement. Although mathematical literacy was one of the domains in PISA 2015, there were a lot of variables about it. All variables related to mathematical literacy were used at the beginning of the study. The variables in the data set were reduced to 13 variables (except the mathematical literacy dependent variable) by using the GainRatioAttributeEval, InfoGainAttributeEval, and Chi-SquareAttributedEval methods. The results mainly show that thirteen independent variables were important for classifying successful students in PISA 2015 using these three methods.

After selecting fourteen variables (students' gender, grade level, parental support questionnaire, test anxiety, achievement motivation questionnaire, enjoy cooperation questionnaire, sense of belonging school questionnaire, learning time in mathematics class periods, out of school study time in mathematics, father's education, mother's education, highest education of parents and mathematical literacy), we compare continuous and categorical variables in terms of classifying students according to mathematics performance as successful or unsuccessful.

Gender was an important variable in terms of classifying successful students in PISA 2015 mathematical literacy in recent study. This result consistent with the previous literature. Gender was seen as one of the important variables affecting mathematical literacy (Else-Quest, Hyde, & Linn, 2010; Fuchs & Wößmann, 2008; Liu & Wilson, 2009; Hyde et al., 1990; Reilly et al., 2017). Different aged males surpassed females on many domestic and international mathematic assessments such as PISA, TIMSS, the National Assessment of Educational Progress (NAEP) and Scholastic Aptitude Test (SAT) (Liu & Wilson, 2009: p. 165). Hyde, Fennema and Lamon (1990) found those girls showed a slight superiority in their meta-analysis by using 100 studies' results. There was no big gender difference in the meta-analysis conducted by Else-Quest, Hyde, and Linn (2010) regarding two major international data sets, the 2003 PISA and the TIMSS, representing 493495 students aged 14-16, to estimate the size of gender differences in mathematics achievement across 69 nations, but male reported more positive mathematics attitudes than females. Gender differences play a role in mathematics achievement, in conjunction with self-efficacy beliefs, attitudes and other related factors (Reilly et al., 2017). When we investigated literature, the gender achievement gap has been declining since 1980 (Feingold, 1988), but gender difference in mathematics can not be ignorable (Hyde et al., 1990; Liu & Wilson, 2009). It is thought that gender differences in mathematical literacy are not based on biological origin. Socio-cultural factors may cause gender difference in success. Within the framework of inclusive education, the content of mathematics courses should be arranged in a way that attracts the attention of both genders. Necessary precautions for gender equality should be taken by policy makers.

Grade level was significantly associated with achievement in PISA mathematics test consistent with several research (Ammermüller, 2004; Fuchs & Wößmann, 2008; Gilleece et al., 2010). However, it should be noted that the sample of PISA 2015 Türkiye mostly consists of 10th grade students. Family background (MISCED, FISCED, HISCED in PISA) effect was important in the PISA and other studies (Anıl, 2009; Chevalier & Lanot, 2002; Fuchs & Wößmann, 2008; Põder et al., 2016, Yayan & Berberoglu, 2004). As parents' education level increased, success in each area of PISA 2000 increased (Fuchs & Wößmann, 2008).

Similar to our study, Ramesh, Parkavi and Ramar (2013) found that parents' education was an important variable in predicting student performances using data mining methods. Kılıç Depren, Aşkın, and Öz (2017) found that parents' highest education level was important factor in mathematic achievement in their EDM research with using TIMSS 2011 data. HISCED found as important variable for reaching science proficiency in another EDM research with using PISA 2006 US national sample (Liu & Whitford, 2011). Mother's education level (MISCED) and father's education level (FISCED) were significant variables in Bezek Güre, Kayri, and Erdoğan's (2020) recent EDM study. Parallel with the researches in the literature, one of the most important variables was found the educational level of the parents in our research. It is thought that the increase in the education level of the parents may cause an increase in the awareness of raising children. The education level of the society should increase so that educated individuals can be raised in the next generation. It is also found that as the education level of the parents increases, the success in education will increase.

In our study, the mean and median scores of the students who were successful and unsuccessful taken from the anxiety questionnaire were found similar. However, anxiety appeared to be an important variable in classifying successful and unsuccessful students. As a result of research conducted in the literature for many years, a negative relationship was found between math anxiety and mathematics achievement (D'Agostino et al., 2022; Gunderson, Park, Maloney, Beilock, & Levine, 2018; Sherman & Wither, 2003; Wu et al., 2014; Tocci & Engelhard, 1991). Similarly, this finding was supported by meta-analysis studies (Hembree, 1990; Ma, 1999; Zhang et al., 2019). Generally, high math anxiety causes low math success. Aksu and Güzeller (2016) found the anxiety was an important variable in classifying students in terms of mathematics achievement in their research using PISA 2012 data. D'Agostino et al. (2022) supported these results in their research which showed a negative association between anxiety and academic achievement using PISA 2015.

Motivation was another variable affecting success consistent with our research (Gunderson et al., 2018; Singh et al., 2002). Keller et al. (2022) emphasized the importance of achievement motivation in their study with top-performing math students in 82 countries using PISA data. Singh et al. (2002) found that motivation and academic time had positive effects on mathematics and science achievement in their research. Learning time in mathematics (ST059Q02TA, MMINS), learning time per week in total (ST060Q01NA, TMINS), out of school study time in mathematics (ST071Q02NA, OUTHOURS) were found important variables in this research. Particularly academic time spent on homework had the strongest impact on achievement in established structural equation modeling in the other research (Singh, Granville, & Dika, 2002). Students should be motivated to learn mathematics and high anxiety situations should be eliminated in education.

Lee and Stankov (2018) aimed to identify non-cognitive constructs that best predict students' mathematics achievement in PISA based on 43 variables. There were common variables like test anxiety (ST118, ANXTEST), learning time in mathematics (ST059Q02TA, MMINS), learning time per week in total (ST060Q01NA, TMINS), out of school study time (ST071Q02NA, OUTHOURS), highest education of parents (HISCED), etc. with our study among these variables.

As the level of belonging to school decreases, it is often associated with low success; but some studies have found no association (Wilms, 2003). In PISA 2003, Finnish students have the lowest sense of belonging to school in Nordic countries, while students' attitude and cognitive performance towards school are quite high (Linnakylä & Malin, 2008). There are different findings in the literature on belonging to school affects success.

Enjoy cooperation is also an important variable in our research. Bratti, Checchi, and Filippin (2011) investigated the impact of collaborative and competitive attitudes on mathematics literacy success using PISA 2003 data. While individual competitive attitudes are positively correlated with test scores, individual collaborative attitudes correlate negatively with test scores. When this situation is handled with peer attitudes, the situation is reversed. Grade, gender, and highest education of parents (HISCED) were taken into consideration in this research like in our research. Masci, Johnes, and Agasisti (2018) tried to find the factors affecting the PISA 2015 mathematics achievement of nine countries by applying machine learning with a two-step methodology. They revealed that the most influential variables affecting mathematics achievement in their research are gender, parental education, motivation, anxiety, and socioeconomic index. It was also another variable that they took into account the enjoy cooperation in their research. Similarly, Slavin (1983) found that collaborative learning is an effective factor in student achievement. The ability to cooperate, which is among the 21st century skills, should be encouraged in students' educational environments.

Hertel and Jude (2016) considered parents as strong stakeholders in education and emphasized that parental support is an important variable in student motivation and success (such as separating high- and low-performing students). In our study, another important variable appears as perceived parental support. EDM methods carried out in this study only identify the classification of mathematical literacy with important variables; these variables do not put forward any temporal relations or causality.

Our research has two dimensions: determining the variables affecting success in mathematical literacy and comparing data mining methods. Data mining techniques were used to classify students as successful and unsuccessful according to their mathematical literacy in this research. Comparison of data mining methods (Multilayer Perceptron, J48, Support Vector Machine, Naïve Bayes) with F-measure, Precision, Recall, MCC, and ROC performance criteria was conducted. With the variables used in this study, the performances obtained from Multilayer Perceptron, J48, Support Vector

Machine and Naïve Bayes methods according to five criteria were found above 0.64. All three methods classified the successful or unsuccessful students in terms of mathematics literacy sufficiently. When the dataset was analyzed with 10-fold Cross Validation option, there is no method gave the best results in terms of all criteria in this research. There is no study comparing these four methods with various criteria in education. However, in addition to these methods, there was a study (Firdausi, Erwin, & Nugroho, 2010) comparing the performances of k-Nearest Neighbors (kNN), Naïve Bayes methods according to recall, precision and accuracy criteria in behavior-based malware detection. In the Firdausi, Erwin, and Nugroho (2010)'s study, J48 showed the best performance according to all three criteria. When the comparisons of data mining methods in education were examined, different results were obtained on different data sets. Ramesh, Parkavi, and Ramar (2013) found that the Multilayer Perception method was more appropriate than the Naïve Bayes, Sequential minimal optimization (SMO), J48, and REPTREE methods when considering the accuracy value for predicting students' achievement with 500 records. Yukselturk et al. (2014) compared four classification algorithms (k-Nearest Neighbor (k-NN), Naïve Bayes (NB), Decision Tree (DT), and Neural Networks (NN)) using 10-fold cross-validation technique for predicting 189 dropout students. Yukselturk et al. (2014) trained and tested the data set using the 10-fold cross validation method, similar to our research, and found the performance of decision trees according to the accuracy criterion as 79.7%. Decision tree methods were found more sensitive than others in their study. In Kaur, Singh, and Josan (2015)'s study conducted on 152 students, they classified student performance using Multilayer Perception, Naïve Bayes, SMO, J48 and REPTree algorithms. Multilayer Perception performed best in small sample according to accuracy, and F-measure. Costa et al. (2017) found that their data mining techniques (Decision Tree via J48, Support Vector Machine, Neural Network, and Naïve Bayes) were sufficiently cultivated for identifying students' academic failures early. Generally, the J48 and Support Vector Machine showed the best performance. Kılıç Depren, Aşkın, and Öz (2017) classified mathematical success with using TIMSS 2011 8th Grade Turkey sample's data considering 11 variables based on "two Decision Tree algorithms (Random Forests and J48), a Bayesian Network Algorithm (Naïve Bayes), an Artificial Neural-Networks algorithm (Multilayer Perceptron), and the Logistic Regression" and found that Logistic Regression performed the best. In the study of Kılıç Depren, Aşkın, and Öz (2017), multilayer perceptron algorithms and J48 performed similarly to our research according to precision and f-measure criteria ($\approx$ 0.77). In terms of MCC ($\approx$0.48), it is seen that the results of both methods are in line with our study of Kılıç Depren et al. (2017). Bezek Güre, Kayri, and Erdoğan (2020) found that Random Forest performed better than Multilayer Perceptron with PISA 2015 mathematical literacy data. Koyuncu and Gelbal (2020) tested the performances of K-Nearest Neighborhood, Naïve Bayes, Logistic Regression, and Neural Networks analyzes at different sample sizes using PISA 2012 data. They found that Naïve Bayes performed well even with small sample size. All of these results show that there is no single best method for EDM in all conditions.

Multilayer Perceptron, J48, Support Vector Machine, and Naïve Bayes methods were used in this research. Researchers can perform similar research using at least two data mining methods. The other data mining methods and other criteria can be used in future EDM studies by using other large-scale assessments or other educational data. A limitation of our research is a sample size. Therefore, comparing the performance of EDM methods can be handled with a bigger sample. Future research may replicate the analysis on other countries' PISA data. Comparative studies can be conducted with data from other countries. As a result of the research, 13 independent variables affecting student performance in mathematical literacy were discussed. Similar studies can be carried out by revealing the variables that affect mathematics achievement through other data sets.

REFERENCES

Aksu, G. (2018). *PISA başarısını tahmin etmede kullanılan veri madenciliği yöntemlerinin incelenmesi. [Investigation of data mining methods used for estimating PISA success]*. (Publication No. 515513) [Doctoral dissertation, Hacettepe University]. Council of Higher Education Thesis Center, Türkiye.

Aksu, G., & Güzeller, C. O. (2016). Classification of PISA 2012 Mathematical Literacy Scores Using Decision-Tree Method: Turkey Sampling. *Education and Science*, *41*(185), 101-122. https://doi.org/10.15390/EB.2016.4766

Aldowah, H., Al-Samarraie, H., & Fauzy, W. M. (2019). Educational data mining and learning analytics for 21st century higher education: A Review and Synthesis. *Telematics and Informatics*, *37*, 13-49**.** doi:10.1016/j.tele.2019.01.007

Ammermüller, A. (2004). PISA: What makes the difference? Explaining the gap in PISA test scores between Finland and Germany. *Explaining the Gap in PISA Test Scores Between Finland and Germany*, ZEW Discussion Papers, No. 04-04, Zentrum für Europäische Wirtschaftsforschung (ZEW), Mannheim.

Anıl, D. (2009). Factors effecting science achievement of science students in programme for international students' achievement (PISA) in Turkey. *Education and Science, 34* (152), 87-100.

Baker, R. S., & Yacef, K. (2009). The state of educational data mining in 2009: A review and future visions. *Journal of Educational Data Mining*, *1*(1), 3-17. https://doi.org/10.5281/zenodo.3554657

Baker, R. S. (2010). Data mining for education. *International Encyclopedia of Education*, *7*(3), 112-118.

Bakhshinategh, B., Zaiane, O. R., El Atia, S., & Ipperciel, D. (2018). Educational data mining applications and tasks: A survey of the last 10 years. *Education and Information Technologies*, *23*(1), 537-553. https://doi.org/10.1007/s10639-017-9616-z

Bezek Güre, Ö. B., Kayri, M., & Erdoğan, F. (2020). Analysis of factors effecting PISA 2015 mathematics literacy via educational data mining. *Education and Science*, *45*(202), 393-415. https://doi.org/10.15390/EB.2020.8477

Bramer, M. (2013). *Principles of data mining (2nd ed.).* Springer-Verlag.

Bratti, M., Checchi, D., & Filippin, A. (2011). Should you compete or cooperate with your schoolmates?. *Education Economics*, *19*(3), 275-289. https://doi.org/10.1080/09645292.2011.585021

Bresfelean, V. P., Bresfelean, M., Ghisoiu, N., & Comes, C. A. (2008, June). Determining students' academic failure profile founded on data mining methods. In *ITI 2008-30th International Conference on Information Technology Interfaces* (pp. 317-322).

Bousbia, N., & Belamri, I. (2014). Which contribution does EDM provide to computer-based learning environments?. In *Educational data mining* (pp. 3-28). Springer.

Bunkar, K., Singh, U. K., Pandya, B., & Bunkar, R. (2012, September). Data mining: Prediction for performance improvement of graduate students using classification. In *2012 Ninth International Conference on Wireless and Optical Communications Networks (WOCN)* (pp. 1-5).

Büyükkıdık, S., Bakırarar, B., & Bulut, O. (2018, September). Comparing the performance of data mining methods in classifying successful students with scientific literacy in PISA 2015. *6th Internatıonal Congress on Measurement and Evaluatıon in Educatıon and Psychology,* Prizren, Kosova.

Cassady, J. C., & Johnson, R. E. (2002). Cognitive Test Anxiety and Academic Performance. *Contemporary Educational Psychology, 27*(2), 270–295*.* https://doi.org/10.1006/ceps.2001.1094

Culler, R. E., & Holahan, C. J. (1980). Test anxiety and academic performance: The effects of study-related behaviors. *Journal of Educational Psychology, 72*(1), 16–20. https://doi.org/10.1037/0022-0663.72.1.16

Costa, E. B., Fonseca, B., Santana, M. A., de Araújo, F. F., & Rego, J. (2017). Evaluating the effectiveness of educational data mining techniques for early prediction of students' academic failure in introductory programming courses. *Computers in Human Behavior*, *73*, 247-256. https://doi.org/10.1016/j.chb.2017.01.047

Chevalier, A., & Lanot, G. (2002). The relative effect of family characteristics and financial situation on educational achievement. *Education Economics*, *10*(2), 165-181. https://doi.org/10.1080/09645290210126904

D’Agostino, A., Schirripa Spagnolo, F., & Salvati, N. (2022). Studying the relationship between anxiety and school achievement: evidence from PISA data. *Statistical Methods & Applications*, *31*(1), 1-20. https://doi.org/10.1007/s10260-021-00563-9

Else-Quest, N. M., Hyde, J. S., & Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: a meta-analysis. *Psychological Bulletin*, *136*(1), 103-127. https://doi.org/10.1037/a0018053

Feingold, A. (1988). Cognitive gender differences are disappearing. *American Psychologist, 43*, 181–191. doi: 10.1037/0003-066X.43.2.95

Firdausi, I., Erwin, A., & Nugroho, A. S. (2010, December). Analysis of machine learning techniques used in behavior-based malware detection. In *2010 second international conference on advances in computing, control, and telecommunication technologies* (pp. 201-203). IEEE. https://doi.org/10.1109/ACT.2010.33

Fraenkel, J. R., Wallen, N. E., & Hyun, H. H. (2012). *How to design and evaluate research in education* (8th ed.). Mc Graw Hill Higher Education.

Fuchs, T., & Wößmann, L. (2008). What accounts for international differences in student performance?: A re-examination using PISA data. In *The economics of education and training* (pp. 209-240). Physica-Verlag HD.

Gamazo, A., & Martínez-Abad, F. (2020). An exploration of factors linked to academic performance in PISA 2018 through data mining techniques. *Frontiers in Psychology*, *11*, 575167. https://doi.org/10.3389/fpsyg.2020.575167

Gilleece, L., Cosgrove, J., & Sofroniou, N. (2010). Equity in mathematics and science outcomes: Characteristics associated with high and low achievement on PISA 2006 in Ireland. *International Journal of Science and Mathematics Education*, *8*(3), 475-496. https://doi.org/10.1007/s10763-010-9199-2

Gunderson, E. A., Park, D., Maloney, E. A., Beilock, S. L., & Levine, S. C. (2018). Reciprocal relations among motivational frameworks, math anxiety, and math achievement in early elementary school. *J. Cogn. Dev. 19*, 21–46. https://doi.org/10.1080/15248372.2017.1421538

Han, J., & Gao, J. (2008). Research challenges for data mining in science and engineering. In Kargupta, H., Han, J., Philip, S. Y., Motwani, R., & Kumar, V. (Eds.), *Next generation of data mining* (pp. 27-52). Chapman and Hall/CRC Press.

Hämäläinen, W., & Vinni, M. (2010). Classifiers for educational technology. In Romero, C., Ventura, S., Pechenizkiy, M., & Baker, R. S. (Eds.) *Handbook on educational data mining (pp. 57-71.)*. CRC Press.

Hand, D., Mannila, H., & Smyth, P. (2001). *Principles of data mining (Adaptive computation and machine learning).* MIT Press.

Hembree, R. (1990). The nature, effects, and relief of mathematics anxiety. *Journal for Research in Mathematics Education, 21* (1), 33–46. https://doi.org/10.5951/jresematheduc.21.1.0033

Hertel, S., & Jude, N. (2016). Parental support and involvement in school. In *Assessing Contexts of Learning* (pp. 209-225). Springer.

Hyde, J. S., Fennema, E., & Lamon, S. J. (1990). Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin*, *107*(2), 139-155. https://doi.org/10.1037/0033-2909.107.2.139

Kaur, P., Singh, M., & Josan, G. S. (2015). Classification and prediction based data mining algorithms to predict slow learners in education sector. *Procedia Computer Science*, *57*, 500-508. https://doi.org/10.1016/j.procs.2015.07.372

Kılıç Depren, S., Aşkın, Ö. E., & Öz, E. (2017). Identifying the classification performances of educational data mining methods: a case study for TIMSS. *Educational Sciences: Theory & Practice*, *17*(5), 1605–1623. https://doi.org/10.12738/estp.2017.5.0634

Keller, L., Preckel, F., Eccles, J. S., & Brunner, M. (2022). Top-performing math students in 82 countries: An integrative data analysis of gender differences in achievement, achievement profiles, and achievement motivation. *Journal of Educational Psychology*, *114*(5), 966-991. https://doi.org/10.1037/edu0000685

Koyuncu, İ., & Gelbal, S. (2020). Comparison of data mining classification algorithms on educational data under different conditions. *Journal of Measurement and Evaluation in Education and Psychology*, *11*(4), 325-345. https://doi.org/10.21031/epod.696664

Lee, J., Kao, H. A., & Yang, S. (2014). Service innovation and smart analytics for industry 4.0 and big data environment. *Procedia Cirp*, *16*, 3-8. https://doi.org/10.1016/j.procir.2014.02.001

Lee, J., & Stankov, L. (2018). Non-cognitive predictors of academic achievement: Evidence from TIMSS and PISA. *Learning and Individual Differences*, *65*, 50-64. https://doi.org/10.1016/j.lindif.2018.05.009

Linnakylä, P. & Malin, A. (2008). Finnish students' school engagement profiles in the light of PISA 2003, *Scandinavian Journal of Educational Research, 52*(6), 583-602. https://doi.org/10.1080/00313830802497174

Liu, O. L., & Wilson, M. (2009). *Gender differences in large-scale math assessments: PISA trend 2000 and 2003. Applied Measurement in Education, 22(2), 164–184.* https://doi.org/10.1080/08957340902754635

Liu, X., & Whitford, M. (2011). Opportunities-to-learn at home: profiles of students with and without reaching science proficiency. *Journal of Science Education and Technology, 20*(4), 375–387. https://doi.org/10.1007/s10956-010-9259-y

Ma, X. (1999). A meta-analysis of the relationship between anxiety toward mathematics and achievement in mathematics. *Journal for Research in Mathematics Education, 30* (5), 520–540. https://doi.org/10.2307/749772

Martínez-Abad, F. (2019). Identification of factors associated with school effectiveness with data mining techniques: Testing a new approach. *Frontiers in Psychology*, *10*, 1-13. https://doi.org/10.3389/fpsyg.2019.02583

Martínez-Abad, F., Gamazo, A., & Rodríguez-Conde, M.-J. (2020). Educational data mining: identification of factors associated with school effectiveness in PISA assessment. *Studies in Educational Evaluation, 66,* 2-10. https://doi.org/10.1016/j.stueduc.2020.100875

Masci, C., Johnes, G., & Agasisti, T. (2018). Student and school performance across countries: A machine learning approach. *European Journal of Operational Research*, *269*(3), 1072-1085. https://doi.org/10.1016/j.ejor.2018.02.031

OECD (2016). *PISA 2015 results in focus*. OECD Publishing. Retrieved https://www.oecd.org/pisa/pisa-2015-results-in-focus.pdf

OECD (2017a). *PISA 2015 technical report*. Organisation for Economic Co-Operation and Development, PISA, OECD Publishing.

OECD (2017b). *PISA 2015 assessment and analytical framework: science, reading, mathematic, financial literacy and collaborative problem solving*, PISA, OECD Publishing. https://doi.org/10.1787/9789264281820-en

Põder, K., Lauri, T., Ivaniushina, V., & Alexandrov, D. (2016). Family background and school choice in cities of Russia and Estonia: Selective agenda of the Soviet past and present. *Studies of Transition States and Societies*, *8*(3), 5-28.

Ramesh, V., Parkavi, P., & Ramar, K. (2013). Predicting student performance: a statistical and data mining approach. *International Journal of Computer Applications*, *63*(8), 35-39.

Reilly, D., Neumann, D. L., & Andrews, G. (2019). Investigating gender differences in mathematics and science: Results from the 2011 Trends in Mathematics and Science Survey. *Research in Science Education, 49*(1), 25-50. https://doi.org/10.1007/s11165- 017-9630-6

Romero, C., & Ventura, S. (2007). Educational data mining: A survey from 1995 to 2005. *Expert Systems with Applications*, *33*(1), 135-146. https://doi.org/10.1016/j.eswa.2006.04.005

Romero, C., & Ventura, S. (2010). Educational data mining: A review of the state of the art. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, *40*(6), 601-618. https://doi.org/10.1109/TSMCC.2010.2053532

Romero, C., & Ventura, S. (2013). Data mining in education. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *3*(1), 12-27. https://doi.org/10.1002/widm.1075

Romero, C., Ventura, S., Pechenizkiy, M., & Baker, R. S. (Eds.). (2010). *Handbook of educational data mining*. Data mining and knowledge discovery series. Chapman and Hall/CRC Press.

Saarela, M., Yener, B., Zaki, M. J., & Kärkkäinen, T. (2016). Predicting math performance from raw large-scale educational assessments data: a machine learning approach. In *JMLR Workshop and Conference Proceedings;* JMLR: W&CP, 48.

Sherman, B. F., & Wither, D. P. (2003). Mathematics anxiety and mathematics achievement. *Mathematics Education Research Journal*, *15*(2), 138-150. https://doi.org/10.1007/bf03217375

Siemens, G., & d Baker, R. S. (2012, April). Learning analytics and educational data mining: towards communication and collaboration. In *Proceedings of the 2nd international conference on learning analytics and knowledge* (pp. 252-254). ACM.

Singh, K., Granville, M., & Dika, S. (2002). Mathematics and science achievement: Effects of motivation, interest, and academic engagement. *The Journal of Educational Research*, *95*(6), 323-332. https://doi.org/10.1080/00220670209596607

Slavin, E. R. (1983). When does cooperative learning increase student achievement?. *Psychological Bulletin*, *94*(3), 429-445.

Taş, U. E., Arıcı, Ö., Ozarkan, H. B., & Özgürlük, B. (2016). *PISA 2015 national report.* Ministry of Education Publishing.

Tocci, C. M., & Engelhard Jr, G. (1991). Achievement, parental support and gender differences in attitudes toward mathematics. *The Journal of Educational Research*, *84*(5), 280-287. https://doi.org/10.1080/00220671.1991.10886028

Wine, J. (1971*). Test anxiety and direction of attention. Psychological Bulletin, 76(*2), 92–104. https://doi.org/10.1037/h0031332

Witten, I. H., & Frank, E., (2005). *Data mining: Practical machine learning tools and techniques (Second Edition)*. Morgan Kaufmann Publishers, San Francisco, CA.

Wu, S. S., Willcutt, E. G., Escovar, E., and Menon, V. (2014). Mathematics achievement and anxiety and their relation to internalizing and externalizing behaviours. *Journal of Learning Disabilities, 47*, 503–514. https://doi.org/10.1177/0022219412 473154

Yayan, B., & Berberoglu, G. (2004). A re-analysis of the TIMSS 1999 mathematics assessment data of the Turkish students. *Studies in Educational Evaluation*, *30*(1), 87-104. https://doi.org/10.1016/j.stueduc.2004.03.005

Yukselturk, E., Ozekes, S., & Kılıç Türel, Y. (2014). Predicting dropout student: an application of data mining methods in an online education program. *European Journal of Open, Distance and e-learning*, *17*(1), 118-133. https://doi.org/10.2478/eurodl-2014-0008

Zhang, J., Zhao, N., & Kong, Q. P. (2019). The relationship between math anxiety and math performance: A meta-analytic investigation. *Frontiers in Psychology*, *10*, 1-17. https://doi.org/10.3389/fpsyg.2019.01613