International Journal of Assessment Tools in Education
2023, Vol. 10, No. 3, 413–433

https://doi.org/10.21449/ijate.1218659

Published at https://ijate.net/ https://dergipark.org.tr/en/pub/ijate *Research Article*

Examining the factors affecting students' science success with Bayesian networks

Hasan Aykut Karaboğa [1,*], **Ibrahim Demir** [2]

[1]Amasya University, Faculty of Education, Department of Educational Sciences, Amasya, Türkiye
[2]Turkish Statistical Institute, Ankara, Türkiye

ARTICLE HISTORY

Received: Dec. 14, 2022
Revised: May 11, 2023
Accepted: Aug. 11, 2023

Keywords:
Bayesian Networks,
Educational Data Mining,
PISA,
Science Success.

Abstract: Bayesian Networks (BNs) are probabilistic graphical statistical models that have been widely used in many fields over the last decade. This method, which can also be used for educational data mining (EDM) purposes, is a fairly new method in education literature. This study models students' science success using the BN approach. Science is one of the core areas in the PISA exam. To this end, we used the data set including the most successful 25% and the least successful 25% students from Turkey based on their scores from Program for International Student Assessment (PISA) survey. We also made the feature selection to determine the most effective variables on success. The accuracy value of the BN model created with the variables determined by the feature selection is 86.2%. We classified effective variables on success into three categories; individual, family-related and school-related. Based on the analysis, we found that family-related variables are very effective in science success, and gender is not a discriminant variable in this success. In addition, this is the first study in the literature on the evaluation of complex data made with the BN model. In this respect, it serves as a guide in the evaluation of international exams and in the use of the data obtained.

1. INTRODUCTION

The world is constantly changing with technological developments. In today's technology-oriented society, especially success in science is directly related to understanding and applying basic scientific knowledge and ensuring the scientific progress of the country by utilizing science and technology in daily life (OECD, 2019a, 2019b). People in 21st century, have to solve a continuous series of daily problems for living in the today's world (Gilbert et al., 2000). International exams such as Program for International Student Assessment (PISA) and Trends in International Mathematics and Science Study (TIMSS) measure and compare the science successes of the countries.

PISA has an important impact on educational systems and policies (Deng & Gopinathan, 2016). A functional and well-structured educational system is the way to increase achieving future goals set by a country (Sağlam & Aydoğmuş, 2016). The development levels of societies are closely related to the education their students receive. Quality education increases career opportunities, affects economic and cultural development and helps people to increase their

*CONTACT: Hasan Aykut Karaboğa h.aykut.karaboga@amasya.edu.tr Amasya University, Faculty of Education, Department of Educational Sciences, Amasya, Türkiye

e-ISSN: 2148-7456 /© **IJATE** 2023

social status. Determining the factors that enhance the quality of education and improving those factors influences the international success of a country.

International exams measure many factors that affect student success. These exams enable educational authorities in countries to not only assess students' achievements against basic educational standards but also compare the results with those of other countries (Gamazo & Martínez-Abad, 2020; Schleicher, 2019). Assessments in international exams help us explore the relationships between student achievement and students themselves, as well as between student achievement and both schools and education systems. By identifying the factors that influence student success, stakeholders can take necessary steps to increase the levels of low achievement in education (Aşkın & Öz, 2020). It is important to determine the functioning or problematic parts of the education system according to scientific data. Thus, shaping future education policies according to the available data will increase the quality of education (Üstün et al., 2020). Researchers have conducted numerous modeling studies using data derived from international exams. Furthermore, the literature has explored various studies examining the factors that influence academic success (Altun & Kalkan, 2019; Carnoy et al., 2015; Chen et al., 2019; Gamazo & Martínez-Abad, 2020; Karakoç Alatlı, 2020, 2020; Kilic Depren, 2018; Kiray et al., 2015; Kjærnsli & Lie, 2004; Lee & Shute, 2010; Rastrollo-Guerrero et al., 2020; Sebastian et al., 2017; Sheldrake et al., 2017; Sirin, 2005; Tang & Zhang, 2020; Topçu et al., 2015; Torrecilla Sánchez et al., 2019; Yip et al., 2004; Yıldırım, 2012). However, researchers face difficulty in modeling the complex relationships (Kiray et al., 2015; Lee & Shute, 2010) between the variety of factors that impact success (Martínez Abad & Chaparro Caso López, 2017). Choosing the appropriate modeling strategy ensures that the findings are a source for educational systems (Aşkın & Öz, 2020).

Successful modeling results were obtained using EDM. The biggest advantages of EDM methods are that they can work with complex related data sets and do not have restrictive statistical assumptions such as variance homogeneity and linearity (Sinharay, 2016). EDM, the use of classical data mining techniques in education (Baker & Yacef, 2009; Romero & Ventura, 2010; Shin & Shim, 2021), provides practical information for educational policy makers and researchers in increasing success (Peña-Ayala, 2014; Romero & Ventura, 2010). In this study, Bayesian Networks (BN), an EDM approach used in few studies in the educational literature, was preferred to model the relationships of variables affecting high and low science success scores. BN create graphical models of the dependency relations of all variables (Nielsen & Jensen, 2009). Unlike other machine learning models, BN enables queries which explicitly reveal variables' cause-effect relationships (Pearl, 2014). So, BN provides an advantage over other machine learning methods in revealing complex relationships (Karaboga et al., 2021).

1.1. Research Problem

According to the PISA 2018 results, Turkey increased its performance in all fields compared with the 2015 results, but remained below the average of the OECD countries. Based on the 2018 PISA results, Turkey has an average of 468 points in science and OECD average is 489 points. The country became 39th in science with this score. However, Turkey ranked 54th among 72 countries in science in PISA 2015. According to the OECD's report, Turkey was the only country to experience improvement in all three areas, in despite of the number of students in the 15-year-old group increased significantly between 2003 and 2018 (MEB, 2019).

Despite the significant structural changes made in the Turkish education system, the desired level of success could not be reached in international exams such as PISA. For this reason, in order to achieve a higher performance, factors influencing success should be determined and Turkey should focus on areas of improvement to obtain better results from international exams for years to come. We used the BN model to identify key factors for improving students' science success in international exams, as we believe it is an effective tool for this purpose.

BNs are used in different applications in education literature. Researchers have used BN for various purposes in EDM such as predicting student proficiency levels (Almond & Mislevy, 1999; Desmarais & Baker, 2012), predicting course performance (Xing et al., 2021; Zwick & Lenaburg, 2009), smart classroom applications (Saini & Goel, 2019), evaluation of intelligent tutoring systems (Ramírez-Noriega et al., 2021), student knowledge assessment system (Levy, 2016; Millán et al., 2013; Xing et al., 2021), cognitive diagnostic modeling (Almond et al., 2007) and educational assessment (Culbertson, 2016). But, only one BN study using the PISA data to investigate the relationship of influential factors with mathematics achievement conducted by Tingir and Almond (2017) was found. However, we could not find any study in the literature on science achievement.

This study seeks to contribute to the literature on modeling science success by using BN with pre-determined variables. Initially, a BN model was developed with these variables, but it was found to be difficult to interpret the model structure and the interrelationships between variables. In order to create a more comprehensible and practical model, we utilized feature selection to identify the most effective variables and reduce their total number. It is noteworthy that BN with feature selection has not been previously applied to PISA data to determine the factors influencing science success and their interrelationships. As such, we conducted a comparative analysis of two distinct BN models for science success using PISA 2018 Turkey data, with one model including all variables and the other model including only selected features. This study is the first of its kind to examine the combined use of feature selection and BN modeling in evaluating science success. The BN model we developed evaluates student science success in PISA 2018 by modeling complex relationships among science-effective variables. Compared to other statistical models, BN offers advantages by transforming complex relationships into interpretable knowledge. We argue that our research will have significant implications for evaluating studies using educational data sets.

1.2. Research Focus

In this study, the following research questions guided the study to investigate the factors affecting the science success of Turkish students in PISA 2018:

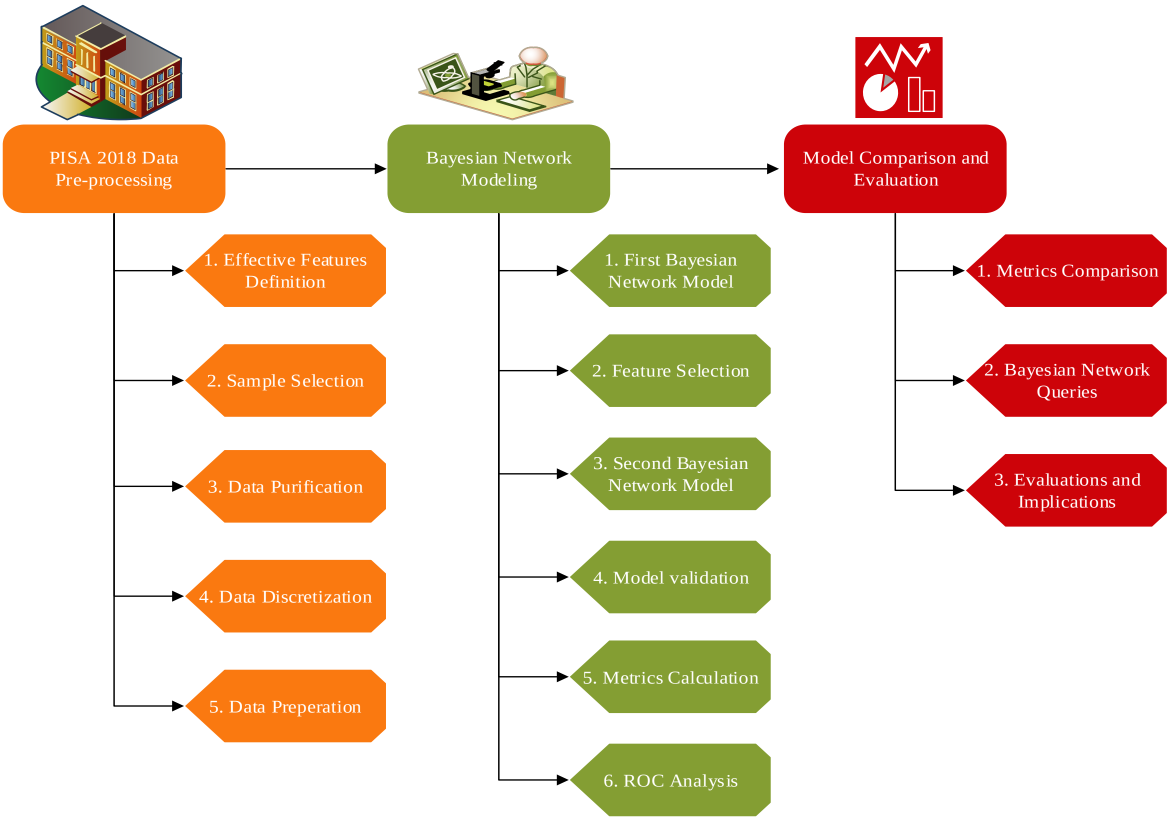
1. What are the factors affecting students' science success?
2. Do the factors affecting students' science success differ according to their high and low success levels? If so, what are the factors that increase success?
3. Is there any performance difference between the two models in terms of science success?
4. Are there any performance differences between male and female students in models?
5. What measures will raise the success level of Turkish students in science?

2. METHOD

In this section, we summarized the sample and explain the steps of the BN design with feature selection methods used to model science success with success-related feature interactions. We describe the implementation stages of the study in Figure 1.

In the first step, we preprocessed the data for BN modelling. The second step, we created the first BN model after the determination of the data related to science success. In the next step, we evaluated the relations in the BN model and obtained results. Considering the obtained results, the next step was feature selection. After this step, we created a second BN model with the selected features. In the last step, we compared the obtained results for these two models and evaluated their implications.

Figure 1. *Graphical representation of research process.*



2.1. Research Sample

The OECD has conducted PISA every 3 years since 2000. The PISA test consists of school, student, and teacher questionnaires. We employed student's school, family and individual evaluations in our study. In the questionnaire data set, besides demographic variables such as gender and age of the student, there are also indexes such as socio-economic status, family wealth, highest parental occupational status which were constructed through Item-Response Theory (OECD, 2019a).

In statistical analysis, missing values should be excluded from the dataset. Thus, 4276 students who remained after the missing values were eliminated. After that, 2138 students representing the most successful 25% and the least successful 25% were included in the analysis. In the analysis, personal variables such as the students gender and study time as well as school and family variables were used. In order to work with BN, the data was discretized (Nojavan et al., 2017; Yang & Webb, 2002). Modeled variables are shown in Table 1.

Table 1. *Student-related model variables*

	Code	Label	Variable Type
Q1	ATTLNACT	Attitude toward school	*Student*
Q2	BEINGBULLIED	Student experience of being bullied	*Student*
Q3	BELONG	Sense of belonging to school	*Student*
Q4	CLSIZE	Class size	*School*
Q5	COMPETE	Competitiveness	*Student*
Q6	CREACTIV	The index of creative extracurricular activities at school	*School*
Q7	EDUSHORT	The scale of the shortage of educational material	*School*
Q8	EMOSUPS	Parents' emotional supports perceived by students	*Parental*
Q9	ESCS	Economic, social and cultural status index	*Parental*
Q10	EUDMO	Eudaemonia: meaning in life	*Student*
Q11	GFOFAIL	The general fear of failure	*Student*
Q12	HISCED	Highest parental education	*Parental*
Q13	HISEI	Highest parental occupational status index	*Parental*
Q14	IC150Q03HA	Digital devices using time during science lessons (In a typical week)	*School*
Q15	ICTHOME	ICT available at home	*Parental*
Q16	MASTGOAL	Mastery goal orientation	*Student*
Q17	PARED	Highest parental education in years of schooling index	*Parental*
Q18	PERCOMP	Perception of competitiveness at school	*Student*
Q19	PERCOOP	Perception of cooperation at school	*Student*
Q20	RESILIENCE	Resilience	*Student*
Q21	ST004D01T	Student gender	*Student*
Q22	STAFFSHORT	The scale of staff shortage	*Student*
Q23	STRATIO	The student-teacher ratio	*School*
Q24	STUBEHA	Student behavior hindering learning	*Student*
Q25	SWBP	Subjective well-being: Positive effect	*Student*
Q26	TEACHBEHA	Teacher behavior hindering learning	*School*
Q27	TEACHINT	Perceived teacher interest	*School*
Q28	TMINS	Total learning time (minutes per week)	*Student*
Q29	SUCCESS	Science Success (lowest 25% and highest 25%)	*Dependent*

2.2. Bayesian Network

Bayesian Networks (BN) are statistical models that graphically display the common probability distributions of variables in addition to their dependency relations of variables (Nielsen & Jensen, 2009). In a BN model, variables are represented as nodes and relationships between variables are represented as edges. Edges are oriented as one-way arrows and indicate the structure of the network. Structure of the BN is specified as DAG (Directed Acyclic Graph) (Neapolitan, 2009). The established DAG structure can be used to make inferences on the parameters of the model using mathematical equations. In other words, the most important feature of BN is the ability to update the probabilities of each node in the entire model with new information (Sener et al., 2019).

As a graphical model, the DAG structure is shown as ***G=(A, B)***, where ***A*** is the set of nodes and ***B*** is the set of edges that provides the nodes' connections. In a BN -containing the variable ***M***-each node ***X*** is associated with the conditional probability distribution of the corresponding variable considering its parents. The conditional probability of a node is given in Equation 1. This probability value is called conditional probability distribution when the $pa(X_i)$ values of the X node are given.

$$p\left(X|pa\left(X_i\right)\right) \qquad \left(i=1,\ldots,M;M\in A\right) \tag{1}$$

The joint probability distribution calculated for the $(X_1,\ldots,X_M)$ nodes in the whole model is given in Equation 2.

$$p\left(X_1,\ldots,X_M\right)=\prod_{i\in A}p\left(X_i|pa\left(X_i\right)\right) \tag{2}$$

The contribution of variables to the model originates from the conditional probability values, which are calculated when $pa(X_i)$ is given.

2.3. Feature Selection

Feature selection methods play an important role in machine learning, particularly in situations where the number of features is high relative to the number of observations. The feature selection aims to identify the most relevant and informative subset of features, which can improve the model's accuracy, reduce overfitting, and enhance interpretability. In this study, we used correlation-based feature selection named the CFS subset algorithm.

Correlation is one of the most important indicators showing the relationship between two variables. One popular feature selection algorithm is the Correlation-based Feature Selection (CFS) algorithm. CFS subset algorithm was introduced by Hall (1999a). CFS is a filter method that evaluates the features based on their correlation with the class variable and with each other. CFS aims to identify features that are highly correlated with the class variable while minimizing redundancy among the features. The CFS algorithm works by calculating a merit score for each feature, which is based on the correlation between the feature and the class variable, as well as the correlation between the feature and the other features. The merit score is used to rank the features, and a subset of the top-ranked features is selected. CFS is effective in improving the performance of machine learning algorithms by reducing the number of irrelevant and redundant features. This algorithm selects features with low correlation between them and high correlation between class tags (Hall, 2000). The CFS selection coefficient - the equation is the standardized Pearson correlation of all variables- was calculated for each subset (Hall, 1999b).

2.4. Classification Criteria

In this study, accuracy, F-Measure, Mean Absolute Percentage Error (MAPE), Kappa (κ), Root Mean Square Error (RMSE), and ROC area were used in the evaluation of model performance. Accuracy is the overall correct classification rate in the positive and negative cluster which is one of the most common performance measure (Ferri et al., 2009). F-measure is the harmonic mean of correctly classified positive and negative values (Hossin & Sulaiman, 2015). The Kappa coefficient deals with the prediction performance of an algorithm. The closer the Kappa coefficient is to 1, the higher predictive performance of the model. The MAPE value could measure the difference between the expected and predicted results. The MAPE value of models with high predictive performance converges to zero. The RMSE is a quadratic metric that measures the magnitude of error by finding the distance between predicted and actual values. RMSE is a measure of how far these errors are propagated.

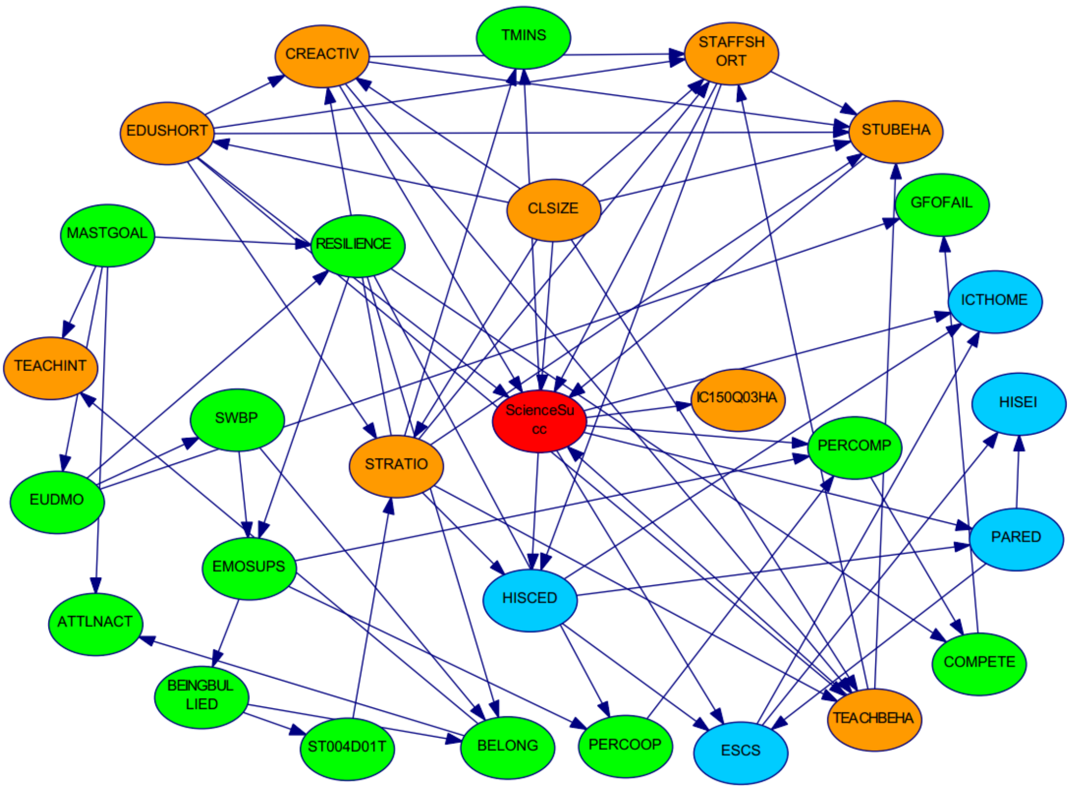
The ROC area namely 'Area under the receiver operating curve (AUC) value' measures the ability of the model to avoid errors during class estimation. The AUC is closely related to specificity and sensitivity values. This value is a measure used in conjunction with the ROC curve to show whether a perfect classification has been made (Marsland, 2015).

3. RESEARCH RESULTS

It is not only variables related to students themselves that affect their science success but family and school-related variables are closely related to their success (Kilic Depren, 2018). Variables are grouped under three sub-headings: variables about the student themselves, variables about his/her family, and variables about his/her school. In the first stage, these variables are discrete to be used in Bayesian networks. Therefore, we preferred the quantile discretization method commonly used in discretizing (Lima, 2014; Ropero et al., 2018).

We utilized academic version of the GeNIe program for BN modeling and we preferred k-fold cross-validation method for model evaluation (BayesFusion, 2017). The quartile values were used to discretize the variables. Thus, the variables were represented in 4 different ways from Q1 to Q4 (very low, low, high, very high). The greedy tick thinning algorithm was used in the analysis. In this technique, the data set is divided into k parts; k-1 parts of the data are used for training and the other part of the data are used for testing (Wong, 2015). Finally, we obtained the classification success performance by calculating the mean error of the k tests pieces (Karaboga et al., 2021). In this study, the k value was taken as 10.

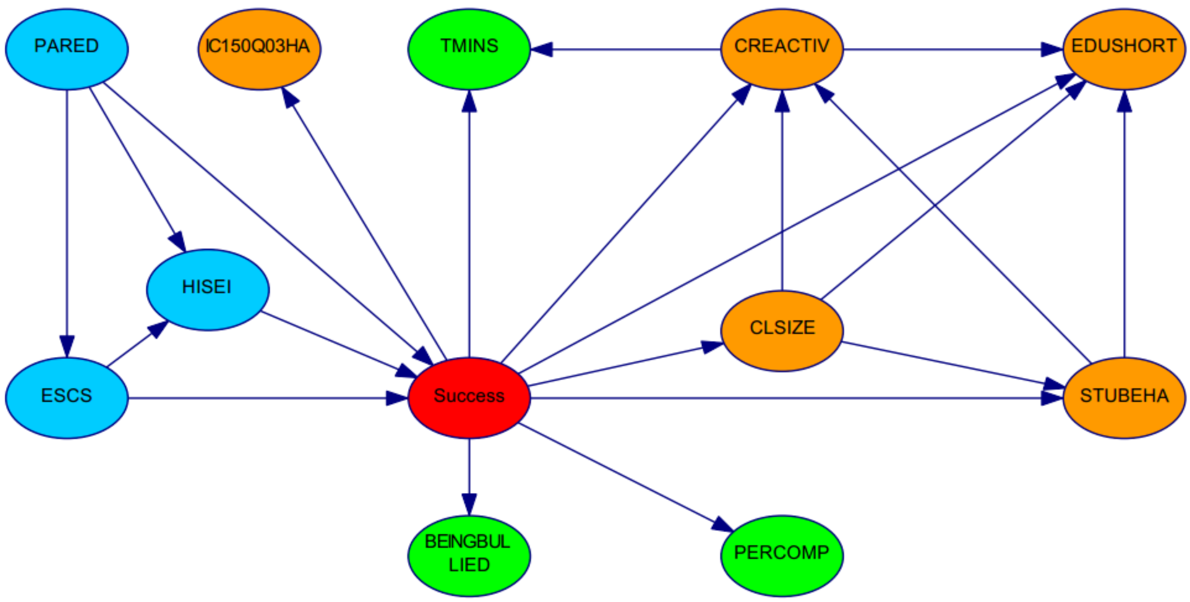
Figure 2. *BN model with 28 variables.*



In the first step, we constructed a BN model with 28 variables that affect science success. The variables were divided into 3 groups in the model: blue group variables are the student's family related variables, green group variables are the student's individual variables, and orange group variables are the student's school related variables. As a result, we obtained 89.4% accuracy from the model shown in Figure 2. However, the model and relationships of the variables were quite complex to understand and interpret.

As the Parsimony principle requires (Zhang, 1992), we reduced the number of variables by using expert knowledge and feature selection for simplifying the complex model structure suggested by the algorithm as well as making it more meaningful. In the second step, we selected 11 variables due to feature selection performed using the CFS subset algorithm. In the last step, we reconstruct the model with 11 effective variables. The final model is shown in Figure 3. The performance of this model is also close to the first model (Accuracy = 86.2%).

Figure 3. *BN Model with 11 Variables.*



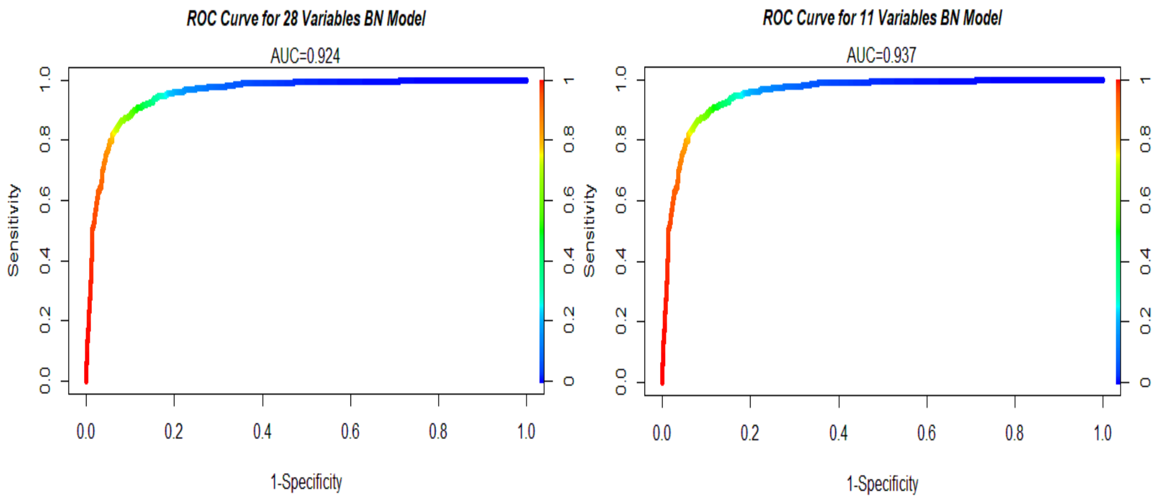
The reduced model produced a more meaningful with fewer variables. The comparison results of the models are shown in Table 2. When the models are compared, the success variable's prediction performance of the models is close to each other.

Table 2. *Model comparison results.*

	BN with 28 variables			BN with 11 variables		
	Female	Male	Overall	Female	Male	Overall
Accuracy	0.859	0.840	0.849	0.863	0.862	0.862
F-Measure	0.859	0.840	0.849	0.863	0.862	0.862
Kappa	0.717	0.679	0.699	0.726	0.723	0.725
RMSE	0.376	0.400	0.388	0.370	0.372	0.371
MAPE	10.664	11.319	10.992	10.617	9.916	10.267

We applied models separately for male and female students to investigate model performance differences, and no differences were found in terms of evaluation criteria. In the first and second models, we observed that male and female students differed by approximately 1% according to the MAPE value. In the literature, gender is effective on science success (Aşkın & Öz, 2020; Harker, 2000; Kilic Depren, 2020; Kjærnsli & Lie, 2004; Reilly et al., 2019; Torrecilla Sánchez et al., 2019; Yip et al., 2004). In this study, however, gender was not an effective variable on science success and prediction performance.

Figure 4. *ROC curves of BN models.*



ROC curves of the BN models obtained with ROCR package (Sing et al., 2005) are given in Figure 4. It was understood that the second model produced 0.937 AUC in the prediction of students' science success. AUC of the second model is better than that of the first model. In a BN model, if we know the value of the any factors, we can build scenarios to predict the student's high and low success probabilities with this new knowledge. The success prediction scenarios of the student-based variables are given in Table 3.

Table 3. *Success prediction scenarios with student-based variables.*

	Evidence	SUCCESS	
		Very Low	Very High
BEING BULLIED	Very Low	0.731	0.269
	Low	0.498	0.502
	High	0.317	0.683
	Very High	0.613	0.387
PERCOMP	Very Low	0.612	0.388
	Low	0.449	0.551
	High	0.499	0.501
	Very High	0.402	0.598
TMINS	Very Low	0.653	0.347
	Low	0.578	0.422
	High	0.348	0.652
	Very High	0.665	0.335

In Table 3, we examined low and high science success according to the values of the variables of peer bullying (BEING BULLIED), perceived competition (PERCOMP), and total studying time in minutes (TMINS). When perceived bullying is very low, the probability of low success is 0.731 and the probability of high success is 0.269. Also, in the case of very high perceived bullying, the probability of low science success is 0.613 and the probability of high science success is 0.387. On the other hand, it is seen that successful students are more likely to be exposed to bullying. In other words, unsuccessful students fail not because of being bullied but because of other reasons. It is understood that successful students are exposed to more intense peer bullying. Considering perceived competitiveness, the probability of low success in the case of the perceived low competitiveness is 0.612, whereas it is calculated as 0.402 in the case of high competitiveness. It is seen that perceived competitiveness increases high success (0.598). Also, too much or too little studying of the student affects success negatively. We found that studying time above average positively affects science success (0.652).

Table 4. *Success prediction scenarios with parental variables.*

	Evidence	Success	
		Very Low	*Very High*
ESCS	Very Low	0.699	0.301
	Low	0.634	0.366
	High	0.522	0.478
	Very High	0.229	0.771
HISEI	Very Low	0.668	0.332
	Low	0.619	0.381
	High	0.534	0.466
	Very High	0.249	0.751
PARED	Low	0.617	0.383
	Moderate	0.680	0.320
	High	0.549	0.451
	Very High	0.257	0.743

The relationship between the student's family-related variables and science success is shown in Table 4. We observed that when the student's ESCS value is low, their success is also low, and when the student's ESCS value is high, their success is also high. Considering the index highest parental occupational status (HISEI) value, we found that if this value is too low, the success is also low (0.668) and that if high, the science success is very high. Finally, when we investigated the relationship between education level of family (PARED) and science success, we revealed that the student's science success was low (0.617) in the case of a low level of parental education, and high when the level of parental education was very high (0.743).

The relationship between the student's school-related variables and science success is given in Table 5. As seen in the table, the probability of science success is quite low in classes with fewer than 25 students (0.169). It is seen that the ideal class size is between 31 and 35. In schools where no creative activities (CREATIV) are carried out, the probability of students' science success is low (0.756). Choir and music events are generally held in Turkish schools. Therefore, no positive effect of these activities on success has been observed. However, artistic activities had a very positive effect on students' science success (0.706). In other words, artistic activities carried out at school should play an active role in increasing student success. When the dataset is examined in detail, science high schools and private high schools have more artistic activities and more successful students. In addition, it is observed that the families of the students in these schools are educated and the lack of educational materials is less.

Shortage of educational material (EDUSHORT) also has a negative impact on science success. In this sample, we observed 3 parts of shortage: low, high, and very high. The probability of science success increases (0.610) when the shortage is low. However, students show low success when the lack of teaching and learning materials is very high (0.715). Digital device use in lessons positively affects science success. Moreover, we have seen that using digital devices for at least 60 min weekly in science lessons increases the students' science success (0.804). Students who declare that they do not work are more likely to fail (0.701). It could be stated that supporting the course with a digital device in science lessons increases the student's learning and thus the possibility of high success.

Student behavior hinder learning negatively influence success. As a result of the study, when students have fewer disruptive behaviors, their success probability increases (0.773), and when students display too many disruptive behaviors, science success is quite low (0.779).

Table 5. *Success prediction scenarios with school-related variables.*

Evidence		Success	
		Very Low	*Very High*
CLSIZE	Less than 25	0.831	0.169
	Between 26-30	0.432	0.568
	Between 31-35	0.368	0.632
	Between 36-50	0.700	0.300
	More than 50	0.468	0.532
CREACTIV	None	0.756	0.244
	Art club activities	0.294	0.706
	Band orchestra choir	0.612	0.388
	School play musical	0.664	0.336
EDUSHORT	Low	0.390	0.610
	High	0.486	0.514
	Very high	0.715	0.285
IC150Q03HA	I don’t study	0.701	0.299
	No time	0.622	0.378
	Between 1-30 min	0.664	0.336
	Between 31-60 min	0.449	0.551
	More than 60	0.196	0.804
STUBEHA	Very Low	0.227	0.773
	Low	0.330	0.670
	High	0.610	0.390
	Very High	0.779	0.221

4. DISCUSSION

Primarily, this is the first BN study that has been conducted with this dataset. Although different data mining methods were used in previous studies, BN was not used to model science success. Unlike rule-based machine learning such as support vector machines, logistic regression and artificial neural networks, it enables queries which explicitly reveal cause-effect relationships between variables (Pearl, 2014). Besides, the posterior probabilities are updated with each new information, allows more accurate estimations (Korb & Nicholson, 2010). Hence, modeling with BN provides an advantage over other machine learning methods in revealing complex relationships (Karaboga et al., 2021). BN, which is widely used in a variety of fields, has been used in a small number of studies in the field of education (Almond et al., 2015; Culbertson, 2016; Reichenberg, 2018). However, BN is more advantageous than other methods with its ability to model students in the field of education (Levy, 2016; Lytvynenko et al., 2019;

Sinharay, 2006) and to evaluate the model quickly (Kenekayoro, 2018; Kustitskaya et al., 2020; Millán et al., 2013; Nguyen & Do, 2009).

Essential improvements in the education system are vital to enhance students' success. Therefore, educators, researchers, and government agencies should prioritize research for identifying factors to improve success. Especially, enhancing science success is considered as a key to the scientific and economic progress of countries (Sjøberg, 2019). Students' individual, family-related and school-related factors are effective on science success (Beese & Liang, 2010; Kiray et al., 2015; Lee & Shute, 2010; Yıldırım, 2012). PISA aims to help explain the differences in student performance by collecting data on students' successes, as well as collecting each student, family and personal information (Beese & Liang, 2010). The effect of interaction among these variables, which are normally effective separately, on success has been investigated using the advantages of BN. In this study, we used the dataset of Turkey obtained from the PISA 2018 survey. In the first step of the study, we discretized the variables.Then, we constructed a dataset that included the most successful 25% and the least successful 25% students. As a consequence, we examined the effects of the factors which influence science success by creating a model with 28 variables. The most effective variables determined with the CFS subset algorithm were BEING BULLIED, PERCOMP, TMINS, ESCS, HISEI, PARED, CLSIZE, CREACTIV, EDUSHORT, IC150Q03HA, and STUBEHA. A more effective model was obtained with these 11 determined variables.

Bullying is a type of violence which disrupts school climate and harms students' physical or mental states (Fry et al., 2018; Wachs et al., 2019). The student's success is low in the case of high perceived bullying (Clarke & Kiselica, 1997; Jan, 2015; Sudrajad et al., 2020). Successful students are exposed to more intense bullying than unsuccessful students.. Also, perceived high competitiveness increases success (Karataş & Ergin, 2018; Muñoz-Merino et al., 2014; OECD, 2020). Less disruptive behaviors of the students increase their science success. On the contrary, in schools with too many disruptive behaviors, the science success decreases (Ertem, 2021; Özdemir et al., 2019). We observed that too much or too little study of the student has a negative impact on success.

The increase in parents' socio-economic and cultural status increases the students' science success. We found that if the student's parental occupational status is too low, their success is also low, and that a very high status of parental occupation correlates with students' high science success. Similarly, students with a low level of family education have low science success, and when their family's education level is very high, their science success is very high. As a result, the economic and socio-cultural status of the student's family, their educational background, and occupational status are effective upon students' science success. According to the literature, it is clear that students having families with high educational, and sociaconomic status are more successful (Gamazo & Martínez-Abad, 2020; Lee & Shute, 2010; Sirin, 2005; Topçu et al., 2015; Yıldırım, 2012). The high science success of these students is related to their awareness of science and education. Children of educated families are also conscious about science education (O'Connell, 2019).

The class sizes of the students who participated in the survey were generally more than 50. Classes are smaller in vocational schools located in small settlements. The success level of students studying in those schools is generally low (Suna et al., 2020). The probability of success is quite low in classes where there are fewer than 25 students. Classes in Anatolian high schools are also larger than those in other schools. Medium-sized classes are provided in science high schools and private colleges. It is stated in the literature that as the classes get smaller, the success increases, but this effect is low (Borland et al., 2005; Hanushek & Woessmann, 2017; Hattie, 2005). Also, reducing class size is quite costly (Ehrenberg et al., 2001; White, 2018). Because of that, educators should determine the ideal class size, considering the situation of the students and the school (Borland et al., 2005; Wößmann, 2005). This study reveals that the ideal class size is between 31 and 35.

Generally, choir and musical events are held in Turkish schools. In these schools where there are no other activities, it is impossible to encourage students to participate in different activities. Artistic activities other than music should have a positive impact on students' science success. Extracurricular artistic school activities play an active role in increasing student success (Tang & Zhang, 2020). That's because, according to the hidden curriculum (Margolis, 2001), extracurricular activities ensure a rise in success by increasing concentration and motivation (Stearns & Glennie, 2010).

Another effective factor on science success is the lack of educational material shortage (Altun & Kalkan, 2019; Archibald, 2006). There may be a lack of educational materials at schools in various disadvantaged regions. Particularly, in socioeconomically disadvantaged regions, schools cannot fill those deficiencies by getting support from families (van der Berg, 2008). In schools where there exist few artistic activities, educational material shortage such as digital devices for lessons is higher. The use of digital devices in lessons has also been identified as a variable that positively affect success. Accordingly, we have seen that the use of digital devices in science classes increases the probability of student success. Supporting science lessons with digital devices boosts student success by facilitating their learning (Bingimlas, 2009; Chen et al., 2019; Odell et al., 2020).

Apart from the studies we mentioned, studies have been conducted on factors affecting science success such as teachers, school, school curriculum (Cansiz & Cansiz, 2019; Tatar et al., 2016). Numerical content of science subjects and the intensity of curriculum are important predictors of science success (Tatar et al., 2016). If we identify the factors affecting student success, we will guide the reforms that need to be made in the curriculum to increase students' low success level (Topçu et al., 2016).

5. CONCLUSIONS and IMPLICATIONS

Science literacy requires students to explain various phenomena scientifically, design and evaluate the scientific method, and interpret the findings scientifically (OECD, 2019a). The relationship of students' background of knowledge and skills with other variables obtained is one of the main indicators of PISA (MEB, 2019; Schleicher, 2019). To sum up, PISA evaluates how students could use their scientific content knowledge in their daily life by combining methodological and epistemic knowledge (OECD, 2019a, 2019c). In this respect, science literacy examines whether students could go beyond the school curriculum.

Based on the PISA 2018 results, Turkish students scored lower than the OECD average. Although some progress has been made compared to previous years, this progress is inadequate. The main purpose of the Turkish science curriculum is to raise science-literate students (MEB, 2018). However, science literate individuals do not grow as the curriculum aims. Hence, it is necessary to explore how students could improve their ability to use information and interpret it in real life.

Even though the most significant source of student success is internal motivation (Augustyniak et al., 2016), school and family variables are also important. In particular, opportunities provided to students by their families, and schools are a major key to success. Low-income families are a significant issue here. Nonetheless, no short-term solution to this problem exists. Instead, it is required to raise awareness in cooperation with the families of students and to organize activities that will encourage them to study. Providing an optimum studying environment and ideal teaching and learning materials will be encouraging for students.

Increasing opportunities in schools will also increase student success. Computer-assisted classes have demonstrated significant potential in enhancing students' problem-solving abilities, particularly in the domain of science education (Bayrak & Bayram, 2010; Chang, 2002). Schools must take measures to prevent peer bullying and disruptive student behaviors that hinder learning. To achieve this, collaborative efforts between schools and families are crucial in devising various policies aimed at safeguarding students. The study time of students should be maintainedat a sufficient level. Too little or too much work should harm student success. Competition and cooperation among students should be encouraged through various activities to increase science success. Motivating students to study more emerges as a key factor in attaining a long-term success rate. For this purpose, education politicians should prepare a rich curriculum based on experiments and observations to have more fun in science lessons.

PISA's science literacy qualifications are almost non-existent within the scope of Turkish science curricular outcomes (Cansiz & Cansiz, 2019). The curriculum is not sufficient to raise scientifically literate individuals. To raise individuals who research, question and use 21st century information and technologies, changes should be made and implemented in education systems. Thus, we can use the assessments obtained using the BN model to increase students' science success in future exams. We proved that the results of this study will provide effective clues for innovations in the educational system. We hope that it will be a useful model for the evaluation of international exams and contributions to educational systems not only for Turkey but also for all OECD countries.

This study is not without limitations. First, the study was conducted with PISA data only from Turkey. Nevertheless, students from different countries could also be analyzed to make the study comparative. Besides, models could be made more successful by combining BN algorithms with newly developed machine learning methods. In addition, different results can be obtained by repeating this study for different data sets.

Acknowledgments

Acknowledgments of people, grants, and funds should be placed in a separate section before the References. If the study has been previously presented at a conference or a scholarly meeting, it should be mentioned here.

Declaration of Conflicting Interests and Ethics

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

Authorship Contribution Statement

Hasan Aykut Karaboga: Investigation, Resources, Methodology Visualization, Software, Formal Analysis, and Writing-original draft. **Ibrahim Demir**: Supervision, and Validation.

Orcid

Hasan Aykut Karaboga https://orcid.org/0000-0001-8877-3267
Ibrahim Demir https://orcid.org/0000-0002-2734-4116

REFERENCES

Almond, R.G., DiBello, L.V., Moulder, B., & Zapata-Rivera, J.-D. (2007). Modeling Diagnostic Assessments with Bayesian Networks. *Journal of Educational Measurement*, *44*(4), 341–359. https://doi.org/10.1111/j.1745-3984.2007.00043.x

Almond, R.G., & Mislevy, R.J. (1999). Graphical Models and Computerized Adaptive Testing. *Applied Psychological Measurement*, *23*(3), 223-237. https://doi.org/10.1177/01466219922031347

Almond, R.G., Mislevy, R.J., Steinberg, L.S., Yan, D., & Williamson, D.M. (2015). *Bayesian Networks in Educational Assessment*. Springer.

Altun, A., & Kalkan, Ö.K. (2019). Cross-national study on students and school factors affecting science literacy. *Educational Studies*, 1-19. https://doi.org/10.1080/03055698.2019.1702511

Archibald, S. (2006). Narrowing in on Educational Resources That Do Affect Student Achievement. *Peabody Journal of Education*, *81*(4), 23-42. https://doi.org/10.1207/s15327930pje8104_2

Aşkın, Ö.E., & Öz, E. (2020). Cross-National Comparisons of Students' Science Success Based on Gender Variability: Evidence From TIMSS. *Journal of Baltic Science Education*, *19*(2), 186–200. https://doi.org/10.33225/jbse/20.19.186

Augustyniak, R.A., Ables, A.Z., Guilford, P., Lujan, H.L., Cortright, R.N., & DiCarlo, S.E. (2016). Intrinsic motivation: An overlooked component for student success. *Advances in Physiology Education*, *40*(4), 465–466. https://doi.org/10.1152/advan.00072.2016

Baker, R.S.J.d, & Yacef, K. (2009). The State of Educational Data Mining in 2009: A Review and Future Visions. *Journal of Educational Data Mining*, *1*(1), Article 1. https://doi.org/10.5281/zenodo.3554657

BayesFusion, L. (2017). GeNIe modeler user manual. *BayesFusion, LLC, Pittsburgh, PA*.

Bayrak, B.K., & Bayram, H. (2010). The effect of computer aided teaching method on the students' academic achievement in the science and technology course. *Procedia - Social and Behavioral Sciences*, *9*, 235–238. https://doi.org/10.1016/j.sbspro.2010.12.142

Beese, J., & Liang, X. (2010). Do resources matter? PISA science achievement comparisons between students in the United States, Canada and Finland. *Improving Schools*, *13*(3), 266–279. https://doi.org/10.1177/1365480210390554

Bingimlas, K.A. (2009). Barriers to the Successful Integration of ICT in Teaching and Learning Environments: A Review of the Literature. *Eurasia Journal of Mathematics, Science & Technology Education*, *5*(3), 235–245.

Borland, M.V., Howsen, R.M., & Trawick, M.W. (2005). An investigation of the effect of class size on student academic achievement. *Education Economics*, *13*(1), 73–83. https://doi.org/10.1080/0964529042000325216

Cansiz, N., & Cansiz, M. (2019). Evaluating Turkish science curriculum with PISA scientific literacy framework. *Turkish Journal of Education*, *8*(3), Article 3. https://doi.org/10.19128/turje.545798

Carnoy, M., Khavenson, T., & Ivanova, A. (2015). Using TIMSS and PISA results to inform educational policy: A study of Russia and its neighbours. *Compare: A Journal of Comparative and International Education*, *45*(2), 248-271. https://doi.org/10.1080/03057925.2013.855002

Chang, C.-Y. (2002). Does Computer-Assisted Instruction + Problem Solving = Improved Science Outcomes? A Pioneer Study. *The Journal of Educational Research*, *95*(3), 143–150. https://doi.org/10.1080/00220670209596584

Chen, J., Zhang, Y., Wei, Y., & Hu, J. (2019). Discrimination of the Contextual Features of Top Performers in Scientific Literacy Using a Machine Learning Approach. *Research in Science Education*. https://doi.org/10.1007/s11165-019-9835-y

Clarke, E.A., & Kiselica, M.S. (1997). A systemic counseling approach to the problem of bullying. *Elementary School Guidance & Counseling*, *31*(4), 310–325.

Culbertson, M.J. (2016). Bayesian Networks in Educational Assessment: The State of the Field. *Applied Psychological Measurement*, *40*(1), 3-21. https://doi.org/10.1177/0146621615590401

Deng, Z., & Gopinathan, S. (2016). PISA and high-performing education systems: Explaining Singapore's education success. *Comparative Education*, *52*(4), 449-472. https://doi.org/

10.1080/03050068.2016.1219535
Desmarais, M.C., & Baker, R.S.J.d. (2012). A review of recent advances in learner and skill modeling in intelligent learning environments. *User Modeling and User-Adapted Interaction*, *22*(1), 9–38. https://doi.org/10.1007/s11257-011-9106-8
Ehrenberg, R.G., Brewer, D.J., Gamoran, A., & Willms, J.D. (2001). Class Size and Student Achievement. *Psychological Science in The Public Interest*, *2*(1), 30.
Ertem, H.Y. (2021). Examination of Turkey's PISA 2018 reading literacy scores within student-level and school-level variables. *Participatory Educational Research*, *8*(1), 248–264. https://doi.org/10.17275/per.21.14.8.1
Ferri, C., Hernández-Orallo, J., & Modroiu, R. (2009). An experimental comparison of performance measures for classification. *Pattern Recognition Letters*, *30*(1), 27–38. https://doi.org/10.1016/j.patrec.2008.08.010
Fry, D., Fang, X., Elliott, S., Casey, T., Zheng, X., Li, J., Florian, L., & McCluskey, G. (2018). The relationships between violence in childhood and educational outcomes: A global systematic review and meta-analysis. *Child Abuse & Neglect*, *75*, 6–28. https://doi.org/10.1016/j.chiabu.2017.06.021
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*. https://doi.org/10.3389/fpsyg.2020.575167
Gilbert, J.K., Boulter, C.J., & Elmer, R. (2000). Positioning Models in Science Education and in Design and Technology Education. In J.K. Gilbert & C.J. Boulter (Eds.), *Developing Models in Science Education* (pp. 3-17). Springer Netherlands. https://doi.org/10.1007/978-94-010-0876-1_1
Hall, M.A. (1999a). *Correlation-based Feature Selection for Machine Learning* [PhD Thesis]. The University of Waikato.
Hall, M.A. (1999b). *Feature selection for discrete and numeric class machine learning* [Working Paper]. Computer Science, University of Waikato. https://researchcommons.waikato.ac.nz/handle/10289/1033
Hall, M.A. (2000). *Correlation-based feature selection of discrete and numeric class machine learning* [Working Paper]. University of Waikato, Department of Computer Science. https://researchcommons.waikato.ac.nz/handle/10289/1024
Hanushek, E.A., & Woessmann, L. (2017). School Resources and Student Achievement: A Review of Cross-Country Economic Research. In M. Rosén, K. Yang Hansen, & U. Wolff (Eds.), *Cognitive Abilities and Educational Outcomes* (pp. 149–171). Springer International Publishing. https://doi.org/10.1007/978-3-319-43473-5_8
Harker, R. (2000). Achievement, Gender and the Single-Sex/Coed Debate. *British Journal of Sociology of Education*, *21*(2), 203–218. https://doi.org/10.1080/713655349
Hattie, J. (2005). The paradox of reducing class size and improving learning outcomes. *International Journal of Educational Research*, *43*(6), 387-425. https://doi.org/10.1016/j.ijer.2006.07.002
Hossin, M., & Sulaiman, M.N. (2015). A Review on Evaluation Metrics for Data Classification Evaluations. *International Journal of Data Mining & Knowledge Management Process*, *5*(2), 01–11. https://doi.org/10.5121/ijdkp.2015.5201
Jan, A. (2015). Bullying in Elementary Schools: Its Causes and Effects on Students. *Journal of Education and Practice*, 15.
Karaboga, H.A., Gunel, A., Korkut, S.V., Demir, I., & Celik, R. (2021). Bayesian Network as a Decision Tool for Predicting ALS Disease. *Brain Sciences*, *11*(2), Article 2. https://doi.org/10.3390/brainsci11020150
Karakoç Alatlı, B. (2020). Investigation of Factors Associated with Science Literacy Performance of Students by Hierarchical Linear Modeling: PISA 2015 Comparison of

Turkey and Singapore. *TED Education and Science Magazine*. https://doi.org/10.15390/EB.2020.8188

Karataş, H., & Ergin, A. (2018). Üniversite Öğrencilerinin Başarı Odaklı Motivasyon Düzeyleri [Achievement-Oriented Motivation Levels of University Students]. *Hacettepe University Journal of Education*, 1–20. https://doi.org/10.16986/HUJE.2018036646

Kenekayoro, P. (2018). An Exploratory Study on the Use of Machine Learning to Predict Student Academic Performance: *International Journal of Knowledge-Based Organizations*, *8*(4), 67–79. https://doi.org/10.4018/IJKBO.2018100104

Kilic Depren, S. (2018). Prediction of Students' Science Achievement: An Application of Multivariate Adaptive Regression Splines and Regression Trees. *Journal of Baltic Science Education*, *17*(5), 887–903. https://doi.org/10.33225/jbse/18.17.887

Kilic Depren, S. (2020). Determination of the Factors Affecting Students' Science Achievement Level in Turkey and Singapore: An Application of Quantile Regression Mixture Model. *Journal of Baltic Science Education*, *19*(2), 247-260. https://doi.org/10.33225/jbse/20.19.247

Kiray, S.A., Gok, B., & Bozkir, A.S. (2015). Identifying the Factors Affecting Science and Mathematics Achievement Using Data Mining Methods. *Journal of Education in Science, Environment and Health*, *1*(1), 28. https://doi.org/10.21891/jeseh.41216

Kjærnsli, M., & Lie, S. (2004). PISA and scientific literacy: Similarities and differences between the nordic countries. *Scandinavian Journal of Educational Research*, *48*(3), 271–286. https://doi.org/10.1080/00313830410001695736

Korb, K.B., & Nicholson, A.E. (2010). *Bayesian Artificial Intelligence*. CRC Press.

Kustitskaya, T.A., Kytmanov, A.A., & Noskov, M.V. (2020). Student-at-risk detection by current learning performance indicators using Bayesian networks. *ArXiv:2004.09774 [Stat]*. http://arxiv.org/abs/2004.09774

Lee, J., & Shute, V.J. (2010). Personal and Social-Contextual Factors in K–12 Academic Performance: An Integrative Perspective on Student Learning. *Educational Psychologist*, *45*(3), 185–202. https://doi.org/10.1080/00461520.2010.493471

Levy, R. (2016). Advances in Bayesian Modeling in Educational Research. *Educational Psychologist*, *51*(3–4), 368–380. https://doi.org/10.1080/00461520.2016.1207540

Lima. (2014). Heuristic Discretization Method for Bayesian Networks. *Journal of Computer Science*, *10*(5), 869–878. https://doi.org/10.3844/jcssp.2014.869.878

Lytvynenko, V., Savina, N., Voronenko, M., Doroschuk, N., Smailova, S., Boskin, O., & Kravchenko, T. (2019). Development, Validation and Testing of the Bayesian Network of Educational Institutions Financing. *2019 10th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS)*, *1*, 412–417. https://doi.org/10.1109/IDAACS.2019.8924307

Margolis, E. (2001). *The Hidden Curriculum in Higher Education*. Psychology Press.

Marsland, S. (2015). *Machine Learning: An Algorithmic Perspective* (Second Edition). CRC press.

Martínez Abad, F., & Chaparro Caso López, A.A. (2017). Data-mining techniques in detecting factors linked to academic achievement. *School Effectiveness and School Improvement*, *28*(1), 39–55. https://doi.org/10.1080/09243453.2016.1235591

MEB. (2018). *Fen Bilimleri Dersi Öğretim Programı [Science Course Curriculum]*.. Talim ve Terbiye Kurulu Başkanlığı, Ankara. https://mufredat.meb.gov.tr/Dosyalar/201812312311937-FEN%20B%C4%B0L%C4%B0MLER%C4%B0%20%C3%96%C4%9ERET%C4%B0M%20PROGRAMI2018.pdf

MEB. (2019). *PISA 2018 Turkiye Ön Raporu [PISA 2018 Turkey Preliminary Report]*. Milli Eğitim Bakanlığı. http://www.meb.gov.tr/meb_iys_dosyalar/2019_12/03105347_PISA_2018_Turkiye_On_Raporu.pdf

Millán, E., Descalço, L., Castillo, G., Oliveira, P., & Diogo, S. (2013). Using Bayesian networks to improve knowledge assessment. *Computers & Education*, *60*(1), 436–447. https://doi.org/10.1016/j.compedu.2012.06.012

Muñoz-Merino, P.J., Molina, M.F., Muñoz-Organero, M., & Kloos, C.D. (2014). Motivation and Emotions in Competition Systems for Education: An Empirical Study. *IEEE Transactions on Education*, *57*(3), 182–187. https://doi.org/10.1109/TE.2013.2297318

Neapolitan, R.E. (2009). *Probabilistic methods for bioinformatics: With an introduction to Bayesian networks*. Morgan Kaufmann/Elsevier.

Nguyen, L., & Do, P. (2009). Combination of Bayesian Network and Overlay Model in User Modeling. In G. Allen, J. Nabrzyski, E. Seidel, G.D. van Albada, J. Dongarra, & P.M.A. Sloot (Eds.), *Computational Science – ICCS 2009* (pp. 5–14). Springer. https://doi.org/10.1007/978-3-642-01973-9_2

Nielsen, T.D., & Jensen, F.V. (2009). *Bayesian Networks and Decision Graphs*. Springer Science & Business Media.

Nojavan A., F., Qian, S.S., & Stow, C.A. (2017). Comparative analysis of discretization methods in Bayesian networks. *Environmental Modelling & Software*, *87*, 64–71. https://doi.org/10.1016/j.envsoft.2016.10.007

O'Connell, M. (2019). Is the impact of SES on educational performance overestimated? Evidence from the PISA survey. *Intelligence*, *75*, 41-47. https://doi.org/10.1016/j.intell.2019.04.005

Odell, B., Galovan, A.M., & Cutumisu, M. (2020). The Relation Between ICT and Science in PISA 2015 for Bulgarian and Finnish Students. *EURASIA Journal of Mathematics, Science and Technology Education*, *16*(6). https://doi.org/10.29333/ejmste/7805

OECD. (2019a). *PISA 2018 Assessment and Analytical Framework*. OECD. https://doi.org/10.1787/b25efab8-en

OECD. (2019b). *PISA 2018 Results (Volume I): What Students Know and Can Do*. OECD. https://doi.org/10.1787/5f07c754-en

OECD. (2019c). *PISA 2018 Results (Volume II): Where All Students Can Succeed*. OECD. https://doi.org/10.1787/b5fd1b8f-en

OECD. (2020). *Do boys and girls have similar attitudes towards competition and failure?* (PISA in Focus 105; PISA in Focus, Vol. 105). https://doi.org/10.1787/a8898906-en

Özdemir, E., Cansiz, M., Cansiz, N., & Üstün, U. (2019). Türkiye deki Öğrencilerin Fen Okuryazarlığını Etkileyen Faktörler Nelerdir PISA 2015 Verisine Dayalı Bir Hiyerarşik Doğrusal Modelleme Çalışması. *Hacettepe University Journal of Education*, 1–16. https://doi.org/10.16986/HUJE.2019050786

Pearl, J. (2014). *Probabilistic reasoning in intelligent systems: Networks of plausible inference* (Revised Second Printing). Morgan Kaufmann. https://doi.org/10.1016/B978-0-08-051489-5.50002-3

Peña-Ayala, A. (2014). Educational data mining: A survey and a data mining-based analysis of recent works. *Expert Systems with Applications*, *41*(4), 1432-1462. https://doi.org/10.1016/j.eswa.2013.08.042

Ramírez-Noriega, A., Juárez-Ramírez, R., Leyva-López, J.C., Jiménez, S., & Figueroa-Pérez, J.F. (2021). A Method for Building the Quantitative and Qualitative Part of Bayesian Networks for Intelligent Tutoring Systems. *The Computer Journal*, *bxab124*. https://doi.org/10.1093/comjnl/bxab124

Rastrollo-Guerrero, J.L., Gómez-Pulido, J.A., & Durán-Domínguez, A. (2020). Analyzing and Predicting Students' Performance by Means of Machine Learning: A Review. *Applied Sciences*, *10*(3), 1042. https://doi.org/10.3390/app10031042

Reichenberg, R. (2018). Dynamic Bayesian Networks in Educational Measurement: Reviewing and Advancing the State of the Field. *Applied Measurement in Education*, *31*(4), 335–

350. https://doi.org/10.1080/08957347.2018.1495217

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. (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

Ropero, R.F., Renooij, S., & van der Gaag, L.C. (2018). Discretizing environmental data for learning Bayesian-network classifiers. *Ecological Modelling*, *368*, 391–403. https://doi.org/10.1016/j.ecolmodel.2017.12.015

Sağlam, A.Ç., & Aydoğmuş, M. (2016). Gelişmiş ve Gelişmekte Olan Ülkelerin Eğitim Sistemlerinin Denetim Yapıları Karşılaştırıldığında Türkiye Eğitim Sisteminin Denetimi Ne Durumdadır? [When the Supervision Structures of the Education Systems of Developed and Developing Countries are Compared, How is the Supervision of the Turkish Education System?]. *Uşak Üniversitesi Sosyal Bilimler Dergisi*, *9*(1), Article 1. https://doi.org/10.12780/uusbd.50788

Saini, M.K., & Goel, N. (2019). How Smart Are Smart Classrooms? A Review of Smart Classroom Technologies. *ACM Computing Surveys*, *52*(6), 130:1-130:28. https://doi.org/10.1145/3365757

Schleicher, A. (2019). PISA 2018: Insights and Interpretations. In *OECD Publishing*. OECD Publishing.

Sebastian, J., Moon, J.-M., & Cunningham, M. (2017). The relationship of school-based parental involvement with student achievement: A comparison of principal and parent survey reports from PISA 2012. *Educational Studies*, *43*(2), 123–146. https://doi.org/10.1080/03055698.2016.1248900

Sener, E., Karaboga, H.A., & Demir, I. (2019). Bayesian Network Model of Turkish Financial Market from Year-to-September 30th of 2016. *Sigma: Journal of Engineering & Natural Sciences / Mühendislik ve Fen Bilimleri Dergisi*, *37*(4), 1493–1507.

Sheldrake, R., Mujtaba, T., & Reiss, M.J. (2017). Science teaching and students’ attitudes and aspirations: The importance of conveying the applications and relevance of science. *International Journal of Educational Research*, *85*, 167-183. https://doi.org/10.1016/j.ijer.2017.08.002

Shin, D., & Shim, J. (2021). A Systematic Review on Data Mining for Mathematics and Science Education. *International Journal of Science and Mathematics Education*, *19*(4), 639–659. https://doi.org/10.1007/s10763-020-10085-7

Sing, T., Sander, O., Beerenwinkel, N., & Lengauer, T. (2005). ROCR: Visualizing classifier performance in R. *Bioinformatics*, *21*(20), 3940-3941. https://doi.org/10.1093/bioinformatics/bti623

Sinharay, S. (2006). Model Diagnostics for Bayesian Networks. *Journal of Educational and Behavioral Statistics*, *31*(1), 1–33.

Sinharay, S. (2016). An NCME Instructional Module on Data Mining Methods for Classification and Regression. *Educational Measurement: Issues and Practice*, *35*(3), 38–54. https://doi.org/10.1111/emip.12115

Sirin, S.R. (2005). Socioeconomic Status and Academic Achievement: A Meta-Analytic Review of Research. *Review of Educational Research*, *75*(3), 417–453. https://doi.org/10.3102/00346543075003417

Sjøberg, S. (2019). The PISA-syndrome – How the OECD has hijacked the way we perceive pupils, schools and education. *Confero: Essays on Education, Philosophy and Politics*, *7*(1), 12–65.

Stearns, E., & Glennie, E.J. (2010). Opportunities to participate: Extracurricular activities' distribution across and academic correlates in high schools. *Social Science Research*, *39*(2), 296–309. https://doi.org/10.1016/j.ssresearch.2009.08.001

Sudrajad, K., Soemanto, Rb., & Prasetya, H. (2020). The Effect of Bullying on Depression, Academic Activity, and Communication in Adolescents in Surakarta: A Multilevel Logistic Regression. *Journal of Health Promotion and Behavior*, *5*(2), 79–86. https://doi.org/10.26911/thejhpb.2020.05.02.02

Suna, H.E., Tanberkan, H., & Özer, M. (2020). Changes in Literacy of Students in Turkey by Years and School Types: Performance of Students in PISA Applications. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, *11*(1), 76-97. https://doi.org/10.21031/epod.702191

Tang, X., & Zhang, D. (2020). How informal science learning experience influences students' science performance: A cross-cultural study based on PISA 2015. *International Journal of Science Education*, *42*(4), 598–616. https://doi.org/10.1080/09500693.2020.1719290

Tatar, E., Tüysüz, C., Tosun, C., & Ilhan, N. (2016). Investigation of Factors Affecting Students' Science Achievement According to Student Science Teachers. *International Journal of Instruction*, *9*(2), 153–166.

Tingir, S., & Almond, R. (2017). Using Bayesian Networks to Visually Compare the Countries: An Example from PISA. *Journal of Education*, *4*(3), 11.

Topçu, M.S., Arıkan, S., & Erbilgin, E. (2015). Turkish Students' Science Performance and Related Factors in PISA 2006 and 2009. *The Australian Educational Researcher*, *42*(1), 117–132. https://doi.org/10.1007/s13384-014-0157-9

Topçu, M.S., Erbilgin, E., & Arikan, S. (2016). Factors Predicting Turkish and Korean Students' Science and Mathematics Achievement in TIMSS 2011. *EURASIA Journal of Mathematics, Science and Technology Education*, *12*(7). https://doi.org/10.12973/eurasia.2016.1530a

Torrecilla Sánchez, E.M., Olmos Miguélañez, S., & Martínez Abad, F. (2019). Explanatory factors as predictors of academic achievement in PISA tests. An analysis of the moderating effect of gender. *International Journal of Educational Research*, *96*, 111–119. https://doi.org/10.1016/j.ijer.2019.06.002

Üstün, U., Özdemir, E., Cansiz, M., & Cansiz, N. (2020). Türkiye'deki Öğrencilerin Fen Okuryazarlığını Etkileyen Faktörler Nelerdir? PISA 2015 Verisine Dayalı Bir Hiyerarşik Doğrusal Modelleme Çalışması [What are the Factors Affecting Students' Science Literacy in Turkey? A Hierarchical Linear Modeling Study Based on PISA 2015 Data]. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, *35*(3), Article 3.

van der Berg, S. (2008). How effective are poor schools? Poverty and educational outcomes in South Africa. *Studies in Educational Evaluation*, *34*(3), 145-154. https://doi.org/10.1016/j.stueduc.2008.07.005

Wachs, S., Bilz, L., Niproschke, S., & Schubarth, W. (2019). Bullying Intervention in Schools: A Multilevel Analysis of Teachers' Success in Handling Bullying from the Students' Perspective. *The Journal of Early Adolescence*, *39*(5), 642-668. https://doi.org/10.1177/0272431618780423

White, H. (2018). Small Class Size Has at Best a Small Effect on Academic Achievement. Plain Language Summary. In *Campbell Collaboration*. Campbell Collaboration. https://eric.ed.gov/?id=ED610283

Wong, T.-T. (2015). Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation. *Pattern Recognition*, *48*(9), 2839-2846. https://doi.org/10.1016/j.patcog.2015.03.009

Wößmann, L. (2005). Educational production in Europe. *Economic Policy*, *20*(43), 446–504. https://doi.org/10.1111/j.1468-0327.2005.00144.x

Xing, W., Li, C., Chen, G., Huang, X., Chao, J., Massicotte, J., & Xie, C. (2021). Automatic Assessment of Students' Engineering Design Performance Using a Bayesian Network Model. *Journal of Educational Computing Research*, *59*(2), 230-256. https://doi.org/10.1177/0735633120960422

Yang, Y., & Webb, G.I. (2002). A Comparative Study of Discretization Methods for Naive-Bayes Classifiers. *In Proceedings of PKAW 2002: The 2002 Pacific Rim Knowledge Acquisition Workshop*, 159–173.

Yip, D.Y., Chiu, M.M., & Ho, E.S.C. (2004). Hong Kong Student Achievement in OECD-PISA Study: Gender Differences in Science Content, Literacy Skills, and Test Item Formats. *International Journal of Science and Mathematics Education*, *2*(1), 91–106. https://doi.org/10.1023/B:IJMA.0000026537.85199.36

Yıldırım, S. (2012). Teacher Support, Motivation, Learning Strategy Use, and Achievement: A Multilevel Mediation Model. *The Journal of Experimental Education*, *80*(2), 150–172. https://doi.org/10.1080/00220973.2011.596855

Zhang, P. (1992). On the Distributional Properties of Model Selection Criteria. *Journal of the American Statistical Association*, *87*(419), 732-737. https://doi.org/10.1080/01621459.1992.10475275

Zwick, R., & Lenaburg, L. (2009). Using Discrete Loss Functions and Weighted Kappa for Classification: An Illustration Based on Bayesian Network Analysis. *Journal of Educational and Behavioral Statistics*, *34*(2), 190-200. https://doi.org/10.3102/1076998609332106