Analysis and Evaluation of Factors Influencing Student Success with Explainable Artificial Intelligence Models

Manuscript received April 2, 2024; revised May 19, 2024

Cem Özkurt
AI and Data Science Research and Application Center
Department of Computer Engineering
Sakarya University of Applied Sciences
Sakarya, Turkey
cemozkurt@subu.edu.tr

***Abstract*—The research aims to explain socio-economic factors affecting student success using interpretable artificial intelligence models and to promote the use of this technology in the development of educational policies. Initially, existing studies on factors determining student success have been examined. The dataset includes socio-economic and personal variables such as parental education, family economic status, student gender, and study duration. Using interpretable artificial intelligence models like InterpretML, analyses have been conducted on this dataset, and the results obtained from examining factors influencing student success have been evaluated. This study aims to contribute to shaping educational policies more effectively through the use of artificial intelligence.**

Keywords*— *explainable artificial intelligence, InterpretML, machine learning, data preprocessing

I. Introduction

Examining the factors affecting student success is of critical importance in developing strategic decision-making processes in education [1]. Among these factors are socio-economic and personal variables such as parental education, family economic status, and the student's gender. Analyzing these factors that determine student success can enable more effective shaping of educational policies. Understanding these factors is particularly important for efforts aimed at improving students' academic performance.

The issue of school dropout poses a significant barrier in the field of education [2]. To address this problem, the use of advanced techniques such as Explainable Artificial Intelligence (XAI) is important. These techniques can be utilized to predict students' tendencies to drop out of school and contribute to shaping educational policies accordingly.

The novelty of this study lies in focusing on explaining the factors affecting student success by classifying them with interpretable artificial intelligence models. In this regard, the analysis of factors influencing student success has been conducted using interpretable AI models such as interpretML, and the results have been evaluated.

It investigates the factors influencing the mathematics achievement of students in grades 6-8. Factors such as school type, family income level, study time, attitude toward mathematics, and attendance at private tutoring classes are associated with students' mathematics achievements [3].

The relationship between students' socio-economic levels and their success in the Transition from Primary to Secondary Education System (TEOG) mathematics exam was examined using data obtained from secondary schools in Altındağ district of Ankara province [4]. A study conducted on senior undergraduate students at Süleyman Demirel University examined the effect of factors such as family income, educational background, and place of residence on students' academic success [5].

The relationship between peers' socioeconomic status (SES) and individual academic achievement is examined. Student SES is measured using participation in the federal free/reduced-price lunch program and parents' education and occupational backgrounds. Particularly, the socioeconomic status of peers has a significant and substantial effect on individual academic achievement [6].

Socioeconomic status (SES) is a composite measure of an individual's or family's economic and social position relative to others, based on income, education, and occupation. When analyzing a family's SES, both parents' education and occupation as well as the individual's income are considered. In the literature, it has been shown that parents' socioeconomic status can significantly contribute to obtaining good grades at universities [7].

A study conducted on students attending middle schools in Burdur using CHAID analysis examined the relationship between students' academic achievement and socioeconomic variables such as parent's education level, family income, and parents' employment status [8]. It examines methods for predicting student academic performance using artificial neural networks. A literature review concluded that artificial neural networks are an effective tool for predicting student success in higher education [9]. It addresses predicting student performance in engineering mathematics courses using a type of artificial intelligence model called Extreme Learning Machines (ELM). It was found that ELM outperformed other models in predicting student performance [10].

II. Materials And Methodology

A. Dataset

The dataset used in the study encompasses various socio-economic and personal factors that influence students' academic achievements. The dataset is presented in two separate files: "student-mat.csv" (for the Mathematics course) and "student-por.csv" (for the Portuguese language course). This dataset contains diverse demographic and educational-related information such as students' schools, genders, age groups, types of residential addresses, family sizes, parents'

 EISSN: 2961-872X ● PISSN: 2961-8924 ● DOI: 10.56741/esl.v3i03.556

cohabitation status, education levels, occupations, reasons for school choice, guardian identities, travel durations to school, weekly study durations, past class failures, educational supports, social and personal factors, family relationships, leisure activities, alcohol consumption, health statuses, and absenteeism numbers. Additionally, it includes first-period (G1), second-period (G2), and final (G3) grades in Mathematics and Portuguese courses, serving as fundamental outcome variables to measure students' academic performance. The dataset comprises 1044 observations and 32 attributes. An example sample of the dataset is provided in Fig. 1.

	school	sex	age	address	famsize	Pstatus	Medu	Fedu	Mjob	Fjob	...	famrel	freetime	goout	Dalc	Walc	health	absences
0	GP	F	18	U	GT3	A	4	4	at_home	teacher	...	4	3	4	1	1	3	6
1	GP	F	17	U	GT3	T	1	1	at_home	other	...	5	3	3	1	1	3	4
2	GP	F	15	U	LE3	T	1	1	at_home	other	...	4	3	2	2	3	3	10
3	GP	F	15	U	GT3	T	4	2	health	services	...	3	2	2	1	1	5	2
4	GP	F	16	U	GT3	T	3	3	other	other	...	4	3	2	1	2	5	4
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...

Fig. 1. Dataset Example

B. *Explainable AI (XAI)*

Explainable AI (XAI) is an approach aimed at making the decisions and inferences of artificial intelligence systems understandable and traceable by humans. XAI develops methods to explain the inner workings of complex AI models, allowing us to understand why decisions are made, how input data is processed, and how results are obtained. As a result, users can confidently accept the decisions of AI systems and question the logic behind them when necessary. XAI is recognized as an important tool for developing reliable and transparent AI systems in various fields, from medical diagnostics to financial risk analysis, autonomous vehicles, and security systems.

InterpretML is a library that enables us to understand the behaviors, decisions, and predictions of machine learning models. This library provides various techniques and tools to explain and interpret model predictions, making the inner workings of complex artificial intelligence models more understandable. InterpretML offers interpretability techniques for decision trees, support vector machines, gradient boosting models, and many other machine learning algorithms. These techniques include feature importance ranking, cluster analysis, decision tree explanations, and classification uncertainty estimation, among others. As a result, researchers, developers, and users can gain a deeper understanding of the behaviors and decisions of machine learning models and use them with confidence.

III. RESULTS

In this study, various preprocessing steps were applied to prepare the data set for the InterpretML algorithm. These steps are intended to ensure the effective use of categorical variables in machine learning models. Specifically, array-type categorical variables present in the dataset were converted to integer types to increase the processing efficiency of the models. This transformation process primarily included variables such as students' gender, age range, and parents' education level. These adjustments allow machine learning algorithms to better understand the data and produce more accurate results. Especially algorithms that work with numerical values can work more effectively and more reliable analysis results can be obtained thanks to this transformation. Preprocessing performed on the data set creates a solid foundation for the analysis process. An example of the dataset prepared for the InterpretML algorithm is presented in Fig. 2.

	school	sex	age	address	famsize	Pstatus	Medu	Fedu	Mjob	Fjob	...	romantic	famrel	freetime	goout	Dalc	Walc	health	absences	Lecture_Notes	Course_Pass_Status
0	0	0	18	0	1	1	4	4	0	4	...	0	4	3	4	1	1	3	6	28.333333	0
1	0	0	17	0	1	0	1	1	0	2	...	0	5	3	3	1	1	3	4	26.666667	0
2	0	0	15	0	0	0	1	1	0	2	...	0	4	3	2	2	3	3	10	41.666667	0
3	0	0	15	0	1	0	4	2	1	3	...	1	3	2	2	1	1	5	2	73.333333	1
4	0	0	16	0	1	0	3	3	2	2	...	0	4	3	2	1	2	5	4	43.333333	0
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...

Fig. 2. Example of data set to be used in the InterpretML algorithm

The confusion matrix, which is a tool used to evaluate the performance of a classification model, is presented in Fig. 3 by the InterpretML algorithm applied to the dataset. This matrix represents the values of the correct and predicted class labels of the model. While the rows represent the actual classes, the columns show the predicted classes. The matrix enables visualization of correctly and incorrectly classified examples and allows for the calculation of various metrics to evaluate the performance of the model.

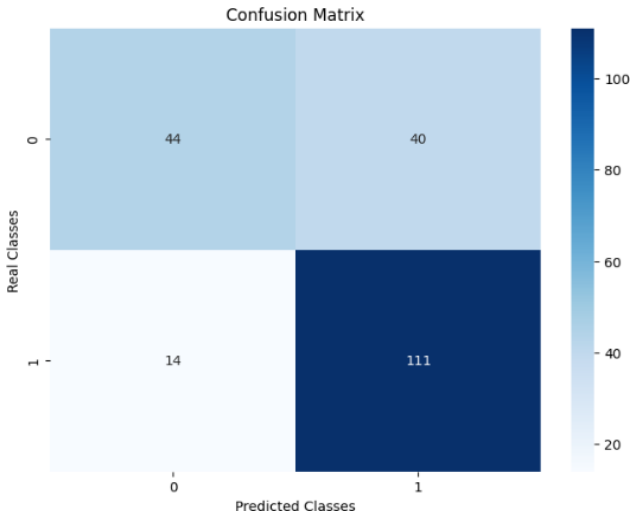


Fig. 3. InterpretML's complexity matrix

Remark:

- True Positive (TP): Represents the case where the model accurately predicts the positive class.
  TP = 111
- True Negative (TN): Represents the case where the model accurately predicts the negative class.
  TN = 14
- False Positive (FP): Represents the case where the model incorrectly predicts the positive class.
  FP = 40
- False Negative (FN): Represents the case where the model incorrectly predicts the negative class.
  FN = 44

The precision metric is a performance measure that evaluates how many of the samples predicted as positive by a classification model are positive. This metric is calculated based on the True Positive (TP) and False Positive (FP) values. A high precision value indicates that most of the samples predicted as positive are indeed

 EISSN: 2961-872X ● PISSN: 2961-8924 ● DOI: 10.56741/esl.v3i03.556

positive, with few false positive predictions. The calculation of the precision metric is shown in Equation 1.

$$Precision = \frac{TP}{TP + FP} \quad (1)$$

In Equation 1, it is used to determine how effective the model is in terms of accuracy and reliability. Especially in cases of imbalanced class distributions (for example, situations where there are rare classes), the precision metric can help to more accurately evaluate the model's performance. A high precision value can reduce the likelihood of false alarms by the model, providing a more reliable experience for end users.

The recall metric is a performance measure that evaluates how many of the true positives a classification model correctly predicts. This metric is calculated based on the True Positive (TP) and False Negative (FN) values. A high recall value indicates that the model correctly predicts most of the true positives, with few false negative predictions. The recall metric is particularly important in areas of vital importance, such as medical diagnosis, and in situations where missing true positives is significant. The calculation of the recall metric is shown in Equation 2.

$$Recall = \frac{TP}{TP + FN} \quad (2)$$

In Equation 2, it is used to evaluate how sensitive the model is and the likelihood of missing true positives. A high recall value indicates that the model tends to detect more true positives, but this can also lead to an increase in false positives. Therefore, by evaluating precision and recall metrics together, the performance of the model can be more comprehensively understood.

The accuracy metric is a performance measure that evaluates the ratio of correctly predicted samples by a classification model. This metric is calculated based on True Positives (TP) and True Negatives (TN) values. A high accuracy value indicates that the model has a high ability to make correct predictions with few incorrect predictions. The accuracy metric is effective when all classes are correctly classified and can be used in balanced class distributions. However, in cases of imbalanced class distributions (for example, situations where there are rare classes), the accuracy metric may be inadequate because incorrect predictions of rare classes can misleadingly inflate the overall accuracy value. The calculation of the accuracy metric is shown in Equation 3.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (3)$$

Equation 3, is used to evaluate the accuracy degree of the model and is commonly used to understand the overall performance of the model.

The F1 metric is the harmonic mean of a classification model's precision and recall metrics, aiming to balance the impact of both false positives and false negatives. This metric is calculated based on True Positive (TP), False Positive (FP), and False Negative (FN) values. A high F1 value indicates that both precision and recall metrics are high, and the model reduces both false positives and false negatives. The F1 metric is particularly useful in cases of imbalanced class distributions (for example, situations where there are rare classes), as precision and recall metrics may be inadequate in these scenarios. The calculation of the F1 metric is shown in Equation 4.

$$F1 = \frac{Precision \times Recall}{Precision + Recall} \quad (4)$$

Equation 4, is used to combine the precision and recall metrics in a balanced manner to comprehensively evaluate the performance of the model.

The values of Accuracy, Precision, Recall, and F1 metrics for the InterpretML method are provided in Table 1.

TABLE 1
INTERPRETML'S CLASSIFICATION METRICS

	Accuracy	Precision	Recall	F1
InterpretML	0.742	0.735	0.888	0.804

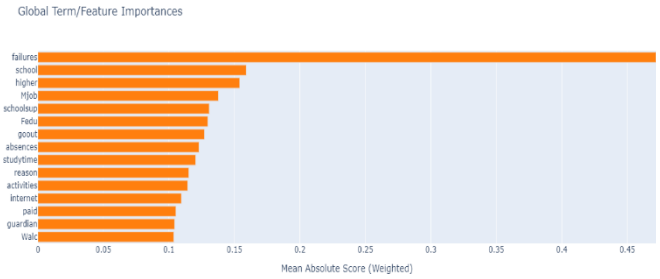


Fig. 4. InterpretML global output

Designed to visualize the factors influencing student performance, Fig. 4 indicates that the X-axis represents each column representing a factor affecting student performance, while the Y-axis indicates the importance of each column on that factor's impact on student achievement. Among the significant factors highlighted are the student's past frequency of failures ('failures') and the influence of the school ('school') they attend. Other significant factors include the student's future educational aspirations ('higher'), parental employment status ('Mjob' and 'Fedu'), frequency of going out ('goout'), number of absences ('absences'), and study time ('studytime'). The least important factors are noted to be activities ('activities'), internet usage ('internet'), and paid employment status ('paid').

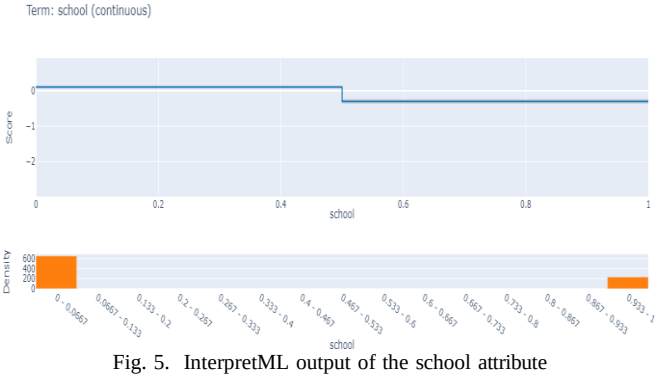


Fig. 5. InterpretML output of the school attribute

In Fig. 5, the Score (student's school) parameter represents the contribution to the outcome over continuous time, while the Density section indicates the number of

occurrences of the school (student's school) parameter values over continuous time. The highest number of occurrences, 619 times, was observed when the school parameter value fell within the range of 0 to 0.0667.

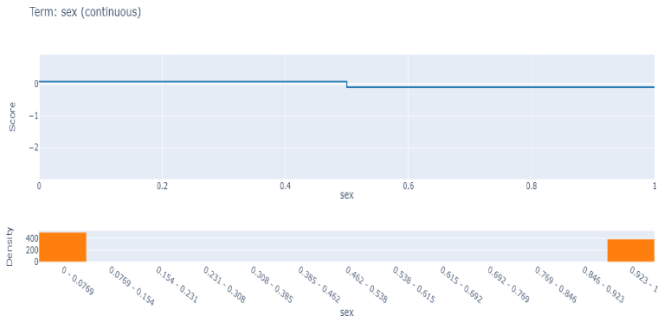


Fig. 6. InterpretML output of the sex attribute

In Fig. 6, the Score (student's gender) parameter represents the contribution to the outcome over continuous time, while the Density section indicates the number of occurrences of the school (student's gender) parameter values over continuous time. The highest number of occurrences, 464 times, was observed when the school parameter value fell within the range of 0 to 0.0769.

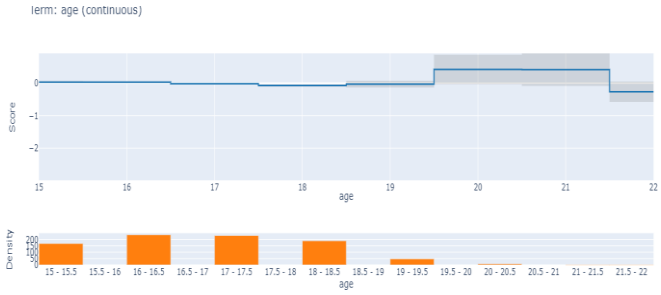


Fig. 7. InterpretML output of the age attribute

In Fig. 7, the Score (student's age) parameter represents the contribution to the outcome over continuous time, while the Density section indicates the number of occurrences of the school (student's age) parameter values over continuous time. The highest number of occurrences, 239 times, was observed when the school parameter value fell within the range of 17 to 17.5.

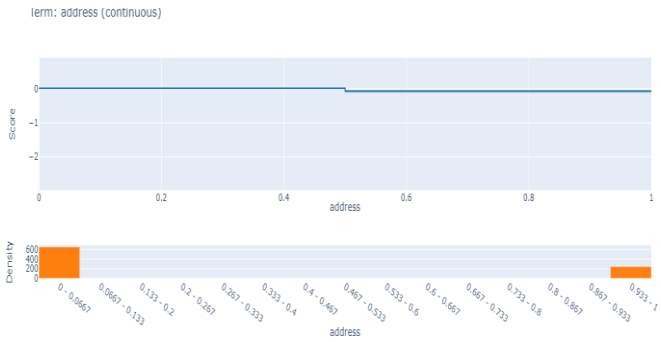


Fig. 8. InterpretML output of the address attribute

In Fig. 8, the Score (student's home address type) parameter represents the contribution to the outcome over continuous time, while the Density section indicates the number of occurrences of the school (student's home address type) parameter values over continuous time. The highest number of occurrences, 602 times, was observed when the school parameter value fell within the range of 0 to 0.0667.

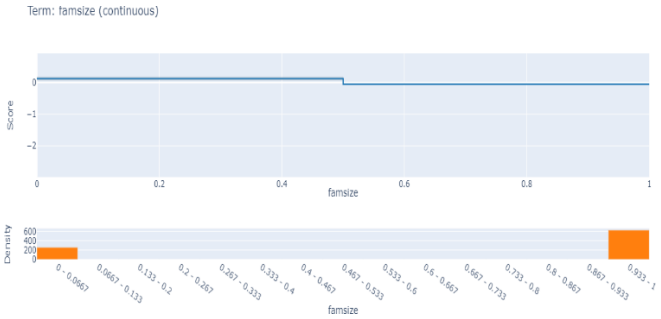


Fig. 9. InterpretML output of the fam size attribute

In Fig. 9, the Score (student's home address type) parameter represents the contribution to the outcome over continuous time, while the Density section indicates the number of occurrences of the school (student's home address type) parameter values over continuous time. The highest number of occurrences, 596 times, was observed when the school parameter value fell within the range of 0.933 to 1.

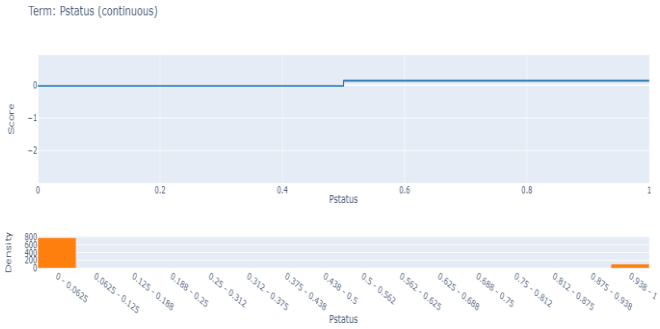


Fig. 10. InterpretML output of the Pstatus attribute

In Fig. 10, the Score (parents' cohabitation status) parameter represents the contribution to the outcome over continuous time, while the Density section indicates the number of occurrences of the school (parents' cohabitation status) parameter values over continuous time. The highest number of occurrences, 736 times, was observed when the school parameter value fell within the range of 0 to 0.0625.

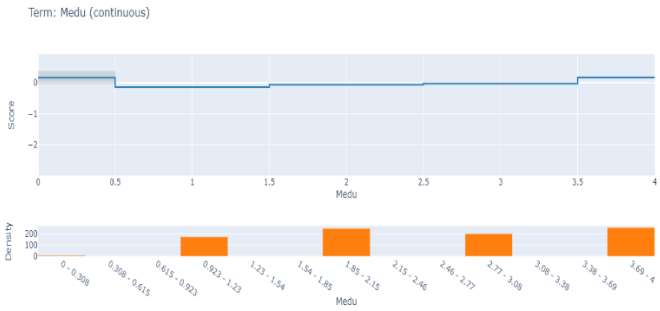


Fig. 11. InterpretML output of the Medu attribute

EISSN: 2961-872X ● PISSN: 2961-8924 ● DOI: 10.56741/esl.v3i03.556

In Fig. 11, the Score (mother's education) parameter represents the contribution to the outcome over continuous time, while the Density section indicates the number of occurrences of the school (mother's education) parameter values over continuous time. The highest number of occurrences, 230 times, was observed when the school parameter value fell within the range of 1.85 to 2.15.

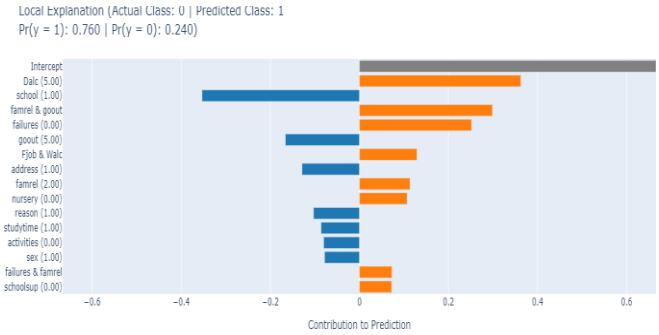


Fig. 12. InterpretML output of a random sample from the dataset

When examining the impact on student academic performance in Fig. 12, it was found that when weekday alcohol consumption (Dalc) is very high (5), it is positively effective at 0.362, while when the student's school (school) is Mousinho da Silveira (1), it is negatively effective at 0.353. The quality of family relationships (famrel) together with going out with friends (goout) has a positive effect of 0.298. The absence of course failures (failures) is positively effective at 0.251, whereas when going out with friends (goout) is very high (5), it is negatively effective at 0.166. When combined with the father's occupation (Fjob) and weekend alcohol consumption (Walc), it is positively effective at 0.129. When the student's home address type (address) is rural (1), it is negatively effective at 0.128. The quality of family relationships (famrel) being poor (2) has a positive effect of 0.113, and not attending nursery school (nursery) has a positive effect of 0.106. Choosing this school for a course (reason) has a negative effect of 0.102. Weekly study time (studytime) being less than 2 hours (1) has a negative effect of 0.086. The absence of extracurricular activities (activities) is negatively effective at 0.080, and gender (sex) being male (1) is negatively effective at 0.078. The combination of course failures (failures) and the quality of family relationships (famrel) together have a positive effect of 0.073, and the absence of extra educational support (schoolsup) is positively effective at 0.072.

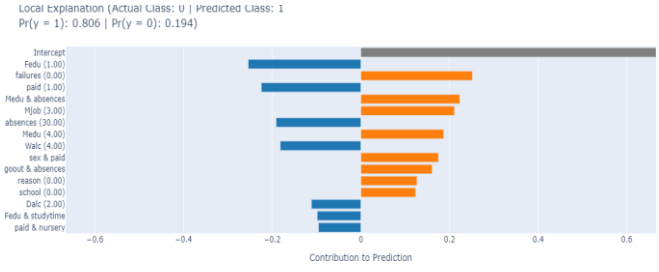


Fig. 13. InterpretML output of a random sample from the dataset

When examining the impact on student academic performance in Fig. 13, it was found that when the father's education level is primary education graduate (1), it is effective negatively at 0.254, while the absence of course failures (0) is effective positively at 0.251. The decision to take private tutoring (1) has a negative effect of 0.224, while the mother's education level being higher education (4) and the number of absences 30 have a positive effect of +0.223 and +0.211 respectively. When the mother's occupation is civil service (3), it is effective positively at 0.211, and when the number of absences is 30, it is effective negatively at 0.191. When the mother's education level is higher education (4), it is effective positively at 0.186, and when weekend alcohol consumption is high (4), it is effective negatively at 0.181. The combination of gender and the decision to take private tutoring has a positive effect of 0.174, and the combination of going out with friends and the number of absences has a positive effect of 0.160. Choosing this school due to its proximity to home (0) has a positive effect of 0.126, and when the student's school is Gabriel Pereira (0), it has a positive effect of 0.123. When alcohol consumption on weekdays is low (2), it has a negative effect of 0.111. The combination of the father's education level and weekly study time has a negative effect of 0.098, and the combination of the decision to take private tutoring and reading in kindergarten has a negative effect of 0.095.

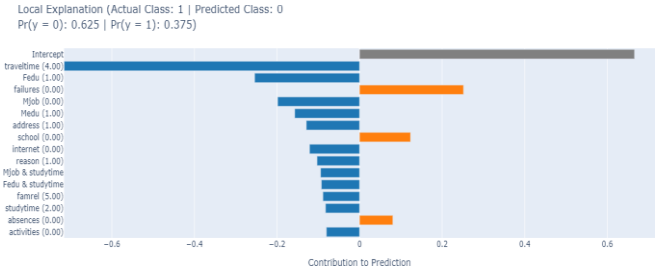


Fig. 14. InterpretML output of a random sample from the dataset

When examining the impact on student academic performance in Fig. 14, it was found that when the travel time from home to school (traveltime) is more than 1 hour (4), it is effective negatively at 0.715, while when the father's education level (Fedu) is primary education graduate (1), it is effective negatively at 0.254. The absence of course failures (failures) has a positive effect of 0.251, and when the mother's occupation (Mjob) is a homemaker (0), it is effective negatively at 0.198. When the mother's education level (Medu) is primary education graduate (1), it is effective negatively at 0.156, and when the student's home address type (address) is rural (1), it is effective negatively at 0.128. When the student's school (school) is Gabriel Pereira (0), it is effective positively at 0.123. The absence of internet access at home (internet) (0) is effective negatively at 0.121. Choosing this school for a course (reason) has a negative effect of 0.102. When combined with the mother's occupation (Mjob) and weekly study time (studytime), it is effective negatively at 0.094. When combined with the father's education level (Fedu) and weekly study time (studytime), it is effective negatively at 0.092. When the quality of family relationships (famrel) is excellent, it is effective negatively at 0.088, and when the weekly study time (studytime) is between 2-5 hours (2), it is effective negatively at 0.082. Since there are no absences (absences), it is effective positively at 0.080, and the absence of extracurricular activities (activities) has a negative effect of 0.080.



Fig. 15. InterpretML output of a random sample from the dataset

When examining the impact on student academic performance in Fig. 15, it was found that when the student's school (school) is Mousinho da Silveira (1), it has a negative effect of 0.353, while the absence of course failures (failures) has a positive effect of 0.251. When the mother's occupation (Mjob) is a homemaker (0), it has a negative effect of 0.198. When the mother's education level (Medu) is higher education (4), it has a positive effect of 0.186, and when the father's education level (Fedu) is higher education (4), it has a positive effect of 0.149. The type of home address (address) being rural (1) has a negative effect of 0.128. The reason for choosing this school (reason) being proximity to home (0) has a positive effect of 0.126, while the absence of internet access at home (internet) (0) has a negative effect of 0.121. Going out with friends (goout) being low (2) has a negative effect of 0.116. When combined with the mother's occupation (Mjob) and weekly study time (study time), it has a negative effect of 0.094. Having a family size (family size) of less than or equal to 3 has a positive effect of 0.089. When the quality of family relationships (famrel) is excellent, it has a negative effect of 0.088, and when the weekly study time (study time) is between 2-5 hours (2), it has a negative effect of 0.082. When weekend alcohol consumption (Walc) is moderate (3), it has a negative effect of 0.081, and the absence of extracurricular activities (activities) has a negative effect of 0.080.

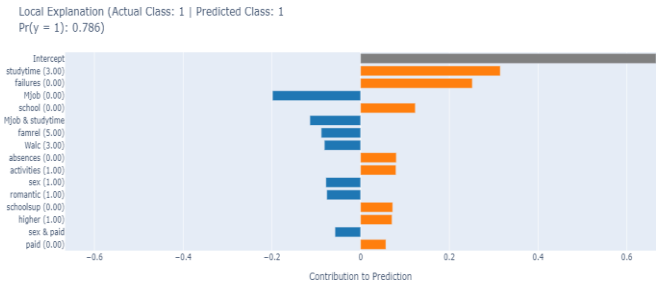


Fig. 16. InterpretML output of a random sample from the dataset.

When examining the impact on student academic performance in Fig. 16, it was determined that when the weekly study time falls between 5-10 hours (study time), the effect is positively determined as 0.314. In the absence of course failures (0), the effect was calculated positively as 0.251. When the mother's occupation is a homemaker (0), the effect is observed negatively as 0.198. When the student's school is Gabriel Pereira (0), the effect is determined positively as 0.123. When combined with the mother's occupation and weekly study time (study time), the effect is calculated negatively as 0.114. When the quality of family relationships (famrel) is excellent, the effect is observed negatively as 0.088. When weekend alcohol consumption (Walc) is moderate (3), the effect is determined negatively as 0.081. Since there are no absences (absences), the effect is calculated positively as 0.080. When extracurricular activities (activities) are present, the effect is observed positively as 0.080. When the gender (sex) is male (1), the effect is determined negatively as 0.078. When in a romantic relationship (romantic), the effect is calculated negatively as 0.076. In the absence of extra educational support (schoolsup), the effect is determined positively as 0.072. When desiring higher education (higher) is yes (1), the effect is calculated positively as 0.070. When combined with gender (sex) and paid tutoring status (paid), the effect is determined negatively as 0.057. Since there is no paid tutoring status (paid), the effect is observed positively as 0.056.

IV. Conclusions

In this study, the factors affecting student success have been explained using interpretable artificial intelligence models. Particularly, the impact of socio-economic and personal variables on student achievement has been discussed. Additionally, the contribution of these factors to shaping educational policies has been addressed. The analysis of factors influencing student success has been conducted by examining variables that play a significant role in determining students' mathematics achievements. These variables include school type, family income level, study time, attitude toward mathematics, and participation in private tutoring, among others. The results obtained highlight important factors affecting student success. Among these factors are the student's history of past failures, the influence of the school they attend, and family-related factors such as parents' employment status and education level, frequency of social outings, absenteeism, and study time. The analyses conducted emphasize the importance of a detailed examination of factors influencing student success. It is particularly crucial to discuss the effects of these factors on learning outcomes supported by theoretical frameworks.

V. References

[1] E. Aslanargun, S. Bozkurt, ve S. Sarıog˘lu, "Sosyo Ekonomik Deg˘is¸kenlerin O¨ g˘rencilerin Akademik Bas¸arısı U¨ zerine Etkileri," *Us¸ak U¨ niversitesi Sosyal Bilimler Dergisi*, cilt. 9, sayı. 27/3, ss. 201–234, 2016.

[2] E. Melo, I. Silva, D. G. Costa, C. M. D. Viegas, ve T. M. Barros, "On the Use of eXplainable Artificial Intelligence to Evaluate School Dropout," *Education Sciences*, cilt. 12, sayı. 12, makale numarası 845, 2022. [Online]. Available: https://www.mdpi.com/2227-7102/12/12/845

[3] E. Savas¸, S. Tas¸, ve A. Duru, "Matematikte O¨ g˘renci Bas¸arısını Etkileyen Fakto¨rler," *I˙no¨nu¨ U¨ niversitesi Eg˘itim Faku¨ltesi Dergisi*, cilt. 11, sayı. 1, ss. 113–132, 2010.

[4] S. Yavuz, M. Odabas¸, ve A. O¨ zdemir, "O¨ g˘rencilerin Sosyoekonomik Du¨zeylerinin TEOG Matematik Bas¸arısına Etkisi," *Journal of Measure- ment and Evaluation in Education and Psychology*, cilt. 7, sayı. 1, ss. 85–95, 2016.

[5] E. Tomul ve G. Polat, "The effects of socioeconomic characteristics of students on their academic achievement in higher education," *American Journal of Educational Research*, cilt. 1, sayı. 10, ss. 449–455, 2013.

[6] S. J. Caldas ve C. Bankston, "Effect of School Population Socioeconomic Status on Individual Academic Achievement," *The Journal of Educational Research*, cilt. 90, no. 5, ss. 269-277, 1997. DOI: 10.1080/00220671.1997.10544583.

[7] S. Saifi ve T. Mehmood, "Effects of socioeconomic status on students achievement," *International Journal of Social Sciences and Education*, cilt. 1, no. 2, ss. 119-128, 2011.

[8] E. O¨ under ve S. Uyar, "CHAID Analysis to Determine Socioeconomic Variables That Explain Students' Academic Success,"

 EISSN: 2961-872X ● PISSN: 2961-8924 ● DOI: 10.56741/esl.v3i03.556

Universal Journal of Educational Research, cilt. 5, sayı. 4, ss. 608–619, 2017.

[9] Y. Baashar, G. Alkawsi, A. Mustafa, A. A. Alkahtani, Y. A. Alsariera, A. Q. Ali, W. Hashim, ve S. K. Tiong, "Toward Predicting Student's Academic Performance Using Artificial Neural Networks (ANNs)," *Applied Sciences*, cilt. 12, sayı. 3, makale numarası: 1289, 2022.

[10] R. C. Deo, Z. M. Yaseen, N. Al-Ansari, T. Huy, T. A. M. Langlands, ve L. Galligan, "Modern Artificial Intelligence Model Development for Undergraduate Student Performance Prediction: An Investigation on Engineering Mathematics Courses," *IEEE Access*, cilt. 8, ss. 136697-136724, 2020.