Special Issue
of the Infocommunication Journal

Finding the contextual impacts on Students' Mathematical performance using a Machine Learning-based Approach

Zakaria Khoudi*, Mourad Nachaoui, and Soufiane Lyaqini

***Abstract*—An extensive dataset for examining Moroccan eighth-grade pupils' mathematical prowess was made available by the 2019 Trends in Mathematics and Science Study (TIMSS). The TIMSS 2019 public dataset contained 8390 Moroccan students, who were the subject of this research. Based on how well they could solve mathematical problems, the participants were split into 3108 high achievers and 5282 poor achievers in the mathematics phase of the exam. This study aimed to pinpoint the essential environmental elements affecting eighth-grade pupils' mathematical abilities. In order to do this, the research used cutting-edge machine learning methods, particularly the efficient distributed gradient boosting toolkit XGBoost. From a vast collection of 700 possible components, this strategy proved critical in identifying the most relevant variables. These factors included a broad spectrum of components at the student, teacher, and school levels. After a thorough investigation, 12 critical contextual factors distinguishing between arithmetic prodigies and average performers were successfully found. The discovery of these critical characteristics has significant implications for future instructional efforts, especially in improving high school pupils' mathematical proficiency. Knowledge of these factors may assist educators and policymakers in creating focused interventions and pedagogical approaches that enhance mathematics performance and comprehension. This research emphasizes how complex mathematics accomplishment is and how crucial it is to approach educational planning holistically. Identifying and addressing these critical environmental elements can significantly enhance students' mathematics achievements at a crucial juncture in their academic development.**

***Index Terms*—Contextual factors, Machine learning, Mathematics performance, Moroccan students, TIMSS 2019**

I. Introduction

THE International Association for the Evaluation of Educational Achievement (IEA) created the international research project known as TIMSS, which is conducted every four years. For further information, check [4]. TIMSS aims to track mathematical and scientific success trends and investigate the effects of various educational factors, curricula, and resources on student accomplishment. The most recent worldwide comprehensive study of math performance was TIMSS 2019. Data from 250.000 students, 30.000 teachers, and 8.000 administrators were used to evaluate grade 8 mathematics in the 39 participating nations, including Morocco. International achievement exams, particularly math and science, have frequently shown the low performance of Moroccan students. Lower test scores are thought to reflect lower levels of educational quality and human capital [18], which have a significant and negative impact on worker outcomes in the labor market as well as the nation's economic growth and human development [15], [20], [24], [1]. As a result, Morocco was ranked lower in the rankings of participating countries in the most recent International Mathematics and Science Study TIMSS 2019. Researchers may now examine the many variables that affect student accomplishment thanks to the widely accessible, comprehensive, and policy-relevant indicators evaluated by TIMSS 2019. This research can aid in the decision-making of policymakers, teachers, and students.

This research investigates the potential causes of Moroccan eighth-graders poor mathematical performance in TIMSS 2019 by using machine learning (ML) methods. Because there is currently a shortage of research on applying machine learning (ML) techniques for international assessment studies, we look into characteristics of the teacher, school, and student variables that have a high significance in predicting the mathematics test results of Moroccan college students. Researchers may compare the outcomes in terms of student accomplishment thanks to TIMSS, which also offers information on the impact of policies and practices in the educational systems of each participating country [25]. Standard statistical techniques, including regression analysis, multilevel modeling, and component analysis, are frequently used when examining student performance in science and mathematics. [19], [28], [31], [34], [37]. However, as previously indicated, these techniques have inherent shortcomings, mainly when dealing with strangely distributed data. Additionally, it might be challenging to make accurate forecasts [30] when the situation is highly complicated. Machine learning techniques are effective at finding outliers [10] and may be used to get around some of the drawbacks of these conventional approaches. This project aims to incorporate all TIMSS-provided student, teacher, and school variables in 3 models to identify significant predictors of students' mathematical success using machine learning methods such as SVM, XGBoost, and Random forests.

The remaining sections are arranged as follows: section II displays related works to this research, section III provides an overview of the data, and an explanation of ML approaches,

Zakaria Khoudi and Mourad Nachaoui are with the Equipe Mathématiques et Interactions, Faculté des Sciences et Techniques, Sultan Moulay Slimane University, Beni Mellal, Morocco.
(E-mail: zakaria.khoudi@usms.ma; m.nachaoui@usms.ma)
Soufiane Lyaqini is with the LAMSAD Laboratory, Ecole Nationale des Sciences Appliquees, Hassan First University, Settat, Morocco.
(E-mail: lyaqini.soufiane@gmail.com)
* Whom correspondence should be addressed

DOI: 10.36244/ICJ.2024.5.2

section IV shows and analyzes estimate findings, and section VI wraps up the article.

II. Related Works

Most studies described above were created to collect data on assessing education systems. Researchers may compare/improve educational policies and their outcomes in terms of success by using educational data, which includes assessments of various nations' educational performance using tools like TIMSS. Modeling the TIMSS data using machine learning techniques was the subject of much significant research. Finding the variables affecting students' mathematical achievement was the goal of Hammouri [14]. A student's success in mathematics is determined by their attitude, accomplishment, confidence in their mathematical abilities, and sense of the importance of mathematics, according to the findings produced using the TIMSS dataset. Liu and Meng [21] investigated the TIMSS 2003 dataset and used these variables to compare the mathematical awareness of high and poor achievers in East Asia and America. For the TIMSS 2011 data, Askin and Gokalp [2] looked at the variables that affect students' performance in their academic endeavors. They employed LR, and ANN approaches to evaluate the performance of prediction and classification. The confidence of the students was shown to be the most important component. Topçu, Erbilgin utilized data from the TIMSS 2011, and Arkan [34] to look at what influences Turkish and Korean students' achievement in science and mathematics. The educational ramifications of their results were also highlighted. Data from TIMSS 2011 were subjected to DT, NB, LR, and ANN applications by Klç-Depren, Askin, and Oz [11]. They sought to identify the top algorithm for categorizing Turkish eighth-graders based on various performance indicators for their proficiency in mathematics. Filiz and Oz [12] used the TIMSS 2015 scientific data to apply the EDM approach. They discovered the elements that contribute most to scientific accomplishment. Baranyi and Gilanyi [3], Chmielewska [8], [9] unveil the notion of "Mathability". This idea highlights how artificial intelligence and human cognitive capacities may be integrated, mainly when comprehending and solving mathematical problems. They demonstrate how improvements in information and communications technologies may improve human mathematic abilities. The authors highlight the possibility of integrating cognitive processes and technology tools in improving mathematical abilities by demonstrating math ability in complicated problem-solving, such as utilizing Maple to solve mathematical equations and inequalities. This study is in line with comprehending the variables that affect students' performance in mathematics, especially as it relates to how technological interventions and cognitive capacities can work together to enhance learning outcomes in mathematics. This is an essential area of research for studies that use machine learning techniques to predict and analyze academic performance. Contribute most to the advancement of science.

Since little research in Morocco focuses on contextual factors impacting children's mathematics competency, this study fills the knowledge gap by utilizing machine learning methods to examine the impact of the student, teacher, and school variables on high and poor-performing children. Two research concerns are addressed in this study:

- Can context help Moroccan eighth-graders with high mathematics levels and those with low mathematics levels be distinguished? If so, what are they?
- What contextual elements should the ideal feature set have that would influence Moroccan pupils' mathematics competence, both high and low?

III. Material and Methods

A. Data description

The study's dataset was obtained from TIMSS 2019 database conducted by the IEA (International Association for the Evaluation of Educational Achievement) https://timss2019.org/international-database/, and it used Moroccan student's grade 8 data files after combining the student, teacher, and school data using IDB Analyzer (version 5.0) offered by IEA, initially 8458 entries almost equal in terms of the number of males and females (Figure 1) with 700 variables (Figure 2) were made.

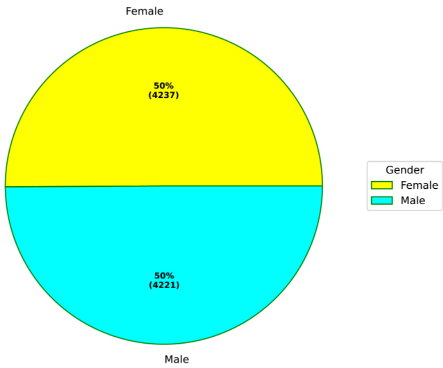

Fig. 1. Gender in the dataset

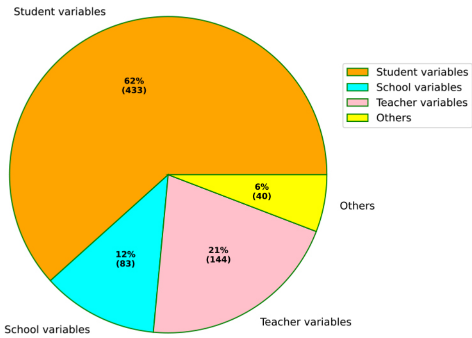

Fig. 2. Structure of dataset variables

Special Issue
of the Infocommunication Journal

These variables are divided into four parts below:

- The variables beginning with BC****** refer to a school background (e.g., BCBG16I, BCBG19, BCBG15D, Etc.), a total of 83 variables.
- The variables beginning with BS****** refer to a student's background (e.g., BSBG10, BSBG11B, BSBG13C, BSBM19B, Etc.), a total of 433 variables.
- The variables beginning with BT****** refer to a teacher background (e.g., BTBM17CA, BTBM17CC, BTBG12F, Etc.), a total of 144 variables.
- IDs, the number of teachers, weights and file maintenance are the remaining variables, totalling 40.

1) Students background: As preliminary observations on student background variables:

- We have randomly chosen a sample of 2000 students. Also, eight variables have been chosen. There is a clear association between teaching clarity across all scientific subjects (possible link to instructional quality at the school level). (Figure 3)
- A wide range in the number of books in the house, more than 50 per cent of students have a few books at home. (Figure 4)
- The most significant level of parental education varies greatly, although low level "Some Primary, Lower Secondary or No School" is the most typical. (Figure 5)

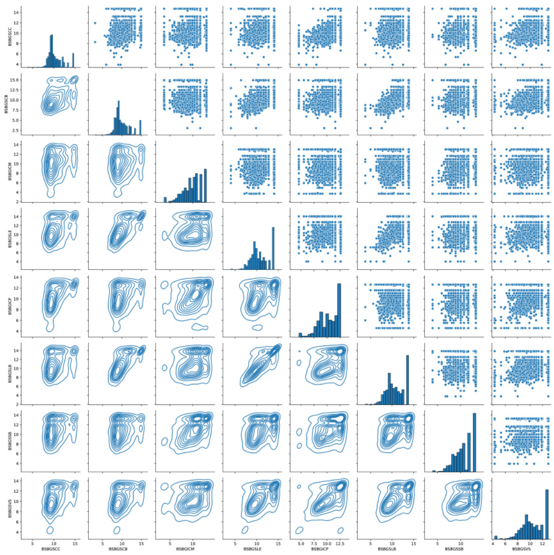

Fig. 3. Correlations between a few of the demographic scores for students' scores

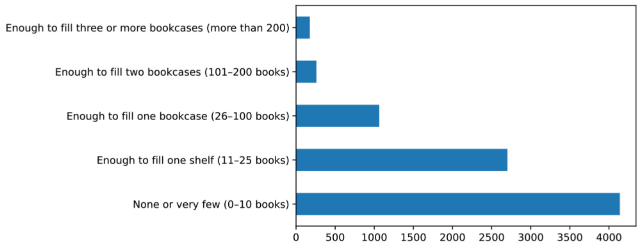

Fig. 4. Each student's home library's number of books

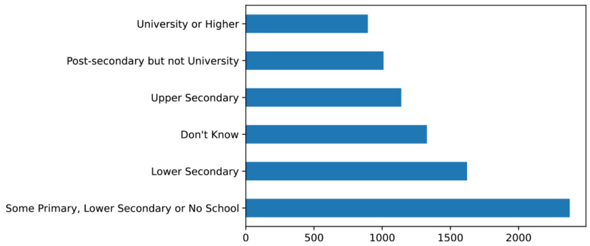

Fig. 5. The highest degree of education attained by the parents of the student

2) Teachers background: The dataset contains 260 math teachers divided into 251 high schools; therefore, each student is taught by one teacher. As an initial finding on the background factors of math teachers are

- Roughly Normal age distribution among teachers, the majority of teachers are young. Their ages range from 25 to 39. (Figure 6)
- Bachelor's degrees are where the majority of education for teachers is found. (Figure 7)
- There are noticeably more male than female teachers. Males outnumber females by more than 50 per cent. (Figure 8)

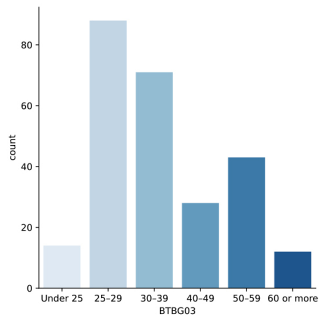

Fig. 6. The age range of math instructors

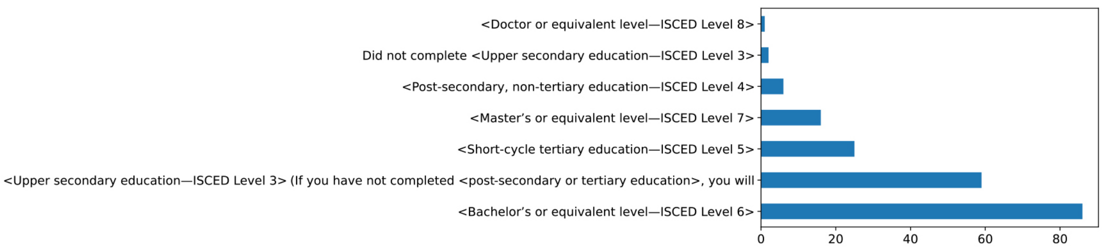

Fig. 7. The frequency of math instructors' academic successs

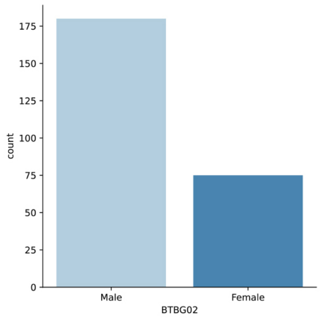

Fig. 8. Math instructors' gender distribution

3) Schools background: The initial dataset includes 251 private and public Moroccan high schools, randomly selected from rural and urban areas. Figure 9 highlights the relationship between discipline, academic competencies, and shortage of equipment in these high schools. As preliminary notices, we found:

- There does not seem to be a relationship between school disciplinary issues and an emphasis on academic performance or a lack of resources.
- The lack of resources for math and science is closely connected because schools seem to have difficulties with both subjects. Meanwhile, some other schools do not have such difficulties.
- But few of these spectrums have little problems with discipline and resources.

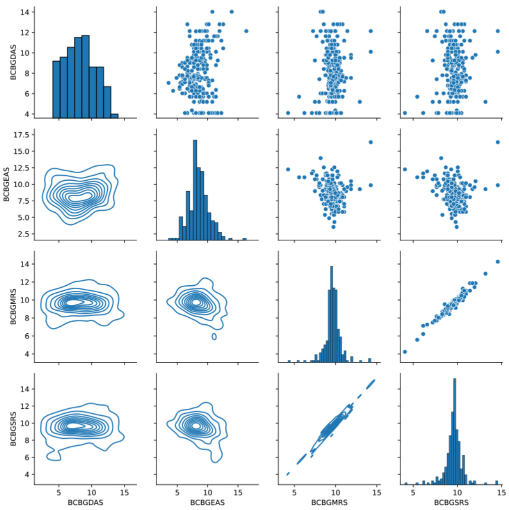

Fig. 9. Shortages, academics, and disciplinary relationships

B. Response variable

TIMSS awards 5 PVs (plausible values) to a student's academic achievement. PVs are then classified into five levels: 1 (Below Low), 2 (Low), 3 (Intermediate), 4 (High), and 5 (Extremely High) (Advanced), and these classed variables are referred to as categorical benchmark variables using the five categories of benchmark variables and a majority vote, a single class was created for each student's math achievement. For example, if a student's benchmark variables (BSMIBM01 through BSMIBM05) were 3, 1, 1, 1, 3, the student's class was coded as 1. After combining the 8458 Moroccan 8th graders using the IDB Analyzer (version 5.0), 68 of them had ties (Figure 10). After removing the ties, the final sample for this study was 8390 students.

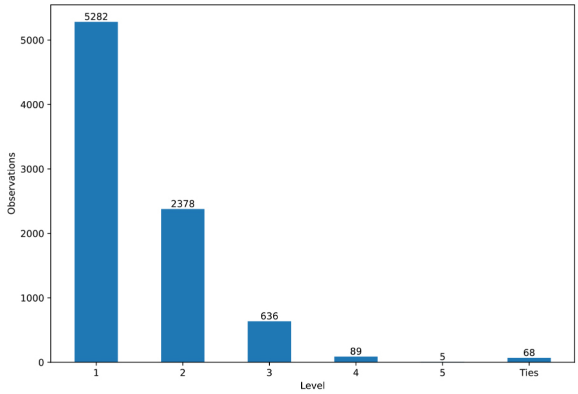

Fig. 10. TIMSS 2019 Moroccan 8th Grade Math Majority Vote Results

Because the proportions of the levels were severely imbalanced, the final four levels were compacted in this study. The study's response variable was whether the student attained the 'Low' level (Level 1: coded as 0) or not (Levels 2, 3, 4, 5, coded as 1). There were 5282 and 3108 students in each group (Figure 11).

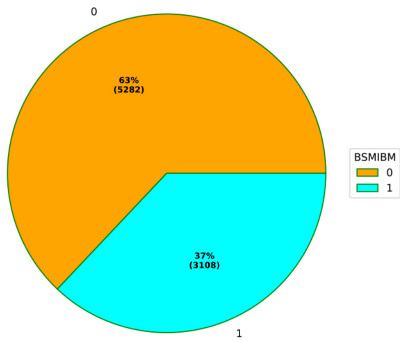

Fig. 11. Response variable categories

C. Data Pre-processing

The combined dataset originally had 700 variables, and 300 of them were eliminated, as shown below:

- 41 variables relevant to IDs (e.g., IDPOP, IDSTUD, IDCLASS, Etc.), weights (e.g., total school weight, total class weight Etc.), file maintenance (e.g., VERSION, IDB Identifier, Etc.), and redundant variables (e.g., ITSEX) were removed from the explanatory variable pool.
- 61 variables with 100% missingness (e.g., BSBS22D, BSBS22E, BSDGSEC, BSDGSCS, etc.) were eliminated.
- 29 numerical scale score variables of category indices for constructs such as 'Students Like Learning Mathematics Lessons' (e.g., BSBGSLM, BSBGICM, BSBGDML, Etc.) were deleted because numerical scale scores and categorical indices transmit essentially the same information.

- 95 benchmark variables and PVs were removed, except for the newly generated benchmark variable, BSMIBM, which functioned as the study's response variable. Including these academic performances, factors would have dominated the model, providing little meaningful information for predicting students' math proficiency.
- Missing values include omitted or invalid, logically inapplicable, and not administered. The "omitted" replies resulted from respondents' carelessness or refusal to answer the question. Because the educational systems of the TIMSS participating nations were so diverse that the replies were mainly "not applicable-administered". Each variable's missing rate was determined, and 74 variables with missing rates of more than 10% were deleted from the dataset.

The cleaned dataset had 400 variables from 8390 students, including 205 students, 122 teachers, and 73 school variables (Figure 12). Only 11 were continuous variables with hours and numbers as replies (e.g., BTBG11, BTBM14, BCBG18, BTDMNUM, whereas the rest 389 were Likert-type scaled.

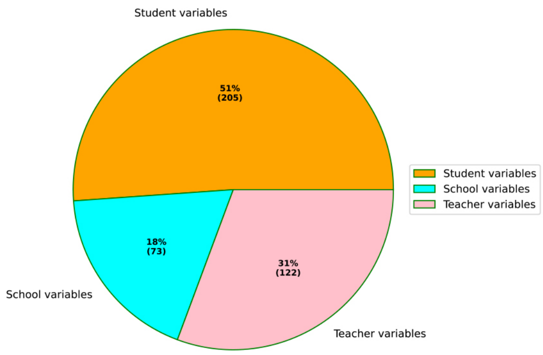

Fig. 12. The final structure of dataset variables

For missing data imputation, we used the SimpleImputer function from sklearn.impute package, which successfully retained all 8390 observations with 400 variables. The SimpleImputer begins by replacing missing data with mode values. The 8390 observations are almost equal in terms of the number of males and females (Figure 13).

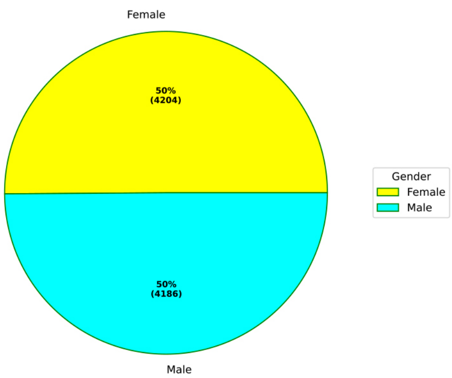

Fig. 13. Gender in the dataset after cleaning

D. Classification Models

After pre-processing and preparing the real datasets, we obtained knowledge of the data set. Important variables affecting students' mathematical achievement may be predicted using a variety of classifiers. Our experiment uses the supervised models Support Vector Machine (SVM), Random Forest, and XGBoost since no machine learning approach consistently produces the best predictions.

SVM is a method for creating classifiers [16], [22],[27]. Its objective is to create a judgment boundary between two groups that enables labels to be predicted from one or more feature vectors. The hyperplane, a judgment boundary, is angled to be as close as feasible to one of the classes' closest data points. The closest-together points form help vectors.

Random forest [5] is a Machine Learning method for handling classification and regression issues. It is based on ensemble learning, a method that combines several classifiers to provide answers to complicated problems. Many decision trees are used in it. With bagging or bootstrap aggregation, the algorithm's created forest is trained. We selected it due its frequent use for feature selection in a data science workflow [26], [13], [17]. The average impurity reduction derived from every decision tree in the forest may determine how important a feature is, whatever the data's linear or non-linear nature. It is true (linearly inseparable).

XGBoost [7] is an application of gradient-boosted decision trees designed to quickly and accurately resolve various data science challenges. When dealing with the bias-variance tradeoff, boosting algorithms are pretty helpful. Boosting handles both the elements of bias and variance, in contrast to bagging algorithms that solely correct for excessive variance in a model. One of the top machine learning models at the moment. Due of its speed, effectiveness, and scalability [36], [6], [7]. Extracting feature significance from the XGBoost model is simple compared to other machine learning models since it is effectively an ensemble of decision trees. Because of this, we decided to use XGBoost in this research to determine feature relevance.

The observations were divided into random training and test data sets, with 7:3 being the standard ratio. The test data was utilized for generalizing the model once it had been evaluated using the training data.

In particular, the response variable BSMIBM was used as a stratifying variable to maintain the ratio of "Low" to "Others" in the training and test datasets. The student counts for the training and test data sets are shown in (Table I) for each level.

TABLE I
TRAINING AND TEST DATA

	0(Low : 63%)	1(Others : 37%)
Data(n = 8390)	5286	3104
Training data (n = 5873)	3700	2173
Test data (n = 2517)	1586	931

The ability to distinguish between instances and controls is a feature of XGboost. The optimal values for the hyperparameters of the best model are found using grid search with 10-fold Cross Validation [29] within the train split. The recommended hyperparameter settings for XGBoost are learning_rate = 0.3, max_depth = 10, and n_estimators = 400. Then, we take the critical characteristics out of XGboost. The primary function of XGBoost, a component of its Python library, is used to assess the applicability of each feature. By adding a feature, the method determines the average training loss decrease for each splitting.

IV. Experimental results

The results reported in this section are based on the three classification models previously covered in the "Classification Models" section.

A. Testing and evaluation

We employed a variety of assessment metrics, or indicators, to assess and comprehend the performance of the models. A confusion matrix [32] is a simple technique to see how well a model works. For our purposes, the positive class represents high-level students, while the negative class represents low-level students.

TABLE II
Confusion matrix in general

Actual Positive	True Positive (TP)	False Negative (FN) Type I error
Actual Negative	False Positive (FP) Type II error	True Negative (TN)
	Predicted Positive	Predicted Negative

We assess the effectiveness of our prediction models using four ML model assessment indicators [35]. Accuracy and F1-score are the two measures used to evaluate classifier performance. The percentage of all correct predictions serves as a measure of accuracy. The method for calculating accuracy is shown in this equation:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

True Positive is abbreviated as TP. True negative is TN, whereas false positive is FP and FN for "false negative". The following equation represents the F1-score. For a classification task, it reflects the harmonic means of the values for accuracy (Predicted features) and recall (sensitivity):

$$F1 = \frac{2 * (Precision * Recall)}{Precision + Recall}$$

Area Under Curve (AUC) [23]: The typical rule of thumb is using a probability threshold of 0.5 for classification predictions in binary classification issues. Using a different threshold might be preferable in a few circumstances when this one would not hold. The most popular way to display a binary classifier's performance at various thresholds is through a Receiver Operating Characteristic (ROC) curve. The True Positive Rate is plotted against the False Positive Rate to achieve it. The false-positive rate is computed as (1 - Specificity). The ROC plot enables us to calculate the Area Under the Curve by estimating the probability that a classifier will score a randomly picked positive instance higher than a randomly selected negative one (commonly abbreviated as simply the AUC).

B. Results

The results of our performance assessment have been presented in Table III, Figures (14,15,16,17,18,19) so that we can identify the key indicators of students' mathematical competency. Our investigation employed three different models: SVM, Random Forests, and XGBoost. However, the scores for these two models' accuracy might have been higher, indicating they may have had tremendous success fixing the present problem. XGBoost models, on the other hand, produced excellent results, with an accuracy of **81.92** and an F1- Score of **74.42**. These findings suggest that XGBoost models might be better suited to identify the ideal proportion of crucial contextual variables that affect eighth-grade kids' mathematical aptitude.

TABLE III
Results of performance assessment

Models	Accuracy	F1-Score
SVM	65.87	16.03
Random Forest	79.18	65.39
XGBoost	**81.92**	**74.42**

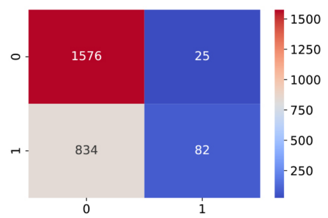

Fig. 14. SVM confusion matrix

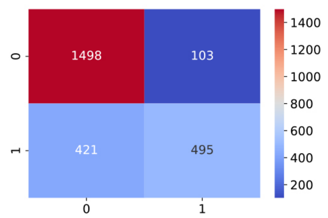

Fig. 15. Random Forests confusion matrix

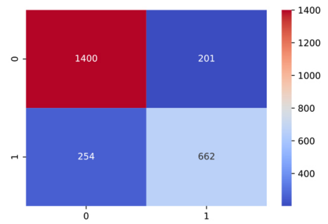

Fig. 16. XGboost confusion matrix

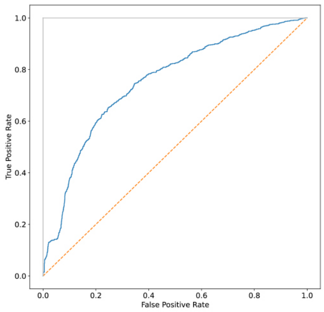

Fig. 17. SVM AUC-ROC Curve

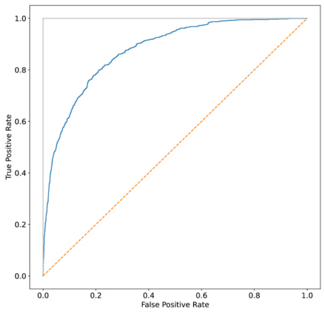

Fig. 18. Random Forests confusion matrix

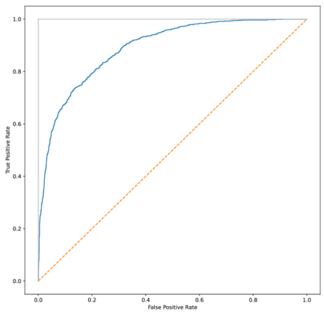

Fig. 19. XGboost confusion matrix

C. Feature selection

A machine learning approach known as feature selection involves selecting a smaller sample of critical features from a more extensive range of unimportant or irrelevant data while minimizing information loss. It is often used when dealing with many characteristics and few instances. Feature selection may accelerate processing and increase prediction accuracy by decreasing the quantity of data that the algorithm must evaluate [33].

In machine learning, "feature importance" is used to pinpoint the features or variables in a dataset that significantly affect the target variable. This method is essential for understanding the interactions and behavior of variables inside a model and aids in identifying the variables that are most important for result prediction.

In our work, feature selection assigns a score to each input feature in a particular model using measurements of feature relevance as criteria. Higher scores imply features that significantly impact the prediction model, representing each feature's relative relevance. There are other ways to calculate feature significance, but XGBoost yields the best results. The significance of several characteristics is shown in Figure 20 from XGBoost, which shows how well the model predicts feature importance. The top 12 traits are highlighted in the graphic, which lists them in decreasing order of importance.

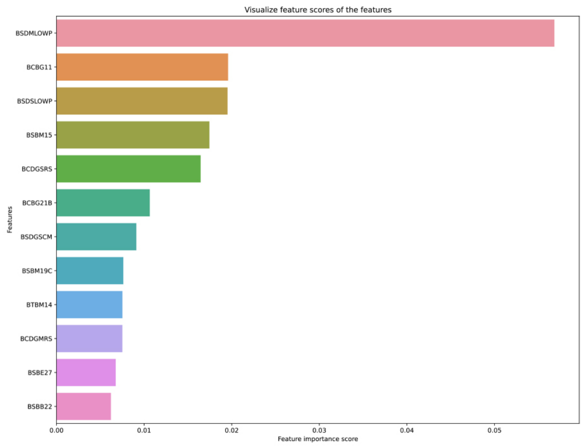

Fig. 20. XGBoost Features Importance

V. Discussion

The experiment's findings demonstrate that XGBoost offers the essential components required for the analysis. The differences between BSDMLOWP, BCBG11, BSDSLOWP, BSBM15, BCDGSRS, BCBG21B, BSDGSCM, BSBM19C, BTBM14, BCDGMRS, BSBE27, and BSBB22 are in the sequence in which they are provided. When determining the mathematical competency of Moroccan students, these 12 qualities provide the best selections for predictor variables since they regularly appear in XGBoost and have high relevance. Table IV provides descriptions of these salient traits. These essential characteristics may be used as predictor variables, according to the descriptions that have been given.

TABLE IV
Description of essential characteristics

Feature	Description	Type
BSDMLOWP	MATHEMATICS ACHIEVEMENT TOO LOW FOR ESTIMATION	1: Yes; 2: No
BCBG11	CLASSROOM LIBRARIES	1: Yes; 2: No
BSDSLOWP	SCIENCE ACHIEVEMENT TOO LOW FOR ESTIMATION	1: Yes; 2: No
BSBM15	WORK ON YOUR OWN	1: Every or almost every lesson; 2: About half the lessons; 3: Some lessons; 4: Never
BCDGSRS	INSTRUCTION AFFECTED BY SCIENCE RESOURCE SHORTAGE	1: Not Affected; 2: Affected; 3: Affected A Lo
BCBG21B	QUALIFICATIONS IN EDUCATIONAL LEADERSHIP/ ISCED 7	1: Yes; 2: No
BSDGSCM	STUDENT CONFIDENT IN MATHEMATICS	1: Very Confident in Mathematics; 2: Somewhat Confident in Mathematics; 3: Not Confident in Mathematics
BSBM19C	MATHEMATICS NOT MY STRENGTH	1: Agree a lot; 2: Agree a little; 3: Disagree a little; 4: Disagree a lot
BTBM14	TIME SPENT MATH INSTRUCTION MINUTES	1 to 1800
BCDGMRS	INSTRUCTION AFFECTED BY MATHEMATICS RESOURCE SHORTAGE	1: Not Affected; 2: Affected; 3: Affected A Lot
BSBE27	HOW OFTEN CONDUCT EXPERIMENTS IN EARTH SCIENCE	1: At least once a week; 2: Once or twice a month; 3: A few times a year; 4: Never
BSBB22	HOW OFTEN CONDUCT EXPERIMENTS IN BIOLOGY	1: At least once a week; 2: Once or twice a month; 3: A few times a year; 4: Never

To go over the outcomes in light of the two study-related research concerns:

- **Q1: Can context help Moroccan eighth-graders with high mathematics levels and those with low mathematics levels be distinguished? If so, what are they?**

Context may assist in separating Moroccan eighth-graders who do well in mathematics from those who do not. According to our research, the following contextual factors are crucial for drawing this distinction:

1) Student factors: It was discovered that seven different student-related factors significantly affected mathematics ability. These include kids' accomplishment levels in science and math, their self-perceptions of their mathematical prowess, the regularity with which they work independently in math class, and their confidence in mathematics, among other things. The study showed that variables about students were more critical in influencing academic achievement than those about instructors and educational institutions. This highlights how important it is to tailor instructional tactics to each learner's unique requirements and features.
2) School and Teacher Variables: The impact of teacher and school characteristics was also assessed in this research. These include classroom libraries, educational leadership credentials, and the effects of scientific and math resource shortages.
3) Teacher Variables: An assessment of the amount of time spent on math teaching was part of the research, and the results showed how important it is to students' academic achievement.
4) Overall Effect and Educational Implications: According to the study, variables that affect students significantly influence their academic achievement more than those that affect educators and educational institutions. Given the present focus on teacher and school accountability in education, this emphasizes the need for more research to examine and identify attributes of teachers and schools.

the following contextual factors help differentiate Moroccan eighth-graders with high and low mathematics achievement levels: confidence in mathematics, frequency of independent math work, perception of math as a strength or weakness, time spent on math instruction, the impact of resource scarcity, and overall mathematics achievement level. These elements provide a thorough grasp of the pupils' mathematical aptitude.

- **Q2: What contextual elements should the ideal feature set have that would influence Moroccan pupils' mathematics competence, both high and low?**

Various contextual factors that include many facets of the student's educational environment, personal attitudes, and accessible resources should be included in the optimal feature set for impacting Moroccan students' high and poor mathematical competency. The following components are essential, according to the TIMSS 2019 study:

1) Student Confidence in Mathematics (BSDGSCM): This variable assesses students' confidence in their aptitude for mathematics. Students' motivation, enthusiasm, and achievement in mathematics may all be significantly impacted by developing their confidence. Students with differing degrees of mathematical proficiency might benefit from tailored interventions and teaching strategies.
2) Frequency of Independent Math Work (BSBM15): Students' quantity of independent math work in the classroom is significant. Students' critical thinking and quantitative skills may be improved by promoting self-directed learning and problem-solving.
3) Viewing Mathematics as a Weakness (BSBM19C): Students' performance may be impacted by how they see themselves in mathematics, specifically if they consider math a personal weakness. Addressing this perspective and making it a strength may be helpful for kids with limited mathematical proficiency.

4) Time Spent on Math Teaching (BTBM14): The amount of time dedicated to math teaching is crucial. Ensuring enough time for teaching may enhance students' mathematical understanding and performance. Educators and schools must take this into account when developing a curriculum.
5) The Effect of Mathematical Resource Scarcity (BCDGMRS): The quality and quantity of mathematical resources greatly influence how well students are taught mathematics. Improving students' comprehension and interest in mathematics requires addressing the need for more resources, particularly for those with lower competency levels.
6) Mathematics Achievement Level (BSDMLOWP): This variable aids in identifying pupils whose competency in the subject may be too low to evaluate reliably. Giving these pupils focused assistance may help them become more proficient in mathematics.

Including these contextual factors in instructional techniques and interventions may produce a more customized and successful strategy for raising Moroccan students' mathematical proficiency. It entails concentrating on each student's attitudes, confidence, learning environment, and content and teaching.

The study's findings have implications for educational policy and practice as they may help educators and policymakers create focused interventions and resources that will enhance student outcomes, especially in mathematics. Having a better understanding of the main performance-influencing variables may aid in the development of instructional and support systems that are more successful.

Several vital elements emerge in the concept of "Mathability" as defined by Baranyi and Gilanyi [3], along with the field of cognitive infocommunications. Firstly, cognitive infocommunications seeks to understand the coevolution of artificial and natural cognitive processes, shedding light on how students learn and understand mathematical concepts and how artificial systems like machine learning algorithms can be utilized to assess and predict student performance. Secondly, Baranyi and Gilanyi's notion of "Mathability" refers to the capability to simulate and enhance human mathematical skills, aligning closely with the study's focus on factors affecting students' math performance. Insights into learning environments, achievement levels, and student confidence can inform the development of artificial systems or educational aids replicating and augmenting these aspects. Thirdly, applying these concepts in educational technology can lead to creating tools that support personalized learning, addressing individual student needs and boosting overall math skills. In summary, viewing the study's findings on factors impacting the mathematical performance of Moroccan eighth-graders through the lenses of mathability and cognitive infocommunications can pave the way for more effective teaching methodologies and technological aids that nurture and support mathematical competence.

This research examined 700 student, teacher, and school features from the TIMSS 2019 to predict the mathematical performance of Moroccan eighth-graders. It identified the top 12 variables using XGBoost. Only a small number of teacher and school features were also found by XGBoost, even though all seven variables were related to students, indicating factors affecting students had a more significant impact on their academic performance than those affecting teachers and schools. This was expected, but more research is needed to study and pinpoint teacher and school features, especially in the current climate of increased focus on teacher and school responsibility. Education researchers must provide teachers and schools with research-based assistance since they are now more than ever held accountable for their student's academic success. That was this empirical study's primary goal.

VI. Conclusion

This research thoroughly examines important factors affecting Moroccan students' mathematical skills, including student, teacher, and school characteristics. It highlights twelve key predictor characteristics, such as a poor estimate of math performance, individual math work, and insufficient science and math teaching resources. These variables are essential markers as there is a strong correlation between them and pupils' arithmetic performance. Additionally, the research reveals other noteworthy characteristics that provide insightful information to scholars, professionals, and decision-makers who want to improve Moroccan students' mathematical skills by implementing methods and interventions specific to these impactful elements.

References

[1] Noam Angrist, Simeon Djankov, Pinelopi K Goldberg, and Harry A Patrinos. Measuring human capital using global learning data. *Nature*, 592(7854):403–408, 2021.

[2] Oykum Esra Askin and Fulya Gokalp. Comparing the predictive and classification performances of logistic regression and neural networks: a case study on timss 2011. *Procedia-Social and Behavioral Sciences*, 106:667–676, 2013.

[3] Peter Baranyi and Attila Gilányi. Mathability: Emulating and enhancing human mathematical capabilities. In *2013 IEEE 4th international conference on cognitive Infocommunications (CogInfoCom)*, pages 555–558. IEEE, 2013.

[4] Antoine Bodin and IREM de Franche-Comté. Ce qui est vraiment évalué par pisa en mathématiques. ce qui ne l'est pas. un point de vue français. *Bulletin de l'APMEP*, 463:240–265, 2006.

[5] Leo Breiman. Random forests. *Machine learning*, 45(1):5–32, 2001.

[6] Jason Brownlee. *XGBoost With python: Gradient boosted trees with XGBoost and scikit-learn*. Machine Learning Mastery, 2016.

[7] Tianqi Chen and Carlos Guestrin. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pages 785–794, 2016.

[8] Katarzyna Chmielewska. From mathability to learnability. In *2019 10th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*, pages 513–516. IEEE, 2019.

[9] Katarzyna Chmielewska. Towards "learnability". *Acta Polytechnica Hungarica*, 19(1), 2022.

[10] Karel Dejaeger, Frank Goethals, Antonio Giangreco, Lapo Mola, and Bart Baesens. Gaining insight into student satisfaction using comprehensible data mining techniques. *European Journal of Operational Research*, 218(2):548–562, 2012.

[11] Serpil Kılıç Depren, Öykü̈m Esra Aşkın, and Ersoy Öz. Identifying the classification performances of educational data mining methods: A case study for timss. Educational Sciences: *Theory & Practice*, 17(5), 2017.

[12] Enes Filiz and Ersoy Öz. Finding the best algorithms and effective factors in classification of turkish science student success. *Journal of Baltic Science Education*, 18(2):239–253, 2019.

[13] Baptiste Gregorutti, Bertrand Michel, and Philippe Saint-Pierre. Correlation and variable importance in random forests. *Statistics and Computing*, 27(3):659–678, 2017.

[14] Hind Hammouri. Attitudinal and motivational variables related to mathematics achievement in jordan: Findings from the third international mathematics and science study (timss). *Educational research*, 46(3):241–257, 2004.

[15] Eric A Hanushek and Dennis D Kimko. Schooling, labor-force quality, and the growth of nations. *American economic review*, 90(5):1184–1208, 2000.

[16] Marti A. Hearst, Susan T Dumais, Edgar Osuna, John Platt, and Bernhard Scholkopf. Support vector machines. *IEEE Intelligent Systems and their applications*, 13(4):18–28, 1998.

[17] Mia Huljanah, Zuherman Rustam, Suarsih Utama, and Titin Siswantining. Feature selection using random forest classifier for predicting prostate cancer. In *IOP Conference Series: Materials Science and Engineering*, volume 546, page 052 031. IOP Publishing, 2019.

[18] Zakaria Khoudi, Mourad Nachaoui, and Soufiane Lyaqini. Identifying the contextual factors related to the reading performance of moroccan fourth-grade students from a machine learning-based approach. *Education and Information Technologies*, pages 1–27, 2023.

[19] Serpil Kilic and Öyküm Esra Askin. Parental influence on students' mathematics achievement: The comparative study of turkey and best performer countries in timss 2011. *Procedia-social and behavioral sciences*, 106:2000–2007, 2013.

[20] Jong-Wha Lee and Robert J Barro. Schooling quality in a cross–section of countries. *Economica*, 68(272):465–488, 2001.

[21] Shujie Liu and Lingqi Meng. Re-examining factor structure of the attitudinal items from timss 2003 in cross-cultural study of mathematics self-concept. *Educational Psychology*, 30(6):699–712, 2010.

[22] S Lyaqini, M Nachaoui, and A Hadri. An efficient primal-dual method for solving non-smooth machine learning problem. *Chaos, Solitons & Fractals*, 155:111754, 2022.

[23] Caren Marzban. The roc curve and the area under it as performance measures. *Weather and Forecasting*, 19(6):1106–1114, 2004.

[24] David Mitch. Education matters. Global schooling gains from the 19th to the 21st century. By Robert J. Barro and Jong-Wha Lee. New York: Oxford University Press, 2015. pp. xi, 289, 34.95, hardcover. *The Journal of Economic History*, 76(4):1271–1273, 2016.

[25] Ina VS Mullis, Michael O Martin, Pierre Foy, and Alka Arora. *TIMSS 2019 international results in mathematics*. ERIC, 2020.

[26] Md Mursalin, Yuan Zhang, Yuehui Chen, and Nitesh V Chawla. Automated epileptic seizure detection using improved correlation-based feature selection with random forest classifier. *Neurocomputing*, 241:204–214, 2017.

[27] Mourad Nachaoui, Soufiane Lyaqini, and Marouane Chaouch. Indicating if water is safe for human consumption using an enhanced machine learning approach. *Statistics, Optimization & Information Computing*, 11(1):70–81, 2023.

[28] Oliver Neuschmidt, Juliane Barth, and Dirk Hastedt. Trends in gender differences in mathematics and science (timss 1995–2003). *Studies in Educational Evaluation*, 34(2):56–72, 2008.

[29] Isaac Kofi Nti, Owusu Nyarko-Boateng, and Justice Aning. Performance of machine learning algorithms with different k values in k-fold cross-validation. *J. Inf. Technol. Comput. Sci*, 6:61–71, 2021.

[30] Muhammad A Razi and Kuriakose Athappilly. A comparative predictive analysis of neural networks (nns), nonlinear regression and classification and regression tree (cart) models. *Expert systems with applications*, 29(1):65–74, 2005.

[31] James B Schreiber. Scoring above the international average: A logistic regression model of the timss advanced mathematics exam. 2000.

[32] Thomas R Shultz, Scott E Fahlman, Susan Craw, Periklis Andritsos, Panayiotis Tsaparas, Ricardo Silva, Chris Drummond, PL Lanzi, J Gama, RP Wiegand, et al. Confusion matrix. *Encyclopedia of machine learning*, 61(8):209–209, 2011.

[33] Jiliang Tang, Salem Alelyani, and Huan Liu. Feature selection for classification: A review. *Data classification: Algorithms and applications*, page 37, 2014.

[34] Mustafa Sami Topçu, Evrim Erbilgin, and Serkan Arıkan. Factors predicting turkish and korean students' science and mathematics achievement in timss 2011. 2016.

[35] Ž Vujović et al. Classification model evaluation metrics. *International Journal of Advanced Computer Science and Applications*, 12(6):599–606, 2021.

[36] Yan Wang and Xuelei Sherry Ni. Axg boostrisk model via feature selection and bayesian hyper-parameter optimization. *arXiv preprint ArXiv:*1901.08433, 2019.

[37] Ludger Wößmann. Educational production in east asia: The impact of family background and schooling policies on student performance. *German Economic Review*, 6(3):331–353, 2005.

Zakaria Khoudi has received master's degree in information and communication technology from Abdelmalek Essaadi University. Now, he is a PhD student in computer science at Sultan Moulay Slimane University. His research interests are in data science, machine learning and data mining.

Mourad Nachaoui has received his PhD in 2011 from Nantes University. He is an professor at Sultan Moulay Slimane University. He was a guest professor at the Laboratory of Mathematics Jean Leray, Nantes University, from 2013 to 2019. His main research interests are applied mathematics, inverse problems, data sciences and computer sciences.

Soufiane Lyaqini has received his PhD from Sultan Moulay Slimane University. Now, he is an associate professor at Hassan 1 University. His research interests are in machine learning, data mining and optimization.