

Investigating the Predictive Factors of Secondary Education Mathematics Performance

Maria Elena Garcia-Ruiz, Ammar Ali-Khan

Maria Elena Garcia-Ruiz, Department of Mathematics Education, Faculty of Science and Mathematics, Universidad Complutense de Madrid (UCM), 28040 Madrid, Spain; .

Abstract: Several socioeconomic, environmental, ethnic, family, and educational factors influence an individual's academic performance and can determine their school performance in mathematics. Mathematical competence is one of the skills that allow students to build visions of the future from performance in the present. However, the perception that students have of mathematics, in addition to the teacher–student relationship, the classroom, gender, teaching–learning, and motivation are crucial factors for achieving an optimal academic performance and preventing school failure. The aim of the present study was: (1) to examine which variables of the dimensions “Learning Mathematics” and “School Environment” significantly contribute to the marks in the second quarter and quantify their relative importance; (2) to determine the optimal algorithm model for predicting the maximum gain in students' marks in the second quarter and quantifying it; and (3) to analyze the maximum gain in terms of gender. A total of 2018 high school students in Melilla were included in this cross-sectional study. Mathematical learning and the school environment were assessed using a validated 14-item questionnaire. Gain lift was employed to quantify the improvement in students' performance. The role of the classroom and teacher–student relationship had a greater influence on mathematics scores than affinity indicators, teaching, study time, teaching resources used, study aids, and motivation.

1. Introduction

The importance of mathematics transcends mere knowledge and is considered one of the key competencies for personal and professional development [1]. The learning and development of this competence in Spain appear in the school curriculum regulated by the Organic Law 8/2013, 9th December, for the improvement of educational quality (LOMCE), as well as the ECD 65/2015 of 21st January, which describes the relationships between competences, contents, and criteria for the evaluation of primary, high school, and post-16 education.

This policy also sets guidelines for improving students' academic performance. This is generally affected by socioeconomic, environmental, family, ethnic, and educational factors, among others [2]. Significant associations between academic performance and interpersonal relationships in the educational environment have also been noted, with the teacher–student relationship being the

most relevant [3]. On the other hand, the classroom climate, perceptions of learning mathematics, and personal motivation also play an important role in students' academic success [4–6].



Therefore, this study focuses on the analysis of several academic factors that influence the academic performance of students, such as learning mathematics, interpersonal relationships within the classroom, and the perception and motivation of students. In this context, an analysis of students' academic performance in mathematics is proposed using an innovative methodological approach that allows the maximum gain in academic performance in mathematics to be quantified through predictive calculations of the data collected in the study.

1.1. Mathematics and Its Learning

Mathematical competence is one of the skills that allow students to build visions of the future from performance in the present [7]. This competence also favors entrepreneurship, regardless of later studies that the students will pursue [8]. Despite the importance of mathematics as a fundamental competence for personal and professional development, it has been demonstrated that students in general have a low capacity for understanding and justifying basic concepts in mathematics [9]. These results are partly explained by a fear of mistakes, due to negative connotations at a personal and social level [10].

To improve these results, practical activities are recommended, because they facilitate meaningful learning and encourage interest in Science, Technology, Engineering, and Mathematics (STEM) [11]. However, students' perception of science practices decreases with age, affected by the demands for good grades in higher courses [12]. In addition, it has been shown that interest in STEM higher education greatly depends on the mathematics skills acquired during high school [13]. This competence decreases more rapidly among teenage girls, affected in part by gender stereotypes [13]. On the other hand, students' perception of mathematics affects their academic performance [14,15]. Although their overall assessment is not negative, students avoid higher STEM studies since the other options appear to be less rigorous and more accessible to them [16].

With regard to gender, there is a tendency to order and manage mathematical knowledge better in the case of girls, and girls also tend to seek help in the case of difficulty to a greater extent than boys [17]. Therefore, teachers must have effective feedback tools to enable them to know the interest and mathematical understanding of their students [18]. The perception of students of their teachers, classroom, classmates, and learning in the different subjects are factors that affect their motivation and academic performance [19,20].

From the teacher's point of view, the most appropriate teaching methodology should be considered, depending on the available resources, subject matter, and type of students, favoring their motivation in learning [21]. On the student side, it is essential for them to be involved in the design of powerful learning environments [22]. In addition, students who abandon school need to create effective learning environments [23]. The teacher–student relationship [24] and motivational aspects should also be taken into account, as everything necessarily impacts students and their academic performance [5,19].

1.2. Interactions in the Educational Environment

Classroom interactions and disruptive behavior among high school students are other important elements that impact academic achievement [25,26]. These factors may justify absenteeism in school and early school failure associated with academic failure [26]. Therefore, it is important to evaluate the classroom as a learning environment, articulate interventions aimed at its improvement [27], and address pedagogical aspects of mathematics learning in the psychosocial environment [28]. In addition, it should be noted that the effects of classmates in ethnically mixed classrooms are important for educational development during adolescence [29].

From a constructivist pedagogical perspective, errors are considered natural elements of learning processes. In the classroom, however, making mistakes is not only a personal

event, but also a social event, and is a major disincentive for students of mathematics [10]. This affects interest in higher STEM studies [13].

For all of the above, it is necessary to create and validate a scale to evaluate the climate that both teachers and students create in the classroom [30] because of its influence on the academic success of the students [31]. These tools are required especially among groups at risk of social exclusion, and exposed to early school failure [32].

The scientific literature postulates that the teacher–student relationship is another important aspect that affects students academically [33,34]. This relationship depends on the students' perception of the subject taught by the teacher and the difficulty as perceived by the teacher [35]. These perceptions significantly influence interest in the subject matter and condition the student's future academic career [36,37]. However, a recent study states that the effects of teachers on student outcomes are generally not significant for most [38]. In contrast, the influence of the teacher may reduce opportunities perceived by students at risk of social exclusion [39]. In this sense, good communication between the teacher and student is advisable, in order to know about and improve the mathematical competence of the student through attractive didactic strategies [18,40,41].

1.3. Motivation

The role of motivation and effort should not be forgotten, as these factors influence the learning of mathematics [6]. Beyond capacity, traits related to perseverance, such as determination and self-discipline, are associated with adaptive educational outcomes [42]. The scientific literature provides evidence of significant positive correlations between self-motivation and students' socio-cognitive perceptions of STEM activities in academic outcomes [43,44].

On the other hand, when analyzing the motivation to study, it has been observed that an external stimulus, such as a qualification or title, causes superficial cognitive affect that does not last. Conversely, internal stimuli, such as the desire for personal improvement, the idea of social mobility, the creative spirit of research, or commitment to a virtual object of learning, cause deep and permanent cognitive affects [6]. In association with this, the role of the teacher in implementing effective intervention strategies among an unmotivated student body is noteworthy [45]. Therefore, motivation can and should be encouraged from an attractive and participatory methodological and curricular proposal for students [46]. Therefore, it is important to change the perception of students towards failure. Adequate motivation changes the perception of students towards mathematics and promotes entrepreneurship [7,8].

1.4. Justification

According to the 2019 National School Board Report, the Autonomous City of Melilla has the highest early school failure rate in the country, with a value of 29.5% compared to Spain (17.9%) and the European Union (10.6%). Nationally, the average rate is 17.9%, 14.0% for girls and 21.7% for boys. In Melilla, gender-based percentages are 26.9% for girls and 31.9% for boys [47]. These values are far removed from the 2009 European Union forecasts for all member states. In these forecasts, it was established as an objective that “by 2020, the percentage of 15-year-olds with insufficient competence in reading, mathematics and science should be less than 15%.”

In addition, the National School Board Report shows that in the Autonomous City of Melilla, the ratio for secondary school classrooms is 29.6 and is 28.2 in post-16 education, which are both well above the national averages of 25.3 and 26.1, respectively [47]. These data suggest the need for the construction of new schools in the city and invite us to analyze the underlying academic problem from a broad perspective, with the participation of the entire educational community.

In this context, due to the high rate of school failure and the low results in mathematics exhibited by Melilla high school and post-16 education students, there is a need to evaluate the academic performance of mathematics students through the analysis of several indicators of the dimensions “*B. Learning Mathematics*” and “*C. School Environment*” that could be related to it.

Considering this, this study aims to achieve the following objectives: (1) To examine which variables of the dimensions “*B. Learning mathematics*” and “*C. School Environment*” significantly contribute to the scores of the second quarter and quantify their relative importance; (2) to determine the optimal algorithm model for predicting the maximum gain in students’ grades in the second quarter and quantifying it; and (3) to analyze maximum earnings based on gender. The following research questions are also raised, related to the qualifications obtained in the second trimester in mathematics: PI1: Could actions be taken to improve results? If yes, PI2: How could this be evaluated? PI3: How much could be improved? PI4: Are there gender differences?

2. Materials and Methods

To achieve the objectives set out in this study, a quasi-experimental method was used, without pre-post and without a control group. The selection of the study population considered the following inclusion criteria: (1) Residing in the Autonomous City of Melilla and (2) studying high school and post-16 education during the 2018/2019 academic year. With the above criteria, the population of students in the city amounts to 5875 students, of whom 50.84% are girls. Students from all schools and all levels participated in the study, avoiding any kind of bias. The total sample consisted of 2018 students (53.40% girls), of which 417 were first secondary school, 473 were second secondary school, 394 were third secondary school, 417 were fourth secondary school, 233 were first post-16 education, and 84 were second post-16 students.

Questionnaires were completed during school hours using Google Forms. This procedure ensured the reliability of data collection. In addition, all subjects who voluntarily participated in this study had been previously informed of the nature of the study. In order to access the classrooms, the Provincial Directorate of Melilla of the Ministry of Education and Vocational Training was authorized.

The instrument was designed ad hoc with 135 Likert-scale closed items with four graduated levels (1, none; 2, a little; 3, enough; and 4, a lot) for most items. For its creation, procedures for reviewing literature and determining dimensions were followed, which consisted of drafting of the items, expert content validation, construct validation, and reliability analysis [48]. The questionnaire was previously validated by expert judgement (content validity), both at the drafting level and at the level of adequacy of the items. A test questionnaire was also administered to detect aspects of improvement. Subsequently, with the complete data matrix, the instrument was validated through the Kaiser–Guttman and Tucker Lewis Index criteria, with a score of 1052.

This study took the results of 14 items of the dimensions “*B. Learning Mathematics*” and “*C. School*

Environment” from the general questionnaire. To analyze students’ academic performance, the second trimester math grade was used as a variable (the questionnaire was approved in the third evaluation). This is because grades are considered to be one of the most reliable predictors for analyzing academic performance [49]. The final version of the questionnaire used in this study is shown in Appendix A. Table 1 lists the different dimensions, their indicators, and the corresponding analyzed items.

2.1. Inferential Statistics

2.1.1. Global Adjustment of the Model

Prior to the adjustment of the multiple linear regression model (Table 2), linearity was calculated, since, if two independent variables were correlated, this would have affected the model. None of the variables had a Variance Inflation Factor (VIF) greater than 10, indicating that linearity was not a problem in the data.

The results of Table 2 indicate that the model is significant ($F(14,2024) = 21.38, p < 0.05$); that is, it predicts the scores obtained by the students in the second quarter quite well. However, the model only explains 12.27% of the variability of the response (R^2). In other words, 87.72% of the variation in the scores of the second quarter cannot be explained with the variables used. Therefore, there must be other variables that also influence the scores. It can also be observed that the variables which have a significant contribution among the analyzed dependent variables are TGM, CLB, PAV, and PCA. The intercept was also significant, which means that, if the score for all the independent variables is 0, the subjects will have a dependent variable score of 0.66.

Table 1. Relationship between dimensions, indicators, and items.

Dimension	Indicators	Cod e	Items
B. Learning of Mathematics	B.1 Affinity	TG M	Do you like mathematics?
	B.4 Teaching	PM E	Does your mathematics teacher explain things to you well in class?
	B.5 Study Time	LJM	From Monday to Thursday, how many hours do you dedicate to the study of mathematics?
	B.6 Resources	UL T	Do you use the course book to study mathematics?
		UA E	Do you use the notebook or class notes to study mathematics?
	B.7 Study Assistance	AP M	Does your academy or tutor help you when learning mathematics?
	B.8 Motivation:	MA M	Do your friends motivate you to study mathematics?
		MF M	Does your family motivate you to study mathematics?
MP M		Does your teacher motivate you to study mathematics?	
C. School Environment	C.1 Educational Center	ILS	Do you think your school is a place where you feel protected and safe?
	C.2 Room	CL B	Do you think your classroom is a good place to learn?
	C.3 Classmates	TCE	Do you work with your classmates while learning mathematics?
	C.4 Relationship with teachers	PA V	Do you think your teacher thinks highly of and values you?

PCA Does your mathematics teacher create a suitably respectful environment in class?

Note: Likert-scale for study variables: LJM: 1, none; 2, less than 1 h; 3, from 1 to 2 h; and 4, more than 2 h. For all other items, (1, none; 2, a little; 3, enough; and 4, a lot).

Table 2. Adjustment of the linear regression model.

Residuals:				
Min	1Q	Median	3Q	Max
-2.720	-1.160	-0.096	1.006	3.340
Coefficients				
	Estimate	Std. Error	T Value	Pr(> t)
Intercept	0.661314	0.134186	4.928	8.97×10^{-7} ***
TGM	0.333891	0.023150	14.423	$<2 \times 10^{-16}$ ***
PME	0.017346	0.029916	0.580	0.562102
LJM	0.029923	0.026354	1.135	0.256334
ULT	0.001705	0.026223	0.065	0.948162
UAE	0.036658	0.035496	1.033	0.301860
APM	-0.04208	0.026612	-1.582	0.113906
MAM	-0.02595	0.024325	-1.067	0.286073
MFM	0.051622	0.033720	1.531	0.125943
MPM	-0.00669	0.026851	-0.249	0.803236
ILS	0.024922	0.024280	1.026	0.304806
CLB	0.077882	0.024528	3.175	0.001520 **
TCE	-0.03188	0.024563	-1.298	0.194403
PAV	0.094211	0.027647	3.408	0.000668 ***
PCA	0.072070	0.030587	2.356	0.018559 *

Signif. codes: 0 '***'; 0.001 '**'; 0.01 '*'; 0.05 '.'; 0.1 ' '; Residual standard error: 1.295 on 2024 degrees of freedom; Multiple R-squared: 0.1288. Adjusted R-squared: 0.1228 F-statistic: 21.38 on 14 and 2024 DF, p -value $< 2.2 \times 10^{-16}$.

The relative importance of the predictors (Figure 1) was then determined using the “relaimpo” package.

According to previous methods (Figure 1), the most important variables of the study were TGM and PAV.

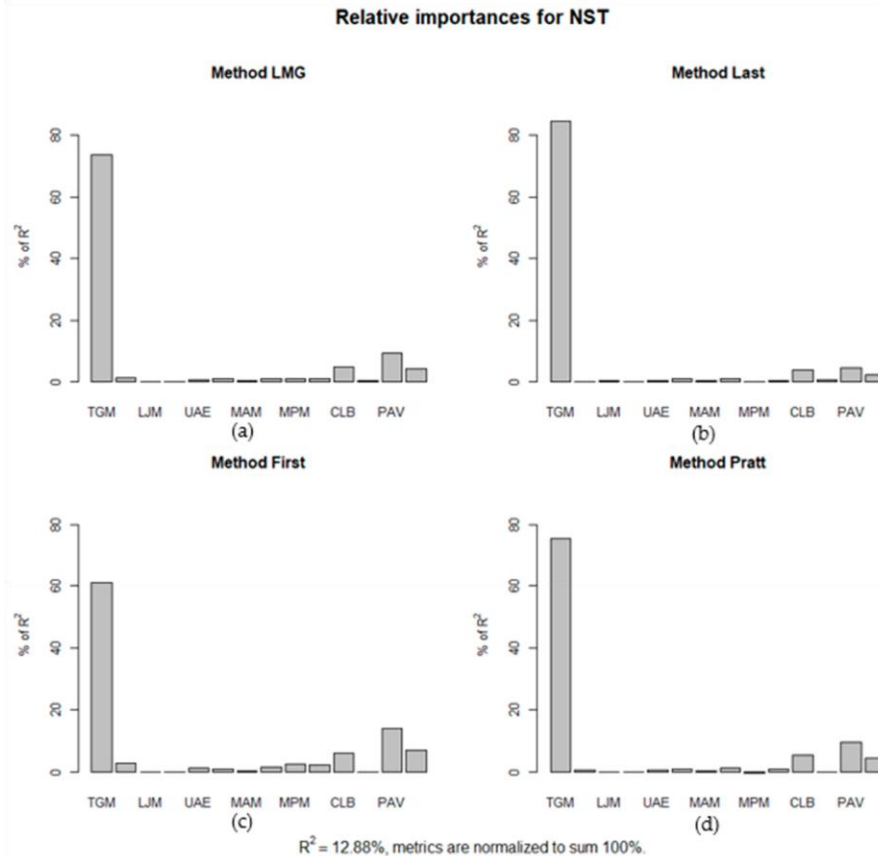


Figure 1. Relative importance of the research predictors: (a) The variance decomposition metric. Method LMG (Lindeman, Merenda, and Gold); (b) method last; (c) method first; and (d) Pratt method.

2.1.2. Parameters of the Model

The ANOVA tells us that the model is generally good; however, it does not indicate anything about the contribution of each variable or term in the model. In the results, the “coefficients” section can be seen, which contains the parameters of the model (the beta values) and their significance. The values of the parameters are as follows:

- $b_0 = 8.97 \times 10^{-7}$, implies that, if all the independent variables took the value 0, the participants’ scores would be 8.97×10^{-7} .
- b_1 (TGM) = 0.33; b_2 (PME) = 0.01; b_3 (LJM) = 0.02; b_4 (ULT) = 0.001; b_5 (UAE) = 0.04; b_6 (APM) = -0.04; b_7 (MAM) = -0.03; b_8 (MFM) = 0.05; b_9 (MPM) = -0.01; b_{10} (ILS) = 0.02; b_{11} (CLB) = 0.07; b_{12} (TCE) = -0.03; b_{13} (PAV) = 0.09; and b_{14} (PCA) = 0.07.

Beta values (b_1 – b_{14}) represent the slope of the model and the change in response associated with each unit of change in the predictor and imply the following: (1) If students’ scores increase in terms of TGM, PME, LJM, ULT, UAE, ILS, CLB, PAV, and PCA, they will increase their scores compared to their second quarter scores; and (2) participants’ scores in the second quarter increase as their scores in APM, MAM, MPM, and TCE decrease.

Based on the above and replacing the values of the regression Equation (1), the following model can be obtained:

$$\text{NST} = 0.66 + 0.33 * \text{TGM} + 0.07 * \text{CLB} + 0.09 * \text{PAV} + 0.07 * \text{PCA}, \quad (1)$$

This was optimized through sequential replacements using Python (Figure 2).

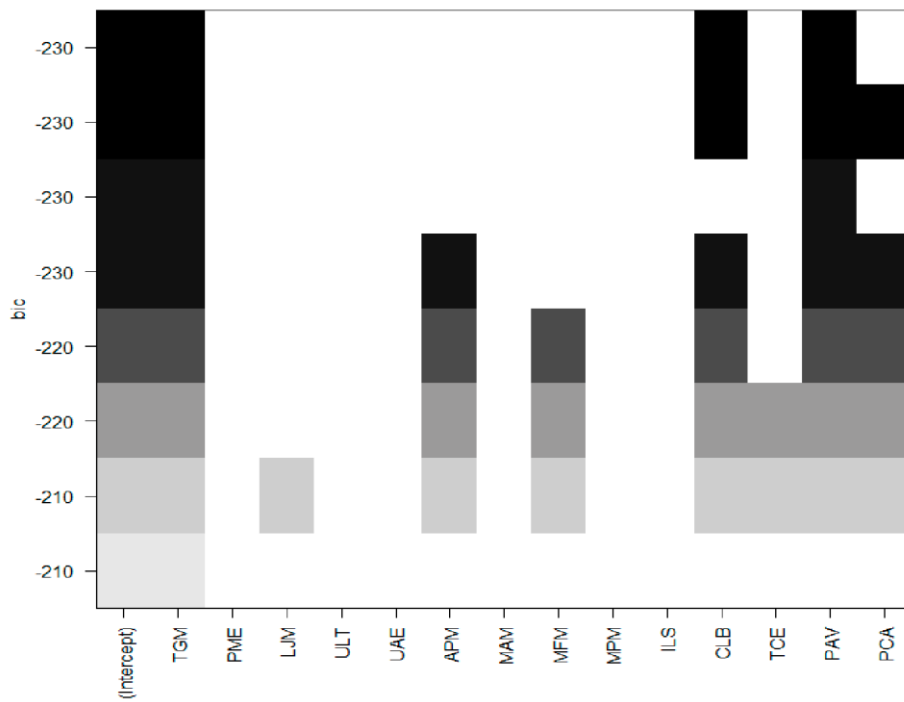


Figure 2. Optimum regression model using sequential replacements.

According to Figure 2, it can be seen that the best regression model would be the one whose Bayesian information criterion (BIC), was the best. In this case, only the intercept and the TGM variable could be selected to predict the scores of the students in the second trimester.

2.1.3. Marginal Averages Predicted by the Model and Estimated with the Data

The marginal means predicted by the linear regression model were as follows: TGM = 1.5; CLB = 1.6; PAV = 1.7; and PCA = 1.8. In other words, TGM has the lowest performance. Taking into account that the mean scores in the previous variables were 1.5, 1.6, 1.7, and 1.8, it can be concluded that the configured linear regression model is good. Next, the diagnosis of the model was analyzed.

Figure 3a corresponds to the residuals vs. adjusted (predicted) values. The residuals are a part of the dependent variable (VD) not explained by the model, and an estimation of the model error term. Under the assumptions of the linear regression model, the variability of the residuals should not change in the range of the VD, and therefore, we should not observe any pattern in the residuals. Values that are outliers are indicated with a label indicating the observation number. Secondly, the normal quantile-quantile of the standardized residues figure can be observed. In the linear regression model, we assume that the errors follow a normal distribution, so we expect the figure of the residuals to follow a straight line. Deviations from this line may mean that errors do not follow a normal distribution. Figure 3c is similar to the first figure but uses the square root of the standardized residues. As in the first case, there should be no pattern in the residues. Figure 3d shows Cook's distances, which is a statistic that tries to identify the points that have more influence. Generally, these are points that are distant from the rest, both in the VD and the independent variable (VI). The high values are labeled with the numbers of observations and indicate that it is necessary to study these.

As a consequence of these annotations, it was necessary to calculate the gain and lift plots by running an automatic regression model with the H2o package. Gain Lift is an application that selects the best model among more than 100 algorithms to classify the different subjects and predict the quantitative increase.

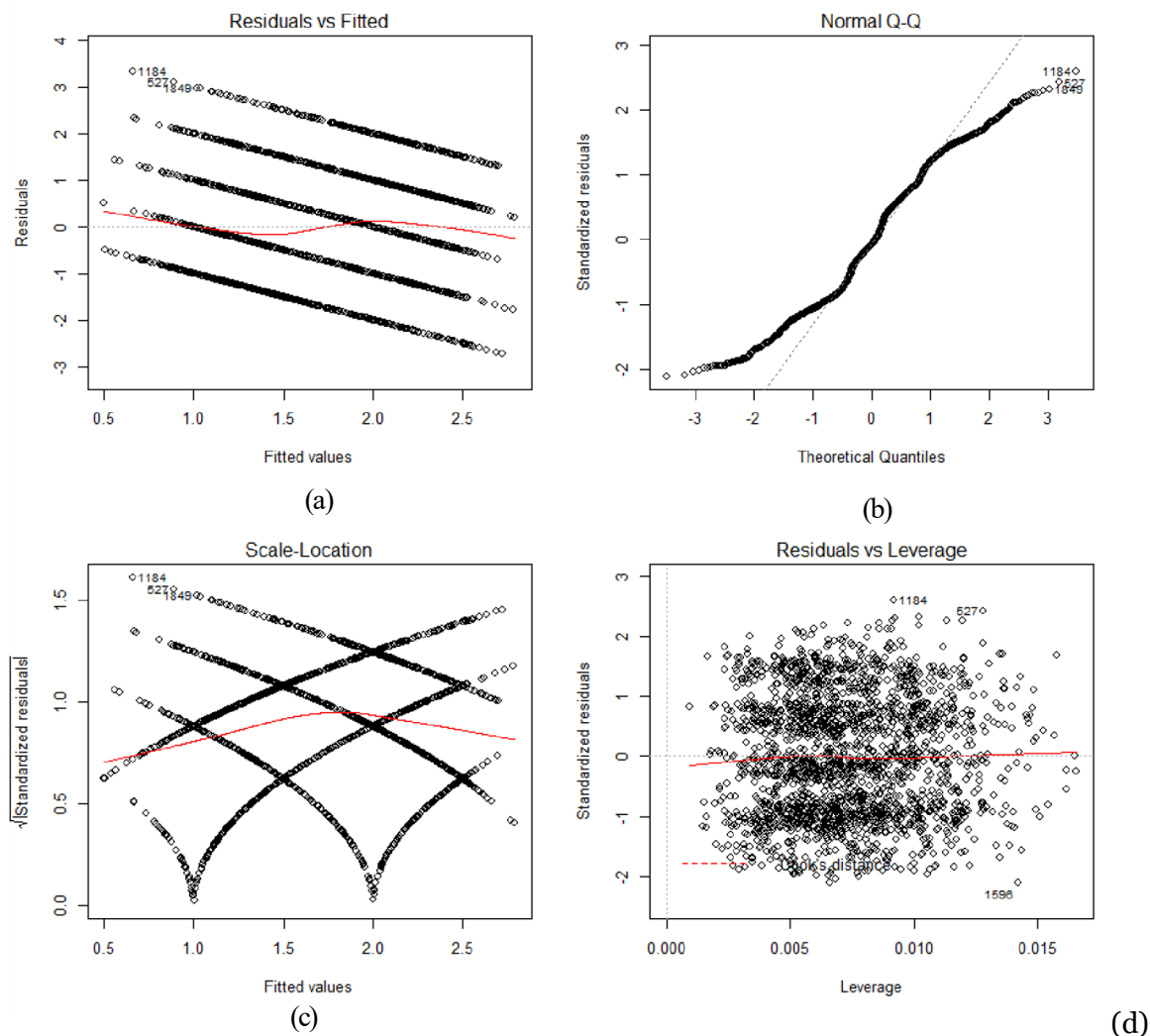


Figure 3. Diagnostics of the linear regression model: (a) Residuals vs. fitted values; (b) Normal Quantil-Quantil; (c) scale-location; and (d) residuals vs. leverage values.

The steps followed were, first, to select the target variable Second Quarter Notes (NST) representative of academic performance. The NST was binary, since it is not possible to develop a gain plot and lift plot if the dependent variable does not take only two values; if the subject had passed, it was assigned the value of 1, and otherwise, the value 0 was assigned. Then, the value matrix was fragmented into two; 85% of the data were taken for the training matrix and 15% for the testing matrix.

Before constructing the gain and lift plot, the following steps were followed in cascade; the H2o algorithm in R was used, keeping the same previous matrices, but in this case, the testing matrix was split into two, one for testing and the other for validation. The main objective was to determine the accuracy of the algorithm and the area under the curve. In the first case, the accuracy of the algorithm was 73.3% and the area under the curve was 56%, which means that more than 50% of the cases would be correctly predicted through the automatic algorithm used. Of all the selected algorithms, H2o indicated that the model that best fit the study data was the “XGBoost” for the total sample. In contrast, the predictive model was different in the gender-differentiated analysis, with the automatic algorithm “Gradient Boosting Machine” being the most optimal among those used for the subsample formed by the girls, and “Stacked Ensembles” for the subsample formed by the boys.

3. Results

The results show that the items of the study correlated differently with the NST variable (second quarter grades). There were insignificant positive correlations between academic performance and PME (Does your math teacher explain well in class?), LJM (From Monday to Thursday, how many hours do you dedicate to the study of mathematics?), ULT (Do you use the textbook to study mathematics?), UAE (Do you use the notebook or class notes to study math?), and ILS (Do you think your school is a place where you feel protected and safe?)

The variables that correlated significantly with NST were TGM (Do you like mathematics?), CLB (Do you think your classroom is a good place to learn?), PAV (Do you think your math teacher appreciates and values you?), and PCA (Does your mathematics teacher create an appropriate climate of respect in the classroom?). The regression Equation (1) answers the first research question (PI1: Could actions be taken to improve the results?) and shows that the variable that best explains the NST scores is TGM (Do you like mathematics?). However, the model only explains 12% of the variability of the answer.

("Second quarter grades" = $0.66 + 0.33 * \text{Do you like mathematics?} + 0.07 * \text{Do you think your classroom is a good place to learn?} + 0.09 * \text{Do you think your mathematics teacher appreciates and values you?} + 0.07 * \text{Does your mathematics teacher create an appropriate climate of respect in classroom?}$)

On the other hand, there is a small negative correlation between academic performance and APM

(Does your academy or tutor help you to learn mathematics?), MAM (Do your friends motivate you to study mathematics?), MPM (Does your teacher motivate you to study mathematics?), and TCE (Do you learn math by working with your classmates?). In other words, academic performance in mathematics correlates negatively with external motivation and support.

To answer the research question PI2 (*How could it be evaluated?*), a statistical tool called Gain Lift was implemented, which is also used in business studies, but is innovative in educational research and appropriate for answering the research questions posed and facilitating the achievement of the objectives of this study. This instrument predicts, in a quantitative way, the maximum gain of the grades obtained in the second quarter, taking into account the variables of the study involved. To do this, it selects the optimal model from more than 100 algorithms to classify the different subjects and predict the quantitative increase.

The research question PI3 (*How much could it be improved?*) is answered in the subsequent analysis of Figures 4–8.

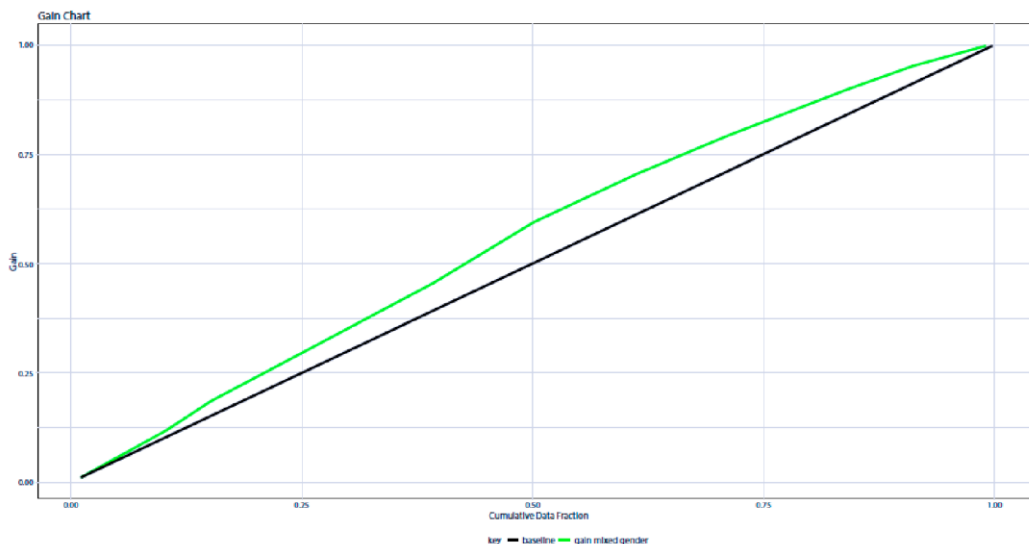
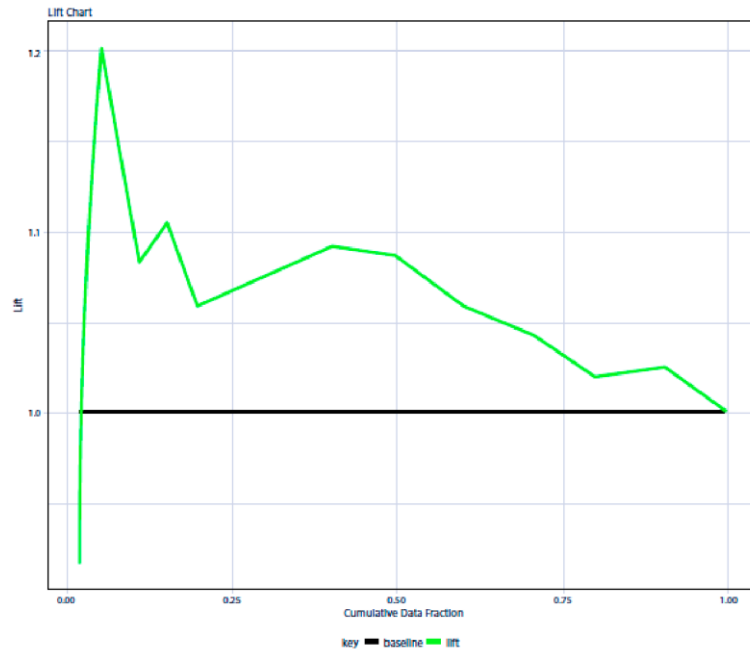
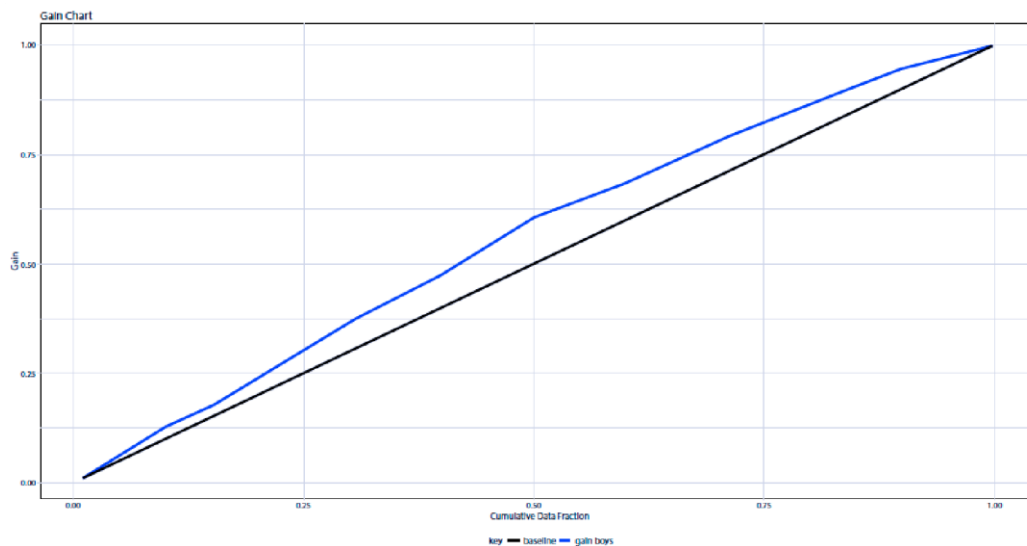


Figure 4. Gain chart for the mixed student sample.**Figure 5.** Lift chart for the mixed student sample.**Figure 6.** Gain chart for the male students' sub-sample.

The results of Figure 4 show that if “n” students are randomly selected, the “XGBoost” algorithm predicts that 25% of the subsample would improve their second quarter grades by 18% at most. The same result would be obtained if 50% of the students of a subsample “n” were randomly selected. In contrast, for 75% of the subsample, the model predicts a maximum gain in second quarter grades of 10%. To achieve these gains, appropriate interventions in the school setting are recommended. This affirmation is based on the fact that the analysis of correlations of the present study shows that the variables that had a significant impact on the second trimester’s grades corresponded to the dimension “School Environment” and the predictors with a greater weight were a positive perception towards the subject TGM and towards the teacher PAV.

The lift chart allows this gain to be quantified in a complementary way. More specifically, the lift chart results show that if “n” students were randomly selected, 50% of the sub-sample would improve their second quarter scores by 0.4 points if appropriate interventions were carried out in the school environment.

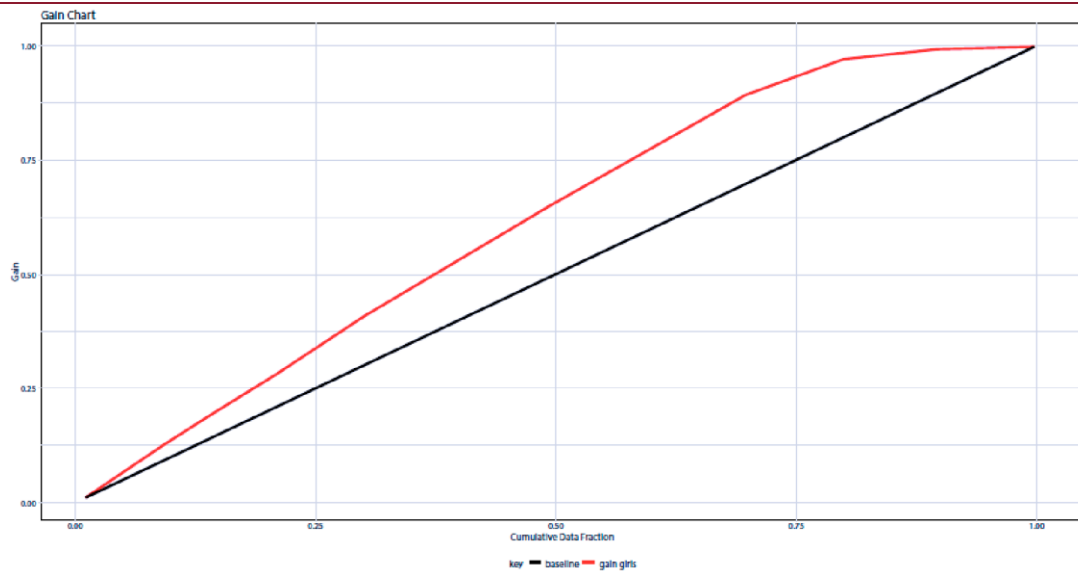


Figure 7. Gain chart for the female students' sub-sample.

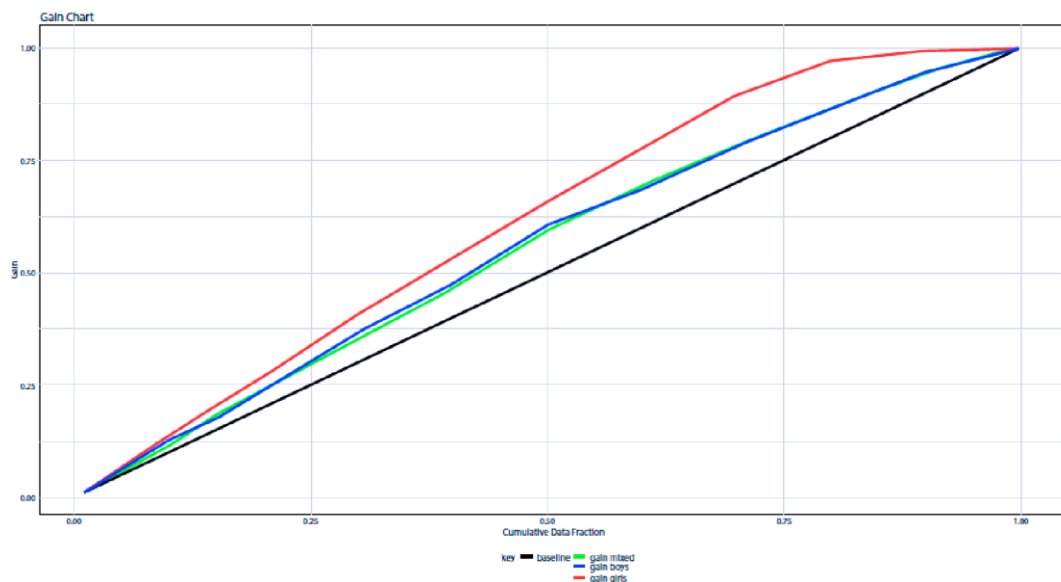


Figure 8. Gain chart comparison of the mixed student sample and the gender-differentiated subsamples.

In Figures 4 and 5, it can be seen that if appropriate interventions are used in the school environment, contemplated in the study variables, students can improve their performance (their grades in the second quarter). Metrically, if the scores with respect to the dimensions analyzed (Do you like mathematics? Do you think your classroom is a good place to learn? Do you think your mathematics teacher appreciates and values you? Does your math teacher create an appropriate climate of respect in classroom?) were optimal, the scores measured in the second quarter would have increased by 0.4 points.

The overall results show a small gain in students' academic achievement in mathematics, if appropriate interventions are applied in the school environment. The following questions arise from the results: Would similar results be expected among girls and boys? Which gender would have the greatest gain? These are encompassed in the research question IP4: Are there gender differences? To answer all of these questions, the gender-differentiated results in Figures 6–8 were analyzed.

The results of the gain chart for the subsample of boys show similar results to those obtained in the global mixed figure analyzed above. It can be observed that if a subsample of

“n” students is randomly selected, the “Stacked Ensembles” algorithm predicts that 25% of the subsample would have a maximum gain of 22% in their scores in the second quarter, if appropriate interventions were carried out in the school environment. The same result would be obtained if 50% of the students were randomly selected. For this percentage, the maximum gain would be obtained. On the other hand, the maximum gain would be 10% for 75% of the selected sub-sample (in this case, the gain is reduced by 54%, according to the other cases).

The gain chart results for girls indicate that if a subsample of “n” students is randomly selected, the “Gradient Boosting Machine” algorithm predicts that 25% would have a maximum gain of 36%, achieving the optimal gain in their second quarter grades if appropriate interventions were carried out in the school setting. The percentages would decrease if we randomly selected 50% of the students whose maximum gain would be 33% and 25% if we selected 75% of the sample (in this scenario, the gain is reduced by 30% from the maximum gain).

A comparative gender analysis (Figure 7) showed similar results for the total sample and for the subsample formed by the boys. The explanation for this lies in the nature of the gain model. It selects, from 100 predictive algorithms, the one that best fits the study sample. For the total sample, the optimal algorithm used was “XGBoost”. For the subsample of girls, the optimal algorithm was “Gradient Boosting Machine”, and for the subsample of boys, the optimal algorithm was “Stacked Ensembles”.

It can be seen from Figure 7 that if a subsample “n” was randomly selected, 25% of both boys and girls would achieve the maximum gain. Figure 7 shows a difference in favor of girls over boys of 14 percentage points.

The percentages would decrease for both genders if a random subsample of 50% was selected. In that case, girls would achieve a gain of about 33% and boys would maintain the maximum gain of 22%. In this case, the difference between the two genders is 11 percentage points.

In 75% of the subsample, it could be observed that the gain for girls was still high, at around 25%. In contrast, the gain for boys was considerably reduced to 10%. The difference between the two amounts to 15 percentage points, resulting in the highest percentage difference between girls and boys. On the other hand, in this assumption, both genders reach the highest percentage decrease with respect to their own maximum gains. In the case of boys, this is 55% and in the case of girls, it is approximately 31%.

4. Discussion

In this study, it was observed that there are specific actions that have a quantifiable impact on the academic performance of students. Some factors of the dimension “Mathematics Learning” have been analyzed, such as indicators of affinity, teaching, study time, didactic resources employed, study aids, and motivation. As for the dimension “School Environment”, we have collected items from the indicators educational center, classroom, partners, and relationship with teachers.

In the results of the present study, it could be observed that the variables that had a significant influence on the ratings obtained in the second quarter (NST) were TGM, PAV, CLB, and PCA. Among these, the ones with less weight were the perceptions of the climate in the classroom (CLB) and the teacher’s dynamizing role in the classroom (PCA). According to previous methods, the variables of the study that had a more significant impact on the academic performance were positive perceptions of the learning of mathematics (TGM) and the mathematics teaching figure (PAV). The regression Equation (1) allows the weight of each item in the NST variable to be quantified.

This shows that the variable that best explains the scores obtained in NST is TGM and its weight is four times the average of the other variables in the equation. Therefore, a positive

perception of the subject is the predictor with the greatest impact on the scores of students in mathematics. The rest of the variables have a similar weight in the equation, of less than 10%.

The scientific literature postulates that the teacher–student relationship significantly influences interest in the subject [50], conditions the academic trajectory of the students [36,37], and affects the academic performance of the students [33,34]. This relationship depends on the students' perception of the subject taught by the teacher and the difficulty perceived by the teacher [35]. In addition, the influence of the teacher can reduce the differences in opportunities perceived by students at risk of social exclusion [39].

The effect of the classroom on academic performance is also reflected in the regression Equation (1). Additionally, previous findings have established that the climate and interactions in the classroom between students and teachers condition the teaching–learning process [25,26]. Therefore, several authors consider the creation and validation of a scale to be necessary for evaluating the climate in the classroom [30], because of its influence on the academic performance of students [31]. These instruments would help evaluate the climate in classrooms and the influence of disruptive students, who are exposed to early school failure [32]. This assessment is required to articulate interventions for improvement [27].

Another remarkable aspect of the regression Equation (1) is the combination of items. It is also possible to appreciate a direct relationship between a positive perception towards the subject, the climate in the classroom, and the figure of the teacher. Therefore, the teacher, as a moderating agent of the climate in the classroom, can significantly influence students and encourage a positive perception of mathematics that will result in an improvement in academic results.

Several authors have highlighted the ability of teachers to motivate students through effective intervention strategies [45]. To do this, teachers must promote good bilateral communication with students to improve their mathematical competence [18,40,41] and stimulate their motivation via a methodological and curricular proposal that is attractive and participatory for students [46].

On the other hand, there are positive but weak correlations between academic performance and PME (Does your math teacher explain well in class?), LJM (From Monday to Thursday, how many hours do you dedicate to the study of mathematics?), ULT (Do you use the textbook to study mathematics?), UAE (Do you use the notebook or class notes to study mathematics?), and ILS (Do you think your school is a place where you feel protected and safe?). In other words, performance correlates positively with students' perceptions of the teacher and the educational environment.

These results can be explained by the erroneous perception that students have of mathematics as a subject that is practiced, but not studied. In addition, the fear of making mistakes discourages and limits student participation in class and their learning of mathematics [10]. Limited learning of mathematics conditions interest in STEM [13].

On the other hand, there are negative correlations between academic performance and APM (Does your academy or tutor help you to learn mathematics?), MAM (Do your friends motivate you to study mathematics?), MPM (Does your teacher motivate you to study mathematics?), and TCE (Do you learn mathematics by studying with your classmates?). In other words, academic performance in mathematics correlates negatively with external motivation and support.

In relation to the findings of the present study, it is observed that external motivations not sustained over time generate initial expectations that are not very effective in terms of academic performance [51]. In contrast, the role of internal motivation has a significant impact on students [19]. In this sense, several authors maintain that the internal motivation of mathematics students modifies their perception of the subject and affects their academic performance [6]. To this end, the use of appropriate teaching strategies is recommended to

awaken the students' interest and participation in STEM [52]. These actions would help to reduce the high rate of early school dropouts in the Autonomous City of Melilla [47].

On the other hand, the Gain Lift algorithm has been used to determine the maximum gain in mathematical performance. To quantify this, the joint effect of the variables set out in Table 1 was analyzed. The results indicate that if appropriate interventions are employed in the school setting, a greater gain can be obtained. The figures show that the gain decreases from 50% of the chosen random subsample. It was also observed that the performance gain among girls is significantly greater than that among boys for any subsample taken. Similar findings postulate that girls possess a greater capacity to manage and order mathematical knowledge [17]. In contrast, several authors argue that this competition decreases significantly during adolescence, affected in part by gender stereotypes [13]. As a result, the number of female students pursuing higher education in STEM is significantly lower than that of males [13,53].

Although the linear regression model used is significant, the results indicate that 87.72% of the variation in the second quarter grades cannot be explained by the variables used in this study. In this sense, the mean scores of the gain in the mathematics grades amount to 0.4 points. This is consistent with the results of Timmermans (2018), which state that, although some teachers have substantial effects on students, the overall effects for most students are insignificant. Therefore, there must be other variables that have a significant influence, such as the socioeconomic status, family [2,54,55], or the possible negative [56,57] and positive [58] effects of technologies, on the academic performance of mathematics students.

The main strength of this study, on the other hand, is the use of Gain Lift to predict the maximum gain in mathematics scores if effective interventions are employed in the school environment. Moreover, the regression Equation (1) used for the analysis of the variables allowed us to determine the effect that each has on the grades of mathematics students. Another strength is that the model can be extended to other subjects and educational levels with similar approaches.

On the other hand, some limitations of the present study should be recognized. First, due to its cross-sectional design, no causal conclusions can be drawn. Secondly, the inaccurate nature of the Gain Lift model only predicts the maximum level of the possible gain in the grades of mathematics students.

5. Conclusions

This paper has achieved all the initial objectives proposed and all the research questions raised have been answered.

It is concluded that the grades obtained by students can be optimized if adequate intervention strategies are applied in the classroom. The results empirically confirm that certain pedagogical actions have a positive effect on students and, accordingly, on their academic performance. The innovative character of this study lies in the ability to quantify the maximum level of grade gain with the Gain Lift algorithm. With this instrument, it is observed that the gain in mathematics grades for girls is significantly greater than the gain obtained for boys, regardless of the subsample chosen.

The items focused on the "School Environment" dimension, such as the role of the classroom, and the teacher–student relationship, have had a greater influence on the mathematics grades than the items of the "Mathematics Learning" dimension based on the indicators affinity, teaching, study time, didactic resources used, study aids, and motivation. However, the low rates of gain indicate that there are other important factors affecting the performance of students in mathematics which are not covered in this study. Therefore, other variables, such as socioeconomic, family, or technological status,

which may affect mathematics students, should be analyzed. Considering this, a wide range of lines of research should be pursued in order to delve deeper into strategies that can be applied to analyze school failure, especially linked to the area of mathematics, which is of great concern to the educational community worldwide.

References

1. Gracia-Bafalluy, M.; Puyuelo-San Clemente, M. Mathematical Learning and Its Difficulties in Southern European Countries. In *International Handbook of Mathematical Learning Difficulties*; Springer: Cham, Switzerland, 2019; pp. 165–182.
2. Martínez-Abad, F.; Gamazo, A.; Rodríguez-Conde, M.-J. Educational Data Mining: Identification of factors associated with school effectiveness in PISA assessment. *Stud. Educ. Eval.* **2020**, *66*, 100875. [[CrossRef](#)]
3. Warwick, P.; Vrikki, M.; Karlsen, A.M.F.; Dudley, P.; Vermunt, J.D. The role of pupil voice as a trigger for teacher learning in Lesson Study professional groups. *Camb. J. Educ.* **2019**, *49*, 435–455. [[CrossRef](#)]
4. Ozhan, S.C.; Kocadere, S.A. The Effects of Flow, Emotional Engagement, and Motivation on Success in a Gamified Online Learning Environment. *J. Educ. Comput. Res.* **2020**, *57*, 2006–2031. [[CrossRef](#)]
5. Prieto, J.M.A. A systematic review about gum cation, motivation and learning in high school. *Teor. Educ.* **2020**, *32*, 73–99. [[CrossRef](#)]
6. Rojo Robas, V.; Maria Madariaga, J.; Domingo Villarroel, J. Secondary Education Students' Beliefs about Mathematics and Their Repercussions on Motivation. *Mathematics* **2020**, *8*, 368. [[CrossRef](#)]
7. Levrini, O.; Tasquier, G.; Branchetti, L.; Barelli, E. Developing future-scaffolding skills through science education. *Int. J. Sci. Educ.* **2019**. [[CrossRef](#)]
8. Garrido-Yserte, R.; Crecente-Romero, F.; Gallo-Rivera, M.-T. The relationship between capacities and entrepreneurial intention in secondary school students. *Econ. Res. Ekon. Istraz.* **2019**. [[CrossRef](#)]
9. Cipkova, E.; Karolcik, S.; Scholzova, L. Are secondary school graduates prepared for the studies of natural sciences? –Evaluation and analysis of the result of scientific literacy levels achieved by secondary school graduates. *Res. Sci. Technol. Educ.* **2019**. [[CrossRef](#)]
10. Kaefer, J.; Kuger, S.; Klieme, E.; Kunter, M. The significance of dealing with mistakes for student achievement and motivation: Results of doubly latent multilevel analyses. *Eur. J. Psychol. Educ.* **2019**, *34*, 731–753. [[CrossRef](#)]
11. Husein, N.A.; Rahman, N.F.A.; Phang, F.A.; Abu Bakar, A.H.; Ahmad, R.R.; Kasim, N.; Duralim, M. The Relationship between Mentoring in Students' Perception towards STEM Education. *Adv. Sci. Lett.* **2018**, *24*, 72–73. [[CrossRef](#)]
12. Sharpe, R.; Abrahams, I. Secondary school students' attitudes to practical work in biology, chemistry and physics in England. *Res. Sci. Technol. Educ.* **2019**. [[CrossRef](#)]
13. Barth, J.M.; Kim, H.; Eno, C.A.; Guadagno, R.E. Matching Abilities to Careers for Others and Self: Do Gender Stereotypes Matter to Students in Advanced Math and Science Classes? *Sex. Roles* **2018**, *79*, 83–97. [[CrossRef](#)]
14. McCurdy, M.; Skinner, C.; McClurg, V.; Whitsitt, L.; Moore, T. Bonus rewards for everyone: Enhancing mathematics performance with supplemental interdependent group contingencies. *Prev. Sch. Fail.* **2020**, *64*, 77–88. [[CrossRef](#)]

15. Tovar, E.F.; Bercht, M. Explorando la dimensión afectiva entre el estudiante y el conocimiento matemático mediado por las TIC. *RENOTE* **2016**, *14*. [[CrossRef](#)]
16. Hine, G. Teachers' perceptions on declining student enrolments in Australian senior secondary mathematics courses. *Issues Educ. Res.* **2018**, *28*, 635–654.
17. Gasco Txabarri, J. Diferencias en el uso de estrategias en el aprendizaje de las matemáticas en enseñanza secundaria según el sexo. *Cuad. Investig. Educ.* **2017**, *8*, 47–59. [[CrossRef](#)]
18. Sánchez-Matamoros, G.; Fernández, C.; Llinares, S. Relationships among prospective secondary mathematics teachers' skills of attending, interpreting and responding to students' understanding. *Educ. Stud. Math.* **2019**, *100*, 83–99. [[CrossRef](#)]
19. Boeheim, R.; Knogler, M.; Kosel, C.; Seidel, T. Exploring student hand-raising across two school subjects using mixed methods: An investigation of an everyday classroom behavior from a motivational perspective. *Learn. Instr.* **2020**, *65*. [[CrossRef](#)]
20. Lazarides, R.; Dicke, A.-L.; Rubach, C.; Eccles, J.S. Profiles of Motivational Beliefs in Math: Exploring Their Development, Relations to Student-Perceived Classroom Characteristics, and Impact on Future Career Aspirations and Choices. *J. Educ. Psychol.* **2020**, *112*, 70–92. [[CrossRef](#)]
21. Tse, W.S.; Choi, L.Y.A.; Tang, W.S. Effects of video-based flipped class instruction on subject reading motivation. *Br. J. Educ. Technol.* **2019**, *50*, 385–398. [[CrossRef](#)]
22. Plackle, I.; Konings, K.D.; Struyven, K.; Libotton, A.; van Merriënboer, J.J.G.; Engels, N. Powerful learning environments in secondary vocational education: Towards a shared understanding. *Eur. J. Teach. Educ.* **2019**. [[CrossRef](#)]
23. Welters, R.; Lewthwaite, B.; Thomas, J.; Wilson, K. Re-engaged students? Perceptions of mainstream and flexible learning environments—A “semi-quantitative” approach. *Int. J. Incl. Educ.* **2019**, *23*, 1315–1331. [[CrossRef](#)]
24. Hershkovitz, A.; Baruch, A.F. La relación profesor-alumno y la comunicación en Facebook: Percepciones de los alumnos. *Comun. Rev. Científica Iberoam. Comun. Educ.* **2017**, 91–101. [[CrossRef](#)]
25. Lerang, M.S.; Ertesvag, S.K.; Havik, T. Perceived Classroom Interaction, Goal Orientation and Their Association with Social and Academic Learning Outcomes. *Scand. J. Educ. Res.* **2019**, *63*, 913–934. [[CrossRef](#)]
26. Rathmann, K.; Herke, M.; Hurrelmann, K.; Richter, M. Class Climate, Academic Well-Being and Self-Rated Health among School Children in Germany: Findings of the National Educational Panel Study (NEPS). *Gesundheitswesen* **2018**, *80*, 334–343. [[CrossRef](#)]
27. Quijada, A.; Ruiz, M.A.; Huertas, J.A.; Alonso-Tapia, J. Development and validation of the School Climate Questionnaire for Secondary and High School Teachers (SCQ-SHST). *Ann. Psicol.* **2020**, *36*, 155–165. [[CrossRef](#)]
28. Adnan, M.; Isa, N.H. Relationship between The Psychosocial Learning Environment to the Achievement of Form Four Additional Mathematics. *J. Mech. Contin. Math. Sci.* **2019**, 176–184. [[CrossRef](#)]
29. Keller, T.; Takacs, K. Peers that count: The influence of deskmates on test scores. *Res. Soc. Stratif. Mobil.* **2019**, *62*. [[CrossRef](#)]
30. Fernández-Rio, J.; Cecchini, J.A.; Merino-Barrero, J.A.; Valero-Valenzuela, A. Perceived Classroom Responsibility Climate Questionnaire: A new scale. *Psicothema* **2019**, *31*, 475–481. [[CrossRef](#)]

31. Johnson, S.L.; Reichenberg, R.E.; Shukla, K.; Waasdorp, T.E.; Bradshaw, C.P. Improving the Measurement of School Climate Using Item Response Theory. *Educ. Meas. Issues Pract.* **2019**. [CrossRef]
32. Eugene, D.R. A Multilevel Model for Examining Perceptions of School Climate, Socioeconomic Status, and Academic Achievement for Secondary School Students. *J. Educ. Stud. Placed Risk* **2019**. [CrossRef]
33. Holzberger, D.; Praetorius, A.-K.; Seidel, T.; Kunter, M. Identifying effective teachers: The relation between teaching profiles and students' development in achievement and enjoyment. *Eur. J. Psychol. Educ.* **2019**, *34*, 801–823. [CrossRef]
34. Klapproth, F.; Fischer, B.D. Preservice teachers' evaluations of students' achievement development in the context of school-track recommendations. *Eur. J. Psychol. Educ.* **2019**, *34*, 825–846. [CrossRef]
35. Roorda, D.L.; Jorgensen, T.D.; Koomen, H.M.Y. Different teachers, different relationships? Student-teacher relationships and engagement in secondary education. *Learn. Individ. Differ.* **2019**, *75*. [CrossRef]
36. Lazarides, R.; Gaspard, H.; Dicke, A.-L. Dynamics of classroom motivation: Teacher enthusiasm and the development of math interest and teacher support. *Learn. Instr.* **2019**, *60*, 126–137. [CrossRef]
37. Ozdemir, S.B.; Ozdemir, M. How do Adolescents' Perceptions of Relationships with Teachers Change during Upper-Secondary School Years? *J. Youth Adolesc.* **2019**. [CrossRef]
38. Timmermans, A.C.; Rubie-Davies, C.M. Do teachers differ in the level of expectations or in the extent to which they differentiate in expectations? Relations between teacher-level expectations, teacher background and beliefs, and subsequent student performance. *Educ. Res. Eval.* **2018**, *24*, 241–263. [CrossRef]
39. Munter, C.; Haines, C. Mathematics teachers' enactment of cognitively demanding tasks and students' perception of racial differences in opportunity. *Math. Think. Learn.* **2019**, *21*, 155–177. [CrossRef]
40. Uner, S.; Akkus, H. Secondary students' perceptions of their teachers' pedagogical content knowledge: A scale development study. *Teach. Dev.* **2019**. [CrossRef]
41. Van der Kleij, F.M. Comparison of teacher and student perceptions of formative assessment feedback practices and association with individual student characteristics. *Teach. Teach. Educ.* **2019**, *85*, 175–189. [CrossRef]
42. Hagger, M.S.; Hamilton, K. Grit and self-discipline as predictors of effort and academic attainment. *Br. J. Educ. Psychol.* **2019**, *89*, 324–342. [CrossRef] [PubMed]
43. Hagger, M.S.; Hamilton, K. Motivational predictors of students' participation in out-of-school learning activities and academic attainment in science: An application of the trans-contextual model using Bayesian path analysis. *Learn. Individ. Differ.* **2018**, *67*, 232–244. [CrossRef]
44. Hodis, F.A. Underpinnings of Expectancies of Success in Mathematics: An Analysis of General, School-Related, and Domain-Specific Motivation Antecedents. *J. Educ. Psychol.* **2018**, *110*, 407–430. [CrossRef]
45. Gil, A.J.; Maria Antelm-Lanzat, A.; Cacheiro-Gonzalez, M.L.; Pérez-Navio, E. School dropout factors: A teacher and school manager perspective. *Educ. Stud.* **2019**, *45*, 756–770. [CrossRef]
46. Vergara, D.; Fernández, M.L.; Lorenzo, M. Enhancing Student Motivation in Secondary School Mathematics Courses: A Methodological Approach. *Educ. Sci.* **2019**, *9*, 83. [CrossRef]
47. Informe 2019 Sobre el Estado del Sistema Educativo: Curso 2017–2018. 2019. Available online: <http://www.educacionyfp.gob.es/dam/jcr:4f35ae94-f996-4ceb-b3f0-21b2e421ec26/i19cee-informe.pdf> (accessed on 17 November 2020).

48. Cisneros-Cohernour, E.J.; Jorquera Jaramillo, M.C.; Aguilar Pereyra, Á.M. Validación de instrumentos de evaluación docente en el contexto de una universidad española. *Voces Silencios. Rev. Latinoam. Educ.* **2012**, *3*, 41–55. [[CrossRef](#)]
49. Barca, A.; Peralbo, M.; Porto, A.; Marcos, J.; Brenlla, J. Metas académicas del alumnado de Educación Secundaria Obligatoria (ESO) y Bachillerato con alto y bajo rendimiento escolar. *Rev. Educ.* **2011**, *354*, 341–368.
50. Martin, A.J.; Collie, R.J. Teacher-Student Relationships and Students' Engagement in High School: Does the Number of Negative and Positive Relationships with Teachers Matter? *J. Educ. Psychol.* **2019**, *111*, 861–876. [[CrossRef](#)]
51. Swope, A.K.; Fredrick, J.W.; Becker, S.P.; Burns, G.L.; Garner, A.A.; Jarrett, M.A.; Kofler, M.J.; Luebbe, A.M. Sluggish cognitive tempo and positive valence systems: Unique relations with greater reward valuation but less willingness to work. *J. Affect. Disord.* **2020**, *261*, 131–138. [[CrossRef](#)] [[PubMed](#)]
52. Mora, H.; Signes-Pont, M.T.; Fuster-Guillo, A.; Pertegal-Felices, M.L. A collaborative working model for enhancing the learning process of science & engineering students. *Comput. Hum. Behav.* **2020**, *103*, 140–150. [[CrossRef](#)]
53. Caspi, A.; Gorsky, P.; Nitzani-Hendel, R.; Zacharia, Z.; Rosenfeld, S.; Berman, S.; Shildhouse, B. Ninth-grade students' perceptions of the factors that led them to major in high school science, technology, engineering, and mathematics disciplines. *Sci. Educ.* **2019**, *103*, 1176–1205. [[CrossRef](#)]
54. Banerjee, P.A. A systematic review of factors linked to poor academic performance of disadvantaged students in science and maths in schools. *Cogent Educ.* **2016**, *3*, 1178441. [[CrossRef](#)]
55. Yu, K.; Martin, A.J.; Hou, Y.; Osborn, J.; Zhan, X. Motivation, Engagement, Academic Buoyancy, and Adaptability: The Roles of Socio-Demographics among Middle School Students in China. *Meas. Interdiscip. Res. Perspect.* **2019**, *17*, 119–132. [[CrossRef](#)]
56. Hu, X.; Gong, Y.; Lai, C.; Leung, F.K.S. The relationship between ICT and student literacy in mathematics, reading, and science across 44 countries: A multilevel analysis. *Comput. Educ.* **2018**, *125*, 1–13. [[CrossRef](#)]
57. Benitez Diaz, L.M.; Sevillano Garcia, M.L.; Vazquez Cano, E. Effects on academic performance in secondary students according to the use of ICT. *Int. J. Educ. Res. Innov.* **2019**, *12*, 90–108.
58. Park, J.-H.; Kim, C.; Ham, J. High-school students' understanding and use of mathematics textbooks.