

Article

Analysis of Psychological Factors Influencing Mathematical Achievement and Machine Learning Classification

Juhyung Park ¹, Sungtae Kim ² and Beakcheol Jang ^{1,*}¹ Graduate School of Information, Yonsei University, Seoul 03722, Republic of Korea; winderland@yonsei.ac.kr² Able Edutech Inc., Seoul 04081, Republic of Korea; iforyou76@yonsei.ac.kr

* Correspondence: bjang@yonsei.ac.kr

Abstract: This study analyzed the psychological factors that influence mathematical achievement in order to classify students' mathematical achievement. Here, we employed linear regression to investigate the variables that contribute to mathematical achievement, and we found that self-efficacy, math-efficacy, learning approach motivation, and reliance on academics affect mathematical achievement. These variables are derived from the Test of Learning Psychology (TLP), a psychological test developed by Able Edutech Inc. specifically to measure students' learning psychology in the mathematics field. We then conducted machine learning classification with the identified variables. As a result, the random forest model demonstrated the best performance, achieving accuracy values of 73% (Test 1) and 81% (Test 2), with F1-scores of 79% (Test 1) and 82% (Test 2). Finally, students' skills were classified according to the TLP items. The results demonstrated that students' academic abilities could be identified using a psychological test in the field of mathematics. Thus, the TLP results can serve as a valuable resource to develop personalized learning programs and enhance students' mathematical skills.

Keywords: machine learning; linear regression; psychological test; mathematical achievement

MSC: 68T01

**Citation:** Park, J.; Kim, S.; Jang, B.

Analysis of Psychological Factors

Influencing Mathematical

Achievement and Machine Learning

Classification. *Mathematics* **2023**, *11*,3380. [https://doi.org/10.3390/](https://doi.org/10.3390/math11153380)[math11153380](https://doi.org/10.3390/math11153380)

Academic Editors: Mingbo Zhao,

Haijun Zhang and Zhou Wu

Received: 18 June 2023

Revised: 25 July 2023

Accepted: 31 July 2023

Published: 2 August 2023

**Copyright:** © 2023 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article

distributed under the terms and

conditions of the Creative Commons

Attribution (CC BY) license ([https://](https://creativecommons.org/licenses/by/4.0/)[creativecommons.org/licenses/by/](https://creativecommons.org/licenses/by/4.0/)

4.0/).

1. Introduction

The EdTech market has been growing steadily with new value created by integrating information technologies, e.g., big data and artificial intelligence (AI), with education [1]. In particular, AI is becoming increasingly influential, especially in the field of mathematics education, because it enables students to develop and improve their mathematical skills [2,3]. MATHia, which was developed by researchers at Carnegie Mellon University, utilizes AI technology to provide feedback and customize learning programs by identifying students' weaknesses. In addition, Woongjin ThinkBig, a South Korean education group, has developed an AI application to teach mathematics and analyze each student's individual learning abilities based on big data and AI to provide personalized learning materials that are appropriate for each student.

Similarly, there is an active movement to provide personalized content recommendations in the mathematics education field, which promotes the need for research on identifying student skills and the factors that affect mathematical achievement. To develop such a personalized learning system, it is important to understand students' psychology because the student's abilities must be assessed to realize personalized learning, and individual psychological factors, e.g., self-confidence and anxiety, can influence the student's abilities [4,5].

Previous studies have shown that psychological factors are significant contributors to academic achievements [4–12]; however, these studies primarily focused on general psychological factors and did not specifically explore the field of mathematical psychology in

predicting student abilities. In addition, most mathematics education institutions have traditionally relied solely on grades as the primary indicator of students' skills, thereby overlooking the potential insights that can be acquired through a more comprehensive approach.

Given these gaps in the existing research, the purpose of this study is to analyze mathematics-related psychological factors that impact students' mathematical achievement. By examining these factors, we seek to gain a deeper understanding of their influence on students' mathematics performance. In addition, our objective is to go beyond the analysis of these mathematics-related psychological factors and advance toward a more proactive approach. In doing so, we intend to uncover prediction models that can classify students' mathematical abilities effectively. To achieve this goal, we employed machine learning classification techniques based on the identified psychological factors. By adopting this methodology, we aim to identify the key psychological factors and harness their potential predictive power. Ultimately, this study aims to contribute to the field by bridging the gap between mathematics and psychology and shed light on the relationship between relevant psychological factors and mathematical achievement.

2. Related Work

Related studies can be broadly categorized into those that reveal psychological factors affecting academic achievement and those that classify or predict students' grades. In terms of identifying the psychological factors that affect academic achievement, a previous study stated that personal traits, e.g., self-confidence, are among the most crucial variables that determine a student's mathematical achievement [4]. Another study proved that Anxiety and depression have been shown to disrupt concentration and reduce academic achievement in high school students [5]. According to one study [6], students' self-efficacy, engagement, and mathematical achievement are positively associated. In addition, a negative relationship between anxiety (from multiple psychological test items) and mathematical achievement has been reported [7]. Students with self-efficacy tend to have more positive emotions, thereby resulting in better academic performance [8].

A previous study that analyzed the aspects of learning motivation reported that intrinsic motivation affects the behaviors and achievement of learners [9]. Another study [10] reported that burnout increases the probability of experiencing psychological and physical disengagement from academic pursuits, which in turn can result in a decline in academic achievement. It has also been found that higher levels of academic self-efficacy were positively associated with greater academic achievement and resilience, thereby indicating a direct relationship between these variables [11]. In addition, a previous study [12] found that various costs, including emotion, effort, opportunity, and ego costs, play a crucial role in predicting mathematical achievement.

In studies on the classification of student grades [13–16], student performance has been classified using machine learning technologies. The classification and regression tree (CART) algorithm and k-nearest neighbors (KNN) techniques have been used to classify the skills of college students attending web-based lectures based on data related to their homework assignments and quizzes [13]. In addition, the average grades of Bulgarian university students were classified according to their admission scores using KNN and decision tree techniques [14]. Another study employed a support vector machine (SVM) to classify college students' academic performances based on Internet usage data [15], and one study [16] employed the random forest to predict the final grades of Malaysia Polytechnics students based on their previous semester's final examination results.

Studies have also investigated grade prediction using machine learning techniques [17–19]. For example, one study performed linear regression to predict academic achievement based on students' backgrounds and past academic scores from the Institute of Aeronautical Engineering in India [17]. In addition, the random forest method has been employed to predict the grade point averages of master students in computer science at ETH Zurich based on their bachelor's grade point averages [18]. In addition, a machine learning-based recom-

mentation system considered the grades of students at the Ho Chi Minh City University of Technology in Vietnam to predict students' future grades [19].

Previous studies have analyzed various variables that influence academic performance; however, there is an identifiable research gap when it comes to psychological factors specifically related to mathematics. To address this gap, this study focuses on the field of mathematics and utilizes a psychological test to identify the factors that influence mathematical achievement.

Furthermore, to the best of our knowledge, no prior study has focused on predicting academic achievements primarily based on psychological factors. Thus, in this study, we employed machine learning techniques to classify students' abilities based on the identified psychological variables. Through these techniques, we aim to provide insights into how psychological factors can be utilized to predict and understand mathematical achievement.

3. Materials and Methods

3.1. Data Description

In this study, we considered 1880 elementary, middle, and high school students who were learning mathematics at Able Edutech Inc., a mathematics EdTech company, from August 2016 to April 2022. Table 1 shows the variables, i.e., the ID (each student's unique number), mathematical achievement (Test 1 and Test 2 scores, which are diagnostic tests), and the Test of Learning Psychology (TLP) scores of each student. To facilitate our research, Able Edutech Inc. coded the students' personal data into a unique ID number to anonymize the data. The TLP is a psychological test about mathematics learning developed by the Korea Learning Psychometric Research Institute and the Yonsei University Cognitive Science Research Center. The TLP results were collected when students first enrolled in the company, and the mathematical achievements were obtained through two tests after attending courses. Regarding the TLP items used in this study, some were difficult to analyze due to limited or missing data on student responses; thus, these items were excluded from the analysis. The five selected psychological test factors were self-efficacy, math-efficacy, learning approach motivation, performance approach motivation, and reliance on academics.

Table 1. Variables considered in this study.

Variables	Description	Range	Count
ID	Student's unique number	1–1880	1880
Mathematical achievement	Test 1 and Test 2 Scores	0–100	2
TLP items (self-efficacy, math-efficacy, learning approach motivation, performance approach motivation, reliance on academics)	Scores of psychological test items	0–100	5

Self-efficacy refers to the belief and confidence in one's ability and the degree of belief one possesses in their ability to perform a certain task. Math-efficacy, similar to self-efficacy, refers to the belief and confidence students have in their mathematical abilities. Learning approach motivation represents the extent to which a student enjoys seeking knowledge and the extent to which they study to obtain knowledge. Reliance on academics measures students' dependence on academics and their awareness of the importance of academics in supporting their studies.

3.2. Data Visualization

Based on the collected data, the number of students sorted by grade level is shown in Figure 1. Most students were fifth-graders (in elementary schools), first-graders (in middle

schools), and first-graders (in high schools). However, in the original data, in terms of a grade-wise analysis, the data for only 464 students were linked to relevant variables, e.g., elementary/middle/high school grades, TLP test results, and mathematics scores. In contrast, the data for 1880 students contained TLP test and mathematics scores when grade information was excluded. Thus, in this study, rather than analyzing the students by grade, we analyzed the data and classified the abilities of all 1880 students using machine learning.

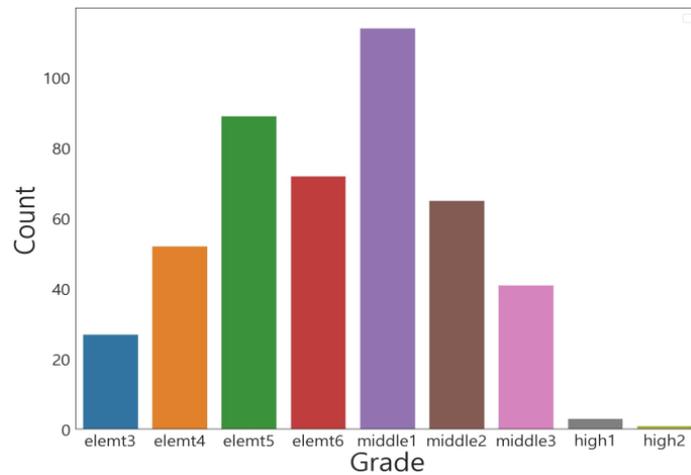


Figure 1. Number of students by grade.

The data on the mathematical achievement comprised the scores for Test 1 and Test 2, which are diagnostic tests for learning mathematics. The Test 1 scores are shown in Figure 2. The results revealed that 202 students were in the 0–10 points range, followed by 65, 53, 45, 32, 30, 18, 13, 5, and 1 students in the 30–40, 20–30, 10–20, 40–50, 50–60, 60–70, 70–80, 90–100, and 80–90 points range, respectively.

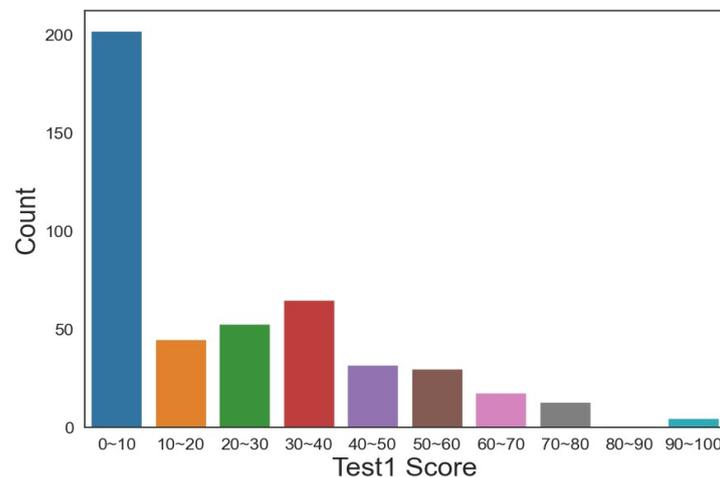


Figure 2. Distribution of Test 1 scores.

The results of Test 2 are shown in Figure 3. Here, 218 students scored in the 0–10-point range. For the remaining scores, 75, 61, 59, 17, 16, 14, 3, 1, and 0 students scored in the 10–20, 20–30, 30–40, 70–80, 40–50, 60–70, 50–60, 90–100, and 80–90.

Then, the number of students was analyzed according to their TLP scores, as shown in Figure 4. The histograms of self-efficacy, math-efficacy, and learning approach motivation exhibit a left-skewed distribution, and their medians are greater than the means. Performance approach motivation and reliance on academies have a symmetric tendency, which is similar to a normal distribution.

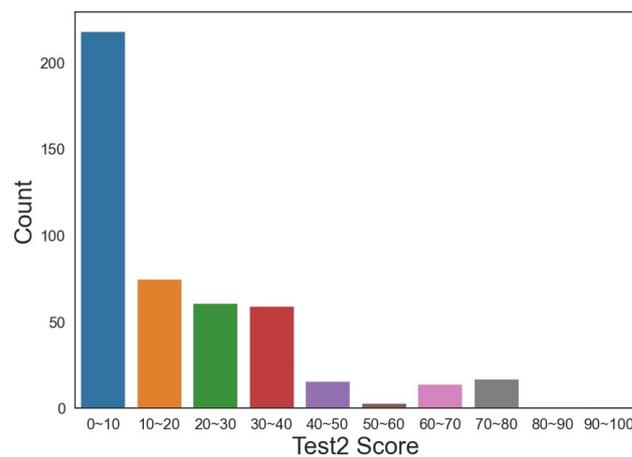


Figure 3. Distribution of Test 2 scores.

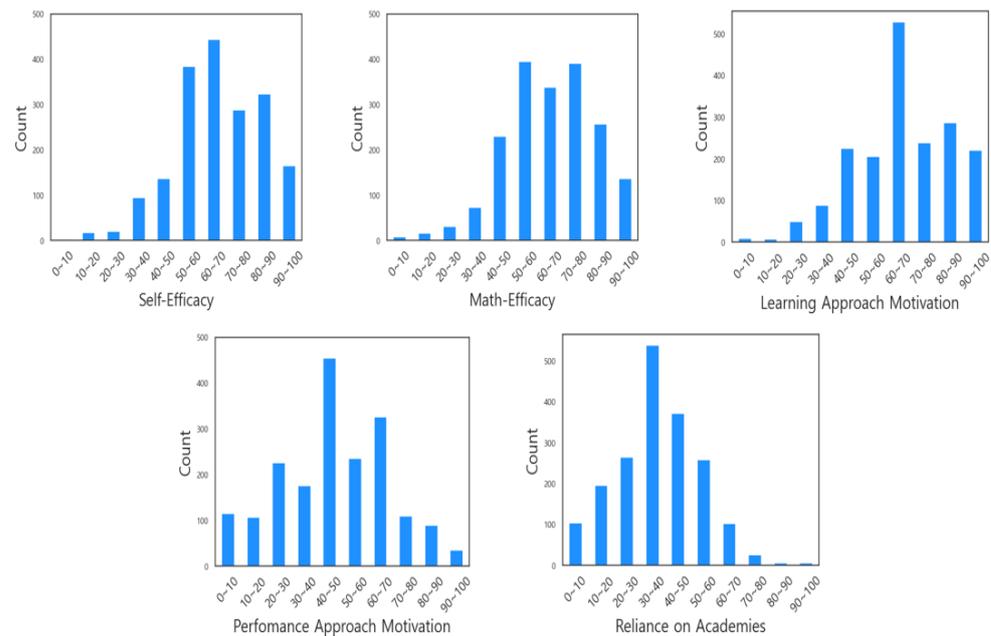


Figure 4. Number of students by TLP.

Specifically, regarding the point ranges of students according to their self-efficacy, 444 students scored in the 60–70-point range. Then, 384, 324, 289, 165, 137, 95, 21, 18, and 3 students scored in the 50–60, 80–90, 70–80, 0–100, 40–50, 30–40, 20–30, 10–20, and 0–10 point ranges, respectively. In terms of math-efficacy, the largest number of students was 395, with a score range of 50–60. Then, 391, 339, 258, 230, 137, 73, 32, 17, and 8 students scored in the 70–80, 60–70, 80–90, 40–50, 90–100, 30–40, 20–30, 10–20, and 0–10 point ranges, respectively. Regarding learning approach motivation, 529 students scored in the 60–70 point range. Then 287, 240, 226, 221, 207, 89, 50, 23, and 8 students scored in the 80–90, 70–80, 40–50, 90–100, 50–60, 30–40, 20–30, 0–10, and 10–20 point ranges, respectively. For performance approach motivation, 455 students scored in the 40–50 point range. Then, 326, 236, 226, 176, 118, 110, 107, 90, and 36 students scored in the 60–70, 50–60, 20–30, 30–40, 0–10, 70–80, 10–20, 80–90, and 90–100 point ranges, respectively. Regarding reliance on academies, most students ($n = 539$) scored in the 30–40 range. Then, 373, 266, 260, 196, 105, 103, 26, and 6 students scored in the 40–50, 20–30, 50–60, 10–20, 0–10, 60–70, 70–80, and both 80–90 and 90–100 point ranges, respectively.

4. Method

The overall workflow of this study is illustrated in Figure 5. In this study, we attempted to identify TLP items that impact mathematical achievement and classify the students' abilities accordingly.

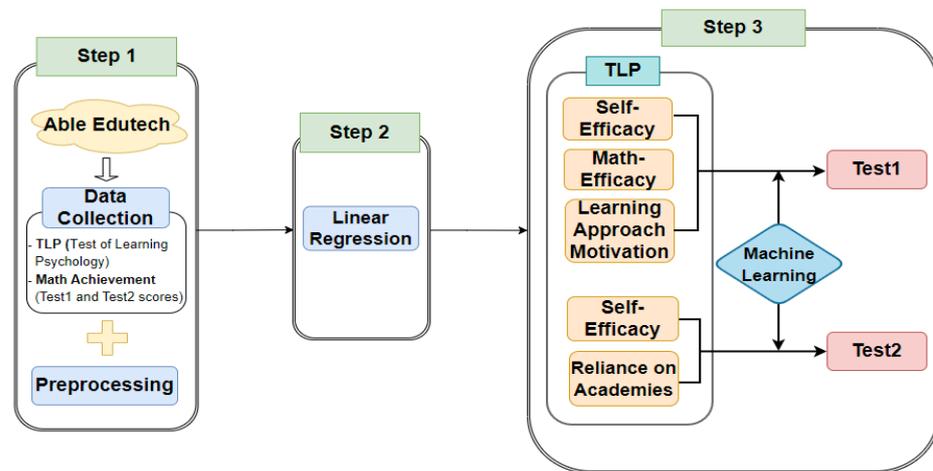


Figure 5. Workflow.

In the first step, we collected and preprocessed the data (TLP items and Test 1/Test 2 scores) provided by Able Edutech, and we selected the variables required for analysis. In the preprocessing stage, we removed missing values. In addition, outliers were observed in the mathematical achievement variable, and according to an agreement with Able Edutech, we replaced these outliers with the maximum value. The dependent variable for machine learning, i.e., mathematical achievement (Test 1 and Test 2), was categorized into high and low levels based on the grades. However, there was a significant data imbalance in all tests; thus, oversampling techniques were utilized to balance the data. The results are shown in Tables 2 and 3.

Table 2. Count of Test 1 classes after oversampling.

Test 1 Classes	Count	Oversampled Count
High level	276	1604
Low level	1604	1604

Table 3. Count of Test 2 classes after oversampling.

Test 2 Classes	Count	Oversampled Count
High level	126	1754
Low level	1754	1754

In the second step, linear regression analysis was conducted to identify the TLP items that have a statistically significant influence on students' mathematical achievement and select relevant variables for machine learning classification.

In the final step, we classified students' mathematical achievements according to these variables using machine learning techniques. Here, we employed various algorithms, i.e., the logistic regression, K-nearest neighbors (KNN), random forest, decision tree, Support vector machine (SVM), gradient boosting machine (GBM), light gradient boosting machine (LGBM), and extreme gradient boosting (XGBoost) algorithms. These algorithms are described in the following.

4.1. Used Algorithms

4.1.1. Linear Regression

Linear regression is a regression analysis technique that models the linear correlation between the dependent variable y and the independent variable x . A linear relationship implies that the independent variable x affects the dependent variable y according to $y = ax + b$, where the slope a and y -intercept b are obtained from the training data. For such a linear relationship, the value of y for the dependent variable can be predicted when a new variable, x , is given. Linear regression results can be easily interpreted and modeled quickly.

4.1.2. Machine Learning Classification Algorithms

Logistic regression is a representative supervised machine learning algorithm used for binary classification tasks. Using a sigmoid function with a value between 0–1, classification is performed based on the probability that an item belongs to a particular category. The sigmoid function is expressed in Equation (1), where e , i.e., Euler’s number, is (2.718281...).

$$\text{Sigmoid Function} = \frac{1}{(1 + e^{-x})} \tag{1}$$

Here, if the input value x is a large negative number, it is set to 0, and if x is a large positive number, it is set to 1. Thus, classification is performed by predicting a probability value between 0–1.

A support vector machine (SVM) is a supervised machine learning algorithm and a powerful classification model [20]. Support vectors are the nearest data points to the decision boundary. The decision boundary is selected in a way that maximizes the distance between them. The objective of the SVM is to find an optimal decision boundary, which can be expressed as follows:

$$w^T x + b = 0. \tag{2}$$

Here, w represents the normal vector to the decision boundary, and b is the intercept. In addition, w^T denotes the transpose of w . The goal of the SVM is to maximize the distance between the support vectors and the decision boundary. Therefore, the SVM can be formulated as the following optimization problem:

$$\text{minimize } \frac{1}{2} \|w\|^2, \text{ subject to } y_i (w^T x_i + b) \geq 1 \text{ (for all data points)}. \tag{3}$$

Here, $\|w\|^2$ represents the norm of w , and y_i represents the class of the data point. The SVM can find the optimal decision boundary by solving this optimization problem. SVM is a model that classifies which side of the boundary the input data belongs to through the decision boundary.

The K-nearest neighbors (KNN) technique is a simple supervised classification algorithm that works to identify classes with a set number of k data that are near to new data among the existing data and classify the new data into a class that has more existing data. The distance calculation is performed based on the Euclidean distance metric:

$$\text{Euclidean distance} = \sum_{i=1}^n (x_i - y_i)^2 \tag{4}$$

Here, x represents a new data point, and y represents all data points in the dataset. Compared with other models, the KNN algorithm is advantageous because it is relatively easier to understand.

A random forest is a type of ensemble learning method wherein results are obtained by collecting the classification results from multiple decision trees constructed during the training process. Random forests are used to solve various problems, e.g., classification and regression tasks. The random forest model was developed according to Breiman’s method [21]. Note that random forest predictions are based on the results of many randomly

generated decision trees; thus, overfitting is reduced, and good generalization performance is demonstrated. Therefore, the random forest algorithm is a fast technique that provides highly accurate results.

A decision tree is a supervised machine learning algorithm that classifies data using classification criteria based on the attributes of each data item. In other words, the decision tree method classifies data by branching based on whether specific criteria are satisfied. The results of decision trees are easy to interpret and understand. Regarding other techniques, regularization or variable creation/removal is required in some cases, whereas decision trees rarely require data processing. Additionally, the decision tree model is rarely affected by outliers, exhibits good stability, and can be applied to both numerical and categorical data.

The gradient boosting machine (GBM), developed by Friedman [22], is a supervised machine learning technique used for regression or classification tasks that reduces residuals through gradient descent and incorporates boosting techniques. Here, boosting refers to combining weak learners to reduce errors and create a strong learner. In other words, a strong prediction model is constructed using an ensemble of weak prediction models, and the subsequent classifier is trained based on the prediction error of the previous weak classifier to compensate for its error. Generally, the GBM method outperforms the random forests method.

The extreme gradient boosting (XGBoost) algorithm, developed by Chen [23] at the University of Washington, improves and extends gradient boosting algorithms to support parallel learning. The XGBoost method comprises a Classification and Regression Tree (CART) model, which is expressed as follows.

$$\hat{y}_i = \sum_{k=1}^K f_k(x_i), f_k \in F \quad (5)$$

Here, \hat{y}_i represents the predicted value of data point x_i , K denotes the number of CARTs used, and f represents the CART models. The objective function for training the CART model is expressed as follows.

$$Obj = \sum_{i=1}^n l(y_i, \hat{y}_i) + \sum_{k=1}^K \Omega(f_k) \quad (6)$$

Here, $l(y_i, \hat{y}_i)$ represents the objective function computed from the true answer y_i and the predicted value \hat{y}_i , and Ω is the regularization function of the model used to prevent overfitting. This algorithm executes faster than the gradient boosting method and provides excellent performance in classification and regression tasks. Additionally, it includes an overfitting regulation function; thus, it exhibits strong durability.

The light gradient boosting machine (LGBM) algorithm was developed by Microsoft. This algorithm performs ranking and classification based on the decision tree algorithm and overcomes the shortcoming of existing models that require a long computation time. This process is achieved by deepening the tree in a leaf-wise manner as opposed to gradient boosting-type trees, which are generally level-wise methods, thereby reducing both time and memory costs.

4.1.3. Machine Learning Evaluation Metrics

In this study, various evaluation metrics were utilized to assess the performance of the machine learning classification algorithms. Accuracy measures the proportion of correctly predicted samples out of the total samples. It represents how well the model classifies the data correctly. Precision is the proportion of true positive predictions (correctly predicted positive samples) to the total positive predictions made by the model. It assesses the accuracy of positive predictions. Recall measures the proportion of true positive predictions to the total actual positive samples. It evaluates the model's ability to find all positive samples. The F1-score is the harmonic mean of Precision and Recall, providing a balance between the two metrics. It is useful when both Precision and Recall are important,

and a higher F1-score indicates a better model performance. The evaluation metrics were calculated as follows:

$$\begin{aligned} \text{Accuracy} &= \frac{TP + TN}{TP + FN + FP + TN} \\ \text{Precision} &= \frac{TP}{TP + FP} \\ \text{Recall} &= \frac{TP}{TP + FN} \\ \text{F1 - Score} &= 2 \times \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \end{aligned}$$

True Positive = TP, True Negative = TN, False Positives = FP, False Negatives = FN.

ROC-AUC is a widely used evaluation metric for binary classification models. It evaluates the model’s performance by plotting the True Positive Rate (Recall) against the False Positive Rate and calculating the area under the ROC curve (AUC). A higher AUC value, ranging from 0 to 1, indicates better model performance.

5. Results

5.1. Linear Regression

Data processing for regression analysis was performed using the statsmodels Python package. Table 4 shows the results of the linear regression analysis with the results of the TLP items as independent variables and the Test 1 and Test 2 scores (mathematical achievement) as dependent variables. The variables are statistically significant ($p < 0.05$).

Table 4. Variables considered in the study.

Dependent Variables	Independent Variables	Coefficient	T-Value	P-Value	Prob (F-Statistics)
Test 1	Self-Efficacy	0.160 ***	3.579	0.000	0.000
	Math-Efficacy	0.190 ***	4.102	0.000	0.000
	Learning Approach Motivation	0.119 **	2.846	0.004	0.004
Test 2	Self-Efficacy	0.074 *	2.216	0.027	0.027
	Reliance on Academies	−0.079 *	−2.214	0.027	0.027

***: $p < 0.001$, **: $p < 0.01$, *: $p < 0.05$.

Table 4 shows that Prob (F-statistics) verifies the significance of the models (<0.05). The significance of the p -value for each dependent variable was confirmed. Here, self-efficacy, math-efficacy, and learning approach motivation had a statistically significant effect on Test 1, with math-efficacy having a relatively greater influence than the other factors. In Test 2, self-efficacy and reliance on academies were identified as statistically significant variables, and reliance on academies exhibited a relatively greater influence than self-efficacy.

5.2. Performance Evaluation of Machine Learning Classification Models

In the linear regression results, we confirmed which TLP variables affect mathematical achievement. Additionally, to classify mathematical achievement, we applied machine learning techniques that enable the classification of students’ mathematical achievements using the identified variables.

We considered the logistic regression, KNN, random forest, decision tree, SVM, GBM, LGBM, and XGBoost, which are machine learning classification models implemented in Python. Using these models, machine learning classification was performed with self-efficacy, math-efficacy, and learning approach motivation as the independent variables for the dependent variable Test 1 and self-efficacy and reliance on academies as the independent variables for Test 2. The training and test data were split at a ratio of 8:2, and the evaluation

results of all machine learning application models were performed with five-fold cross-validation. The evaluation results of the machine learning-based classification of the dependent variable scores (i.e., Test 1 and Test 2) into high and low ranks are shown in Table 5.

Table 5. Evaluation results.

	Model	Accuracy	Precision	Recall	F1-Score	AUC
Test 1	Logistic regression	0.58	0.59	0.58	0.58	0.59
	KNN	0.58	0.57	0.62	0.60	0.60
	Random forest	0.73	0.70	0.74	0.79	0.78
	Decision tree	0.69	0.68	0.77	0.71	0.70
	SVM	0.60	0.62	0.51	0.56	0.63
	GBM	0.65	0.65	0.69	0.67	0.68
	LGBM	0.66	0.65	0.69	0.67	0.70
	XGBoost	0.70	0.68	0.76	0.71	0.74
Test 2	Logistic regression	0.57	0.57	0.55	0.56	0.60
	KNN	0.65	0.62	0.78	0.69	0.67
	Random forest	0.81	0.76	0.88	0.82	0.88
	Decision tree	0.77	0.73	0.87	0.79	0.79
	SVM	0.59	0.57	0.66	0.61	0.62
	GBM	0.68	0.65	0.76	0.70	0.70
	LGBM	0.69	0.66	0.75	0.70	0.73
	XGBoost	0.77	0.73	0.86	0.79	0.80

Machine learning-based classification was performed with the Test 1 score as the dependent variable. The best accuracy obtained by the random forest model was 73%. Then, accuracy values of 70, 69, 66, 65, 60, and 58% were obtained by the XGBoost, decision tree, LGBM, GBM, SVM, and both KNN and logistic regression methods, respectively. The precision of the random forest method was the highest at 70%. Then, 68, 65, 62, 59, and 57%, respectively, for decision tree and XGBoost, GBM and LGBM, SVM, logistic regression, and KNN. The decision tree method obtained the highest recall score of 77%, followed by the XGBoost, random forest, GBM and LGBM, KNN, logistic regression, and SVM with 76, 74, 69, 62, 58, and 51%, respectively. The highest F1-score of 79% was obtained by the random forest method, followed by the decision tree and XGBoost, GBM and LGBM, KNN, logistic regression, and SVM methods with 71, 67, 60, 58, and 56%, respectively.

Then, machine learning-based classification was performed with the Test 2 scores as the dependent variable. The results showed that the best accuracy of 81% was obtained by the random forest model, followed by the decision tree and XGBoost, LGBM, GBM, KNN, SVM, and logistic regression methods with 77, 69, 68, 65, 59, and 57%, respectively. The random forest method obtained the best precision of 76%, followed by the decision tree and XGBoost, LGBM, GBM, KNN, and both logistic regression and SVM methods with 73, 66, 65, 62, and 57%, respectively. The recall of the random forest method was the highest at 88%, followed by the decision tree, XGBoost, KNN, GBM, LGBM, SVM, and logistic regression methods, which obtained recall values of 87, 86, 78, 76, 75, 66, and 55%, respectively. The highest F1-score of 82% was obtained by the random forest method. The F1-scores of the decision tree and XGBoost, GBM and LGBM, KNN, SVM, and logistic regression methods were 79, 70, 69, 61, and 56%, respectively. In terms of the AUC, both Test 1 and Test 2 demonstrated that the random forest method obtained the highest value (Test 1: 0.78, Test 2:

0.88), and the logistic regression method obtained the lowest value (Test 1: 0.59, Test 2: 0.60). Comprehensively, we found that the random forest method obtained the best performance in terms of Test 1 and Test 2 across most of the performance metrics. Thus, we classified the students' abilities by TLP items, resulting in high evaluation values.

6. Discussion

In this study, we identified the psychological factors that influence mathematical achievement and classified students' abilities based on the identified variables. As a result, the following academic implications arise from this research: By introducing the TLP, which is a psychological assessment test specifically designed to measure students' learning psychology in the mathematics field, we investigated the relationship between psychological factors and mathematical achievement. We believe that our findings have significant academic value because they effectively fill an identified gap in previous studies by focusing on a mathematics-centric psychological test. Furthermore, we discovered the influence of new variables, e.g., math-efficacy, learning approach motivation, and reliance on academics, on the mathematical achievements of students. These findings establish a foundation to utilize these relevant variables in predicting mathematical achievement.

The practical implications of this study are summarized as follows. We have identified variables that influence students' abilities and identified machine learning classification algorithms that can be applied in the mathematics education field. Previously, educational institutions, particularly academies, have relied solely on grades to assess students' abilities. However, through our findings, we have demonstrated the ability to predict students' abilities using mathematics-related psychological factors and identified psychological elements that may be lacking in students' mathematical proficiency. Ultimately, we expect that this will enable educational institutions to design effective personalized learning programs to improve students' academic performance by positively transforming their deficient psychological factors in addition to the existing grade-based management system.

7. Conclusions

In this study, we utilized the TLP items to identify their impact on students' mathematical test scores and employed machine learning techniques to classify students' mathematical skills. We believe that our research findings provide two significant insights.

First, the linear regression analysis results indicated that self-efficacy, math-efficacy, and learning approach motivation influenced mathematical achievement in Test 1. In Test 2, self-efficacy and reliance on academics affected mathematical test scores, which measured mathematical achievement. Overall, we have confirmed the influence of self-efficacy on mathematical achievement and demonstrated that the psychological test of mathematical learning can measure these achievements effectively.

Second, by applying several machine learning techniques, we achieved high performance in all performance evaluation indicators (accuracy, precision, recall, and F1-Score), and students' skills were successfully classified based on the TLP items.

These results have practical implications for both educators and psychologists seeking to understand the psychological factors that influence students' mathematical learning. In addition, the results can support the development of personalized study programs based on each student's skills and enhance their mathematical achievements.

In future research, it would be beneficial to obtain psychological test results on mathematical learning from a larger sample of students and collect mathematics scores over a more extended period compared to the data used in the current study. In terms of the TLP items employed in this study, some items had limited or no data on student responses, thereby making it challenging to analyze these items effectively. Thus, these items were excluded; however, by acquiring additional data on these psychological variables, future studies could explore their impact further. Similarly, it would be beneficial to examine the effects of additional variables, e.g., age, gender, grade, and study hours, on mathematical achievement. This expanded analysis is expected to realize more accurate predictions of

students' mathematical skills and provide a more comprehensive understanding of the factors that influence student performance. Furthermore, our findings can be utilized to develop a personalized learning system that incorporates the classification of students' skills. Such a system could recommend relevant mathematics content tailored to individual students' abilities, thereby offering a promising method to improve mathematical learning outcomes.

Author Contributions: J.P.: Conceptualization, writing original draft, review and editing, data curation and analysis, methodology, visualization, project administration. S.K.: Resources, project administration. B.J.: Conceptualization, supervision, funding acquisition. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Ministry of SMEs and Startups grant funded by the Korean government (No. S3246571) and the Yonsei University Research Fund (No. 2023-22-0104).

Data Availability Statement: Data for this study was collected from Able Edutech Corporation. For further assistance, please contact the author.

Conflicts of Interest: The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

- Luan, H.; Geczy, P.; Lai, H.; Gobert, J.; Yang, S.J.H.; Ogata, H.; Baltes, J.; Guerra, R.; Li, P.; Tsai, C.C. Challenges and future directions of big data and artificial intelligence in education. *Front. Psychol.* **2020**, *11*, 580820. [[CrossRef](#)] [[PubMed](#)]
- Hwang, G.-J.; Tu, Y.-F. Roles and research trends of artificial intelligence in mathematics education: A bibliometric mapping analysis and systematic review. *Mathematics* **2021**, *9*, 584. [[CrossRef](#)]
- Bin Mohamed, M.Z.; Hidayat, R.; binti Suhaizi, N.N.; bin Mahmud, M.K.H.; binti Baharuddin, S.N. Artificial intelligence in mathematics education: A systematic literature review. *Int. Electron. J. Math. Educ.* **2022**, *17*, em0694. [[CrossRef](#)]
- Çiftçi, S.K.; Yildiz, P. The Effect of Self-Confidence on Mathematics Achievement: The Metaanalysis of Trends in International Mathematics and Science Study (TIMSS). *Int. J. Instr.* **2019**, *12*, 683–694. [[CrossRef](#)]
- Khesht-Masjedi, M.F.; Shokrgozar, S.; Abdollahi, E.; Habibi, B.; Asghari, T.; Ofoghi, R.S.; Pazhooman, S. The relationship between gender, age, anxiety, depression, and academic achievement among teenagers. *J. Fam. Med. Prim. Care* **2019**, *8*, 799–804. [[CrossRef](#)]
- Olivier, E.; Archambault, I.; De Clercq, M.; Galand, B. Student self-efficacy, classroom engagement, and academic achievement: Comparing three theoretical frameworks. *J. Youth Adolesc.* **2019**, *48*, 326–340. [[CrossRef](#)] [[PubMed](#)]
- Abu-Hilal, M.M. A structural model for predicting mathematics achievement: Its relation with anxiety and self-concept in mathematics. *Psychol. Rep.* **2000**, *86*, 835–847. [[CrossRef](#)] [[PubMed](#)]
- Hayat, A.A.; Shateri, K.; Amini, M.; Shokrpour, N. Relationships between academic self-efficacy, learning-related emotions, and metacognitive learning strategies with academic performance in medical students: A structural equation model. *BMC Med. Educ.* **2020**, *20*, 76. [[CrossRef](#)]
- Tokan, M.K.; Imakulata, M.M. The effect of motivation and learning behaviour on student achievement. *S. Afr. J. Educ.* **2019**, *39*, 1–8. [[CrossRef](#)]
- Madigan, D.J.; Curran, T. Does burnout affect academic achievement? A meta-analysis of over 100,000 students. *Educ. Psychol. Rev.* **2021**, *33*, 387–405. [[CrossRef](#)]
- León Hernández, A.; González Escobar, S.; Arratia López Fuentes NI, G.; Barcelata Eguiarte, B.E. Stress, self-efficacy, academic achievement and resilience in emerging adults. *Electron. J. Res. Educ. Psychol.* **2019**, *17*, 129–148.
- Jiang, Y.; Rosenzweig, E.Q.; Gaspard, H. An expectancy-value-cost approach in predicting adolescent students' academic motivation and achievement. *Contemp. Educ. Psychol.* **2018**, *54*, 139–152. [[CrossRef](#)]
- Romero, C.; Ventura, S.; Espejo, P.G.; Hervás, C. Data mining algorithms to classify students. In Proceedings of the First International Conference on Educational Data Mining (EDM 2008), Montreal, QC, Canada, 20–21 June 2008; pp. 8–17.
- Kabakchieva, D. Student performance prediction by using data mining classification algorithms. *Int. J. Comput. Sci. Manag. Res.* **2012**, *1*, 686–690.
- Xu, X.; Wang, J.; Peng, H.; Wu, R. Prediction of academic performance associated with internet usage behaviors using machine learning algorithms. *Comput. Hum. Behav.* **2019**, *98*, 166–173. [[CrossRef](#)]
- BujangSelamat, A.; Ibrahim, R.; Krejcar, O.; Herrera-Viedma, E.; Fujita, H.; Ghani, N.A.M. Multiclass prediction model for student grade prediction using machine learning. *IEEE Access* **2021**, *9*, 95608–95621.
- Sravani, B.; Bala, M.M. Prediction of student performance using linear regression. In Proceedings of the International Conference for Emerging Technology (INCET), Belgaum, India, 5–7 June 2020; pp. 1–5. [[CrossRef](#)]
- Zimmermann, J.; Brodersen, K.H.; Pellet, J.-P.; August, E.; Buhmann, J.M. Predicting graduate-level performance from undergraduate achievement. In Proceedings of the 4th International Conference on Educational Data Mining, Eindhoven, The Netherlands, 6–8 July 2011; pp. 357–358.

19. Le Mai, T.; Do, P.T.; Chung, M.T.; Thoai, N. Adapting the score prediction to characteristics of undergraduate student data. In Proceedings of the International Conference on Advanced Computing and Applications (ACOMP), Nha Trang, Vietnam, 26–28 November 2019; pp. 70–77. [[CrossRef](#)]
20. Cortes, C.; Vapnik, V. Support-vector networks. *Mach. Learn.* **1995**, *20*, 273–297. [[CrossRef](#)]
21. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
22. Friedman, J.H. Greedy function approximation: A gradient boosting machine. *Ann. Statist.* **2001**, *29*, 1189–1232. Available online: <http://www.jstor.org/stable/2699986> (accessed on 17 June 2023). [[CrossRef](#)]
23. Chen, T.; Guestrin, C. Xgboost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 785–794. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.