Contemporary Mathematics
https://ojs.wiserpub.com/index.php/CM/

Research Article

Mathematical Modelling of Latent Variables of Students' Performance of Public Engineering Universities in Mathematics Using Machine Learning

Ali Asghar[1*], Syed Feroz Shah[1], Asif Ali Shaikh[1], Waqar Sethar[2], Muhammad Tariq[3]

[1]Department of Basic Science and Related Studies, Mehran University of Engineering and Technology, Jamshoro, Sindh, Pakistan
[2]Department of MUISTD, Mehran University of Engineering and Technology, Jamshoro, Sindh, Pakistan
[3]Mathematics Research Center, Near East University, Near East Boulevard, PC: 99138, Nicosia/Mersin 10, Turkey
E-mail: asghar.sangah@faculty.muet.edu.pk

Received: 20 May 2025; **Revised:** 7 June 2025; **Accepted:** 11 June 2025

Abstract: This research paper provides a mathematical modelling and analysis for students' performance in mathematics in public engineering universities of Sindh Pakistan. The research distinguishes and measures the effect of different latent variables such as Institute Environment, Mathematics Anxiety, University Reputation (Students' Opinion) and University Facilities all these latent variables have been found from 36 observed variables and Proposed model is designed. Additionally, Random Forest (RF), Support Vector Machine (SVM), Gradient Boosting Machine (GBM) and Artificial Neural Networks (ANN) are used for comparison and validation of proposed model of machine learning which is designed. Evaluation of these models have been done through Accuracy, precision, recall, F1 score, and MAE, MSE and RMSE have been used. The findings provide further evidence for the contribution of well-being, engagement, and prior success in mathematics achievement. The proposed model provides practical implications for decision-makers and educators to come up with focused policies and interventions to enhance mathematics learning performance of engineering students in Sindh.

***Keywords*:** machine learning, latent variables, students' performance

MSC: 62H25, 68T05, 62D05

Abbreviation

GBM	Gradient Boosting Machine
SVM	Support Vector Machine
ANN	Artificial Neural Network
F1-score	F1 Score
MAE	Mean Absolute Error
MSE	Mean Squared Error
RMSE	Root Mean Squared Error
ITS	Intelligent Tutoring System

DOI: https://doi.org/10.37256/cm.6420257272

EFA	Exploratory Factor Analysis
CFA	Confirmatory Factor Analysis
RF	Random Forest
TP	True Positive
TN	True Negative
FP	False Positive
KMO	Kaiser-Meyer-Olkin Measure
PCA	Principal Component Analysis
PSO	Particle Swarm Optimization

1. Introduction

Test and assignment-based evaluation methods of students' performance have not been very comprehensive and productive in fulfilling the varied requirements of the students. As educational practices become more complex, the prediction of student performance has proven to be of increasing magnitude. To do so, researchers have been relying on Machine Learning (ML) methods, a class of methods that have proved to be useful in dealing with large, complex datasets and revealing previously hidden patterns and connections. When applied to student data, machine learning algorithms can yield accurate, personalized predictions, which are an invaluable aid to educators and administrators striving to improve learning for their students [1].

The focus for this research was prompted by the increasing need for evidence-based decision-making in education. Thanks to a large number of students data available in everything from demographics and academic achievement to extracurricular participation and behaviour, educators and institutions have the power to wield them in new ways they could barely even conceive just a few years ago [2]. The purpose of this research is investigating the possibilities of using machine learning to predict student performance, looking at several other factors besides academic, including social, behavioural.

Meeting these needs helps to promote self-regulation and wellbeing. A sense of efficacy, the belief in one's capabilities, is thought to promote meeting challenges and bolstering resilience, ultimately influencing academic achievement and retention [3]. This idea lies at the heart of establishing adaptive learning strategies. Adolescent self-esteem development plays a central role in adolescents' self-concept and social behaviour, psychological well-being is influenced by family and society. It is important to know the mechanisms behind this phenomenon to help adolescents in this crucial period [4]. Predictive models for at-risk students with standards-based grading supports early detection of needed intervention and increased success. Such models play a crucial role in tailoring education to different learning requirements [5]. Neural Network (NN) and optimization algorithms [6] based Intelligent systems have found application in the development of self-learning systems for education, providing personalized learning capabilities. These flexible systems evolve with student needs, with more positive outcomes. Neural network-based PSO models provide improved evaluations of teaching effectiveness by considering multiple performance indices. This will help ensure the fairness and trustworthiness of education evaluation [7]. Assisted by neural networks, science indicators can be used to evaluate the research performance of higher education institutions in order to obtain useful suggestions for education policy makers. These profiles add to the optimal performance of institutions [8]. The NN-based evaluation approach considers the individual differences, thus providing a customized evaluation indicating the individual capability of each student in the field of engineering education. This method is fairer and more objective to evaluate the student's academic level [9]. Educational data mining employs predictive analytics as a means to predict students who are at risk of attainment of low grades so that intervention and support can be provided when it needs to be [10]. This method produces a better education by not failing. Techniques in educational data mining are being explored for identifying predictors of academic success. One study investigated high school students' national test scores, searching for determinants of their performance. These approaches have been useful in educational research, including for predicting students' success [11]. The quality of education factor has emerged as an issue noted in national reports, with attention to the success of the quality of education. In Colombia, for example, the indicator reflects a positive trend over time in terms of quality of education according to all national achievements. This is the result of focused educational programs and interventions [12]. The impacts of poverty on student academic achievement are similarly an important construct related to student learning. Studies of social lag in Mexican cities have demonstrated that educational inequalities are closely linked to larger socioeconomic ones. Such results highlight the importance of targeted interventions in response to these systemic impediments to educational equity [13]. Institutional change has been vital for the development of backward regions in Europe. Fact-finding aimed at looking at how evolutionary changes in regional institutions could assist in promoting development in traditionally backward regions. On the basis of these findings, this study demonstrates that it is possible that institutional reform has the ability to contribute to the economic development and educational advancement of underdeveloped areas [14]. Policy based on 'triple helix', where universities, industry and government work together, has worked well in some places. In Wales, universities, for instance, have been at the forefront of stimulating critically-needed regional economic development through research partnerships and innovation. The crucial role of universities in regional development is emphasized by this model [15]. In the field of neuropsychiatric research, machine learning technology that essentially can identify patterns and make predictions has already been used. The past decade has seen the emergence of work into using tools such as Google Trends to gain an impression of the level of interest that the public has in different neuropsychiatric disorders, helping provide knowledge as to how such conditions may be understood and managed across the globe [16]. The 'curse of dimensionality' is a famous problem in data mining and machine learning. Traditional methods do not scale well as datasets become larger and more complex. This is especially true in educational data mining where data is in large volume and as a result, it may introduce additional complexity and potential error [17]. Behaviour patterns, like procrastination, as a proxy to predict student performance has been an emerging field of study in educational data mining. Procrastination have been found to be a strong predictor of low and high academic achievement, thereby presenting an opportunity for early intervention in the students at risk [18]. The learning analytics to support teaching and learning have been heavily investigated. An assessment of learning analytics research: its scope, nature and use discovered that learning analytics has developed robustly in using data to improve education. Using analytics, educators can gain understanding over learning behaviours and work to improve performance in a variety of subject matter [19]. This predictive modelling has been instrumental in the early identification of students who need support. A two-level model was established for predicting the risk of underachieving among students, for better higher education outcomes [20]. Such predictive systems assist teachers in concentrating on students requiring additional attention. Transfer learning approaches have been applied for predicting student success which are proved very useful for programming courses. These approaches enable the exploitation of prior knowledge in diverse knowledge domains for prediction of students' academic performance [21]. Another area in which data mining techniques appear promising is in predicting early dropout. A case study with high school students, was conducted and the results showed that data mining can be properly used to predict which students are at risk of quitting the course. This can be useful to predict in time dropout and improve retention rates [22]. SVMs have been the workhorse of machine learning for classification and prediction. Identifying suitable training sets for SVM models is essential for the performance of SVM models and Accuracy (ACC). A detailed survey of such approaches with great impact and success in different areas including education can be found in [23]. Another interesting technique that is also frequently applied in Educational Data Mining (EDM) is the random forest. They have been used in a diversity of domains, intrusion detection among them, but are useful to other domains, such as, student results. Random forests are characterized by robustness to data and suitable for high-dimensional datasets, thus being highly applicable to educational predictions [24]. Random Forest (RF) is a machine learning technique that constructs a set of decision trees to classify datasets. Employment in educational scenarios have had an impact in the development of strong models of student performance prediction. This approach has been shown to well handling large datasets and avoiding overfitting [25]. Support Vector Machines and the Kernel Trick represent fundamental tools to construct predictive models in educational data. By converting data into higher dimensions, they can be used to classify intricate educational data sets, particularly the ones used to predicting student success from multiple features [26]. An investigation of various feature selection techniques in student performance prediction proved that combining multiple methods is a successful way to determine relevant features. These models have led to more accurate models for predicting academic success which stress the necessity of the correct variable selection [27]. It is also applied to the data mining in order to estimate students' GPAs at college graduation. A historical data based model could predict the GPA quite accurately, which will be an important indicator to detect the students who might not be able to graduate successfully [28]. Machine learning algorithm has been employed in education to effectively forecast class outcomes. One paper emphasized the application of predictive models in predicting students' performance based on real-time analysis of data. This methodology is widely used in education institutions to customize courses according to the expected results [29]. In recent works prediction techniques for early warning systems at the undergraduate level have been compared and it is demonstrated different approaches for prediction result in different findings. Educators can use these approaches to hone their early intervention efforts and bolster retention rates so that students get the help they need before getting left behind [30]. Predicting academic performance using behavioral and learning data has been one of the main interests of educational research. One research used these methods to experiment in the classroom to automatically predict academic performance based on student behavior, providing an alternative view on predicting how well a student will do in school rather than relying only on conventional measures [31]. In the area of student performance estimation, the impact of consecutive lessons has also been investigated in research. Through understanding the performance of students across courses, researchers could establish models that more accurately predict the future performance and provide better academic support to students who are at risk of dropping off [32]. The types of math tasks which influence students' success in lower secondary classrooms: A latent class analysis. This method of statistical analysis brought about meaningful results regarding in general skills needed for success in mathematics as an education process-indicated learning strategy [33]. Higher order thinking skills have been defined and studied in education and particularly in mathematics education. This study highlights the significance of developing Thinking Frameworks (TFs) as it has a direct impact on the students' problem solving and critical thinking abilities [34]. Another study investigated the relationship between teachers' participation in professional learning communities, collective teacher efficacy, and cognitive activation on students' mathematics results. The study emphasized the importance of these factors, particularly together and in relation to learning outcomes and demonstrated a strong positive role played by teacher collaboration on student learning [35]. Comparisons between teachers and students' views about quality in maths' education Comparisons about teachers versus students' views of quality in mathematics education research in different countries. The findings on step 1 (sensitivity with regards to the relationship of educational quality and student achievement) provide different perspectives to see things through, and cultural and sociological factors that underlie educational achievement as well [36]. The relationship between instructional supports and cognitive demands has been investigated, particularly within the field of science and math education. This study showed that quality of interaction, rather than task related factors, is important in attention to task and understanding by students in the class [37].

The importance of this study is not only in the prediction of grades. This study employs machine learning algorithms to pinpoint the relative importance of factors that predict student performance, thereby revealing the interactions between variables that are not necessarily straight forward. For example, the upbringing, mental health, study habits and socialization of a student are other external and internal factors that can impact on a student's academic performance. Through focusing on the collective picture of student achievement, this study aims to provide a better sense of the factors affecting student outcomes.

Research Gap is that the Machine learning with latent factors is missing and have not been systematically addressed. There is also gap in the literature for these factors in Public Engineering Universities of Sindh. And there is gap too of usage of modern techniques of Machine Learning to model and predict with more accuracy.

The motivation to choose to focus on students' performance in Mathematics because of its crucial role in engineering education. Failure in Mathematics often leads to overall Engineering Career and delaying degree completion. Failure in Mathematics severely impact students' engineering careers and future opportunities. The study aims to uncover latent factors affecting mathematical performance using machine learning.

We present the different machine learning methods used, the settings of the analyses and the corresponding results in the ensuing sections of this paper. We aim to add value to the growing literature at the intersection of technology in education, by adopting a multi-pronged view of technology usage and its impact on our understanding of and ability to enhance student outcomes.

2. Materials and methods

For Latent Variables Factor analysis, Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA) have been used. And for comparison the results with Proposed model, these Random Forest, Support Vector Machine, Gradient Boosting Machine (GBM) and Artificial Neural Network have been used. And other model evaluation methods like Accuracy, Precision, Recall, F1-score, MAE, MSE and RMSE has been used. Missing data were imputed using machine learning techniques such as K-Nearest Neighbours (KNN) and Random Forest imputation and normalization was performed Z-score standardization. This pre-processing established a reliable foundation for accurate modelling. Models and their evaluation are as under in Figure 1.

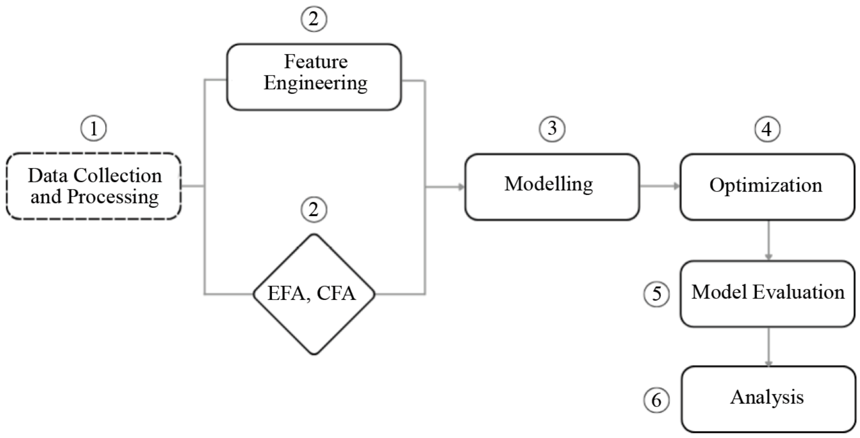

Figure 1. Models and evaluation

2.1 *RF*

Random Forest is an ensemble learning technique that builds multiple decision trees and outputs mode of the classes (classification). It is, we argue, relatively insensitive to overfitting and apt for efficient handling of high-dimensional data. In this work, the Random Forest classifier is learned a set of hyper parameters, such as number of trees, maximum tree depth and are tuned via grid search and cross validation. The default biases of the configuration were fine-tuned to optimize accuracy and bias of the model.

$$\widehat{y} = \frac{1}{N}\sum_{i=1}^{N} f_i(x) \tag{1}$$

where $\widehat{y}$ is predicted value and f_i is outcome of corresponding decision tree and N is number of trees.

2.2 *SVM*

Support Vector Machines (SVM) are a class of classifier that searches for a hyperplane that can separate classes in feature space. In a kernel trick, feature vectors were non-linearly mapped to a space in which the features were linearly separable. We evaluated using a SVM classifier with an Radial Basis Function (RBF) kernel and the hyper-parameters were optimized with a grid search for the regularization parameter (C) and kernel parameter (gamma).

Object is to calculate optimal hyperplane $w.x + b = 0$ where w is weight vector, b is bias term and x is independent feature vector.

Support Vector Machine Mathematically:

$$\min_{w,\, b} \frac{1}{2}\|w\|^2 \text{ subject to } y_i(w.x_i + b) \geq 1 \ \forall i \tag{2}$$

y_i is label of i-th data point (+1 or -1), x_i is feature vector of i-th data point (+1 or -1).

2.3 *GBM*

A gradient boost is an ensemble method which trains trees sequentially such that each tree attempts to correct the errors of the previous one. It is to improve the loss function itself to reduce errors. The model in this paper is trained with XGBoost, an enhanced GBM. Hyperparameters were optimized via cross-validation, including the learning rate, the number of estimators, and the maximum depth of the tree.

Mathematically:

$$F_t(x) = F_{t-1}(x) + \eta.h_t(x) \tag{3}$$

where $F_t(x)$ is prediction at t-th stage, $h_t(x)$ us trained model to fit the residuals, η is learning rate.

2.4 *ANN*

Furthermore, the use of an ANN (a deep learning model) was utilized to predict student achievement. The network was made by an input layer, one hidden layer, and an output layer using the softmax activation function. The number of hidden layer neurons was found empirically and the model was trained through backpropagation with Adam optimiser. The learning rate, dropout rate, and batch size were optimized through grid search.

Mathematically:

$$y = f(W_2.f(W_1.x + b_1) + b_2) \tag{4}$$

where W_1 and W_2 are used as weight matrices for hidden and input layers. b_1 and b_2 are used as bias terms for hidden and input layers, and f is the activation function x and y as independent and dependent feature vector.

2.5 *Proposed model*

The proposed model gives the complex relationships among multiple observed factors influencing students' mathematics performance in public engineering universities. It integrates linear, nonlinear, interaction, and exponential effects of Latent Variables.

$$\Gamma(\Psi) = \alpha_1 X_1 + \alpha_2 X_2^2 + \alpha_3 X_3^3 + \ldots + \alpha_n X_n^k + e^{(\gamma_1 X_{n+1} + \gamma_2 X_{n+2} + \ldots + \gamma_m X_{n+m})}$$

$$+ \delta_1 X_1 X_2 + \delta_2 X_3 X_4 + \ldots + \delta_k X_{n+m-1} X_{n+m} + \varepsilon \tag{5}$$

Proposed Model yields outcome where the strength and direction (positive or negative) of each predictor's (α_i, δ_i and γ_i) effect on combined latent performance variable $\Gamma(\Psi)$ and Identifying nonlinear and interaction effects, to help for finding uncover hidden patterns, e.g to see the combined effect of two variables (e.g., $X_{n+m-1}X_{n+m}$) which significantly impacts performance.

2.6 *Model evaluation methods*

2.6.1 *Accuracy*

The number of correctly predicted instances over the number of all instances in the dataset. It provides an overall feeling of how well the model predicts the true class.

Mathematically:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \times 100 \tag{6}$$

where T, F, P, N stands for True, False, Positive and Negative respectively.

2.6.2 *Precision*

The precision is the true positives divided by the sum of true positives and false positives. This tests the model's capacity to correctly welcome positive instances.

Mathematically:

$$\text{Precision} = \frac{TP}{TP+FP} \times 100 \tag{7}$$

2.6.3 *Recall (%)*

The precision is the true positives divided by the sum of true positives and false negative.

Mathematically:

$$\text{Recall} = \frac{TP}{TP+FN} \times 100 \tag{8}$$

2.6.4 *F1-score*

F1-score is the harmonic mean of precision and recall. It is particularly valuable in imbalanced datasets, which it balances precision and recall in-between.

Mathematically:

$$\text{F1 score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \times 100 \tag{9}$$

2.6.5 *MAE*

In this MAE the absolute difference between the predicted value and actual value is averaged out and return's the average error in a human-perceivable scale i.e. gives that how much the predictions are off on average.

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \widehat{y}_i| \tag{10}$$

where y_i and $\widehat{y}$ are actual and calculated values respectively.

2.6.6 *MSE*

MSE is a function and is similar to MAE, instead of taking absolute differences between the predicted and target values, it squares it. It has high variance in that it is particular susceptible to outliers in the data.

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \widehat{y}_i)^2 \tag{11}$$

2.6.7 *RMSE*

RMSE is simply the square root of the MSE and gives a similar measure in unit. It is another measure of how good the model is at predicting continuous values.

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \widehat{y}_i)^2} \tag{12}$$

3. Results

3.1 *Factor analysis*

Factor analysis serves as a powerful multivariate statistical method designed to uncover the hidden structure underlying a large set of observed variables. By grouping correlated variables into a smaller number of latent variables (or factors), it enables the identification of unobserved dimensions that drive the relationships among the variables. These latent factors are inferred constructs that help explain the observed correlations and variance, providing a simplified representation of the data's complexity. The primary objective of factor analysis is to reduce the number of variables while maintaining the explanatory power, thus facilitating the identification of underlying constructs that govern the data. This technique is particularly useful in fields like psychology, social sciences, and marketing, where it assists in data interpretation, theory development, and model refinement by revealing patterns that might otherwise remain hidden.

In Figure 2, Latent Variables (Blue Circles): The blue circles represent the latent factors in the model. In this diagram, three latent factors are identified: University Facilities, Mathematics' Anxiety, and University Reputation. Observed Variables (Yellow Rectangles): The yellow boxes represent the observed variables (indicators) that are being measured. These variables are hypothesized to be influenced by the corresponding latent factors. Factor Loadings (Arrows with Numbers): The numbers next to the arrows indicate factor loadings, which represent the strength of the relationship between the observed variables and their respective latent factors. Higher values (closer to 1) suggest a stronger relationship, while lower values indicate weaker relationships. For example, the variable "University Facilities" is strongly related to the latent factor "University Facilities" with a loading of 0.789. Relationships between Latent Factors: The arrows between the latent variables (e.g., from University Facilities to Institutional Environment) represent the relationships between these factors. These numbers indicate the strength of these relationships. For example, the relationship between Institutional Environment and University Reputation has a loading of 0.605. Model Structure: The overall structure suggests that several observed variables (such as those related to university facilities, maths anxiety, and university reputation) contribute to the formation of the latent factors, which are interconnected with one another.

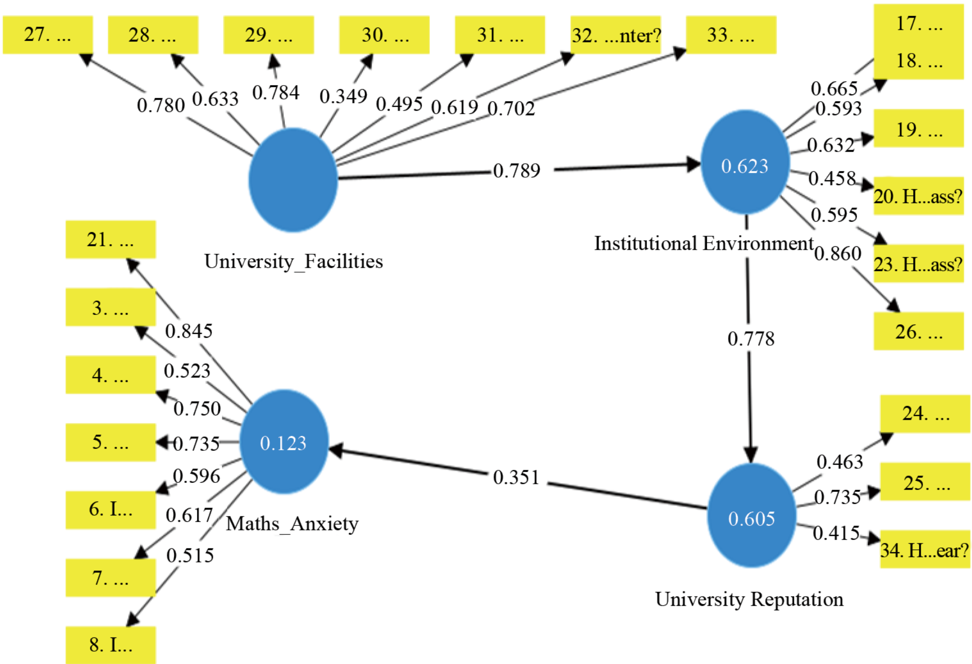

Figure 2. Factor analysis diagram show the latent factors in groups of manifested variables

3.1.1 *Exploratory factor analysis*

Table 1 shows Kaiser-Meyer-Olkin (KMO) Test is a measure of how suited your data is for Factor Analysis. Adequate sample between 0.8 and 1. Bartlett's Test of Sphericity: It tests whether the correlation off-diagonal elements are 0, a prerequisite for factor analysis to work. Average variance for all variables > 0.50 (at least 50% variation in latent variable) [38]. Cronbach's alpha is a way of assessing reliability by comparing the amount of shared variance, or covariance, among the items making up an instrument to the amount of overall variance.

Table 1. EFA for evaluation of reliability and validity of the latent variable

Latent variables	No. of variables	Cronbach α	EFA No. of factor	KMO	Bartlett's test sphericity	Average variance explained	Mean
Inst. Env.	7	0.813	1	0.867	0.000	0.616	3.91
Math Anxt.	7	0.853	1	0.812	0.000	0.542	4.18
Uni. Reput.	4	0.556	1	0.935	0.000	0.509	3.82
Uni. Facili.	7	0.823	1	0.811	0.000	0.509	4.84

3.1.2 *Confirmatory factor analysis*

Table 2 shows Cronbach alpha is > 0.70 except for Students opinion (0.555); Composite reliability for all variables is > 0.70 except student opinion (0.606); Average variance for all variables > 0.50 (at least 50% variation) [38]; Construct reliability is satisfactory based on reliability and AVE values.

Table 2. CFA for reliability and conformity of the latent variable

	Cronbach's alpha	Composite reliability	Composite reliability	Average variance extracted
Instructor role	0.813	0.827	0.805	0.616
Maths anxiety	0.853	0.860	0.843	0.542
University Reput.	0.556	0.606	0.556	0.509
University Facili.	0.823	0.846	0.821	0.509

3.2 *Data processing and feature selection*

Before using machine learning models Data pre-processing and feature selection has been involved for handling missing values through imputation, normalizing data to a common scale, and detecting and addressing outliers to ensure clean and standardized input. Feature selection included correlation analysis to remove redundant features, univariate tests to identify significant predictors, and Recursive Feature Elimination (RFE) to optimize the feature set. Principal Component Analysis (PCA) was applied to reduce dimensionality and address multi collinearity. These steps ensured that the dataset was well-prepared for accurate machine learning predictions.

3.3 *Machine leaning modelling and analysis*

To predict and analysis student performance, several supervised machine learning algorithms were employed, including Random Forest (RF), Support Vector Machine (SVM), Gradient Boosting Machine (GBM), and Artificial Neural Networks (ANN). A comparative analysis of these models was conducted based on accuracy, precision, F1-score, accuracy, MAE, MSE and RMSE. Techniques were used to interpret the comparison with proposed model decisions and predicting the performance.

Table 3. Comparison of models and evaluation models

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	MAE	MSE	RMSE
Proposed Model	94.5	93.4	95.8	94.6	0.52	0.45	0.67
Random Forest (RF)	87.5	88.2	85.8	87.5	0.65	0.62	0.79
Support Vector Machine	85.3	84.5	86.0	85.2	0.70	0.75	0.87
Gradient Boosting Machine	91.0	90.5	91.2	90.8	0.58	0.56	0.75
Artificial Neural Networks	90.2	87.2	89.5	87.1	0.60	0.58	0.76

The performance analysis of the different classifiers, such as RF, SVM, GBM, ANN and the proposed model are compared through accuracy, precision, recall, F1-Score, MAE, MSE and RMSE as shown in the Table 3. The highest level of accuracy (94.5%), precision (93.4%), recall (95.8%), and F1-Score (94.6%) is achieved by the proposed model compared with all models. It obtains the lowest MAE (0.52), MSE (0.45), and RMSE (0.67) meaning its an overall effective approach for prediction accuracy and error minimization. On the other hand, the remaining models (that is, Random Forest (accuracy: 87.5%) and SVM (accuracy: 85.3%)) are also good, but the to some extent they exhibit the inferior performance against proposed model in all the measurements including recall and F1-Score. In general, the proposed model is known by better result in all three measurements, which obviously substantiate the trustworthiness and effectiveness for the task.

3.4 *Figures of modelling*

It is shown in Figure 3, the different models' accuracy were compared. As observed, our proposed model obtains the best accuracy of 94.5%, which is superior to all other methods. The Random Forest model comes next with 87.5% followed by the Support Vector Machine (SVM) with 85.3%. Classification accuracy is the major measure of overall model performance, and the high accuracy of the proposed model obviously underlines its capability to efficiently differentiate the instances. Figure 4 compares the accuracy of the models, where it is a ratio of true positive predictions to all positive predictions of the model. The linked model outperforms with 93.4%, indicating it is better at predicting the positives with less false positives. Other black-box models such as SVM (84.5%) and Random Forest (88.2%) have lower precision implying they have more false positives. Figure 5 selects a better performance comparison on the bases of several metrics: Accuracy, Precision, Recall, and F1-Score. The proposed model is consistently best across all arem as in this case also, showing its overall effectiveness. It is interesting that Abstract Syntax Tree-Graph Attention Network (AST-GAT) has the best recall (95.8%) and F1-Score (94.6%), which shows that AST-GAT can not only detect most of the positive cases but also well balance the precision and recall. The other models are all of lower recall, and F1-Score, therefore the proposed model is the most accurate. Figure 6 compares the models comparing from MAE, MSE and RMSE. The proposed approach assumes again the smallest values in all three measurements, hence identifying it as the most accurate prediction one with the least error. Comparatively, the MAE, MSE, and RMSE of Random Forest and SVM are larger, which means they have greater prediction error than the proposed model.

From Figure 7, the model demonstrates strong overall performance, as indicated by the high number of correctly classified instances along the diagonal of the confusion matrix. The accuracy is particularly notable for classes 0, 3, and 4, where no misclassifications were observed.

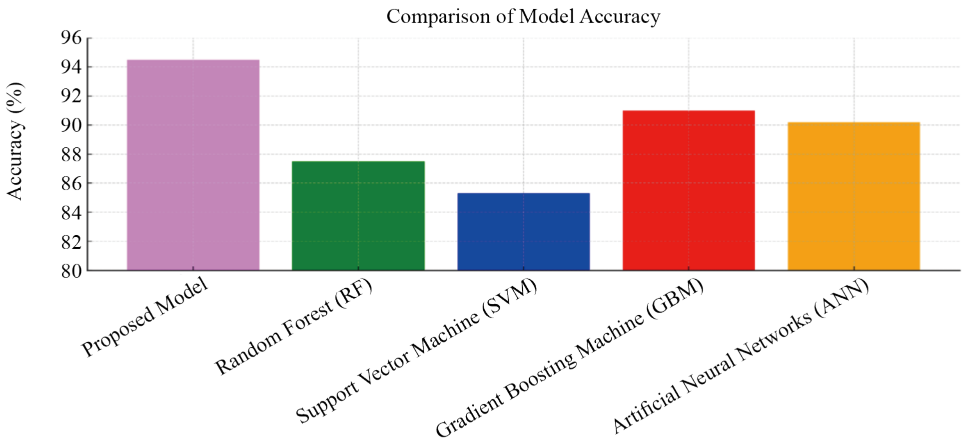

Figure 3. Comparison of accuracy of models

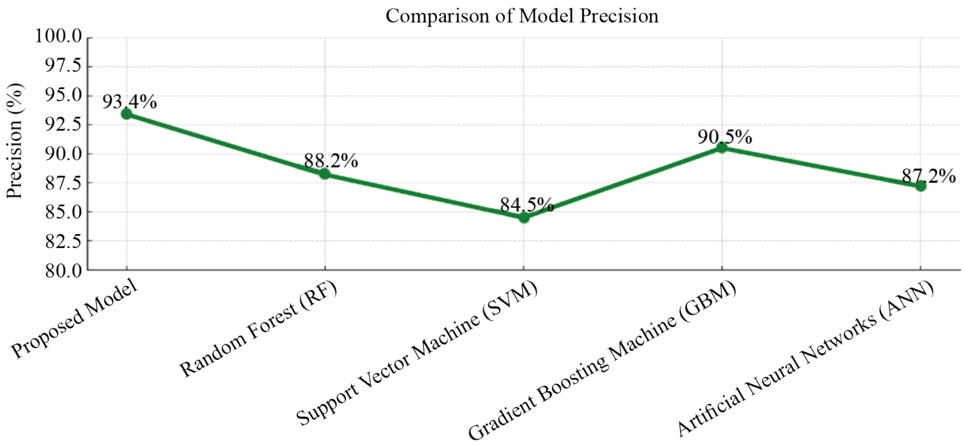

Figure 4. Comparison of precision of models

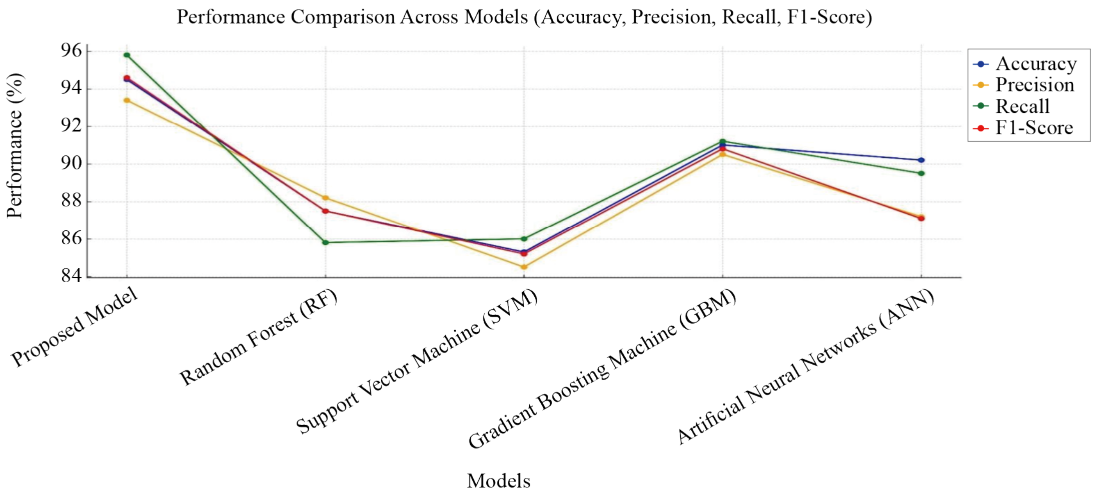

Figure 5. Performance comparison of models using accuracy, precision, recall and F1-Score

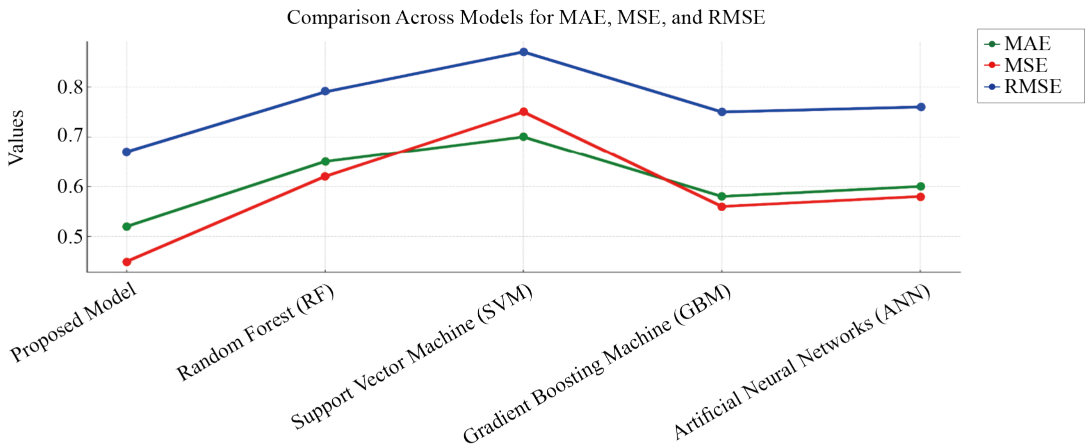

Figure 6. Comparison of models for MAE, MSE and RMSE

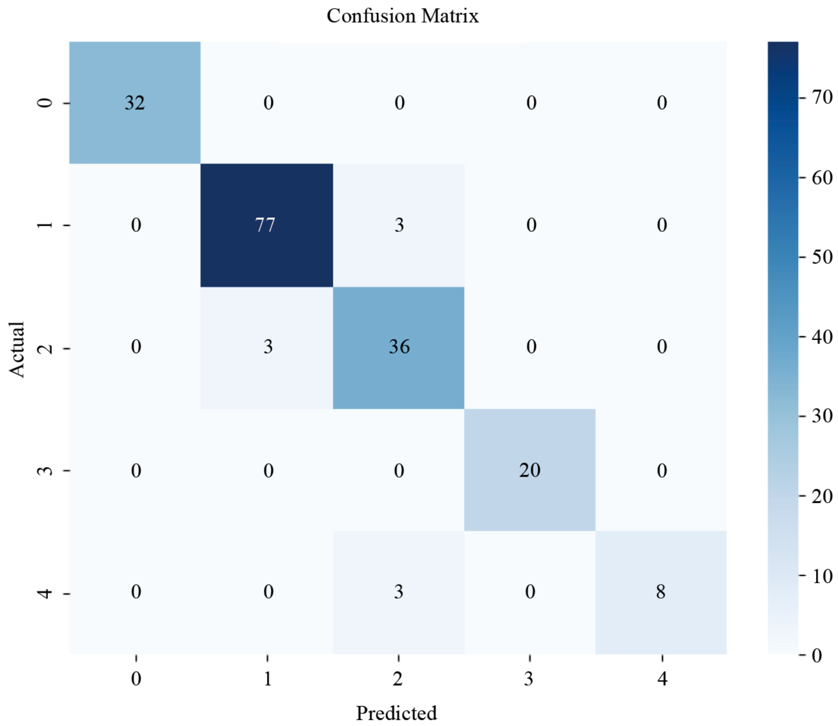

Figure 7. Confusion matrix analysis of proposed model

4. Discussion

Despite the remarkable performance of the proposed model, future research improvements do exist. One of the potential directions is model interpretability and implementations, since most state-of-the-art models are often considered as black boxes, notably deep learning models. Research can entail improving versions of the models to be able allow the end-users to make sense of the way the models arrive at the inferences, particularly important in regulated industries, education, sports, Science and Engineering. An interesting field of research is validation the proposed model with other more complex datasets or real time-situations which is done, evaluation of the scalability and robustness of the model is worthy. Furthermore, hybrid models, which combine different algorithms to obtain the benefits of each, may also be of interest for future research. It would also be interesting to consider how the model generalizes in different subgroups or in situations involving imbalanced data. This would facilitate understanding importance of the models and to study outperforms under these conditions and whether there is the necessity for additional optimization in relation with the challenge of the respective classifier of either dealing with imbalanced distributions or identifying minority classes.

5. Conclusions

The study emphasizes the successful usage of machine learning algorithms for predicting and analyzing the students' performance in the domain of mathematics, especially in the public engineering universities of Sindh, Pakistan. The performance latent variable $\Gamma(\Psi)$ of manifested variables have been remained the most accurate with least errors as compared to the discussed available models. Using predictive models such as Proposed Model RF, SVM, GBM, and ANN, the work explains the capability of these models to predict the final student success with the evidence that consists of several latent factors like academic background, mental health, and institutional support. Based on the evaluation metrics accuracy, precision, recall, and F1-Score, it is obvious that proposed model performs better than the conventional machine learning algorithms in the prediction accuracy, precision, recall, F1-score. The accuracy of the proposed model reaches 94.5%, much higher than those of the RF, SVM and GBM models and also in all the key performance metrics with least errors MAE, MSE and RMSE. The confusion matrix analysis also highlights the model has strong capacity to correctly assign student performance to multiple classes, which further demonstrates its applicability in educational decision-making. The findings indicate that variables such as mental health (maths' anxiety) and University facilities also contribute in a significant way in influencing the academic performance of students, particularly in mathematics. Also, results underscore the predictive significance of past academic achievement and the impact of the institution, both of which are consistent with previous research regarding predictors of student success. As the study utilized multi-variable methodologies, it provides important evidence for educational policy-makers and institutions to consider in order to develop specific systemic interventions to improve student performance in science and mathematics in non-traditional schools. In summary, the study constructs a solid foundation for leveraging machine learning to predict student success with a data-driven method to educational interventions. Such prediction can act as a 'whistle blower' in the hands of the teachers to ring the alarm-bell in the meantime and help the student before it is too late and the academic career is on the verge of getting lost. Further research should seek to incorporate real-time behavior and long-term.

Ethical consideration

The data used in this study was collected with prior approval from academic authorities, following institutional ethical standards. All student information was anonymized to protect privacy, and informed consent was obtained.

Research limitation and constraints

Due to privacy and ethical constraints, our study is aggregated and limited within the ethical boundaries imposed by cultural values and institutional rules.

Acknowledgement

Thanks Mehran University of Engineering and Technology Jamshoro for Financial support for this research paper.

Conflict of interest

The authors declare no competing financial interest.

References

[1] Du Boulay B, Del Soldato T. Implementation of motivational tactics in tutoring systems: 20 years on. *International Journal of Artificial Intelligence in Education*. 2016; 26: 170-182. Available from: https://doi.org/10.1007/s40593-015-0052-1.

[2] Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and wellbeing. *American Psychologist*. 2000; 55: 68-78. Available from: https://doi.org/10.1037/0003-066X.55.1.68.

[3] Bandura A. Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*. 1977; 84: 191-215. Available from: https://doi.org/10.1037/0033-295X.84.2.191.

[4] Rosenberg M, Morri H. *Society and the Adolescent Self-Image*. Princeton University Press; 1965.

[5] Marbouti F, Diefes-Dux HA, Madhavan K. Models for early prediction of at-risk students in a course using standards-based grading. *Computers & Education*. 2016; 103: 1-15. Available from: https://doi.org/10.1016/j.compedu.2016.09.005.

[6] Kose U, Arslan A. Optimization of self-learning in computer science engineering course: an intelligent software system supported by artificial neural network and vortex optimization algorithm. *Computer Applications in Engineering Education*. 2017; 25: 142-156. Available from: https://doi.org/10.1002/cae.21787.

[7] Rashid TA, Ahmad HA. Lecturer performance system using neural network with particle swarm optimization. *Computer Applications in Engineering Education*. 2016; 24(4): 629-638. Available from: https://doi.org/10.1002/cae.21737.

[8] Villasenor EA, Arencibia-Jorge R, Carrillo-Calvet H. Multiparametric characterization of scientometric performance profiles assisted by neural networks: a study of Mexican higher education institutions. *Scientometrics*. 2017; 110: 77-104. Available from: https://doi.org/10.1007/s11192-016-2166-0.

[9] Mustafa HMH, Al-Hamadi A, Al-Ghamdi SA, Hassan MM, Khedr AA. On assessment of students' academic achievement considering categorized individual differences at engineering education (neural networks approach). In: *2013 12th International Conference on Information Technology Based Higher Education and Training (ITHET)*. Antalya: IEEE; 2013. p.1-10. Available from: https://doi.org/10.1109/ITHET.2013.6671003.

[10] Fernandes E, Holanda M, Victorino M, Borges V, Carvalho R, Van Erven G. Educational data mining: Predictive analysis of academic performance of public school students in the capital of Brazil. *Journal of Business Research*. 2019; 94: 335-343. Available from: https://doi.org/10.1016/j.jbusres.2018.02.012.

[11] Timaran-Pereira R, Hidalgo A, Caicedo J, Benavides J. Discovering factors associated with academic performance of high school students in saber 11th test using educational data mining techniques. In: *Proceedings of E-Learn: World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education*. Association for the Advancement of Computing in Education (AACE); 2018. p.188-197.

[12] Ministerio de Educación Nacional [Ministry of National Education]. *Colombia mejora progresivamente en calidad educativa, así lo evidencian los resultados del ISCE [Colombia's educational quality is progressively improving, as evidenced by the ISCE results]*. Available from: http://mineducacion.gov.co [Accessed 12th March 2018].

[13] Maldonado FJB, Meza MVG. El rezago social en áreas metropolitanas de México [Social backwardness in metropolitan areas of Mexico]. *Estudios Económicos*. 2013; 265-297.

[14] Rodríguez-Pose A, Ketterer T. Institutional change and the development of lagging regions in Europe. *Regional Studies*. 2020; 54: 974-986. Available from: https://doi.org/10.1080/00343404.2019.1608356.

[15] Pugh R. Universities and economic development in lagging regions: 'triple helix' policy in Wales. *Regional Studies*. 2017; 51: 982-993. Available from: https://doi.org/10.1080/00343404.2016.1171306.

[16] Sharma M. Research and Google trend for human neuropsychiatric disorders and machine learning: a brief report. *Psychiatria Danubina*. 2021; 33: 354-357. Available from: https://doi.org/10.24869/psyd.2021.354.

[17] Keogh EJ, Mueen A. Curse of dimensionality. In: *Encyclopedia of Machine Learning and Data Mining*. Boston, MA: Springer. 2017. p.314-315. Available from: https://doi.org/10.1007/978-1-4899-7687-1_192.

[18] Hooshyar D, Pedaste M, Yang Y. Mining educational data to predict students' performance through procrastination behavior. *Entropy*. 2020; 22: 12. Available from: https://doi.org/10.3390/e22010012.

[19] Du X, Yang J, Shelton BE, Hung J, Zhang M. A systematic meta-review and analysis of learning analytics research. *Behaviour & Information Technology*. 2021; 40: 49-62. Available from: https://doi.org/10.1080/0144929X.2019.1669712.

[20] Shelton ES, Yang J, Hung JL, Du X. Two-stage predictive modeling for identifying at-risk students. In: *Innovative Technologies and Learning, ICITL 2018*. Springer; 2018. p.578-583. Available from: https://doi.org/10.1007/978-3-319-99737-7_61.

[21] Lagus J, Longi K, Klami A, Hellas A. Transfer-learning methods in programming course outcome prediction. *ACM Transactions on Computing Education*. 2018; 18(4): 1-18. Available from: https://doi.org/10.1145/3152714.

[22] Marquez-Vera C, Cano A, Romero C, Noaman AYM, Fardoun HM, Ventura S. Early dropout prediction using data mining: A case study with high school students. *Expert Systems*. 2016; 33: 107-124. Available from: https://doi.org/10.1111/exsy.12135.

[23] Nalepa J, Kawulok M. Selecting training sets for support vector machines: a review. *Artificial Intelligence Review*. 2019; 52: 857-900. Available from: https://doi.org/10.1007/s10462-017-9611-1.

[24] Resende PAA, Drummond AC. A survey of random forest-based methods for intrusion detection systems. *ACM Computing Surveys*. 2018; 51(3): 1-36. Available from: https://doi.org/10.1145/3178582.

[25] Breiman L. Random forests. *Machine Learning*. 2001; 45: 5-32. Available from: https://doi.org/10.1023/A:1010933404324.

[26] Ashfaque JM, Iqbal A. Introduction to support vector machines and kernel methods. 2019; 1-9. Available from: https://www.researchgate.net/publication/332370436 [Accessed 15th April 2019].

[27] Ma C, Yao B, Ge F, Pan Y, Guo Y. Improving prediction of student performance based on multiple feature selection approaches. In: *ICEBT '17: Proceedings of the 2017 1st International Conference on E-Education, E-Business and E-Technology*. Toronto, ON, Canada: Association for Computing Machinery; 2017. p.36-41. Available from: https://doi.org/10.1145/3141151.3141160.

[28] Tekin A. Early prediction of students' grade point averages at graduation: A data mining approach. *European Journal of Educational Research*. 2014; 54: 207-226.

[29] Pushpa S, Manjunath T, Mrunal T, Singh A, Suhas C. Class result prediction using machine learning. In: *Proceedings of the 2017 International Conference on Smart Technologies for Smart Nation (SmartTechCon)*. Bengaluru, India: IEEE; 2018. p.1208-1212. Available from: https://doi.org/10.1109/SmartTechCon.2017.8358559.

[30] Howard E, Meehan M, Parnell A. Contrasting prediction methods for early warning systems at undergraduate level. *The Internet and Higher Education*. 2018; 37: 66-75. Available from: https://doi.org/10.1016/j.iheduc.2018.02.001.

[31] Villagrá-Arnedo C, Gallego-Duran F, Compan-Rosique P, Llorens-Largo F, Molina-Carmona R. Predicting academic performance from behavioural and learning data. *International Journal of Design & Nature and Ecodynamics*. 2016; 11: 239-249. Available from: https://doi.org/10.2495/DNE-V11-N3-239-249.

[32] Sorour S, Goda K, Mine T. Estimation of student performance by considering consecutive lessons. In: *Proceedings of the 4th International Congress on Advanced Applied Informatics*. Okayama, Japan: IEEE; 2016. p.121-126. Available from: https://doi.org/10.1109/IIAI-AAI.2015.170.

[33] Adlef AK, Ross N, König J, Kaiser G. Types of mathematical tasks in lower secondary classrooms in Germany: Statistical findings from a latent class analysis based on general mathematical competencies. *Educational Studies in Mathematics*. 2023; 114: 371-392. Available from: https://doi.org/10.1007/s10649-023-10254-9.

[34] Lewis A, David S. Defining higher order thinking. *Theory into Practice*. 1993; 32: 131-137. Available from: https: //www.jstor.org/stable/1476693 [Accessed 13th March 2025].

[35] Liu S, Yin H. Opening the black box: How professional learning communities, collective teacher efficacy, and cognitive activation affect students' mathematics achievement in schools. *Teaching and Teacher Education*. 2024; 139: 104443. Available from: https://doi.org/10.1016/j.tate.2023.104443.

[36] Liu X, Yang Hansen K, De Neve J, Valcke M. Teacher versus student perspectives on instructional quality in mathematics education across countries. *Instructional Science*. 2024; 52: 477-513. Available from: https://doi.org/ 10.1007/s11251-023-09652-6.

[37] Prediger S, Erath K, Quabeck K, Stahnke R. Effects of interaction qualities beyond task quality: Disentangling instructional support and cognitive demands. *International Journal of Science and Mathematics Education*. 2024; 22: 885-909. Available from: https://doi.org/10.1007/s10763-023-10389-4.

[38] Hair Jr JF, Sarstedt M, Hopkins L, Kuppelwieser VG. Partial least squares structural equation modeling (PLS-SEM): An emerging tool in business research. *European Business Review*. 2014; 26(2): 106-121. Available from: https: //doi.org/10.1108/EBR-10-2013-0128.