Iranian Journal of Fuzzy Systems
Volume 22, Number 6, (2025), pp. 183-203
Original paper

# A hybrid ANFIS-gradient boosting frameworks for predicting advanced mathematics student performance

M. Zarei [1], M. S. Asgari [2] and N. Ghafoori Adl [3]

[1,2,3] *Department of Mathematics,CT. C., Islamic Azad University, Tehran,Iran*
mahnazzarei1@gmail.com, msasgari@yahoo.com, naser.ghafoori@gmail.com

**Abstract**

This paper presents a new hybrid prediction framework for evaluating student performance in advanced mathematics, thus overcoming the inherent constraints of classic Adaptive Neuro-Fuzzy Inference Systems (ANFIS). To improve predictive accuracy and model interpretability, our method combines ANFIS with advanced gradient boosting techniques, namely XGBoost and LightGBM. The proposed framework integrates fuzzy logic for input space partitioning with localized gradient boosting models as rule outcomes, effectively merging the interpretability of fuzzy systems with the strong non-linear modeling capabilities of machine learning. Comprehensive assessment reveals that both the ANFIS-XGBoost and ANFIS-LightGBM models substantially exceed the traditional ANFIS in various performance parameters. Feature selection, informed by SHAP analysis and XGBoost feature importance metrics, pinpointed essential predictors, including the quality of previous mathematics education and core course grades. Enhancements in regression measures further highlight the effectiveness of the hybrid methodology. The findings indicate that the suggested framework provides a reliable and effective alternative for educational institutions to forecast student success, guide focused interventions, and enhance learning outcomes in advanced mathematical fields.

*Keywords:* Adaptive neuro-fuzzy inference systems, feature selection, hybrid ANFIS-XGBoost model, hybrid ANFIS-LightGBM model.

## 1 Introduction

The precise forecasting of student performance in advanced mathematics is a key and ongoing challenge in educational analytics, with substantial consequences for curriculum enhancement, tailored interventions, and strategic resource distribution [25, 30, 35]. Mathematics, unlike many other disciplines, demonstrates distinctive predictive complexity owing to its hierarchical knowledge structure and abstract conceptual frameworks. Complications arise from its inherently non-linear learning trajectories and heightened vulnerability to cumulative knowledge gaps, rendering traditional linear models and generalized predictive approaches often inadequate [25, 35]. This need for accurate forecasting extends well beyond scholarly interest; it is crucial for educators and institutions seeking to proactively identify at-risk students, implement timely and targeted support strategies, and refine pedagogical practices to improve mathematics learning outcomes [26]. Accordingly, robust and reliable forecasting methods capable of providing actionable insights are essential for supporting both students and instructors and for enhancing the overall quality of mathematics instruction [1].

Predictive analytics offers a transformative framework for education by enabling proactive rather than retrospective evaluation [29]. Through early forecasting of learning trajectories, predictive models facilitate preemptive interventions and personalized learning pathways, significantly improving student outcomes [11, 29]. Such models are particularly valuable in advanced mathematics, where prompt identification of students struggling with abstract reasoning or cumulative concepts can meaningfully enhance instructional effectiveness [11].

Corresponding Author: M. S. Asgari
Received: May 2025; Revised: October 2025; Accepted: November 2025.
https://doi.org/10.22111/ijfs.2025.52017.9178

In response to this growing need, Educational Data Mining (EDM) has emerged as a critical field that applies data-driven and machine learning (ML) techniques to extract actionable insights from educational datasets [6]. A wide range of ML algorithms - including decision trees [24], support vector machines [32], artificial neural networks [3], and gradient boosting methods such as XGBoost [11, 18] and LightGBM [36, 40] - have demonstrated effectiveness in predicting academic performance. However, despite their success, a primary challenge remains: reconciling predictive robustness with model interpretability. This is particularly relevant when dealing with the inherent ambiguity of mathematical skill assessment and subjective distinctions between performance levels (e.g., differentiating "Proficient" from "Advanced" mathematical abilities) [33]. Addressing this challenge requires analytical frameworks capable of capturing both quantitative and qualitative dimensions of mathematical learning.

In modern intelligent systems, methods capable of managing the inherent complexity and uncertainty of real-world educational data are essential. Traditional statistical and machine learning methods often struggle with non-linearities, intrinsic imprecision, and the dual demand for accuracy and interpretability typical of mathematics-related datasets. Hybrid intelligent systems - designed to combine the complementary strengths of different computational paradigms - have emerged as a promising solution to these limitations. Among these, the Adaptive Neuro-Fuzzy Inference System (ANFIS) [20] stands out for its capacity to model uncertainty through fuzzy linguistic rules while preserving interpretability [4, 37]. However, standalone ANFIS frameworks face notable limitations, particularly reduced adaptability and robustness when working with high-dimensional or noisy datasets, due in part to their reliance on pre-defined fuzzy partitions and simple linear consequents [38]. While hybrid models combining ANFIS with gradient boosting techniques have been explored in prior research [8, 31], existing approaches often remain structurally limited, as they do not fully exploit the predictive power of boosting algorithms within the fuzzy rule framework.

To address these limitations, the present study introduces a novel hybrid framework that deeply integrates ANFIS with advanced gradient boosting algorithms. Unlike prior hybridizations that combine the models externally, our approach embeds gradient boosting predictors *within* the consequent layer of ANFIS itself. In conventional ANFIS, each fuzzy rule produces an output of the form

$$y_i = p_{i0} + p_{i1}x_1 + \cdots + p_{in}x_n, \tag{1}$$

which serves as a simple linear approximation. In our framework, this is replaced by a localized gradient-boosted decision-tree model - XGBoost or LightGBM - trained specifically for each fuzzy rule using a weighted loss function guided by its normalized firing strengths. As a result, each rule functions as a specialized "local expert" capable of learning intricate non-linear relationships within its fuzzy region. This architectural innovation enhances both predictive accuracy and interpretability, creating a cooperative ensemble of local models that outperform traditional ANFIS and previously proposed hybrid configurations.

Accordingly, this paper presents an innovative hybrid predictive architecture, consisting of ANFIS-XGBoost and ANFIS-LightGBM models, for forecasting student success in advanced mathematics. The empirical results demonstrate that this framework achieves higher predictive accuracy and improved interpretability compared to conventional ANFIS. By uniting the transparent reasoning of fuzzy logic with the powerful learning capabilities of gradient boosting ensembles, the proposed method offers an effective and interpretable solution for complex educational prediction tasks.

The remainder of this paper is organized as follows: Section 2 presents an in-depth review of the various factors influencing student performance in advanced mathematics, synthesizing prior research to identify key determinants of academic success. Section 3 details the theoretical foundations and methodologies adopted in this study, including the architectures and operational principles of ANFIS, XGBoost, and LightGBM. Section 4 describes the experimental design, covering data collection, dataset characterization, feature selection, and the construction of the hybrid prediction models. In Section 5, we evaluate the performance of the proposed ANFIS-XGBoost and ANFIS-LightGBM models against the conventional ANFIS model and discuss their implications for forecasting student success. Finally, Section 6 summarizes the key outcomes, addresses the study's limitations, outlines future research directions, and highlights the broader significance of our contributions to educational analytics.

## 2 Literature review: Determinants of student performance in advanced mathematics

This section examines the various elements influencing student performance in advanced mathematics. Understanding these characteristics is essential for comprehending why students excel or encounter difficulties in studying mathematics. This research seeks to identify the primary factors influencing student learning and performance in advanced mathematics by a meticulous examination of these variables.

Student attitudes significantly influence their engagement and success in mathematics. Wakhata et al. [39] propose that attitude constitutes an essential aspect of an individual's identity, significantly influencing their motivation and learning efficacy. A positive disposition towards mathematics enhances students' confidence, fosters enjoyment in problem-solving, and cultivates a greater appreciation for the subject, ultimately leading to improved academic performance [19]. Conversely, negative attitudes are frequently associated with increased anxiety, avoidance of mathematics, and lack of attentiveness during math class [22]. Students with a more favorable perception of mathematics are typically expected to achieve higher grades in the subject.

Learning engagement is a significant component influencing academic success. Student engagement in academic activities is a crucial determinant of their performance. "Learning involvement" encompasses various aspects: student behavior, emotional engagement, and cognitive processes. Regular attendance, punctuality, and active class participation are indicative of academic success [13]. "Feeling involved" pertains to the feelings and perspectives students hold on their academic responsibilities, whereas "thinking involved" relates to their capacity to tackle challenging challenges and persist despite difficulties in their studies. The many components of engagement collaborate to enhance students' performance in mathematics by fostering a deeper connection and interest in the topic [2].

Motivation acts as a crucial catalyst in influencing students' learning practices, particularly impacting their readiness to tackle challenging mathematical concepts. It significantly affects their motivation to exert effort in mathematics. Motivation is associated with persistence, self-discipline, and proficiency in problem-solving [17]. Research indicates that self-motivated children typically perform better academically and achieve higher grades in mathematics [16]. Motivation is regarded as a crucial factor that predicts a student's success in mathematics due to its significant and direct impact on learning and outcomes.

Self-confidence and assurance in academic performance are crucial for student success. "Self-confidence in learning" refers to the extent to which a student trusts in their own capability to do activities and achieve desired outcomes. It is a significant determinant of their academic performance. Yesuf et al. [41] assert that pupils with robust self-confidence are more inclined to persist in the face of challenges and have superior problem-solving abilities. Furthermore, self-confidence has been demonstrated to alleviate math anxiety, hence indirectly enhancing pupils' academic performance [reference Ref36]. Students who possess confidence in their learning capabilities are significantly more likely to excel in mathematics.

The responsibility of educators in mathematics instruction is beyond the simple delivery of knowledge, incorporating pedagogical efficacy and the fostering of significant student engagement. Effective pedagogy, proficient classroom management, and instructor competence significantly influence student performance [23]. Furthermore, institutional support and the implementation of empirically validated pedagogical strategies enhance the overall quality of instruction [28]. Studies indicate that positive feedback from teachers enhances students' attention, retention, and overall academic performance [12]. Consequently, educators play a crucial role in facilitating students' success in mathematics.

Another significant aspect in learning mathematics is the learning environment, which includes both home and classroom settings, and exerts a great influence on students' academic progress. An effective and supportive educational environment facilitates student learning and fosters the development of social skills, hence enhancing their engagement in the learning process [5]. Castro et al. [7] assert that familial support significantly contributes to children's mental and emotional development. Conversely, an unsupportive or detrimental home environment might impede students' academic success. A good learning environment that facilitates student learning is critical for improving math outcomes [34].

## 3 Theoretical foundations

This section provides a comprehensive overview of the theoretical and methodological underpinnings that inform our study. We detail the architecture and functionality of the Adaptive Neuro-Fuzzy Inference System (ANFIS) based on Takagi-Sugeno-Kang model (TSK), or simply the Sugeno fuzzy model [38], highlighting its role in integrating fuzzy logic with neural network learning. We then describe gradient-boosted decision trees with a focus on two prominent implementations- XGBoost and LightGBM- which are employed for their efficiency and predictive power.

### 3.1 The adaptive neuro-fuzzy inference system

The Adaptive Neuro-Fuzzy Inference System (ANFIS) is a hybrid computational framework that integrates the interpretability of fuzzy logic with the learning capabilities of neural networks. Widely employed for modeling complex, nonlinear systems under uncertainty, ANFIS serves as a powerful tool for applications ranging from predictive modeling and control systems to decision support. This section provides a comprehensive mathematical and conceptual description of the ANFIS architecture, its operational phases, and its inherent advantages.

ANFIS combines fuzzy logic and neural networks within a five-layer architecture, where each layer performs a specific function. This design allows the system to approximate nonlinear functions by learning from data while preserving the interpretability associated with fuzzy rule-based systems. The following discussion details the architecture and its mathematical formulation.

**Layer 1: Input Variables and Fuzzification** The input to ANFIS consists of $n$ features $x_1, x_2, \ldots, x_n$. Each input $x_j$ is transformed into a fuzzy set using membership functions (MFs) such as Gaussian, triangular, or trapezoidal functions. For the $i$-th rule and $j$-th input, the degree of membership is denoted as $\mu_{A_{ij}}(x_j)$. For example, the Gaussian MF is defined as

$$\mu_{A_{ij}}(x_j) = \exp\left(-\frac{(x_j - c_{ij})^2}{2\sigma_{ij}^2}\right), \tag{2}$$

where $c_{ij}$ is the center and $\sigma_{ij}$ is the spread. This layer outputs the membership degrees for all inputs and rules. Gaussian functions are often preferred for their smooth transitions and advantageous mathematical properties compared to triangular or trapezoidal functions.

**Layer 2: Rule Formulation and Firing Strengths** In this layer, fuzzy rules are formulated by combining the membership degrees from Layer 1. Each rule $i$ is expressed in the form:

$$\text{IF } x_1 \text{ is } A_{i1} \text{ AND } x_2 \text{ is } A_{i2} \text{ AND } \cdots \text{ AND } x_n \text{ is } A_{in}, \text{ THEN } y_i = f_i(\mathbf{x}), \tag{3}$$

where $f_i(\mathbf{x})$ represents the consequent of the $i$-th rule. The firing strength $w_i$ of each rule is computed as the product of the membership degrees:

$$w_i = \prod_{j=1}^{n} \mu_{A_{ij}}(x_j). \tag{4}$$

This product quantifies the degree to which the $i$-th rule is activated.

**Layer 3: Normalization of Firing Strengths** The firing strengths from Layer 2 are normalized to ensure that their sum equals one, thereby reflecting the relative influence of each rule in the final output. The normalized firing strength $\overline{w}_i$ is given by:

$$\overline{w}_i = \frac{w_i}{\sum_{k=1}^{R} w_k}, \tag{5}$$

where $R$ is the total number of rules. This normalization allows the model to operate as a weighted average of the rule outputs.

**Layer 4: Consequent Computation** In this layer, the consequent of each rule is computed as a function of the input variables. Traditionally, the consequent is modeled as a linear function:

$$y_i = p_{i0} + p_{i1}x_1 + p_{i2}x_2 + \cdots + p_{in}x_n, \tag{6}$$

where $p_{ij}$ are the consequent parameters. The output of this layer is the weighted consequent, $\overline{w}_i y_i$, with these parameters learned during training using gradient-based optimization methods.

**Layer 5: Aggregation and Output** The final output of the ANFIS is obtained by aggregating the weighted consequents from all rules:

$$y = \sum_{i=1}^{R} \overline{w}_i y_i. \tag{7}$$

This aggregation results in a crisp output that represents the system's prediction.

**Training ANFIS:** ANFIS utilizes a hybrid learning algorithm that combines gradient descent with least-squares estimation. The training process is divided into two phases:

**1.** Forward Pass: During the forward pass, the consequent parameters $p_{ij}$ are estimated using least-squares regression while keeping the premise parameters ($c_{ij}$ and $\sigma_{ij}$) fixed.

**2.** Backward Pass: In the backward pass, the premise parameters are updated using gradient descent to minimize the error between the predicted and actual outputs.

### 3.2 Gradient-boosted decision trees

Gradient-Boosted Decision Trees (GBDT) constitute a powerful ensemble learning methodology that constructs a predictive model by sequentially combining multiple weak learners, typically decision trees [15]. Each tree in the sequence is trained to correct the residual errors of the previously constructed ensemble, leading to progressive refinement of the predictive function and improved generalization performance. The model is expressed as

$$f_K(\mathbf{x}) = \sum_{k=1}^{K} \eta\, h_k(\mathbf{x}), \tag{8}$$

where $h_k(\mathbf{x})$ denotes the $k$-th weak learner, $\eta$ is the learning rate controlling its contribution, and $K$ is the total number of boosting iterations. This additive structure enables the model to approximate the target function by iteratively fitting trees to the negative gradient of the loss function.

The objective function optimized during training is defined as

$$\mathcal{L}(\phi) = \sum_{i=1}^{n} l(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(h_k), \tag{9}$$

where $l(y_i, \hat{y}_i)$ is a differentiable loss function quantifying the discrepancy between the true label $y_i$ and the prediction $\hat{y}_i$, and $\Omega(h_k)$ is a regularization term that controls the complexity of each tree. A commonly used regularization form is

$$\Omega(h) = \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} w_j^2, \tag{10}$$

where $T$ represents the number of leaves in the tree, $w_j$ the weight associated with leaf $j$, and $\gamma$ and $\lambda$ are regularization parameters.

At each boosting iteration, the objective function is efficiently optimized using a second-order Taylor approximation of the loss:

$$\mathcal{L}^{(t)} \approx \sum_{i=1}^{n} \left[ g_i\, h_t(\mathbf{x}_i) + \frac{1}{2} h_i\, h_t^2(\mathbf{x}_i) \right] + \Omega(h_t), \tag{11}$$

where $g_i = \partial_{\hat{y}^{(t-1)}} l(y_i, \hat{y}^{(t-1)})$ and $h_i = \partial^2_{\hat{y}^{(t-1)}} l(y_i, \hat{y}^{(t-1)})$ are the first- and second-order derivatives of the loss with respect to the model prediction. This formulation provides a principled optimization framework and is a key factor behind the efficiency of modern gradient boosting algorithms.

Contemporary research and application have led to highly optimized implementations of GBDT, notably XGBoost and LightGBM, which are frequently employed in predictive modeling.

**XGBoost.** XGBoost (Extreme Gradient Boosting) is a widely used and highly optimized implementation of GBDT, designed to achieve high accuracy, scalability, and robustness [9]. It employs advanced split-finding strategies, sparsity-aware algorithms for handling missing data, and regularization techniques to control model complexity. The corresponding predictive function is given by

$$f^{\mathrm{XGB}}(\mathbf{x}) = \sum_{k=1}^{K} \eta\, h_k(\mathbf{x}), \tag{12}$$

where each base learner $h_k$, is trained to minimize the second-order approximation of the objective function. The strong regularization framework of XGBoost contributes to its stability and generalization across diverse application domains.

**LightGBM.** LightGBM (Light Gradient Boosting Machine) is another efficient GBDT framework, developed to accelerate training and reduce memory consumption, particularly for large-scale datasets [21]. It introduces a histogram-based split-finding algorithm that discretizes continuous features into bins, thereby lowering computational complexity. Furthermore, it adopts a leaf-wise tree growth strategy, which prioritizes splits that yield the maximum loss reduction, allowing for faster convergence. The corresponding predictive function is defined as

$$f^{\mathrm{LGB}}(\mathbf{x}) = \sum_{k=1}^{K} \eta\, g_k(\mathbf{x}), \tag{13}$$

where $g_k(\mathbf{x})$ represents the base learners constructed using LightGBM's specialized tree-building mechanism.

# 4 Research methodology

This section elucidates the accuracy of our predictions and enhances our understanding of the elements affecting student success in advanced mathematics. We shall initially delineate our data collection methodology, encompassing the subjects of our study, the criteria for their selection, and the instruments employed for information gathering. We also assessed the reliability of our tools with Cronbach's alpha. Subsequently, we will provide a fundamental overview of the data, highlighting its principal statistical characteristics. Subsequently, we will elucidate our criteria for selecting the factors included in our analysis, employing a correlation matrix, SHAP plot, and XGBoost feature selection to justify our choice of these particular predictors. The primary focus of our research is the development of a unified prediction model. This model employs ANFIS in conjunction with sophisticated methods known as gradient-boosted decision trees, notably XGBoost and LightGBM. We integrated these methodologies to leverage the comprehensible characteristics of fuzzy logic alongside the robust predictive capabilities of ensemble learning. This integrated methodology aims to provide more precise forecasts and a deeper comprehension of the factors influencing student performance in advanced mathematics.

## 4.1 Data collection

The data for this study were collected from a cohort of 280 undergraduate students pursuing bachelor's degrees in Information Technology. The participants, in their second to fifth year of study, were intentionally selected based on certain criteria, including the satisfactory completion of final examinations in General Mathematics I, General Mathematics II, and Differential Equations. The study was conducted at the Islamic Azad University in Tehran Province, Iran.

The questionnaire instrument underwent rigorous content validation by a panel of three expert mathematicians to ensure clarity, relevance, and coverage of the targeted constructs. Data collection was carried out using a standardized questionnaire administered both in paper form and electronically via Google Forms to maximize participation and data completeness.

The final dataset comprises 29 distinct attributes designed to capture a broad spectrum of demographic, educational, psychological, and behavioral factors influencing student achievement in advanced mathematics. A comprehensive list of these variables and their corresponding notations is presented in Table 1. To provide a more concrete view of the data structure and the nature of the collected variables, Table 2 displays a representative fragment of the dataset, including attribute names, symbols, and observed categories or numerical value intervals. This table reflects the actual distribution of categorical and numerical features before model development, thereby supporting transparency and reproducibility of the study.

To ascertain the psychometric reliability of the questionnaire, Cronbach's alpha was employed, with a pre-established acceptance threshold of 0.7. The reliability analysis yielded the following Cronbach's alpha coefficients for each factor: Educational Background and Preparation (0.737), Learning Engagement and Study Habits (0.713), Instructor Influence and Teaching Effectiveness (0.776), Learning Resources and Support Systems (0.736), Attitude and Psychological Factors (0.822), and Financial and Psychological Barriers (0.735). These findings affirm the questionnaire's robust reliability as an instrument for assessing the specified constructs, thus substantiating its suitability for subsequent research.

This study investigates students' comprehensive academic performance in mathematics, specifically focusing on three core undergraduate courses: General Mathematics I (`FG_GMath_I`), General Mathematics II (`FG_GMath_II`), and Differential Equations (`FG_Dif`). Student achievement in each course is evaluated based on two primary metrics: the final grade, recorded on a scale ranging from 10 to 20 inclusive (with 10 representing the minimum passing threshold), and the number of attempts required to achieve a passing grade in the final examination. The latter is represented by variables `NA_GMath_I`, `NA_GMath_II`, and `NA_Dif`, each denoting the count of examination attempts undertaken by a student before successfully achieving a passing score. For instance, successful completion on the third attempt is recorded as an attempt count of three, irrespective of the final grade achieved above the passing criterion. To effectively account for the impact of repeated attempts on overall performance, course instructors assigned a composite performance score, ranging from 0 to 100, for each course. This composite score integrated both the final grade attained and the number of attempts required. This methodological approach ensures that students achieving a passing grade on their initial attempt are differentiated in their performance evaluation from those attaining a comparable final grade after multiple attempts. Given the equivalent academic weighting of all three courses, an overall mathematics performance metric was calculated as the arithmetic mean of the individual course performance scores. This aggregate measure provides a holistic and nuanced assessment of student achievement across these foundational mathematics courses.

Table 1: Model features and variables

| Class of Attribute | Symbol | Description |
|---|---|---|
| | *Age* | Age of student at enrollment. |
| | *Gender* | Male (148), Female(132) |
| | *M_Edu* | Educational level of the student's mother. |
| Personal and Demographic data | *F_Edu* | Educational level of the student's father. |
| | *Fam_STEM* | Indicates if the student comes from a STEM-oriented family. |
| | *Job_Field* | Employment status in a field related to the student's area of study. |
| | *Field_Hs* | Student's high school major. |
| | *Math_Per_Hs* | Performance in mathematics-related subjects during high school. |
| Educational Background and Preparation | *Math_Edu_Qual* | Quality of mathematics education received prior to university. |
| | *Know_Math_App* | Awareness of how mathematics is used in the student's field of study. |
| | *Attend_Math* | Frequency of attendance in mathematics classes. |
| | *Study_Time* | Amount of time spent studying mathematics during a week. |
| Learning Engagement and Study Habits | *Attempt_Correct* | Tendency to revisit and correct mistakes in assignments. |
| | *Ask_Class* | Frequency with which a student asks questions during math classes. |
| | *Teach_Rating* | Evaluation of the teacher's effectiveness in explaining mathematical concepts. |
| Instructor Influence and Teaching Effectiveness | *Teach_Encourage* | Level of encouragement and support provided by teachers. |
| | *Edu_Resources* | Availability of resources such as libraries, labs, and math software. |
| | *Supp_Learnres* | Access to additional support like tutoring or extra classes. |
| Learning Resources and Support Systems | *Online_Res* | Utilization of online platforms for mathematics learning. |
| | *Use_AI* | Utilization of artificial intelligence tools for solving math problems. |
| | *Love_Math* | Level of enthusiasm or interest in mathematics. |
| | *Selfconf_Math* | Confidence in one's ability to solve mathematical problems. |
| Attitude and Psychological Factors | *Motiv_Math* | Drive to improve problem-solving skills in mathematics. |
| | *Fam_Encourage* | Level of encouragement and support provided by family. |
| | *Improve_Perf* | Indicator of performance decline over time. |
| | *Fin_Issues* | Presence of financial challenges that may impact academic performance. |
| | *Stress_Fees* | Level of stress related to tuition fee payments. |
| Financial & Psychological Barriers | *Anxiety_Prob* | Degree of anxiety when facing challenging math problems. |
| | *Stress_Exam* | Degree of stress experienced during mathematics examinations. |
| Target | *Total_Performance* | Define in Section 4.1 . |

Table 2: Categories and value intervals of variables used in the model

| **Class of Attribute** | **Symbol** | **Values / Categories** |
|---|---|---|
| | *Age* | Numeric range:18-32 |
| | *Gender* | Female, Male |
| | *M_Edu* | Below high school diploma, High school diploma, Bachelor's degree, Master's degree and above |
| Personal and Demographic Data | *F_Edu* | Below high school diploma, High school diploma, Bachelor's degree, Master's degree and above |
| | *Fam_STEM* | No, Yes |
| | *Job_Field* | Not employed, Job related to field, Job unrelated to field |
| | *Field_Hs* | None,Humanity, Science, Computer, Math |
| | *Math_Per_Hs* | Very Poor, Poor, Average, Good, Excellent |
| Educational Background and Preparation | *Math_Edu_Qual* | Very Poor, Poor, Average, Good, Excellent |
| | *Know_Math_App* | None, Low, Moderate, High, Very High |
| | *Math_App_Fos* | Very Low, Low, Moderate, High, Critical |
| | *Attend_Math* | None, Rarely, Sometimes, Often, Always |
| | *Study_Time* | None, Less than 2 Hours, 2-4 Hours, 5-8 Hours, More than 8 hours |
| Learning Engagement and Study Habits | *Attempt_Correct* | Never, Rarely, Sometimes, Often, Always |
| | *Ask_Class* | Never ask, prefer self-study., only when confused., Sometimes ask, mostly solve myself., Ask Often ask to confirm understanding., Always ask, never hesitate. |
| | *Procrast_Math* | Never, Rarely, Sometimes, Often, I don't do the assignments. |
| | *Teach_Rating* | Very Poor, Poor, Average, High, Very High |
| Instructor Influence and Teaching Effectiveness | *Teach_Encourage* | Very Low, Low, Moderate, High, Very High |
| | *Edu_Resources* | Insufficient access to resources., Limited access with occasional difficulties., Resources are mostly available but sometimes limited, Fully accessible and easy to use. |
| | *Supp_Learnres* | No, Yes |
| Learning Resources and Support Systems | *Online_Res* | Never, Low, Moderate, High, Critical |
| | *Use_AI* | No, Yes |
| | *Love_Math* | Not Interested at All, Not Very Interested, Neutral, Interested, Very Interested |
| | *Selfconf_Math* | Not confident, Low, Acceptable, Good, High |
| | *Motiv_Math* | None, Low, Moderate, High, Critical |
| | *Fam_Encourage* | None, Low, Moderate, High, Very High |
| Attitude and Psychological Factors | *Encourag_Impact* | None, Low, Moderate, High, Critical |
| | *Improve_Perf* | Performance decline, No noticeable improvement, Slight improvement, Moderate improvement, Significant improvement |
| | *Fin_Issues* | None, Low, Moderate, High, Critical |
| | *Stress_Fees* | Never, Rarely, Sometimes, Often, Always |
| Financial & Psychological Barriers | *Anxiety_Prob* | Never, Rarely, Sometimes, Often, Always |
| | *Stress_Exam* | None, Low, Moderate, Very High |
| | *Grade_Truth* | No Answer., Not Reflective., Somewhat Reflective., Mostly Reflective., Completely Reflective. |
| | *FG_GMath_I* | Numeric range:10-20 |
| | *NA_GMath_I* | Numeric range:1-4 |
| | *FG_GMath_II* | Numeric range:10-20 |
| Mathematics Performance Variables (Numerical) | *NA_GMath_II* | Numeric range:1-4 |
| | *FG_DIF* | Numeric range:10-20 |
| | *NA_DIF* | Numeric range:1-3 |
| Target | *Total_Performance* | Numeric range: 20-95 |

## 4.2 Statistical characterization of dataset

In this section, we present an analysis of descriptive statistics encompassing two key aspects of student experience. Firstly, we examine a summary of descriptive statistics that elucidates student perceptions regarding various factors believed to influence their mathematics performance. Secondly, we investigate descriptive statistics pertaining to student achievement specifically in three foundational mathematics courses: General Mathematics I, General Mathematics II, and Differential Equations.

A detailed summary of the descriptive statistics concerning student perceptions of performance-influencing factors is provided in Table 3. This table presents metrics such as the mean, standard deviation (SD), skewness, and kurtosis, offering valuable insights into both the central tendency and the distributional characteristics of student opinions across the different dimensions examined.

For Educational Background and Preparation, the mean score of 2.80 (SD = 0.93) suggests that, on average, students perceive their prior educational experiences and preparedness for advanced mathematics as moderately adequate, neither exceptionally strong nor significantly weak. The slight negative skewness (-0.359) indicates a tendency for slightly more students to rate this factor as above average, suggesting a somewhat positive self-assessment in terms of preparation. The kurtosis value of -0.358, being negative, implies a flatter distribution than a normal distribution, indicating less concentration around the mean and fewer extreme responses. This distribution suggests that while most students cluster around a moderate perception of their preparation, a segment feels well-prepared, possibly due to previous rigorous academic experiences.

Regarding Learning Engagement and Study Habits, the lower mean score of 2.01 (SD = 1.04) points to a general perception of moderate learning engagement and somewhat inconsistent study practices among the student cohort. The skewness of -0.123, close to zero, reveals a relatively symmetrical distribution, suggesting that opinions are fairly balanced around the average level of engagement. However, the negative kurtosis (-0.393), again indicating a flatter distribution, suggests a considerable diversity in students' self-reported engagement and study habits. This variability likely reflects individual differences in learning styles and the capacity to effectively balance academic demands with personal commitments, as highlighted by Nguyen [27].

Concerning Instructor Influence and Teaching Effectiveness, the mean score of 2.03 (SD = 1.12) is also moderate. The skewness of -0.122, near zero, signifies a balanced distribution of student ratings, indicating no strong bias towards overly positive or negative instructor evaluations across the sample. The slightly negative kurtosis (-0.497) further supports the notion that extreme viewpoints are not prevalent. This even distribution might reflect a range of instructional experiences, where some students benefit significantly from effective teaching, while others may not consistently receive the same level of instructional support.

Learning Resources & Support Systems exhibits the lowest mean score at 1.38 (SD = 0.76). The skewness of 0.056, close to zero, again points to a symmetrical distribution of responses. The notably low kurtosis (-0.840), however, suggests a strong consensus among students regarding the perceived inadequacy of available learning resources and support. This collective perception highlights a potential area of concern regarding the sufficiency of academic support systems, improvements to which could positively impact student performance in mathematics [10].

For Attitude & Psychological Factors, the higher mean score of 2.79 (SD = 1.05) indicates generally positive student perceptions in this domain. The substantial negative skewness (-0.765) suggests that a larger proportion of students hold more positive attitudes and psychological profiles concerning mathematics learning. The kurtosis of 0.162, close to zero, indicates an approximately normal distribution, with responses clustered around the mean and fewer extreme opinions. This pattern reinforces the idea that students generally maintain a reasonably positive psychological outlook towards their mathematical abilities, though individual experiences may vary.

Finally, Financial and Psychological Barriers presents a mean score of 2.57 (SD = 1.25), suggesting that students experience moderate levels of challenges in this area. The slight negative skewness (-0.242), combined with a distinctly negative kurtosis (-1.226), emphasizes considerable variability in individual experiences. This variability indicates that while some students face significant financial and psychological obstacles, others encounter fewer such impediments. This variation underscores the need for individualized and context-sensitive support interventions tailored to address the diverse needs of the student population [14]

In summary, these descriptive statistics reveal a diverse landscape of student perceptions. The relatively low evaluations for learning resources and support systems signal a critical area requiring attention, while the marked heterogeneity in financial and psychological barriers suggests that support mechanisms should be personalized to accommodate the varied needs across student demographics. These findings are crucial for informing targeted strategies aimed at enhancing the overall academic experiences and mathematical performance of students in higher education.

Table 4 presents descriptive statistics characterizing student achievement across three mathematics courses: General Mathematics I, General Mathematics II, and Differential Equations. Performance scores, ranging from 0 to 100, are

Table 3: Descriptive statistics of student perceptions of factors related to math performance

| Factors | Mean | SD | Skewness | | Kurtosis | |
|---|---|---|---|---|---|---|
| | | | Statistics | Std. Error | Statistics | Std. Error |
| Educational Background & Preparation | 2.80 | 0.93 | -0.359 | 0.146 | -0.358 | 0.292 |
| Learning Engagement & Study Habits | 2.01 | 1.04 | -0.123 | 0.146 | -0.393 | 0.292 |
| Instructor Influence & Teaching Effectiveness | 2.03 | 1.12 | -0.122 | 0.146 | -0.497 | 0.292 |
| Learning Resources & Support Systems | 1.38 | 0.76 | 0.056 | 0.146 | -0.840 | 0.292 |
| Attitude & Psychological Factors | 2.79 | 1.05 | -0.765 | 0.146 | 0.162 | 0.292 |
| Financial & Psychological Barriers | 2.57 | 1.25 | -0.242 | 0.146 | -1.226 | 0.292 |

Table 4: Descriptive statistics of student perceptions of factors related to three mathematics courses

| Factors | Mean | SD | Skewness | | Kurtosis | |
|---|---|---|---|---|---|---|
| | | | Statistics | Std. Error | Statistics | Std. Error |
| Performance General Mathematics I | 56.17 | 22.37 | 0.381 | 0.146 | -0.943 | 0.292 |
| Performance General Mathematics II | 55.91 | 24.52 | 0.421 | 0.146 | -1.180 | 0.292 |
| Performance Differential Equations | 54.52 | 21.01 | 0.328 | 0.146 | -0.798 | 0.292 |

analyzed. The mean performance scores indicate a moderate level of overall achievement, with averages of 56.17 for General Mathematics I, 55.91 for General Mathematics II, and 54.52 for Differential Equations. These mean values suggest that, on average, student performance is slightly above the midpoint of the scoring range, yet remains within the moderate performance spectrum.

The table presents the descriptive statistics for student performance across three mathematics courses: General Mathematics I, General Mathematics II, and Differential Equations. The data highlights key aspects of student achievement, including the mean, standard deviation (SD), skewness, and kurtosis for each course.

The average performance in the three courses is relatively moderate, with General Mathematics I having a mean of 56.17, General Mathematics II 55.91, and Differential Equations 54.52. These values suggest that students, on average, performed slightly above the midpoint of the possible score range, but still within the moderate performance spectrum.

The standard deviation values indicate substantial variability in student performance within each course. The SD is highest for General Mathematics II (24.52), followed by General Mathematics I (22.37), and Differential Equations (21.01), signaling a broad range of student achievement in all three courses.

All three courses exhibit positive skewness, with values of 0.381 for General Mathematics I, 0.421 for General Mathematics II, and 0.328 for Differential Equations. This indicates that there are a smaller number of students achieving exceptionally high scores, pulling the mean slightly upward.

The negative kurtosis values across all courses ($-0.943$ for General Mathematics I, $-1.180$ for General Mathematics II, and $-0.798$ for Differential Equations) suggest a platykurtic distribution. This means that the performance distributions are flatter, with fewer extreme values compared to a normal distribution.

In conclusion, these findings suggest that while student performance is generally moderate, there is considerable variation within each course. The positive skewness highlights the presence of high-achieving students, and the negative kurtosis indicates that extreme performance values are rare.

## 4.3 Feature selection and justification

A primary issue associated with the implementation of ANFIS is the exponential increase in fuzzy rules as the number of input variables rises. To address this issue and improve model interpretability and computational efficiency, a feature selection approach is considered necessary. By systematically narrowing the input feature space to a concise set of the most relevant predictors, the complexity of the rule base and, consequently, the computing load is significantly reduced. This feature selection process sought to enhance the model's generalization ability, hence diminishing the risk of overfitting and improving forecast accuracy. The amalgamation of a data-driven feature selection methodology with the fuzzy inference system enabled the creation of a hybrid strategy that adeptly reconciles the frequently conflicting goals of model interpretability and predictive accuracy.

Before developing the prediction model, we perform an initial exploratory data analysis using a correlation matrix. This analysis aims to identify and measure the correlations between the different candidate features and the specified goal variable, `Total_Performance`. The computed correlation coefficients clarified the linear relationships among variables, facilitating the distinction between features of significant importance to the outcome variable and those with relatively lower influence. Features exhibiting strong associations with `Total_Performance` were prioritized for model inclusion. In contrast, traits demonstrating weak or maybe redundant links were deemed suitable for elimination or change. The resulting feature correlation matrix is illustrated in Figure 1, with direct (positive) relationships shown in red and inverse (negative) relationships represented in blue. The correlation analysis indicates that the variables `Math_Edu_Qual`, `FG_GMath_I`, `FG_GMath_II`, and `FG_DIF`, which represent students' academic grades in mathematics courses, are the dependent variables with the most robust correlational relationships with `Total_Performance`.

For instance, the feature `Math_Edu_Qual`, denoting the quality of pre-university mathematical education, demonstrates a steady trend in the SHAP Figure 2. Increased values (denoted by yellow points) are predominantly situated on the right side of the zero axis, signifying a strong correlation between advanced mathematics educational degrees and improved expected outcomes. Conversely, lower values (shown by blue dots) are clustered on the left side, implying that worse credentials correlated with less anticipated performance. The bidirectional effect highlighted the critical significance of `Math_Edu_Qual` in the predictive model.

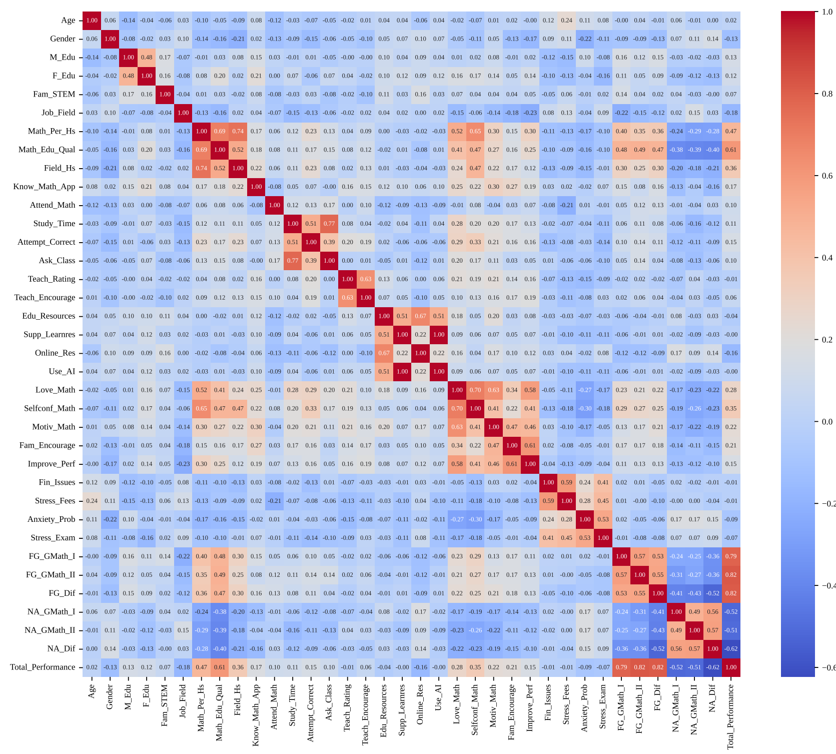



Figure 1: Correlation matrix between different features.

Subsequent analysis, as illustrated in the SHAP Figure 2, indicated that elevated values of variables associated with familial support, self-efficacy in mathematical problem-solving, motivation to enhance problem-solving abilities, and recognition of the practical applications of mathematics in their major all positively influenced the prediction. In contrast, diminished values of these traits produce negative SHAP contributions, hence reducing the overall anticipated outcome.

Furthermore, the SHAP analysis revealed that a greater degree of education achieved by the student's mother positively influenced the anticipated performance. Significantly, a father's low educational attainment was occasionally linked to positive SHAP values, indicating that the influence of paternal education on performance may demonstrate a non-linear relationship or be affected by conditional factors, highlighting the complexity of its effects.

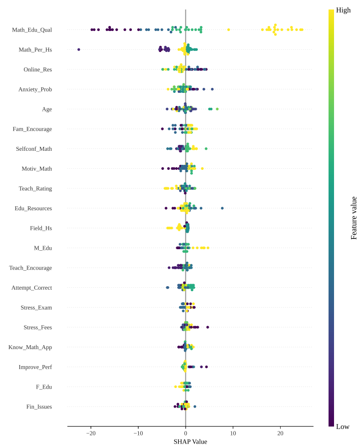



Figure 2: SHAP summary plot of feature importance for factors influencing mathematics performance.

Also, as illustrated in SHAP Figure 3 and corroborated by our overall performance calculations, elevated final test results in General Mathematics I, General Mathematics II, and Differential Equations corresponded with enhanced expected performance, as indicated by positive SHAP values. In contrast, diminished exam scores decreased the anticipated performance, as indicated by negative SHAP values. Moreover, repeated tries at these courses correlated with negative SHAP scores, signifying that frequent attempts detrimentally influenced the total expected outcome.

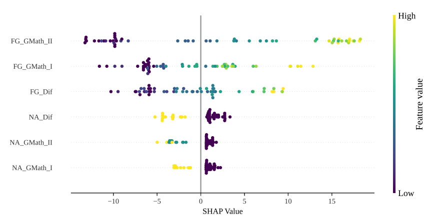



Figure 3: SHAP summary plot of feature importance for course grades in core mathematics

To augment our comprehension of the most critical attributes, we utilize XGBoost's intrinsic feature importance analysis [9] with SHAP analysis. XGBoost feature importance provides a thorough evaluation of each feature's contribution to the model's predicted accuracy for all students. This procedure entails examining the frequency of feature usage to improve the model's precision. This facilitates the recognition of the most salient traits and their differentiation from those of lesser significance. Eliminating characteristics with minimal XGBoost significance is deemed to streamline the model while preserving predictive accuracy. XGBoost's capacity to understand intricate linkages guarantees the consideration of feature interactions, which is beneficial when combined with the fuzzy logic of ANFIS.

XGBoost identifies the most critical traits overall, as depicted in Figure 4. Both approaches, XGBoost and SHAP, identify `Math_Edu_Qual` (the quality of pre-university mathematics education) as a significant predictor. XGBoost demonstrates greater significance than features such as `Supp_Learneres` (learning support resources) and `Anxity_Prob` (math anxiety). This illustrates the vital significance of `Math_Edu_Qual` in the model's decision-making process. Consequently, these methodologies for decision-making regarding feature selection and the simplification of fuzzy rules in ANFIS enhances model interpretability and predictive performance.

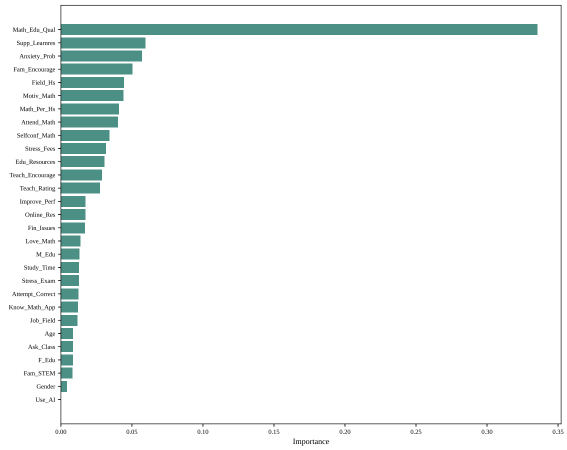



Figure 4: Feature importances for total performance.

## 4.4 Hybrid model implementation

The proposed hybrid neuro-fuzzy-boosting framework is designed to combine the interpretability of fuzzy inference systems with the high predictive power of gradient boosting algorithms, specifically XGBoost and LightGBM. The central idea is to partition the input space into linguistically interpretable fuzzy regions-such as "Very Poor," "Poor," "Medium," "Good," and "Excellent"- and assign a dedicated local gradient boosting model to each fuzzy region. This enables the model to retain the transparent, rule-based reasoning characteristic of fuzzy systems, while leveraging the strong nonlinear approximation capabilities of ensemble learning. As a result, the hybrid approach can capture intricate patterns in data without sacrificing interpretability, making it well-suited for decision support in educational performance assessment.

The implementation begins with data preprocessing and feature selection, as outlined in Section 4.3. Let $N_{\text{in}}$ denote the number of selected input variables used to predict the target variable `Total_Performance`. Categorical features are encoded using the `OrdinalEncoder`, while numerical features are transformed into fuzzy representations through Gaussian membership functions. This fuzzification step converts crisp input values into degrees of membership to linguistic categories, allowing the system to operate in an interpretable manner. For each input variable $x_j$, a Gaussian membership function is defined as

$$\mu_{A_{ij}}(x_j) = \exp\left(-\frac{(x_j - c_{ij})^2}{2\sigma_{ij}^2}\right), \tag{14}$$

where $c_{ij}$ and $\sigma_{ij}$ represent the center and spread of the fuzzy set associated with the $i$-th linguistic label.

Once the fuzzy sets are established, they are combined to construct a fuzzy rule base that partitions the input space into interpretable subregions. A standard rule, for instance, rule (i), for a collection of inputs $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ is structured as follows:

$$\text{Rule}(i) : x_1 \text{ is } A_{i1} \text{ AND } x_2 \text{ is } A_{i2} \text{ AND } \cdots \text{ AND } x_n \text{ is } A_{in} \text{ THEN } \texttt{Total\_Performance} = f_i(\mathbf{x}),$$

where $f_i(\mathbf{x})$ denotes the output of the local gradient boosting model associated with the $i$-th rule.

In a Sugeno-type ANFIS structure, the number of hidden neurons corresponds directly to the number of fuzzy rules, since each rule generates one node in the rule layer. The total number of fuzzy rules $R$ can be expressed as the Cartesian product of the linguistic label sets associated with the selected input features. In this study, the linguistic categories and value ranges used to construct these fuzzy rules are presented in Table 2, which provides a representative fragment of the dataset, including the attribute names and their corresponding linguistic partitions.

For each input vector $\mathbf{x}$, the firing strength of the $i$-th rule is calculated as the product of membership degrees across all input dimensions:

$$w_i = \prod_{j=1}^{n} \mu_{A_{ij}}(x_j). \tag{15}$$

The firing strengths are then normalized to ensure a fair contribution of all activated rules:

$$\overline{w}_i = \frac{w_i}{\sum_{k=1}^{R} w_k}. \tag{16}$$

This weighting mechanism ensures that each fuzzy rule contributes proportionally to the final prediction according to the degree of rule activation, enabling smooth transitions between different fuzzy regions.

Unlike classical ANFIS, which employs simple linear functions in the consequent layer (Section 3.1), the proposed hybrid framework integrates nonlinear predictive models in the form of XGBoost and LightGBM, as discussed in Section 3.2. The predictive functions for these boosting algorithms are given by

$$f^{\mathrm{XGB}}(\mathbf{x}) = \sum_{k=1}^{K} \eta \, h_k(\mathbf{x}), \qquad f^{\mathrm{LGB}}(\mathbf{x}) = \sum_{k=1}^{K} \eta \, g_k(\mathbf{x}), \tag{17}$$

where $h_k$ and $g_k$ denote the base learners, $\eta$ is the learning rate, and $K$ is the number of boosting iterations. Each fuzzy rule is associated with an independent local gradient boosting model $f_i$, which is trained on the full dataset $D = (\mathbf{x}^{(m)}, y^{(m)})_{m=1}^{M}$. During training, each data sample is weighted according to its degree of membership to the corresponding rule:

$$\mathcal{L}_i^{\mathrm{XGB}} = \sum_{m=1}^{M} \overline{w}_i^{(m)} \cdot \left( y^{(m)} - f_i^{\mathrm{XGB}}(\mathbf{x}^{(m)}) \right)^2. \qquad \mathcal{L}_i^{\mathrm{LGB}} = \sum_{m=1}^{M} \overline{w}_i^{(m)} \cdot \left( y^{(m)} - f_i^{\mathrm{LGB}}(\mathbf{x}^{(m)}) \right)^2. \tag{18}$$

This membership-weighted training strategy allows each local model to specialize in its corresponding subregion of the input space, thereby enhancing local predictive accuracy without requiring global model complexity.

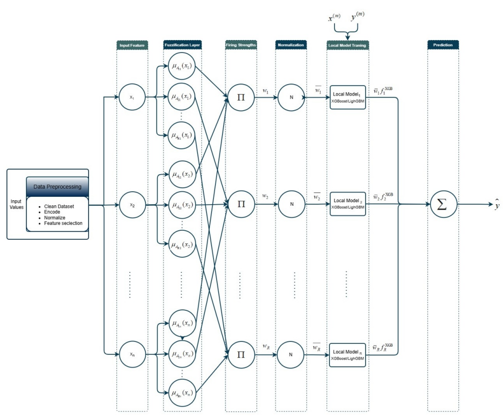



Figure 5: Architecture of the hybrid ANFIS-XGBoost framework.

During inference, a new input vector $\mathbf{x}$ is first fuzzified, and the firing strengths of all fuzzy rules are computed. Each local gradient boosting model $f_i(\mathbf{x})$ generates a prediction corresponding to its fuzzy region. These predictions are then aggregated using fuzzy-weighted averaging:

$$\hat{y} = \sum_{i=1}^{R} \overline{w}_i, f_i(\mathbf{x}). \tag{19}$$

This aggregation mechanism ensures that the final output is a smooth, adaptive combination of local models, capturing both local detail and global structure in the input-output relationship.

Finally, the continuous output $\hat{y}$ is mapped into predefined performance categories to provide interpretable results for practical use. Specifically, the predicted value is assigned to one of the intervals ( 0-35 ), ( 35-50 ), ( 50-70 ), ( 70-85 ), and ( 85-100 ), corresponding to the qualitative labels "Very Poor," "Poor," "Medium," "Good," and "Excellent," respectively. For instance, if $\hat{y} = 76$, the model outputs the label "Good" for `Total_Performance`. This final step transforms quantitative predictions into qualitative assessments, making the output more actionable for educational stakeholders and decision-makers. The architecture is depicted in Figure 5, and the complete training process is provided in Algorithm 1.

**Algorithm 1** Hybrid ANFIS-XGBoost for Predicting Student Performance

**Require:** Dataset with features $x_j$ for $j = 1, \ldots, n$ and target variable $y$; fuzzy parameters $\{c_{ij}, \sigma_{ij}\}$; XGBoost hyperparameters (e.g., learning rate, number of rounds, etc.)

**Ensure:** Predicted student performance $\hat{y}$

Step 1: **Data Preprocessing:** Clean the dataset, encode categorical features (using `OrdinalEncoder`), normalize numerical variables, and perform feature selection.

Step 2: **Fuzzification, Firing Strengths and Normalization:**

**for** $j \leftarrow 1$ **to** $n$ **do**
  **for** $i \leftarrow 1$ **to** $R$ **do**
    Transform $x_j$ into a fuzzy variable via the fuzzy membership functions, $\mu_{A_{ij}}(x_j)$.
  **end for**
**end for**
**for** $i \leftarrow 1$ **to** $R$ **do**
  Formulate the fuzzy rule:

$$\text{Rule } (i): \quad \text{IF } x_1 \text{ is } A_{i1} \text{ AND } x_2 \text{ is } A_{i2} \text{ AND } \cdots \text{ AND } x_n \text{ is } A_{in}, \text{ THEN } y_i.$$

**end for**
**for** $i \leftarrow 1$ **to** $R$ **do**
  Compute the firing strength: $w_i = \prod_{j=1}^{n} \mu_{A_{ij}}(x_j)$.
**end for**
**for** $i \leftarrow 1$ **to** $R$ **do**
  Normalize the firing strengths: $\overline{w}_i = \frac{w_i}{\sum_{k=1}^{R} w_k}$.
**end for**

Step 3: **Local Model Training**

**for** $i \leftarrow 1$ **to** $R$ **do**
  Initialize an empty weighted dataset $D_i$.
  **for** $m \leftarrow 1$ **to** $M$ **do**
    Add sample $(\mathbf{x}^{(m)}, y^{(m)})$ to $D_i$ with weight $\overline{w}_i^{(m)}$.
  **end for**
  Train the XGBoost model $f_i^{\text{XGB}}$ on $D_i$, minimizing the weighted loss:

$$\mathcal{L}_i^{\text{XGB}} = \sum_{m=1}^{M} \overline{w}_i^{(m)} \Big( y^{(m)} - f_i^{\text{XGB}}(\mathbf{x}^{(m)}) \Big)^2.$$

**end for**

Step 4: **Inference and Aggregation:**

- For a new input $\mathbf{x}$, repeat the fuzzification, firing Strengths and normalization steps to compute the normalized weights $\overline{w}_i$ for $\mathbf{x}$.
- Compute the final prediction using the ensemble of pre-trained models,

$$\hat{y} = \sum_{i=1}^{R} \overline{w}_i \, f_i^{\text{XGB}}(\mathbf{x}).$$

Step 5: **Categorical Mapping:**

- **If** $\hat{y} < 35$, **then** assign the label "Very Poor".
- **Else if** $\hat{y} < 50$, **then** assign the label "Poor".
- **Else if** $\hat{y} < 70$, **then** assign the label "Medium".
- **Else if** $\hat{y} < 85$, **then** assign the label "Good".
- **Else then** assign the label "Excellent".

Output: Predicted student performance $\hat{y}$ along with its corresponding categorical label.

# 5 Experimental results and comparative analysis

This section presents a detailed look at how well our proposed methods performed. We will show how they are better than the standard ANFIS method and explain the reasons why we developed these new approaches.

## 5.1 Evaluation metrics and benchmarking

To thoroughly evaluate the predictive capabilities of our models about student performance, we utilize a variety of standard assessment criteria designed for both regression and classification tasks. In regression, which entails forecasting continuous performance scores, we employ Mean Squared Error (MSE), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and R-squared ($R^2$) to quantitatively assess the deviations between anticipated and actual values. These metrics offer a comprehensive framework for assessing the accuracy and dependability of our algorithms in predicting continuous events. Furthermore, having classified the continuous performance scores into specific categories - namely "Very Poor," "Poor," "Medium," "Good," and "Excellent"- we utilize a confusion matrix to assess the classification accuracy of the models. This matrix functions as an essential diagnostic instrument by delineating the distribution of predicted and actual class labels. This section details these measurements and their specific functions in our complete evaluation strategy.

**Mean Squared Error (MSE).** The Mean Squared Error measures the average of the squared discrepancies between the observed values $y^{(m)}$ and the model's predictions $\hat{y}^{(m)}$. The Mean Squared Error (MSE) is formally defined as:

$$\text{MSE} = \frac{1}{M}\sum_{m=1}^{M}(y^{(m)} - \hat{y}^{(m)})^2. \tag{20}$$

A reduced MSE signifies that the model's predictions are, on average, nearer to the actual values.

**Mean Absolute Error (MAE).** The Mean Absolute Error quantifies the average absolute deviation between actual and anticipated values. The Mean Absolute Error (MAE) is defined as follows:

$$\text{MAE} = \frac{1}{M}\sum_{m=1}^{M}\left|y^{(m)} - \hat{y}^{(m)}\right|. \tag{21}$$

Similar to MSE, a reduced MAE score indicates superior predictive performance, as it signifies diminished discrepancies between forecasts and actual observations.

**Root Mean Squared Error (RMSE).** The Root Mean Squared Error is defined as the square root of the Mean Squared Error:

$$\text{RMSE} = \sqrt{\text{MSE}}. \tag{22}$$

RMSE is advantageous since it is articulated in the same units as the target variable, hence enhancing the intuitive comprehension of prediction mistakes.

**Coefficient of Determination** ($R^2$). The $R^2$ measure signifies the fraction of variation in the dependent variable elucidated by the model. It is computed as:

$$R^2 = 1 - \frac{\sum_{m=1}^{M}(y^{(m)} - \hat{y}^{(m)})^2}{\sum_{m=1}^{M}(y^{(m)} - \bar{y})^2}. \tag{23}$$

where $\bar{y}$ represents the average of the observed values. An $R^2$ number approaching 1 indicates that a significant percentage of the variance is accounted for by the model, with values exceeding 0.70 often deemed acceptable for strong predictive efficacy.

In the aforementioned formulations, $y^{(m)}$ and $\hat{y}^{(m)}$ denote the actual and predicted values, respectively, whereas $M$ signifies the total number of observations.

These measures collectively provide a thorough assessment of the model's capacity to reliably predict student performance in both continuous outcomes and category classifications.

## 5.2 Results of hybrid ANFIS-XGBoost and ANFIS-LightGBM models

This section presents the development and evaluation of two hybrid prediction models - Hybrid ANFIS-XGBoost and Hybrid ANFIS-LightGBM - for predicting university students' performance in advanced mathematics courses. The proposed methodology, outlined in Algorithm 1, integrates the adaptive neuro-fuzzy inference system (ANFIS) with advanced gradient boosting techniques to effectively model complex, non-linear relationships between multiple predictive factors.

The comparative performance of the proposed models against the baseline ANFIS and Random Forest is summarized in Tables 5 and 6. Both hybrid models demonstrated a substantial improvement over the standard ANFIS framework. Specifically, Hybrid ANFIS-XGBoost achieved a reduction in RMSE from 3.7566 to 1.9194 and improved $R^2$ from 0.9615 to 0.9853. Similarly, Hybrid ANFIS-LightGBM reduced RMSE from 5.1716 to 2.3835, with an $R^2$ increase from 0.9471 to 0.9888. These gains confirm the enhanced capacity of the hybrid frameworks to capture and generalize intricate patterns in the data.

Additionally, Random Forest yielded the highest $R^2$ score of 0.9924, serving as a strong benchmark. However, the hybrid ANFIS models offer superior interpretability through their fuzzy rule-based structure, which is critical for explaining decision boundaries in educational settings. The deep tree structure of LightGBM (`max_depth = 8`) effectively captures nuanced dependencies between fuzzy rules and input features, while the shallower XGBoost configuration (`max_depth = 5`) with strategic subsampling (`subsample = 0.8`, `colsample_bytree = 0.8`) mitigates overfitting and enhances generalization. Moreover, using 200 boosting iterations in both models strengthened the stability of gradient boosting across fuzzy partitions.

Table 5: Performance comparison of ANFIS, hybrid ANFIS-XGBoost, and random forest.

| Factors | ANFIS | Hybrid ANFIS-XGBoost | Random Forest |
|---|---|---|---|
| MAE | 2.808167 | 1.735601 | 1.239854 |
| MSE | 14.112208 | 3.379612 | 2.761541 |
| RMSE | 3.756622 | 1.919399 | 1.661788 |
| $R^2$ | 0.961541 | 0.985339 | 0.992474 |

Table 6: Performance comparison of ANFIS, hybrid ANFIS-LightGBM, and random forest.

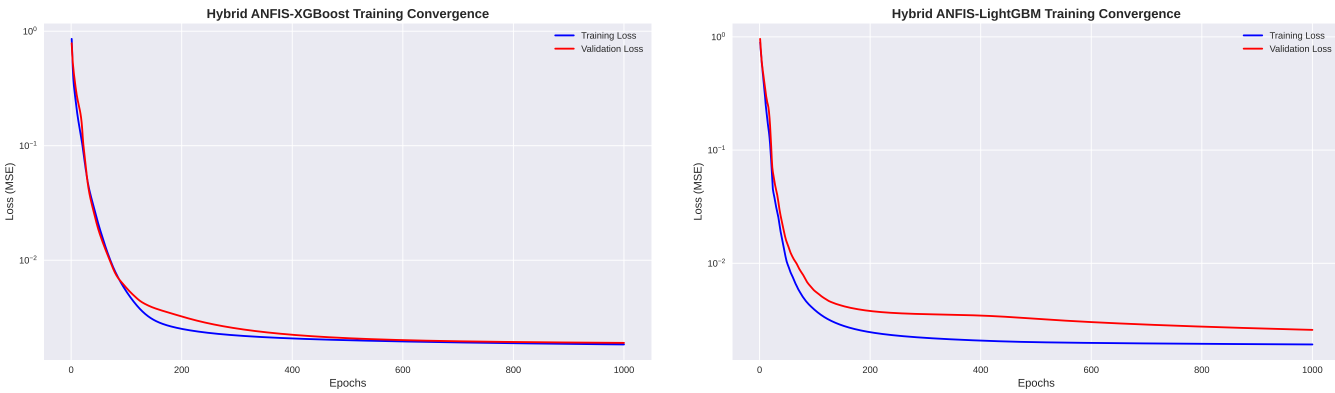
| **Metric** | **ANFIS** | **Hybrid ANFIS-LightGBM** | **Random Forest** |
|---|---|---|---|
| MAE | 3.7067 | 2.7924 | 1.2471 |
| MSE | 16.7449 | 5.4482 | 2.7812 |
| RMSE | 5.1716 | 2.3835 | 1.6677 |
| $R^2$ | 0.9471 | 0.9888 | 0.9924 |

To evaluate the stability and effectiveness of the training process, a convergence analysis was conducted for both the Hybrid ANFIS-XGBoost and Hybrid ANFIS-LightGBM frameworks. The models were trained using a three-way data split (60% training, 20% validation, 20% testing). The validation set was used to monitor model performance during training and prevent overfitting through early stopping. Figure 6 illustrates the evolution of the training and validation loss across epochs, showing a stable downward trend that reflects effective learning.

Figures 7(a) and 7(b) provide a visual comparison of actual student performance against the performance levels forecasted by the conventional ANFIS model and our novel hybrid models, utilizing data excluded from the training set (the test dataset). The graphs indicate that both the ANFIS-XGBoost and ANFIS-LightGBM models more accurately forecast student performance in comparison to the fundamental ANFIS model.

To examine the efficacy of the models in categorizing student performance, we utilize confusion matrices, illustrated in Figures 8 and 9. These matrices evaluate the classification accuracy of the standard ANFIS model and the hybrid models in categorizing students into various performance categories. The confusion matrices unequivocally demonstrate that the hybrid techniques excel in accurately categorizing students' competence levels. They attain a more precise categorization of students and exhibit fewer errors in classification relative to the conventional ANFIS method.

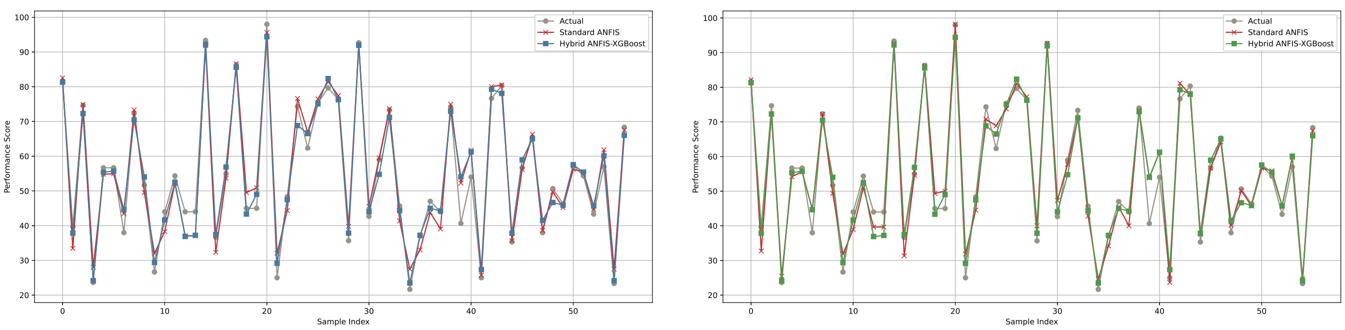
In summary, the integration of ANFIS with LightGBM and XGBoost establishes an effective novel method for predicting student performance. These integrated models demonstrate significant potential as instruments for educa-




(a) Hybrid ANFIS-XGBoost convergence curve

(b) Hybrid ANFIS-LightGBM convergence curve

Figure 6: Convergence analysis of the training and validation loss over epochs for Hybrid ANFIS-XGBoost and Hybrid ANFIS-LightGBM.




(a) ANFIS and hybrid ANFIS-XGBoost prediction

(b) ANFIS and hybrid ANFIS-LightGBM prediction

Figure 7: Comparison of actual performance of students with ANFIS and hybrid ANFIS methods predictions for test data

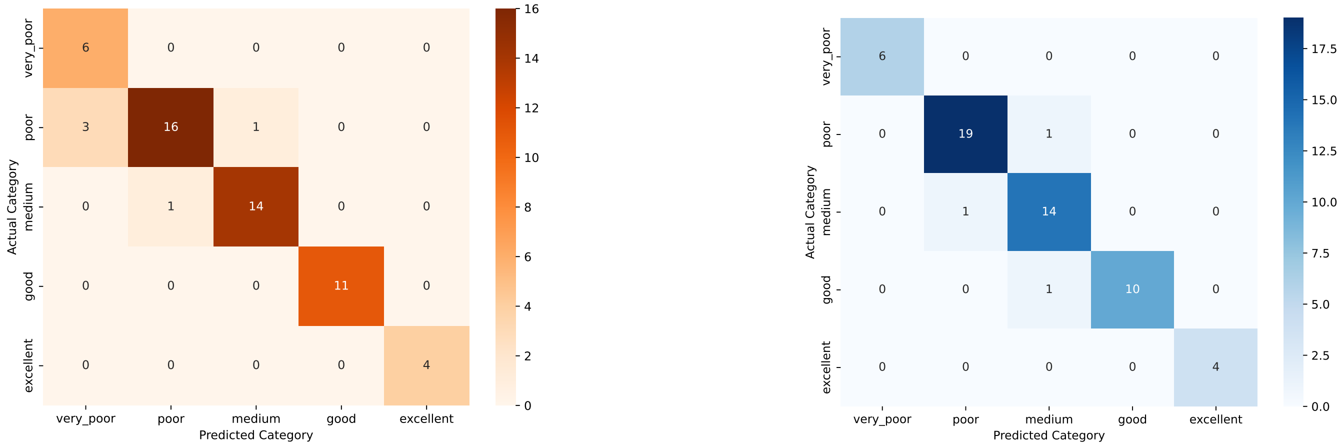
tional institutions seeking to leverage data to enhance student support and deliver assistance when necessary. In the future, enhancing these models by utilizing additional data and incorporating more information about students will be beneficial. This will assess the efficacy of these models across various contexts and for a broader spectrum of pupils.

# 6 Discussion

This work investigated novel approaches to forecasting university students' performance in advanced mathematics by integrating fuzzy logic with robust machine learning techniques. Our primary objective was to develop models that are both precise and user-friendly for educational institutions. This section summarizes our findings, addresses the limits of our study, and proposes avenues for future research.

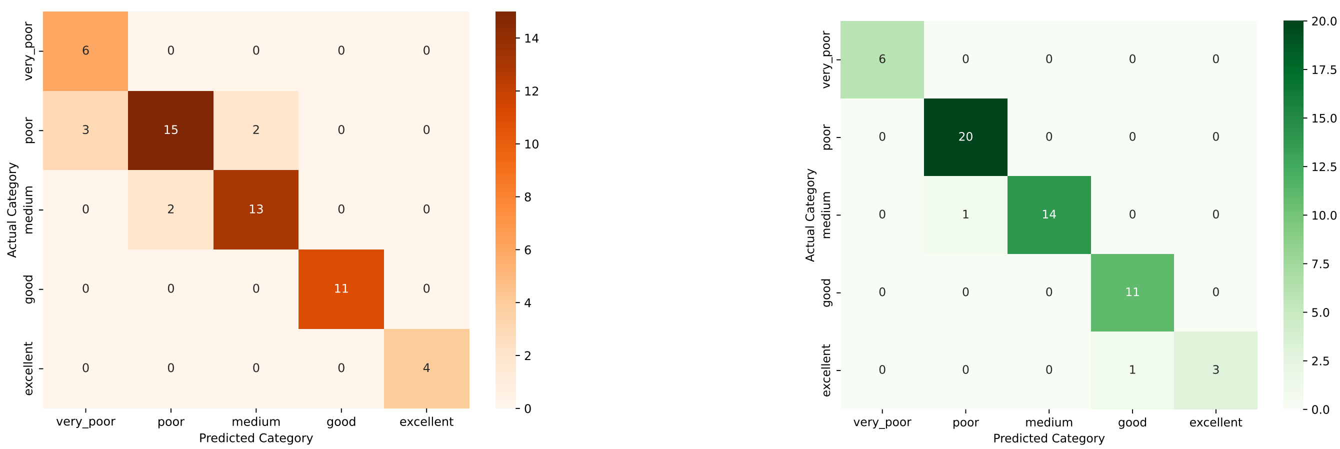
## 6.1 Summary of key findings

Our findings unequivocally demonstrate that the integration of ANFIS with both LightGBM and XGBoost markedly enhances the prediction of student mathematics performance in comparison to the utilization of ANFIS in isolation. The hybrid models, ANFIS-XGBoost and ANFIS-LightGBM, were more effective in elucidating the intricate relationships between various factors influencing student progress in mathematics. Feature selection was an essential process, enabling us to concentrate on the most significant predictors and streamline our models. Through SHAP analysis and XGBoost feature importance, we identified numerous critical parameters that significantly affect student performance. This encompasses the quality of a student's mathematical education prior to university, their grades in essential foundational math courses (General Mathematics I & II, and Differential Equations), familial support, self-efficacy in




(a) ANFIS (b) Hybrid ANFIS-XGBoost

Figure 8: Confusion matrix comparison: ANFIS vs. hybrid ANFIS-XGBoost methods(test data)




(a) ANFIS (b) Hybrid ANFIS-LightGBM

Figure 9: Confusion matrix comparison: ANFIS vs. hybrid ANFIS-LightGBM methods (test data)

mathematics, motivation for improvement, comprehension of mathematical applications, and the educational attainment of the mother. Additionally, participants in our survey noted that learning materials and support systems require enhancement. The hybrid models not only predicted performance scores with greater accuracy but also excelled in categorizing students into several performance classifications.

## 6.2 Constraints

Although our work offers significant insights, it is essential to recognize certain limits. Initially, our data was sourced from a single university in Iran and concentrated on IT students. This indicates that our findings may not be entirely applicable to students in other disciplines or in other nations and educational systems. Although our dataset comprises 280 students, which is enough, an increase in size might enhance the reliability and generalizability of our models. Secondly, we utilized questionnaires to collect student perspectives regarding characteristics such as learning engagement and psychological hurdles. Questionnaire data may be affected by students' self-perceptions and may not consistently represent reality accurately. Moreover, although we evaluated various pertinent characteristics, there may be additional significant elements absent from our analysis, such as particular pedagogical approaches, comprehensive socio-economic contexts, or distinct learning styles, that could enhance predictive accuracy. Ultimately, while our hybrid models offer greater interpretability than solely "black box" machine learning models, they remain more intricate than basic statistical methods, which may pose a challenge for institutions with constrained resources for model deployment and maintenance.

### 6.3 Prospective directions

Our findings and limitations suggest multiple directions for future investigation. A crucial subsequent step is to evaluate these hybrid models using larger datasets from various universities and student demographics, encompassing different degree programs and geographical regions, to ascertain whether our findings are applicable in wider contexts. Subsequent research may broaden the spectrum of features evaluated. Gathering more comprehensive data on students' socio-economic backgrounds, learning behaviors, personality characteristics, and interactions with online learning platforms may yield a more robust dataset and potentially improve predictive precision. Longitudinal studies monitoring students over time would be essential in comprehending how these characteristics and predictions evolve along their academic progression. Exploring practical applications of our models by devising and testing treatments based on the identified important predictors would be advantageous. Interventions designed to augment learning resources or bolster student self-confidence in mathematics could be formulated and assessed. Ultimately, juxtaposing the efficacy of our ANFIS-hybrid models with other sophisticated explainable AI methodologies could enhance our comprehension of the most effective strategies for forecasting and augmenting student success in higher education.

### 6.4 Conclusion

This study effectively constructed and tested hybrid ANFIS models that integrate fuzzy logic with gradient boosting techniques to forecast student performance in advanced mathematics. These models provide enhanced forecast accuracy and interpretability relative to conventional ANFIS. Our findings underscore critical elements affecting student achievement and emphasize the significance of academic preparation and student support networks. The established hybrid models serve as a beneficial resource for educational institutions aiming to utilize data to proactively assist students and enhance mathematics learning outcomes. By comprehending and tackling the elements revealed in this research, educators can create more effective tactics to facilitate student success in higher education.

## Acknowledgement

The authors sincerely thank the editor and the anonymous reviewers for their valuable comments and constructive suggestions, which greatly contributed to improving the clarity, quality, and overall rigor of this manuscript.

## References

[1] A. Abdi, A. Zandipayam, *Prediction of academic performance based on dimensions of academic identity and flourishing among students of the university of medical sciences*, Journal of Medical Education Development, **12**(35) (2019), 19-26. https://doi.org/10.29252/edcj.12.35.19

[2] A. Abín, J. C. Núñez, C. Rodríguez, M. Cueli, T. García, P. Rosário, *Predicting mathematics achievement in secondary education: The role of cognitive, motivational, and emotional variables*, Frontiers in Psychology, **11** (2020). https://doi.org/10.3389/fpsyg.2020.00876

[3] T. Ahajjam, M. Moutaib, H. Aissa, M. Azrour, Y. Farhaoui, M. Fattah, *Predicting students' final performance using artificial neural networks*, Big Data Mining and Analytics, **5**(4) (2022), 294-301. https://doi.org/10.26599/BDMA.2021.9020030

[4] A. Alshanqiti, A. Namoun, *Predicting student performance and its influential factors using hybrid regression and multi-label classification*, IEEE Access, **8** (2020), 203827-203844. https://doi.org/10.1109/ACCESS.2020.3036572

[5] T. J. V. Ark, E. Liebtag, N. McClennen, *The power of place: Authentic learning through place-based education*, ASCD, (2020).

[6] R. S. J. D. Baker, K. Yacef, *The state of educational data mining in 2009: A review and future visions*, Journal of Educational Data Mining, **1**(1) (2009), 3-17. https://doi.org/10.5281/zenodo.3554657

[7] M. Castro, E. Expósito-Casas, E. López-Martín, L. Lizasoain, E. Navarro-Asencio, J. L. Gaviria, *Parental involvement on student academic achievement: A meta-analysis*, Educational Research Review, **14** (2015), 33-46. https://doi.org/10.1016/j.edurev.2015.01.002

[8] P. Chakurkar, D. Vora, *An integrated deep learning and fuzzy logic system for road crack severity analysis, and pedestrian fall risk prediction*, Iranian Journal of Fuzzy Systems, **22**(6) (2025), 103-123. https://doi.org/10.22111/ijfs.2025.52137.9268

[9] T. Chen, C. Guestrin, *XGBoost: A scalable tree boosting system*, Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, (2016), 785-794. https://doi.org/10.1145/2939672.293978

[10] X. Chen, B. Li, A. C. Worthington, *Higher moments and US industry returns: Realized skewness and kurtosis*, Review of Accounting and Finance, **20**(1) (2021), 13-22. https://doi.org/10.1108/RAF-06-2020-0171

[11] F. Cheng, Y. Liu, Y. Jia, *Evaluation of students' performance during the academic period using expert systems with applications: The XGBoost classifier-enhanced AEO hybrid model*, Expert Systems with Applications, **238** (2024), 122136. https://doi.org/10.1016/j.eswa.2023.122136

[12] H. J. Cho, M. R. Melloch, C. Levesque-Bristol, *Enhanced student perceptions of learning and performance using concept-point-recovery teaching sessions: A mixed-method approach*, International Journal of STEM Education, **8**(1) (2021), 1-17. https://doi.org/10.1186/s40594-021-00276-1

[13] P. S. Deri, S. Cahyadi, E. Susiati, *How teacher support affects student engagement in learning math*, Jurnal Neo Konseling, **5**(1) (2023), 21. https://doi.org/10.24036/00713kons2023

[14] A. Diseth, F. K. S. Mathisen, O. Samdal, *A comparison of intrinsic and extrinsic motivation among lower and upper secondary school students*, Educational Psychology, **40**(8) (2020), 961-980. https://doi.org/10.1080/01443410.2020.1778640

[15] J. Friedman, *Greedy function approximation: A gradient boosting machine*, Annals of Statistics, **29**(5) (2001), 1189-1232. https://doi.org/10.1214/aos/1013203451

[16] J. M. Froiland, M. L. Davison, *The longitudinal influences of peers, parents, motivation, and mathematics course-taking on high school math achievement*, Learning and Individual Differences, **50** (2016), 252-259. https://doi.org/10.1016/j.lindif.2016.07.012

[17] M. S. Hagger, K. Hamilton, *Grit and self-discipline as predictors of effort and academic attainment*, British Journal of Educational Psychology, **89**(2) (2018), 324-342. https://doi.org/10.1111/bjep.12241

[18] S. Hakkal, A. A. Lahcen, *XGBoost to enhance learner performance prediction*, Computers and Education: Artificial Intelligence, (2024), Article 100254. https://doi.org/10.1016/j.caeai.2024.100254

[19] S. Hwang, T. Son, *Students' attitude toward mathematics and its relationship with mathematics achievement*, Journal of Education and Elearning Research, **8**(3) (2021), 272-280. https://doi.org/10.20448/journal.509.2021.83.272.280

[20] J. S. R. Jang, *ANFIS: Adaptive-network-based fuzzy inference system*, IEEE Transactions on Systems, Man, and Cybernetics, **23**(3) (1993), 665-685. https://doi.org/10.1109/21.256541

[21] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, T. Y. Liu, *LightGBM: A highly efficient gradient boosting decision tree*, Advances in Neural Information Processing Systems, **30** (2017).

[22] D. Koç, A. N. Elçi, *The effect of mathematical modeling instruction on pre-service primary school teachers' problem-solving skills and attitudes towards mathematics*, Journal of Pedagogical Research, **6**(4) (2022), 111-129. https://doi.org/10.33902/JPR.202217783

[23] M. A. Kraft, *Teacher effects on complex cognitive skills and social-emotional competencies*, Journal of Human Resources, **54**(1) (2017), 1-36. https://doi.org/10.3368/jhr.54.1.0916.8265R3

[24] A. A. S. Maharani, N. Suarna, I. Ali, D. Solihudin, *Application of decision tree algorithm to improve student learning pattern classification model at Sempoa Sip Perjuangan*, Journal of Artificial Intelligence and Engineering Applications, **4**(2) (2025), 1399-1403. https://doi.org/10.59934/jaiea.v4i2.915

[25] C. Márquez-Vera, A. Cano, C. Romero, S. Ventura, *Predicting student failure at school using genetic programming and different data mining approaches with high-dimensional and imbalanced data*, Applied Intelligence, **38** (2013), 315-330. https://doi.org/10.1007/s10489-012-0374-8

[26] S. H. Mercer, L. M. Nellis, R. S. Martínez, M. Kirk, *Supporting the students most in need: Academic self-efficacy and perceived teacher support in relation to within-year academic growth*, Journal of School Psychology, **49**(3) (2011), 323-338. https://doi.org/10.1016/j.jsp.2011.03.006

[27] H. C. Nguyen, *Factors affecting student engagement: A model of how and why students learn*, EDULEARN Proceedings, (2019). https://doi.org/10.21125/edulearn.2019.0578

[28] M. K. Nyatsikor, E. Sosu, P. Mtika, D. Robson, *Teacher characteristics and children's educational attainment in Ghana: Do some teacher characteristics matter more for children attending disadvantaged schools*, Frontiers in Education, **5** (2020). https://doi.org/10.3389/feduc.2020.00162

[29] F. Ofori, E. Maina, R. Gitonga, *Using machine learning algorithms to predict students' performance and improve learning outcomes: A literature-based review*, Journal of Information and Technology, **4**(1) (2020), 33-55.

[30] M. Priyadharshini, B. Vanathi, *Enhance data security with efficient anomaly detection in a real network environment*, Iranian Journal of Fuzzy Systems, **22**(6) (2025), 1-20. https://doi.org/10.22111/ijfs.2025.49590.8748

[31] M. A. Quran, *Efficient and effective anomaly detection in autonomous vehicles: A combination of gradient boosting and ANFIS algorithms*, International Journal of Fuzzy Systems, **27** (2025), 1421-1437. https://doi.org/10.1007/s40815-024-01843-8

[32] F. H. Ramadhan, A. D. Lestari, *Support vector machine on two-class classification problem to determine an otaku*, Journal of Student Research Exploration, **3**(1) (2025), 17-22. https://doi.org/10.52465/josre.v3i1.358

[33] C. Romero, S. Ventura, *Educational data mining: A review of the state of the art*, IEEE Transactions on Systems, Man, and Cybernetics: Part C (Applications and Reviews), **40**(6) (2010), 601-618. https://doi.org/10.1109/TSMCC.2010.2053532

[34] D. Sabanal, M. Gako, H. D. Torre, J. Sabanal, et al, *Predictive model for college students' performance in higher mathematics*, Social Sciences and Humanities Open, **10** (2024), 101134. https://doi.org/10.1016/j.ssaho.2024.101134

[35] T. Sánchez, R. Gilar-Corbi, J. L. Castejón, J. Vidal, J. León, *Students' evaluation of teaching and their academic achievement in a higher education institution of Ecuador*, Frontiers in Psychology, **11** (2020), 233. https://doi.org/10.3389/fpsyg.2020.00233

[36] T. Shen, Y. Zhang, *A performance prediction model for newly enrolled university students based on the fusion of K-Means and LightGBM algorithms*, 2024 3rd International Conference on Cloud Computing, Big Data Application and Software Engineering (CBASE), (2024), 40-44. https://doi.org/10.1109/CBASE64041.2024.10824518

[37] T. O. Soyoye, T. Chen, R. Hill, *Predicting academic performance of university students using adaptive neuro fuzzy inference system (ANFIS)-subtractive clustering algorithm (ANFIS-SC): A case study in the UK*, UK Workshop on Computational Intelligence, (2024), 315-333. https://doi.org/10.1007/978-3-031-78857-4-24

[38] M. Sugeno, K. Tanaka, *Successive identification of a fuzzy model and its applications to prediction of a complex system*, Fuzzy Sets and Systems, **42**(3) (1991), 315-334. https://doi.org/10.1016/0165-0114(91)90110-C

[39] R. Wakhata, S. Balimuttajjo, V. Mutarutinya, *Relationship between students' attitude towards, and performance in mathematics word problems*, PLoS ONE, **19**(2) (2024), e0278593. https://doi.org/10.1371/journal.pone.0278593

[40] C. Wang, L. Chang, T. Liu, *Predicting student performance in online learning using a highly efficient gradient boosting decision tree*, International Conference on Intelligent Information Processing, (2022), 508-521. https://doi.org/10.1007/978-3-031-03948-5-41

[41] Y. M. Yesuf, S. A. Kebede, A. Zewdu, D. M. Gebru, *Predictors of high school students' mathematics self-efficacy in Addis Ababa: The importance of educational expectations*, Frontiers in Psychology, **13** (2023), 1-10. https://doi.org/10.3389/fpsyg.2022.927757