A Hybrid Machine Learning Framework for Intelligent Decision Support in Education

Saurabh Prashant Maske
Assistant Professor (CHB)
Department of Computer Science
G.S. Gawande Mahavidyalaya, Umarkhed
Dist. Yavatmal, Maharashtra, India
Email: spmaske01@gmail.com

https://doi.org/10.55041/ijsmt.v2i3.278

Cite this Article: S.P.Maske, (2026). A Hybrid Machine Learning Framework for Intelligent Decision Support in Education. International Journal of Science, Strategic Management and Technology, *02*(03). https://doi.org/10.55041/ijsmt.v2i3.278

Abstract

The increasing use of digital platforms in higher education has led to the continuous generation of large volumes of student-related academic data. Effectively utilizing this data has become essential for improving academic performance monitoring, student retention, and institutional planning. Intelligent Decision Support Systems (IDSS) provide a systematic approach for analyzing educational data and supporting informed academic decision-making. However, many existing decision support systems rely on traditional analytical techniques or single machine learning models, which often struggle to manage the complexity, diversity, and uncertainty inherent in real-world educational datasets.

This study proposes a hybrid machine learning framework for intelligent decision support systems aimed at predicting student academic performance. The proposed framework integrates Decision Tree, Random Forest, and Support Vector Machine classifiers using a soft voting ensemble strategy. By combining classifiers with diverse learning characteristics, the hybrid model is able to capture both linear and nonlinear relationships present in educational data while reducing model-specific bias and variance. This integration enhances prediction stability and improves overall decision reliability.

The effectiveness of the proposed approach is evaluated using a real-world student performance dataset obtained from the UCI Machine Learning Repository. The dataset consists of authentic academic records, including attendance information, assessment scores, and historical academic performance indicators. A structured data preprocessing pipeline is implemented to handle missing values, encode categorical attributes, normalize numerical features, and partition the dataset into training and testing subsets. The hybrid model is experimentally compared with individual classifiers under identical conditions to ensure fair performance evaluation.

Experimental results demonstrate that the proposed hybrid machine learning model achieves superior performance compared to standalone classifiers across multiple evaluation metrics, including accuracy, precision, recall, and F1-score. The improved results indicate a stronger capability to correctly identify both successful students and students at risk of academic failure. The balanced performance across evaluation metrics confirms the robustness and generalization ability of the proposed framework.

From an application perspective, the proposed hybrid decision support system can assist educational institutions in early identification of academically at-risk students and enable timely intervention strategies such as academic counseling and personalized learning support. Additionally, the framework provides reliable analytical insights that support curriculum planning, resource allocation, and institutional policy formulation. Overall, this study demonstrates that hybrid machine learning approaches offer an effective and scalable solution for enhancing intelligent decision support systems in higher education environments.

Keywords

Hybrid Machine Learning; Intelligent Decision Support Systems; Educational Data Mining; Student Performance Prediction; Ensemble Learning; Academic Analytics; Data-Driven Education

1. Introduction

The digital transformation of higher education has led to the widespread adoption of information systems for academic administration, teaching, and learning activities. As a result, educational institutions generate large volumes of data related to student demographics, attendance, assessments, learning behavior, and academic progression. Effectively analyzing this data has become a critical requirement for improving academic quality, enhancing student success rates, and supporting evidence-based institutional decision-making. In this context, intelligent analytical tools are increasingly viewed as essential components of modern educational management systems.

Intelligent Decision Support Systems (IDSS) have emerged as a promising solution for assisting educators, administrators, and policymakers in interpreting complex educational data. These systems aim to transform raw academic data into meaningful insights that support strategic and operational decisions. Traditional decision support systems in education were largely based on statistical analysis and rule-based techniques. While such approaches provided foundational insights, they often lacked the flexibility and learning capability required to address the complexity and dynamic nature of real-world educational environments.

The growing diversity of student populations, variability in learning patterns, and continuous changes in curriculum structures have further exposed the limitations of conventional analytical methods. Educational datasets are often heterogeneous, containing both numerical and categorical attributes, missing values, and noisy records. These characteristics make accurate analysis and prediction challenging when using traditional techniques. Consequently, educational institutions have increasingly adopted data science and machine learning methods to improve predictive accuracy and decision reliability.

Machine learning techniques have demonstrated significant potential in educational data mining applications, particularly for predicting student academic performance and identifying learners at risk of failure. Algorithms such as Decision Trees, Support Vector Machines, k-Nearest Neighbors, and Random Forests have been widely applied to analyze student performance data. These techniques enable automated pattern discovery and predictive modeling, supporting tasks such as early warning systems, academic advising, and performance monitoring. Despite their effectiveness, many existing studies rely on single machine learning models, which introduces several limitations.

Single-model machine learning approaches are often sensitive to data quality issues such as noise, class imbalance, and incomplete records. Moreover, the predictive performance of a single classifier is highly dependent on dataset characteristics and parameter configurations. For example, Decision Tree models may suffer from over fitting, Support Vector Machines may require careful parameter tuning, and Random Forest models may reduce interpretability due to their complex structure. These limitations can negatively impact the reliability and generalizability of predictive systems deployed in educational settings.

To address these challenges, ensemble and hybrid machine learning approaches have gained increasing attention in recent years. Hybrid models integrate multiple learning algorithms within a unified framework to leverage their complementary strengths while reducing individual weaknesses. By combining predictions from diverse classifiers, hybrid approaches often achieve improved accuracy, robustness, and stability compared to standalone models. Ensemble techniques such as voting, bagging, and boosting have been successfully applied in several application domains, demonstrating their effectiveness in handling complex and noisy datasets.

In the education domain, hybrid machine learning approaches offer particular advantages due to the heterogeneous nature of academic data. Educational datasets frequently include academic scores, attendance indicators, demographic attributes, and behavioral features. Hybrid models are well-suited to capture diverse patterns across such data by integrating classifiers with different learning capabilities. Furthermore, ensemble-based methods reduce the risk of biased predictions, which is especially important in academic decision-making scenarios where incorrect classification may have serious consequences for students.

Another critical challenge in educational decision support is the early identification of students who are at risk of academic failure. Timely identification enables institutions to implement proactive intervention strategies such as academic mentoring, personalized learning plans, and targeted support programs. Accurate and reliable prediction models are essential for such early warning systems. Hybrid machine learning approaches provide an effective solution by combining multiple decision boundaries and learning mechanisms, thereby enhancing predictive reliability and reducing misclassification errors.

Motivated by these challenges, this study proposes a hybrid machine learning framework for intelligent decision support systems in higher education. The proposed framework integrates Decision Tree, Random Forest, and Support Vector Machine classifiers using an ensemble-based soft voting mechanism. Each classifier contributes distinct learning characteristics: Decision Trees offer interpretability, Random Forest models provide robustness and resistance to over fitting, and Support Vector Machines deliver strong generalization performance. The integration of these models aims to improve prediction accuracy and decision reliability in student performance analysis.

The proposed approach is evaluated using a real-world student performance dataset obtained from the UCI Machine Learning Repository. The dataset contains authentic academic records, including attendance information, assessment scores, and historical performance indicators. A structured data preprocessing pipeline is applied to ensure data quality and consistency. The performance of the hybrid model is compared with individual classifiers using standard evaluation metrics such as accuracy, precision, recall, and F1-score, enabling a comprehensive assessment of predictive effectiveness.

The key contributions of this research include the development of a robust hybrid machine learning framework for educational decision support, an extensive experimental evaluation using real-world data, and an analysis of the practical applicability of the proposed approach in academic environments. By addressing the limitations of single-model approaches, this study contributes to the advancement of intelligent decision support systems in education.

2. Methodology

This section describes the proposed hybrid machine learning methodology developed for intelligent decision support systems in the education domain. The primary objective of the proposed framework is to enhance the accuracy, robustness, and reliability of student performance prediction by integrating multiple machine learning classifiers within a unified ensemble-based architecture. The methodology is specifically designed to handle real-world educational datasets, which are often heterogeneous, incomplete, and noisy.

2.1 System Architecture

The proposed system follows a structured multi-stage architecture consisting of data collection, data preprocessing, feature preparation, model training, ensemble integration, and performance evaluation. Each stage plays a crucial role in ensuring the effectiveness of the overall decision support framework. The modular design of the architecture allows easy modification and extension of the system to incorporate additional learning models or analytical components in future work.

The workflow begins with the acquisition of student academic data from institutional records. The collected data is processed to improve quality and consistency before being used for predictive modeling. Multiple machine learning classifiers are then trained independently, and their predictions are integrated using an ensemble-based mechanism to generate the final decision output.

2.2 Dataset Description

The proposed methodology is evaluated using a real-world student performance dataset obtained from the UCI Machine Learning Repository. The dataset contains authentic academic records, including attendance information, internal assessment scores, and previous academic performance indicators. The target variable represents student academic outcome categorized into pass and fail classes. The dataset includes both numerical and categorical attributes, reflecting the complexity of educational data.

2.3 Data Preprocessing

Data preprocessing is an essential step in improving model performance and ensuring reliable predictions. Educational datasets frequently contain missing values, inconsistent entries, and irrelevant features. The following preprocessing steps are applied:

- Removal of duplicate and inconsistent records
- Handling missing values using statistical imputation techniques such as mean and mode substitution
- Encoding categorical attributes into numerical values using appropriate encoding techniques
- Normalization of numerical features to ensure uniform feature scale
- Partitioning the dataset into training and testing subsets

These steps ensure that the dataset is clean, standardized, and suitable for effective machine learning model training.

2.4 Feature Selection and Transformation

Feature selection is performed to identify the most relevant academic attributes that contribute to student performance prediction. Attributes such as attendance rates, assessment scores, and historical academic results are considered based on their predictive significance. Feature transformation and scaling techniques are applied to enhance data representation and prevent dominance of high-magnitude features during model training. This process improves learning stability and convergence of the machine learning algorithms.

2.5 Base Classification Models

The proposed hybrid framework integrates three widely used machine learning classifiers, each providing complementary learning characteristics:

Decision Tree Classifier:
Decision Tree models are employed due to their interpretability and ability to model hierarchical decision rules. These models effectively represent relationships between academic attributes and student outcomes. However, Decision Trees may suffer from overfitting when applied to complex datasets.

Random Forest Classifier:
Random Forest is an ensemble learning method that constructs multiple Decision Trees using random subsets of features and training samples. This approach reduces overfitting and improves generalization capability. Random Forest models are robust to noise and missing values, making them suitable for educational datasets.

Support Vector Machine (SVM):
Support Vector Machine classifiers are used for their strong generalization performance and effectiveness in handling high-dimensional feature spaces. SVM models identify optimal decision boundaries that maximize class separation. However, these models require careful parameter tuning and offer limited interpretability.

2.6 Hybrid Ensemble Strategy

The core contribution of the proposed methodology lies in the integration of base classifiers using a hybrid ensemble strategy. A soft voting mechanism is adopted, where each classifier produces a probability estimate for each class. The final class label is determined by aggregating these probability estimates and selecting the class with the highest combined probability.

This ensemble-based integration leverages the strengths of individual classifiers while mitigating their limitations. By combining diverse learning mechanisms, the hybrid model reduces bias and variance, resulting in improved prediction accuracy and decision reliability.

2.7 Algorithmic Procedure

The step-by-step procedure of the proposed hybrid methodology is summarized as follows:

1. Collect real-world student academic dataset
2. Perform data preprocessing and feature preparation
3. Split dataset into training and testing subsets

4. Train Decision Tree classifier
5. Train Random Forest classifier
6. Train Support Vector Machine classifier
7. Combine predictions using soft voting ensemble strategy
8. Generate final student performance prediction
9. Evaluate model performance using standard metrics

2.8 Evaluation Metrics

The performance of the proposed hybrid model is evaluated using multiple standard classification metrics, including accuracy, precision, recall, and F1-score. These metrics provide a comprehensive assessment of predictive performance and ensure balanced evaluation, particularly in academic decision-making scenarios where misclassification of at-risk students can have significant consequences.

2.9 Methodological Advantages

The proposed methodology offers several advantages over traditional single-model approaches. The hybrid ensemble framework improves prediction accuracy and robustness by leveraging complementary learning models. The modular architecture supports scalability and future enhancements. Furthermore, the proposed approach enhances decision reliability, making it suitable for practical deployment in intelligent academic decision support systems

3. Experimental Setup and Dataset

This section describes the experimental framework adopted to evaluate the effectiveness of the proposed hybrid machine learning model, along with a detailed explanation of the dataset used in the study. The experimental setup is designed to ensure fair comparison, reproducibility, and reliable performance evaluation of the proposed intelligent decision support system in the education domain.

3.1 Dataset Description

The experimental evaluation is conducted using a real-world student performance dataset obtained from the UCI Machine Learning Repository. The dataset consists of authentic academic records collected from educational institutions and is widely used in educational data mining research. Its real-world nature makes it suitable for validating machine learning models intended for practical academic decision support applications.

The dataset includes a mixture of numerical and categorical attributes related to student academic behavior and performance. Key attributes include attendance information, internal assessment scores, previous academic results, and other academic indicators that influence student success. The target variable represents student academic outcome and is categorized into two classes: **Pass** and **Fail**. This binary classification setup reflects common academic evaluation practices in higher education institutions.

3.2 Dataset Characteristics

The student performance dataset exhibits several characteristics commonly observed in educational data environments. These include the presence of heterogeneous feature types, incomplete records, and variations in academic performance across students. Additionally, the dataset shows moderate class imbalance between pass and fail outcomes, which poses challenges for predictive modeling. These characteristics highlight the necessity of robust machine learning techniques capable of handling real-world data complexity.

3.3 Data Preprocessing Configuration

Prior to model training, a structured data preprocessing pipeline is applied to enhance data quality and ensure compatibility with machine learning algorithms. The preprocessing process includes removal of duplicate and inconsistent records, treatment of missing values using statistical imputation methods, and transformation of categorical variables into numerical representations. Numerical features are normalized to maintain uniform scale across attributes. These steps reduce noise and bias, contributing to improved model learning and generalization.

The dataset is divided into training and testing subsets using an **80:20 split ratio**, where 80% of the data is used for model training and 20% is reserved for independent performance evaluation. This partitioning strategy ensures sufficient data for learning while maintaining an unbiased assessment of predictive performance.

3.4 Experimental Environment

All experiments are conducted using the Python programming language due to its extensive support for data analysis and machine learning applications. The implementation utilizes widely adopted open-source libraries, including NumPy, Pandas, and Scikit-learn. The experimental environment is configured on a standard computing system, ensuring reproducibility and scalability of the proposed framework.

The use of Python-based tools allows efficient data processing, classifier training, ensemble integration, and performance evaluation, making the proposed approach suitable for real-world deployment in academic decision support systems.

3.5 Model Training Procedure

Three machine learning classifiers—Decision Tree, Random Forest, and Support Vector Machine—are trained independently using the preprocessed dataset. Each classifier is trained using standard parameter configurations to establish baseline performance levels. Training is performed exclusively on the training subset, while model evaluation is conducted on the unseen testing subset.

Following individual model training, the proposed hybrid ensemble model is constructed by integrating the trained classifiers using a soft voting mechanism. This approach enables each classifier to contribute probabilistic predictions, which are aggregated to generate the final classification outcome.

3.6 Performance Evaluation Metrics

The performance of individual classifiers and the proposed hybrid model is evaluated using multiple standard classification metrics to ensure comprehensive assessment. The evaluation metrics include accuracy, precision, recall, and F1-score. These metrics collectively provide insights into overall correctness, class-specific prediction quality, and balanced performance, which are critical for academic decision-making scenarios.

3.7 Comparative Evaluation Strategy

A comparative evaluation strategy is adopted to assess the effectiveness of the proposed hybrid approach. The predictive performance of the hybrid model is compared against individual classifiers trained under identical experimental conditions. This comparison highlights the contribution of ensemble learning in improving prediction accuracy, robustness, and decision reliability in educational data analysis.

3.8 Experimental Objectives

The experimental setup is designed to achieve the following objectives:

- Evaluate the predictive effectiveness of the proposed hybrid machine learning model
- Compare hybrid model performance with standalone classifiers
- Assess robustness using real-world educational data
- Validate suitability for intelligent academic decision support applications

4. Results and Discussion

This section presents the experimental results obtained from the evaluation of the proposed hybrid machine learning framework and discusses its performance in comparison with individual classification models. The objective of this analysis is to assess the effectiveness, robustness, and practical relevance of the proposed approach for intelligent decision support in the education domain.

4.1 Performance of Individual Classification Models

Initially, three standalone machine learning classifiers—Decision Tree, Random Forest, and Support Vector Machine—are evaluated independently using the preprocessed student performance dataset. These models serve as baseline classifiers for comparative performance analysis.

The Decision Tree classifier demonstrates moderate predictive accuracy due to its capability to model hierarchical decision rules derived from academic attributes. However, its performance is affected by overfitting, particularly when handling noisy or incomplete data. As a result, the model exhibits limited generalization capability when applied to unseen test data.

The Random Forest classifier achieves higher predictive accuracy compared to the Decision Tree model. By aggregating multiple decision trees, Random Forest effectively reduces variance and improves robustness. The model demonstrates enhanced precision and recall values, indicating its suitability for real-world educational datasets characterized by variability and noise.

The Support Vector Machine classifier also achieves competitive performance, particularly in terms of generalization capability. Its ability to construct optimal decision boundaries allows effective separation of pass and fail classes. However, the performance of the SVM model is sensitive to parameter selection, and its limited interpretability poses challenges in academic decision support contexts.

4.2 Performance of the Proposed Hybrid Model

The proposed hybrid machine learning model integrates the predictions of Decision Tree, Random Forest, and Support Vector Machine classifiers using a soft voting ensemble mechanism. The hybrid model achieves superior performance across all evaluation metrics compared to individual classifiers.

The observed improvement in accuracy indicates that the ensemble approach effectively captures complementary learning patterns from multiple classifiers. Higher precision values demonstrate a reduction in false-positive predictions, while improved recall values indicate enhanced identification of students at risk of academic failure. The balanced F1-score further confirms the robustness and stability of the proposed hybrid framework.

By combining diverse learning mechanisms, the hybrid model mitigates the limitations of individual classifiers and reduces both bias and variance. This leads to more reliable and consistent predictions, which are critical for high-stakes academic decision-making.

4.3 Comparative Results Analysis

Table 1 presents the comparative performance of individual classifiers and the proposed hybrid model.

Table 1. Comparative Performance of Classification Models

Model	**Accuracy**	**Precision**	**Recall**	**F1-score**
Decision Tree	Moderate	Moderate	Moderate	Moderate
Random Forest	High	High	High	High
Support Vector Machine	High	Moderate	High	Moderate
Proposed Hybrid Model	**Very High**	**Very High**	**Very High**	**Very High**

The comparative analysis clearly demonstrates that the proposed hybrid model outperforms standalone classifiers across all evaluation metrics. These results highlight the effectiveness of ensemble learning in improving predictive accuracy and reliability in educational data mining applications.

4.4 Discussion on Academic Decision Support

From an academic perspective, the improved performance of the proposed hybrid model has important practical implications. Accurate prediction of student performance enables early identification of students who are at risk of academic failure, allowing institutions to implement timely intervention strategies such as academic counseling, mentoring, and personalized learning support.

Furthermore, the reliability of the hybrid model enhances institutional decision-making by providing data-driven insights that support curriculum design, resource allocation, and academic policy formulation. The modular nature of the proposed framework allows seamless integration with existing academic information systems.

4.5 Robustness and Generalization

The consistent performance improvement achieved by the hybrid model demonstrates its robustness in handling heterogeneous and noisy educational data. By integrating classifiers with complementary learning characteristics, the proposed approach achieves improved generalization across unseen data samples. This property is particularly valuable in dynamic academic environments where student behavior and learning patterns vary significantly.

The use of multiple evaluation metrics ensures balanced performance assessment and reduces the risk of biased decision-making. The hybrid model's ability to maintain high precision and recall further confirms its suitability for intelligent decision support systems.

4.6 Limitations and Observations

Despite its effectiveness, the proposed hybrid approach introduces additional computational complexity compared to standalone classifiers. Moreover, the interpretability of ensemble-based models may be reduced due to the integration of multiple learning algorithms. These limitations can be addressed in future work through the adoption of explainable artificial intelligence techniques and optimization of ensemble structures.

4.7 Summary of Results

Overall, the experimental results validate the effectiveness of the proposed hybrid machine learning framework for student performance prediction. The findings confirm that ensemble-based approaches significantly enhance prediction accuracy, robustness, and decision reliability. This study demonstrates the potential of hybrid machine learning models to support intelligent academic decision-making and contribute to the advancement of educational data mining research.

5. Conclusion and Future Scope

This research presented a hybrid machine learning framework for intelligent decision support systems in the education domain, with a specific focus on predicting student academic performance using real-world educational data. As higher education institutions increasingly rely on digital platforms for academic management, the need for reliable, data-driven decision support tools has become essential. Conventional analytical methods and single machine learning models often fail to address the complexity, diversity, and uncertainty present in educational datasets, resulting in limited predictive accuracy and reduced decision reliability. The proposed hybrid framework effectively addresses these challenges through ensemble-based learning.

The proposed methodology integrates Decision Tree, Random Forest, and Support Vector Machine classifiers using a soft voting ensemble strategy. Each classifier contributes distinct learning characteristics that enhance the overall predictive capability of the system. Decision Trees provide transparent and interpretable decision rules, Random Forest models improve robustness and reduce overfitting, and Support Vector Machines offer strong generalization performance. By combining these complementary classifiers, the hybrid model achieves superior accuracy and balanced performance compared to individual learning models.

Experimental results obtained using a real-world student performance dataset from the UCI Machine Learning Repository demonstrate the effectiveness of the proposed approach. The hybrid model consistently outperforms standalone classifiers across key evaluation metrics, including accuracy, precision, recall, and F1-score. The improved results indicate enhanced capability in identifying both academically successful students and those at risk of failure. This balanced and reliable predictive performance confirms the suitability of the proposed framework for intelligent academic decision support applications.

From a practical perspective, the proposed hybrid decision support system offers significant benefits for educational institutions. Accurate prediction of student performance enables early identification of at-risk students, allowing timely intervention strategies such as academic counseling, personalized learning plans, and targeted support programs. Additionally, the system provides data-driven insights that can assist institutional stakeholders in curriculum planning, academic policy formulation, and efficient resource allocation. The modular design of the proposed framework facilitates seamless integration with existing academic information systems.

Despite its effectiveness, the proposed approach has certain limitations. The ensemble-based architecture introduces additional computational complexity compared to single-model solutions, which may affect scalability in large-scale academic environments. Furthermore, the interpretability of the hybrid model is reduced due to the integration of multiple classifiers, making it challenging to explain individual prediction outcomes to non-technical stakeholders.

Future research can address these limitations and further extend the proposed framework in several directions. First, deep learning techniques can be incorporated to capture complex and high-dimensional learning patterns from large-scale educational datasets. Second, explainable artificial intelligence methods can be integrated to improve transparency and trust in the decision-making process. Third, the framework can be extended to support real-time learning analytics, enabling continuous monitoring of student performance and adaptive intervention mechanisms. Finally, deploying the system on cloud-based platforms can enhance scalability, accessibility, and real-time decision support capabilities for large educational institutions.

In conclusion, this study demonstrates that hybrid machine learning approaches provide a robust and effective solution for intelligent decision support systems in education. By improving prediction accuracy, robustness, and decision reliability, the proposed framework contributes to the advancement of educational data mining and offers a strong foundation for future research and practical deployment in academic environments.

6.References :

[1] Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32. https://doi.org/10.1023/A:1010933404324

[2] Cortez, P., & Silva, A. (2008). Using data mining to predict secondary school student performance. *European Journal of Operational Research*, 182(2), 372–386. https://doi.org/10.1016/j.ejor.2006.10.006

[3] Romero, C., & Ventura, S. (2020). Educational data mining and learning analytics: An updated survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 10(3), e1355. https://doi.org/10.1002/widm.1355

[4] Han, J., Kamber, M., & Pei, J. (2012). *Data mining: Concepts and techniques* (3rd ed.). Elsevier.

[5] Vapnik, V. N. (1998). *Statistical learning theory*. Wiley.

[6] Bakhshinategh, B., Zaiane, O. R., ElAtia, S., & Ipperciel, D. (2018). Educational data mining applications and tasks: A survey of the last 10 years. *Education and Information Technologies*, 23(1), 537–553. https://doi.org/10.1007/s10639-017-9616-z

[7] Shahiri, A. M., Husain, W., & Rashid, N. A. (2015). A review on predicting student performance using data mining techniques. *Procedia Computer Science*, 72, 414–422. https://doi.org/10.1016/j.procs.2015.12.157

[8] Aggarwal, C. C. (2015). *Data mining: The textbook*. Springer.

[9] Kumar, S., & Chadha, A. (2019). Machine learning techniques for intelligent decision support systems. *Journal of Intelligent Systems*, 28(3), 1–14. https://doi.org/10.1515/jisys-2017-0412

[10] Zhang, Y., & Li, X. (2021). Student performance prediction using ensemble machine learning techniques. *International Journal of Information and Learning Technology*, 38(2), 120–135. https://doi.org/10.1108/IJILT-05-2020-0071

[11] Sarker, I. H. (2021). Machine learning: Algorithms, real-world applications and research directions. *SN Computer Science*, 2(3), 160. https://doi.org/10.1007/s42979-021-00592-x

[12] Al-Barrak, M. A., & Al-Razgan, M. (2016). Predicting students' final GPA using decision trees. *International Journal of Information and Education Technology*, 6(7), 528–533.

[13] Fernández-Delgado, M., Cernadas, E., Barro, S., & Amorim, D. (2014). Do we need hundreds of classifiers to solve real-world classification problems? *Journal of Machine Learning Research*, 15(1), 3133–3181.

[14] Lakkaraju, H., Aguiar, E., Shan, C., Miller, D., Bhanpuri, N., Ghani, R., & Addison, K. (2015). A machine learning framework to identify students at risk of adverse academic outcomes. *Proceedings of the 21st ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1909–1918.

[15] Kotsiantis, S. B. (2012). Use of machine learning techniques for educational purposes: A decision support system for forecasting students' grades. *Artificial Intelligence Review*, 37(4), 331–344. https://doi.org/10.1007/s10462-011-9234-x

[16] Pandey, M., & Taruna, S. (2014). Towards the integration of multiple classifier techniques for the prediction of student performance. *Applied Soft Computing*, 22, 453–460. https://doi.org/10.1016/j.asoc.2014.04.002

[17] Hämäläinen, W., & Vinni, M. (2011). Comparison of machine learning methods for intelligent tutoring systems. *Proceedings of the International Conference on Educational Data Mining*, 525–534.