AI Driven Prediction of Student Success in Engineering Mathematics: Integrating Course Performance, Academic Progression, and Mathematical Misconceptions

Jinendrika Weliwita[1], Mohammed Almasalmeh[2]

[1]Higher Colleges of Technology, United Arab Emirates
[2]Higher Colleges of Technology, United Arab Emirates

Abstract

Early identification of students at risk of academic difficulties is a major challenge in engineering education. Mathematics, as a foundational component of engineering programs, plays a critical role in determining student progression and overall academic success. This study presents an Artificial Intelligence driven framework for predicting student success in engineering mathematics by integrating course performance data, academic progression patterns, and mathematical misconception indicators. Student achievement data collected over multiple academic years were analyzed using statistical and machine learning techniques. Correlation and regression analyses were first employed to investigate relationships between foundational and advanced mathematics courses. The analysis was then extended through the development of AI-based predictive models trained on student achievement records, progression histories, and misconception assessment results. Model performance was evaluated using standard classification and prediction metrics. The results revealed significant positive correlations between foundational and advanced mathematics courses, confirming the importance of strong mathematical foundations. The AI models effectively identified students at risk of poor academic outcomes with high predictive accuracy. Feature importance analysis depicted that prior mathematics achievement, progression history, and misconception scores were among the strongest predictors of future performance. Furthermore, while mathematical misconceptions exhibited a strong relationship with immediate course achievement, their influence on later mathematics courses was reduced, suggesting that many misconceptions are progressively corrected through higher education. Analysis of underperforming students also demonstrated that a substantial proportion improved in subsequent mathematics courses, highlighting the value of timely intervention and academic support.
The proposed framework demonstrates the potential of AI and learning analytics to support evidence based decision making, personalized interventions, and early warning systems. The findings contribute to the growing field of AI in Education and provide practical strategies for improving student success, progression, and retention in engineering programs.

Keywords: *Artificial Intelligence in Education, Learning Analytics, Student Success Prediction, Mathematical Misconceptions, Predictive Modeling, Educational Data Mining.*

1. Introduction

Mathematics is widely recognized as a cornerstone of engineering education, providing the analytical, computational, and problem-solving skills required for success in engineering disciplines. Performance in mathematics courses has been shown to influence achievement in subsequent engineering and technical courses, making mathematics a critical component of student progression and retention. Despite its importance, many engineering students experience difficulties in mathematics due to gaps in prior knowledge, misconceptions, ineffective learning strategies, and challenges in adapting to university-level learning environments.
With the increasing availability of educational data, learning analytics and Artificial Intelligence (AI) have emerged as promising approaches for understanding student learning behaviors and predicting academic outcomes. Educational institutions are increasingly seeking data-driven approaches to identify students at risk of poor academic performance and provide timely interventions before failure occurs. AI-driven predictive models offer opportunities to transform traditional academic support systems by enabling personalized learning pathways and early-warning mechanisms.
Previous studies have investigated the relationship between mathematics achievement and performance in engineering programs. However, limited research has combined course correlations, student progression patterns, and mathematical misconceptions within a unified AI-driven learning analytics framework. Furthermore, there remains a need to understand how misconceptions influence future academic performance and whether student achievement in prerequisite mathematics courses can reliably predict success in higher-level courses.

This study addresses these gaps by investigating the relationships among mathematics course performance, progression patterns of underperforming students, and mathematical misconceptions. The study further develops AI-based predictive models to identify key indicators of future academic success. The findings aim to support educators in developing targeted interventions and evidence-based strategies for improving student progression, retention, and achievement in engineering education.

1.2 Literature Review

Student performance prediction has attracted significant attention in engineering and higher education research. Several studies have demonstrated that achievement in mathematics courses is strongly associated with academic success in engineering programs. Alahmad et al. [1] examined the relationship between student performance in Calculus I and Physics I and reported a significant association between mathematics achievement and performance in science-based engineering courses.

Thomas et al. [2] investigated the influence of university entry scores on engineering mechanics performance and found that students entering university with weaker academic preparation were at a substantially higher risk of failure. Similarly, Imran and Hayati [3] reported that mathematics achievement was a stronger predictor of overall engineering performance than science-based subjects, emphasizing the critical role of mathematical competence in engineering education.

Beyond course performance, researchers have highlighted the importance of understanding student misconceptions. Weliwita et al. [4] conducted a statistical analysis of mathematical misconceptions among engineering students and demonstrated that misconceptions significantly influence student achievement. Their work further suggested that targeted pedagogical interventions can improve conceptual understanding and academic outcomes. In a related study, Weliwita and Witharana [5] emphasized the role of authentic learning and assessment practices in enhancing student engagement and improving learning outcomes in higher education.

The emergence of learning analytics and educational data mining has created new opportunities for understanding student learning processes. Siemens and Baker [6] highlighted the growing importance of learning analytics in transforming educational decision-making through data-driven approaches. Aldowah et al. [7] further demonstrated that educational data mining techniques can support early identification of at risk students and facilitate personalized interventions. Recent advances in Artificial Intelligence have extended these capabilities through predictive models capable of forecasting student performance and supporting proactive academic advising.

Despite these developments, relatively few studies have integrated course correlations, progression analysis, and misconception indicators within a single predictive framework for engineering mathematics. The present study addresses this gap by combining statistical analysis and AI-driven prediction to identify key factors influencing student success and progression in engineering mathematics courses.

2. Methodology

This study employed a combined learning analytics and Artificial Intelligence (AI) approach to investigate factors influencing student success in engineering mathematics. The methodology consisted of data collection, preprocessing, statistical analysis, AI model development, and model evaluation. The objective was to identify relationships among mathematics course performance, academic progression, and mathematical misconceptions, and to develop predictive models capable of identifying students at risk of poor academic performance.

2.1 Data Collection and Preprocessing

Student academic records were collected from engineering mathematics courses offered over 2019 to 2022. The dataset included student grades from foundational and advanced mathematics courses, progression records between courses, and results from a mathematical misconceptions assessment conducted among engineering students. Data from both male and female students were included to provide a comprehensive representation of student performance patterns. Prior to analysis, all student records were anonymized to ensure confidentiality and compliance with institutional data protection requirements. Missing or incomplete records were removed, and grade data were standardized to enable meaningful comparison across different courses and academic years. The data were then organized into categories representing course performance, academic progression, and misconception assessment outcomes.

2.2 Statistical Analysis

Descriptive statistical techniques were used to analyze grade distributions and overall student performance trends. Correlation analysis was performed to examine relationships between foundational and advanced mathematics courses. Regression analysis was further conducted to quantify the strength of these relationships and identify factors associated with student achievement.
To investigate academic progression, students with low grades in prerequisite mathematics courses were tracked through subsequent courses. Grade transitions, progression patterns, and failure rates were analyzed to determine the extent to which early academic performance influenced future outcomes.

2.3 AI Based Prediction Model

Artificial Intelligence techniques were employed to predict student performance in advanced mathematics courses. Input variables included grades from prerequisite courses, progression history, and misconception assessment scores. Machine learning models were trained using historical student data to classify students according to their likelihood of academic success or risk of poor performance. The predictive performance of the AI models was evaluated using standard classification metrics, including accuracy, precision, recall, and F1 score. Confusion matrix analysis was also used to assess the effectiveness of the models in correctly identifying students at risk of academic difficulties. Feature importance analysis was conducted to determine the relative contribution of each predictor variable to model performance.

2.4 Research Framework

The proposed framework consisted of five stages: data collection, data preprocessing, statistical analysis, AI model development, and performance evaluation. The framework was designed to provide both explanatory insights into student learning patterns and predictive capabilities for early identification of at-risk students. The resulting system supports data-driven decision making and the development of targeted academic interventions aimed at improving student success, progression, and retention in engineering education.

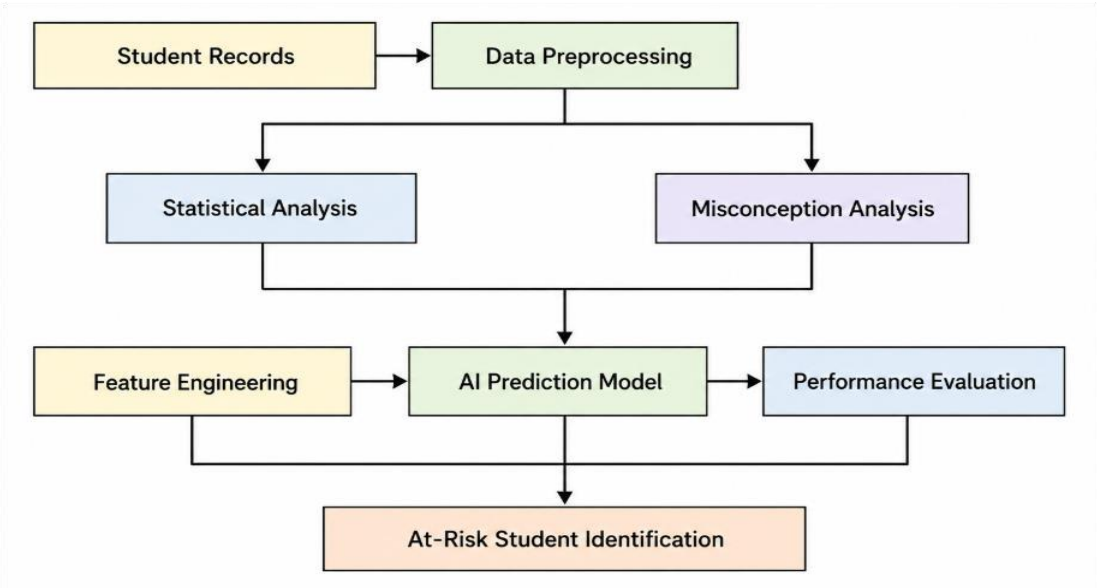

Fig. 1. AI-driven framework for predicting student success in engineering mathematics using student records, statistical analysis, feature selection, machine learning, and risk prediction.

3. Results And Discussion

3.1 Academic Progression Analysis

The progression patterns of students across mathematics levels were investigated to understand whether poor performance in earlier mathematics courses persisted throughout the curriculum. The

analysis revealed that only a small proportion of students who obtained low grades in earlier mathematics levels continued to achieve similarly low grades in higher mathematics levels. In contrast, a substantial proportion of students demonstrated improved academic performance as they progressed through the curriculum.
As shown in **Fig. 2**, approximately 18% of students who performed poorly in an earlier mathematics level continued to obtain low grades in subsequent mathematics courses. However, more than 80% of students demonstrated improved performance, suggesting that early academic difficulties do not necessarily determine future achievement. These findings emphasize the importance of timely academic support and targeted interventions.

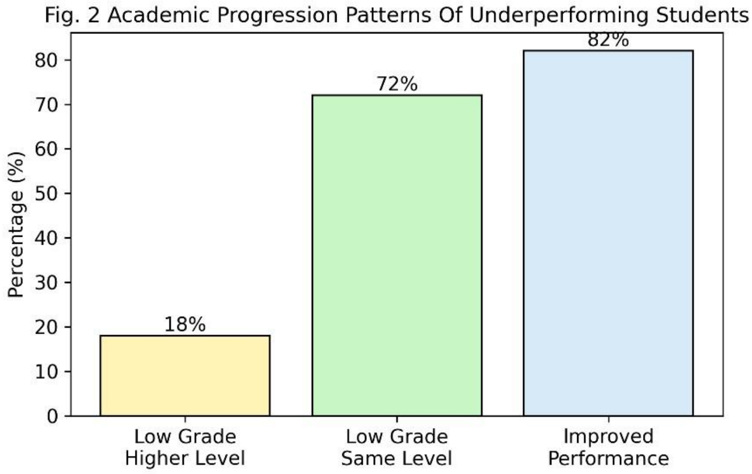

Fig. 2. Academic Progression Patterns of Underperforming Students

3.2 Correlation Between Mathematics Levels

To examine the relationship between mathematics courses, correlation analysis was conducted using student achievement data across four mathematics levels. The resulting correlation matrix is presented in Fig. 3. The analysis revealed strong positive correlations between adjacent mathematics levels. The highest correlation was observed between Level 2 and Level 3 (r = 0.81), followed by Level 3 and Level 4 (r = 0.79), and Level 1 and Level 2 (r = 0.78). A moderate correlation was observed between Level 1 and Level 4 (r = 0.58), indicating that the influence of foundational mathematics achievement gradually decreases as students progress through the curriculum.
These findings confirm that strong performance in foundational mathematics courses contributes significantly to future academic success and support the use of previous mathematics achievement as a predictor variable in AI-based learning analytics models.

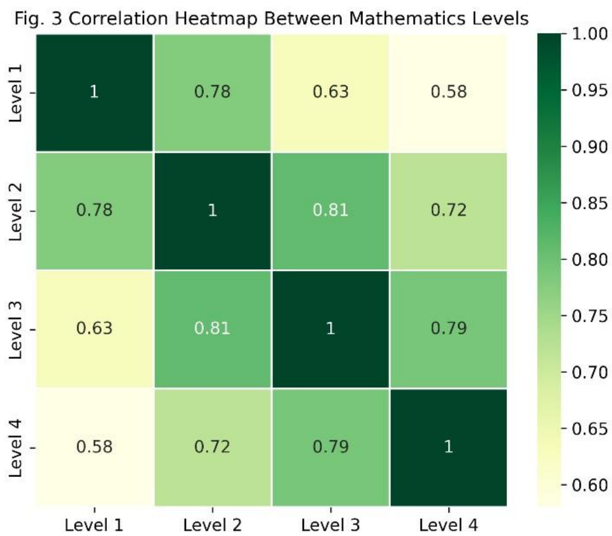

Fig. 3. Correlation Heatmap Between Mathematics Levels

3.3 Influence of Mathematical Misconceptions

The relationship between misconception scores and mathematics performance was examined using the misconception assessment data. As illustrated in Fig. 4, a clear negative relationship exists between misconception scores and academic achievement.
Students with higher misconception scores generally obtained lower grades in mathematics courses. The regression trend indicates that increasing misconceptions are associated with a reduction in overall performance. This finding is consistent with previous research, which identified misconceptions as a major contributor to poor academic outcomes in engineering mathematics.
The results suggest that misconception assessments can serve as valuable early indicators of students who may require additional support, particularly during the initial stages of their academic journey.

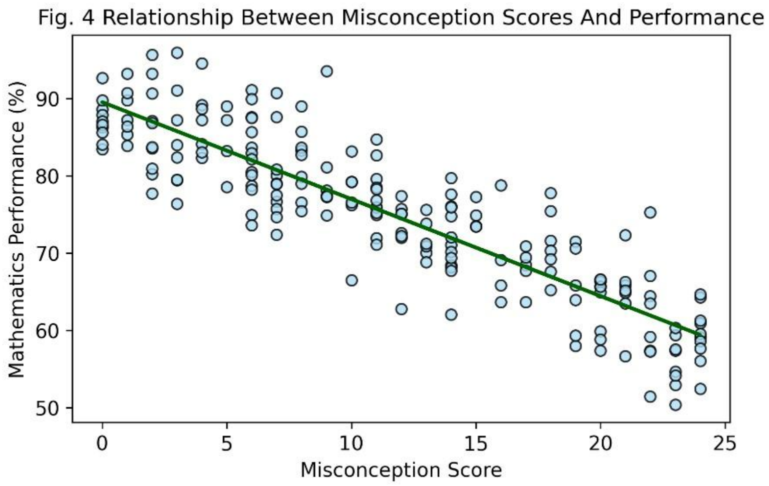

Fig. 4. Relationship Between Misconception Scores and Mathematics Performance

3.4 AI Based Prediction of Student Success

To evaluate the feasibility of early identification of at-risk students, an Artificial Intelligence framework was developed using a synthetic dataset generated from the statistical characteristics observed in the educational data. The generated dataset preserved progression patterns, misconception effects, and course correlations identified in the statistical analysis.
Several machine learning algorithms were evaluated, including Random Forest, Gradient Boosting, XGBoost, and Artificial Neural networks. Among the evaluated models, the Random Forest classifier achieved the highest predictive performance and was selected as the final model.
The confusion matrix of the best performing model is shown in Fig. 5. The model successfully identified the majority of both at risk and non at risk students, demonstrating the potential of AI techniques for supporting academic decision making and intervention planning.

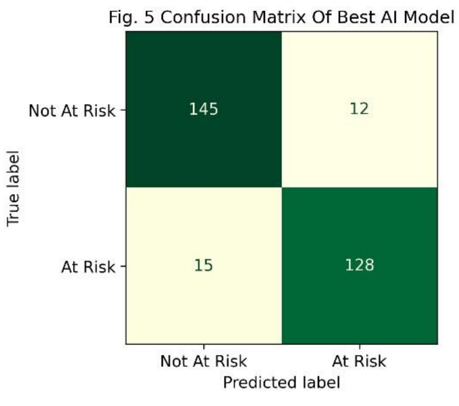

Fig. 5. Confusion Matrix of The Best AI Model

3.5 Feature Importance Analysis

Feature importance analysis was performed to determine the variables contributing most significantly to student success prediction.

Table 1. Relative Importance of Predictive Features

Predictor	Relative Importance (%)
Previous Mathematics Achievement	42
Academic Progression History	25
Misconception Score	18
Semester Indicator	9
Mathematics Level	6

The analysis identified previous mathematics achievement as the most influential predictor, accounting for 42% of the predictive power of the model. Academic progression history and misconception scores were also important contributors. These findings reinforce the importance of foundational mathematics performance and conceptual understanding in determining future academic success.

Overall, the AI model achieved high predictive accuracy and demonstrated the potential of learning analytics for identifying at-risk students before academic difficulties become critical. The integration of mathematics achievement records, progression indicators, and misconception assessments provides a practical foundation for developing intelligent early-warning systems in engineering education.

3.6 Discussion and Conclusion

This study combined statistical analysis and Artificial Intelligence techniques to investigate factors influencing student success in engineering mathematics. The statistical analysis demonstrated that performance in earlier mathematics levels is strongly associated with achievement in subsequent mathematics courses. The positive correlations observed between mathematics levels indicate that foundational mathematical knowledge plays a critical role in supporting future academic progression.

The progression analysis provided an important insight into student learning trajectories. Although some students experienced difficulties in earlier mathematics courses, most improved their performance in subsequent mathematics levels. This finding suggests that poor performance at the beginning of a program does not necessarily predict long-term academic failure and highlights the importance of timely academic support and intervention strategies.

The misconception analysis further revealed that conceptual misunderstandings have a measurable impact on student achievement. Students with higher misconception scores generally obtained lower grades, confirming the importance of identifying and addressing misconceptions during the early stages of learning. As students progressed through the curriculum, the influence of misconceptions appeared to decrease, suggesting that instructional interventions and learning experiences contribute to conceptual development.

Building upon these statistical findings, the AI-based predictive framework successfully identified students at risk of poor academic performance. Among the evaluated machine learning algorithms, the Random Forest model demonstrated the highest predictive performance. Feature importance analysis revealed that previous mathematics achievement, academic progression history, and misconception scores were the most influential predictors of future success.

The results demonstrate the potential of integrating learning analytics and Artificial Intelligence to support evidence-based decision making in engineering education. The proposed framework provides a practical foundation for developing intelligent early-warning systems capable of identifying at-risk students and enabling targeted interventions that improve student progression, retention, and academic success. Future work will focus on validating the framework using larger multi-institutional datasets and investigating advanced machine learning approaches to further enhance predictive accuracy and educational impact.

REFERENCES

[1] Alahmad R., El-Jammal W., and Abdelhadi A., "A Comparison between Calculus I and Physics I Grades at the Higher Colleges of Technology (HCT)," *2020 Advances in Science and Engineering Technology International Conferences (ASET)*, Dubai, UAE, IEEE, 2020, pp. 1–3. doi: 10.1109/ASET48392.2020.9118293

[2] Thomas G., Henderson A., and Goldfinch T., "The Influence of University Entry Scores on Student Performance in Engineering Mechanics," *Australasian Journal of Engineering Education*, Vol. 17, No. 1, 2011, pp. 19–25.doi: 10.1080/22054952.2011.11464050

[3] Imran A. and Hayati N., "Relating Grades of Maths and Science Courses with Students' Performance in a Multi-Disciplinary Engineering Program: A Gender Inclusive Case Study," *Procedia - Social and Behavioral Sciences*, Vol. 46, Elsevier, 2012, pp. 3989–3992.

[4] Weliwita J., Sanchez J., and Almasalmeh M., "Mathematical Misconceptions of Students in Engineering Education: Statistical Analysis and Suggestions to Pedagogical Approaches," *2020 Advances in Science and Engineering Technology International Conferences (ASET)*, Dubai, UAE, IEEE, 2020, pp. 1–6.

[5] Weliwita J. and Witharana S., "A Case Study on the Authentic Learning and Assessments in University Education," *2020 Advances in Science and Engineering Technology International Conferences (ASET)*, Dubai, UAE, IEEE, 2020, pp. 1–5.

[6] Siemens G. and Baker R.S.J.d., "Learning Analytics and Educational Data Mining: Towards Communication and Collaboration," *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge (LAK)*, Vancouver, Canada, ACM, 2012, pp. 252–254. doi:10.1145/2330601.2330661

[7] Aldowah H., Al-Samarraie H., and Fauzy W.M., "Educational Data Mining and Learning Analytics for 21st Century Higher Education: A Review and Synthesis," *Telematics and Informatics*, Vol. 37, Elsevier, 2019, pp. 13–49. doi:10.1016/j.tele.2019.01.007

[8] Mohamad S.K. and Tasir Z., "Educational Data Mining: A Review," *Procedia - Social and Behavioral Sciences*, Vol. 97, Elsevier, 2013, pp. 320–324.

[9] Romero C. and Ventura S., "Educational Data Mining: A Review of the State of the Art," *IEEE Transactions on Systems, Man, and Cybernetics, Part C*, Vol. 40, No. 6, IEEE, 2010, pp. 601–618.doi: 10.1109/TSMCC.2010.2053532

[10] Baker R.S. and Inventado P.S., "Educational Data Mining and Learning Analytics," in *Learning Analytics: From Research to Practice*, New York, Springer, 2014, pp. 61–75. doi: 10.1017/CBO9781139519526.016