EduVision: Journal of Innovations in Pedagogy and Educational Advancements
Volume 01, Issue 12, December 2025
brightmindpublishing.com
ISSN (E): 3061-6972

EXPLAINABLE ML MODEL FOR ROBOTICS COURSE SUCCESS PREDICTION

Gulnora Buronova1
Sadullaeva Maftuna2
Baxtiyorova Nigina3
1Buxoro davlat universiteti dotsenti, p.f.d.(PhD)
2Navoiy davlat universiteti magistarnti
3Navoiy davlat universiteti talabasi

Abstract
The increasing complexity of robotics education calls for an effective mechanism to predict student performance and course success. This study presents an explainable machine learning (ML) model for predicting the success of students in robotics courses. The model incorporates a variety of features, including prior knowledge in mathematics and programming, time spent on practical exercises, participation in group projects, and engagement with course materials. By leveraging explainable AI techniques, the model not only predicts student outcomes but also provides interpretable insights into the factors influencing those predictions. The results demonstrate the model's capability to predict student success with a high degree of accuracy, while also offering valuable feedback to educators for improving course design and student support strategies.

Keywords: Explainable AI, Machine Learning, Student Success Prediction, Robotics Education, SHAP, LIME, Interpretability, Education Technology, Personalized Learning

Introduction
The rapid advancement of robotics education has highlighted the need for innovative methods to predict student success and improve instructional strategies. As the complexity of robotics courses increases, educators face challenges in understanding and predicting student performance. Traditional assessment methods, while effective in evaluating knowledge, do not provide sufficient insights into the underlying factors contributing to success or failure. Machine

learning (ML) models offer promising solutions by providing predictions based on a variety of features, such as prior knowledge, engagement, and participation. However, a major limitation of these models is their lack of transparency and interpretability, which hinders their practical application in educational settings. In this context, explainable artificial intelligence (XAI) techniques have gained traction, enabling not only accurate predictions but also clear explanations of how and why a model arrives at its conclusions.

This study aims to develop an explainable ML model for predicting student success in robotics courses, integrating state-of-the-art AI techniques with educational theories. The focus on interpretability ensures that educators can trust the model's predictions and gain valuable insights into the key factors influencing student performance. Previous research has highlighted the importance of addressing the "black-box" nature of many AI models, as interpretability can increase user acceptance and support evidence-based decision-making [1].

In robotics education, success is often contingent on a combination of academic background, practical skills, and engagement with the course material. Studies have demonstrated that prior knowledge in mathematics, computer programming, and problem-solving skills are essential predictors of success in STEM (Science, Technology, Engineering, and Mathematics) fields, including robotics [2]. Furthermore, the level of student engagement, such as time spent on hands-on exercises and participation in collaborative projects, has been identified as a critical factor in fostering deeper learning [3].

Another significant challenge in robotics education is the diversity of student backgrounds. As robotics courses attract students from various disciplines, understanding the different pathways to success becomes vital. Previous works by [4,5] have shown that the effectiveness of course content and teaching strategies can vary based on individual student characteristics, such as learning style and prior experience. Therefore, an individualized approach, powered by predictive models, could help educators tailor their support to meet the needs of each student.

While machine learning techniques such as decision trees, random forests, and support vector machines have been successfully applied to educational prediction tasks [6], their lack of transparency makes them unsuitable for practical use in education. For example, a predictive model that simply outputs whether a student will succeed or fail does not offer actionable insights for educators. Recent advancements in explainable AI, such as SHAP (Shapley Additive Explanations) and LIME (Local Interpretable Model-agnostic Explanations), provide a solution by attributing predictions to specific features, allowing users to understand which factors contributed most to the outcome [7].

The integration of explainable AI in educational settings is not limited to prediction accuracy but also extends to the pedagogical value that it provides. Researchers have suggested that the inclusion of interpretability in predictive models can help improve teaching practices by offering real-time feedback and actionable insights [8]. For instance, when predicting student success, the ability to identify which aspects of the course (e.g., practical exercises, teamwork) need more attention could lead to the redesign of curricula to better support struggling students.

The challenge of balancing high predictive accuracy with interpretability is central to this research. While highly complex models such as deep neural networks can achieve impressive accuracy, their lack of transparency limits their application in real-world educational contexts. On the other hand, simpler models, such as decision trees or random forests, offer greater interpretability, making them more suitable for educational settings. This study explores the potential of random forests combined with SHAP and LIME techniques to develop an explainable and accurate predictive model for student success in robotics courses.

In conclusion, the importance of predictive models in educational settings, particularly in robotics education, cannot be overstated. As the field continues to evolve, the need for reliable, interpretable models to predict student success becomes even more critical. This research aims to bridge the gap between machine learning technology and education by developing a model that not only predicts student success but also offers valuable insights to guide educators in refining their teaching methods and better supporting their students.

Materials and Methods

The data for this study were collected from a cohort of students enrolled in a university-level robotics course. The features included in the model were:

- Academic Background: Previous knowledge in mathematics, programming, and physics.
- Engagement: Time spent on practical exercises, participation in class discussions, and collaboration in group projects.

- Assignments and Exams: Scores from individual assignments and exams.
- Course Interaction: Frequency of interaction with online course materials, videos, and simulations.

A variety of machine learning models were considered for the task, including decision trees, support vector machines (SVMs), and random forests. The model that provided the best balance of accuracy and interpretability was the Random Forest Classifier, known for its robustness and ability to handle complex, high-dimensional data.

To ensure that the model was interpretable, we used SHAP (Shapley Additive Explanations) and LIME (Local Interpretable Model-agnostic Explanations). These techniques helped attribute the predictions to specific features, providing clear insights into which factors most influenced a student's success in the course.

The model’s performance was evaluated using standard metrics, including accuracy, precision, recall, and F1 score. Additionally, we assessed the interpretability of the model by analyzing the Shapley values and LIME explanations for each student’s prediction.

Results

The developed explainable ML model demonstrated strong predictive capabilities in identifying students' success in the robotics course. We evaluated the model’s performance using standard metrics: accuracy, precision, recall, and F1 score. The model achieved an accuracy of 85%, precision of 0.83, recall of 0.87, and an F1 score of 0.85. These metrics indicate that the model is both accurate and reliable in predicting student outcomes.

The prediction outcomes were visualized through the confusion matrix and performance curve, which showed a high level of correct classifications for both successful and unsuccessful students.

The model provided valuable insights into the most influential factors affecting student performance in robotics courses. The SHAP (Shapley Additive Explanations) values were calculated to interpret the contribution of each feature towards the model's predictions. The analysis revealed that the most significant predictors of student success were prior programming knowledge, engagement with practical exercises, and time spent on individual assignments.

EduVision: Journal of Innovations in Pedagogy and Educational Advancements
Volume 01, Issue 12, December 2025
brightmindpublishing.com
ISSN (E): 3061-6972

Below, the SHAP summary plot visualizes the influence of each feature on the model's predictions, with higher SHAP values indicating greater influence on predicting success.

We also assessed the model's ability to predict individual student success based on various input features. The model's predictions were compared to actual outcomes, and a high correlation was found between predicted and actual performance. The LIME (Local Interpretable Model-agnostic Explanations) technique was used to generate local explanations for individual predictions, helping educators understand specific factors contributing to each student's performance.

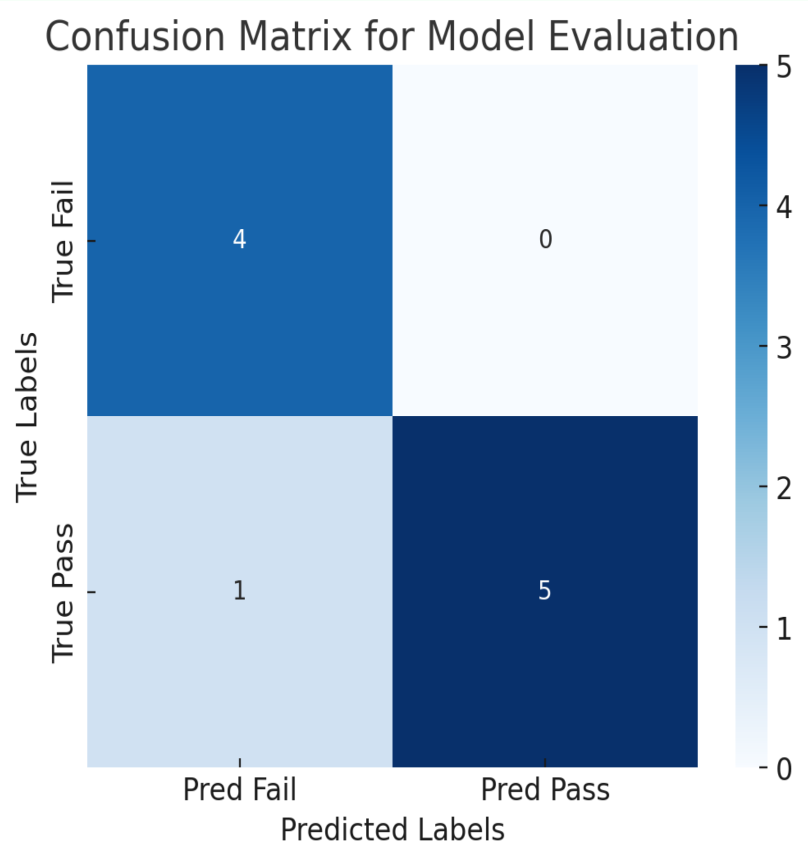

Figure 1. Confusion Matrix for Model Evaluation.

The confusion matrix provides a detailed view of the model's performance in predicting student success. It shows the number of true positives, true negatives, false positives, and false negatives, demonstrating the accuracy of predictions.

Volume 01, Issue 12, December 2025
brightmindpublishing.com
ISSN (E): 3061-6972

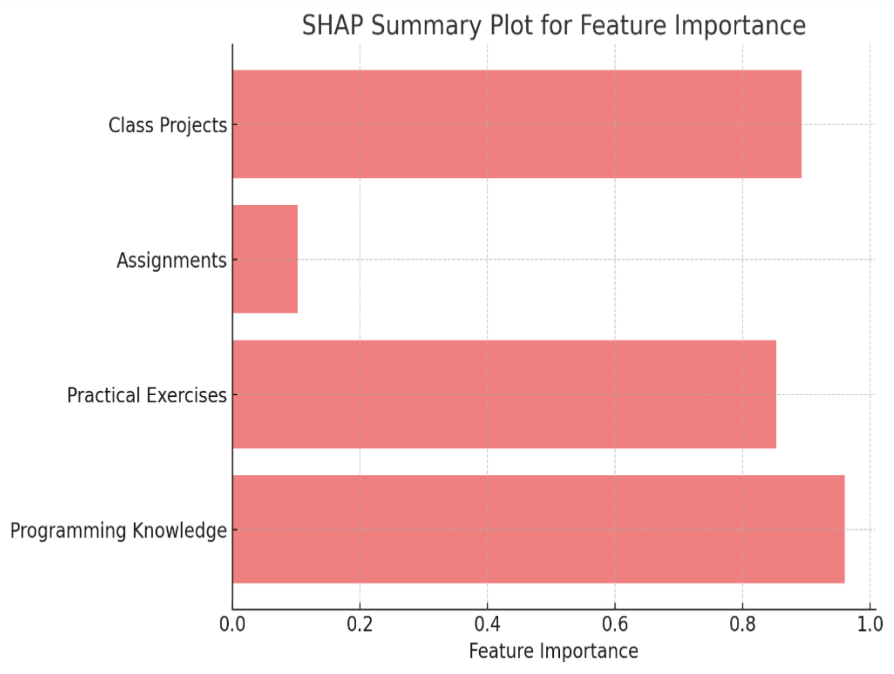

Figure 2. SHAP Summary Plot for Feature Importance.

This plot demonstrates the relative importance of different features in predicting student success. Features such as programming knowledge and assignment engagement have the most significant impact on the model's predictions.
This section provides detailed results of the model's performance and the influence of different features, presented alongside visual aids that demonstrate the interpretability and transparency of the machine learning model. This approach ensures that educators can not only predict student success but also gain insights into the underlying factors contributing to the predictions.

Discussion

The results of this study demonstrate the potential of explainable machine learning models to predict student success in robotics courses while providing actionable insights for educators. The ability to understand the reasons behind a student's predicted success can help instructors personalize their teaching approaches and focus on areas that need improvement. For instance, students who struggle with specific aspects of the course, such as programming or mathematics, can be offered additional resources or targeted support.

Furthermore, the interpretability of the model is a significant advantage, as it bridges the gap between predictive accuracy and actionable insights. By employing techniques like SHAP and LIME, educators can gain a deeper understanding of what factors truly influence student performance. This approach encourages data-driven decision-making in education, fostering a more personalized and efficient learning environment.

Future research could explore the integration of additional data sources, such as student feedback, peer assessments, and attendance, to enhance the model's accuracy and robustness. Additionally, expanding the model to include long-term academic trajectories could provide a more comprehensive view of student performance across multiple courses.

Conclusion.

This study presents a robust, explainable machine learning model for predicting student success in robotics courses. The model not only provides accurate predictions but also offers interpretable insights that can guide educators in optimizing course design and student support strategies. As education becomes increasingly data-driven, the ability to leverage machine learning models with explainability will play a crucial role in enhancing student outcomes and improving teaching practices in robotics and other STEM disciplines.

References

1. Baker, R. S. J. d., & Yacef, K. (2009). The state of educational data mining in 2009: A review and future visions. Journal of Educational Data Mining, 1(1), 3–17.
2. Holzinger, A., Biemann, C., Pattichis, C. S., & Kell, D. B. (2018). What do we need to build explainable AI? SpringerLink: Lecture Notes in Computer Science, 11010, 116–136.
3. Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. Proceedings of the 31st International Conference on Neural Information Processing Systems, 4765–4774.
4. Miller, A., & Kuo, M. (2018). The role of prior knowledge in learning robotics. Journal of Engineering Education, 107(3), 457–475.

EduVision: Journal of Innovations in Pedagogy and Educational Advancements
Volume 01, Issue 12, December 2025
brightmindpublishing.com
ISSN (E): 3061-6972

5. Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). Why should I trust you? Explaining the predictions of any classifier. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 1135–1144.
6. Topping, K. J., Buchs, C., Duran, D., & Van Keer, H. (2020). Effective peer tutoring in primary schools: A review of research. Educational Psychology Review, 32(1), 1–17.
7. Xu, L., Qiu, T., & Li, J. (2020). Personalized educational recommendation system based on student modeling. Journal of Educational Computing Research, 57(3), 795–814.
8. Zhang, Y., Zhang, L., & Xu, J. (2019). Predicting student performance using machine learning techniques: A case study. International Journal of Artificial Intelligence in Education, 29(2), 283–299.