International Research Journal of Education and Technology
Peer Reviewed Journal
ISSN 2581-7795

STUDENT PERFORMANCE PREDICTION USING ML MODEL

KANISHKA T, PREETHI P, MOHAMMED AASIM KHAN I, LAKSHMANA KUMAR P
Department of Artificial Intelligence and Machine Learning, Bannari Amman Institute of Technology, Sathyamangalam, Erode, Tamilnadu, India

---

-

ABSTRACT:

In the evolving landscape of education, online assessments have gained significant traction, offering unparalleled flexibility and accessibility. However, these assessments pose challenges, particularly in accurate grading and identifying students needing additional support. This project proposes a web-based automatic test grading system that leverages decision tree algorithms to address these issues. By automating grading and integrating predictive analytics, the system provides educators with a robust tool to identify at-risk students and deliver tailored interventions effectively. The system operates by collecting and analysing student data, extracting meaningful features, and utilizing decision tree models to predict performance outcomes. Students identified as at-risk are assigned retests featuring simpler questions, designed to reinforce foundational concepts and provide them with an opportunity to improve. The decision tree algorithm is well-suited for this application due to its simplicity, interpretability, and ability to handle diverse data types. Moreover, its transparent decision-making process ensures educators can understand and trust the predictions, facilitating better-targeted support for students. By integrating this system into a web-based platform, educators can streamline test administration and grading while gaining actionable insights into student performance. The predictive and adaptive nature of the solution fosters an inclusive environment by addressing individual student needs, reducing educator workload, and enhancing overall academic outcomes. This innovative approach not only modernizes the assessment process but also contributes to a more equitable and effective educational experience.

Keywords: Online assessments, Automatic test grading, Decision tree algorithm, Student performance prediction , At-risk students, Machine learning in education Adaptive learning, Web-based grading system.

1. INTRODUCTION:

Student assessment refers to the process of measuring student learning and achievement in a particular subject or skill. It involves evaluating student performance, providing feedback, and identifying areas where students may need additional support or instruction. Assessment can take many forms, including written exams, projects, oral presentations, and performance tasks. The main purpose of student assessment is to provide teachers with information about what students know and understand, as well as their strengths and weaknesses. This information can be used to adjust instruction and curriculum to better meet the needs of individual students, and to evaluate the effectiveness of teaching strategies and materials.

Assessment can also be used to motivate students to learn and to help them set goals for their own learning. By providing feedback and recognition for their achievements, students can feel a sense of accomplishment and progress towards their educational goals. Effective student assessment should be aligned with learning objectives and standards, fair and unbiased, and provide clear and specific feedback to students. It should also be ongoing and formative, allowing teachers to adjust instruction and provide additional support as needed.

2. LITERATURE SURVEY:

AUTHOR: Ziling Chen; Gang Cen 2023, Traditional approaches to predicting student performance often overlook the importance of individual characteristics, focusing primarily on common trends within student groups. However, these individual differences play a significant role in shaping academic outcomes and require tailored interventions to support student success. Existing methods also struggle with effectivel analyzing multi-dimensional discrete data, hindering accurate performance predictions. Thus, there is a pressing need for a predictive model that considers both common trends and individual characteristics while addressing the challenges associated with complex data analysis.

Daozong Sunand and Rongxin Luo 2023, Predicting student performance in higher education is essential for effective student management and teaching strategies. However, relying solely on past academic records for prediction may not be ideal due to the limited scope of collected factors and the challenge of maintaining prediction accuracy. Thus, there' a need to develop a model that can predict university student performance by considering multiple factors beyond past academic performance, while also addressing the difficulty of data collection.

Siti Dianah Abdul Bujang 2021, In higher educational institutions, the implementation of predictive analytics has become increasingly crucial for monitoring student academic performance. However, a significant challenge arises in handling imbalanced datasets when predicting student grades. Imbalanced datasets can lead to overfitting and misclassification issues, reducing the predictive accuracy of models. Therefore, there is a need to develop a predictive model that can effectively address imbalanced datasets and provide accurate predictions of student grades.

Abdullah Alshanqiti and Abdallah Namoun2020, In the realm of higher education, accurately understanding and predicting student performance is paramount for program leaders to identify areas for improvement and implement effective interventions. However, existing studies often overlook the identification of key factors influencing student performance, hindering the ability to address underlying issues comprehensively. This gap highlights the need for a holistic approach that not only predicts academic outcomes but also identifies the factors contributing to these outcomes.

The paper employs both clustering and classification techniques in its analysis of student performance. Clustering involves using the T-SNE algorithm for dimensionality reduction and examining the relationship between various factors (admission scores, first-level courses, AAT, GAT) and GPA. Classification experiments utilize different machine learning models to predict student performance, incorporating features such as course grades and admission test scores. The performance of these models is evaluated using various assessment metrics to determine their effectiveness in predicting student outcomes.

3. METHODOLOGY:

To address the identified challenges, the study proposes a two-step methodology. First, it combines the relationship matrix-based bipartite network approach (RMBN) with Louvain clustering to effectively group students based on shared characteristics while accommodating individual differences. This step enhances the understanding of the relationships between different student attributes and their impact on academic performance. Second, the study develops a hybrid neural network model based on a relationship matrix (RMHNN) to overcome the limitations of fitting discrete types of features using traditional algorithms. By incorporating a relationship matrix, the model captures complex relationships between student attributes and their influence on performance, leading to more accurate predictions.

The proposed model employs multi-feature fusion and attention mechanisms to predict university student performance. It analyzes historical academic grades from multiple dimensions to extract relational features using methods such as representation learning and collaborative filtering. The attention mechanism is introduced to automatically extract important features and explore relationships between different dimensional features. The study collects a triplet set of related courses and students' real historical grades, conducts data analysis to prove correlations

between courses, and verifies the effectiveness of different dimensional features through experimentation.

The paper compares the accuracy performance of six machine learning techniques, namely Decision Tree (J48), Support Vector Machine (SVM), Naïve Bayes (NB), K-Nearest Neighbor (kNN), Logistic Regression (LR), and Random Forest (RF), using a real dataset comprising 1282 students' course grades. Additionally, a multiclass prediction model is proposed to address imbalanced multi-classification by integrating oversampling Synthetic Minority Oversampling Technique (SMOTE) with two feature selection methods. The performance of the proposed model is evaluated and compared with other approaches to assess its effectiveness in predicting student grades.

To achieve the objectives, the paper introduces a hybrid regression model that combines collaborative filtering, fuzzy set rules, and Lasso linear regression techniques. This model is designed to predict future grades in various courses, providing a comprehensive understanding of student academic progress. Additionally, an optimized multi-label classifier is proposed predict the qualitative values for the influence of various factors associated with student performance. This classifier leverages an optimized self-organizing map to identify dominant factors impacting performance. The methodology is empirically validated using seven publicly available datasets, allowing for a rigorous evaluation of the proposed approach.

The paper employs both clustering and classification techniques in its analysis of student performance. Clustering involves using the T-SNE algorithm for dimensionality reduction and examining the relationship between various factors (admission scores, first-level courses, AAT, GAT) and GPA. Classification experiments utilize different machine learning models to predict student performance, incorporating features such as course grades and admission test scores. The performance of these models is evaluated using various assessment metrics to determine their effectiveness in predicting student outcomes.

FLOW CHART:

4. PROPOSED SOLUTION:

The proposed solution leverages machine learning, specifically a decision tree algorithm, to predict student performance and provide targeted academic support. The system begins by collecting and preprocessing student data, including test scores, attendance, and participation levels, ensuring data consistency and accuracy. Key features such as question difficulty, response time, and engagement levels are extracted to train the model. The decision tree algorithm is then applied to classify students into performance categories—high-performing, average, or at-risk—based on their learning patterns.

Once at-risk students are identified, the system assigns them adaptive retests with simpler questions to reinforce fundamental concepts and improve their understanding. Educators receive real-time insights through a web-based platform, allowing them to monitor student progress and implement timely interventions. The platform automates test grading, reducing the manual workload for educators while ensuring objective evaluations.

By integrating predictive analytics, the system fosters a personalized learning environment, enabling educators to offer data-driven academic support. This approach not only enhances student learning outcomes but also contributes to a more efficient and inclusive educational system.Overall, the *Placement Training Details Management System* provides an efficient, data- driven, and automated approach to placement training, ensuring students are well-prepared for recruitment opportunities while reducing the administrative burden on faculty and training coordinators.

5. CONCLUSION:

In conclusion, the proposed web-based student performance prediction system utilizing a decision tree algorithm offers a powerful solution to modern educational challenges. By automating test grading and leveraging predictive analytics, the system ensures accurate performance evaluation while identifying at-risk students who need additional support.

The adaptive retesting mechanism provides these students with an opportunity to strengthen their foundational knowledge, promoting a more inclusive learning environment. Educators benefit from real-time insights, enabling them to make data-driven decisions and implement timely interventions. Additionally, the transparency and interpretability of the decision tree model ensure trust in the predictions, making it easier for educators to understand and act upon the results.

The system not only reduces the manual workload of grading but also enhances the overall efficiency of student assessment and academic support. By integrating this intelligent approach into educational institutions, the platform fosters personalized learning experiences, improves student outcomes, and contributes to a fair and effective evaluation system.

As education continues to evolve, leveraging machine learning for performance prediction represents a significant step toward data-driven teaching methodologies, ensuring that students receive the guidance and resources they need to succeed in their academic journey.

6. REFERENCES:

1. G. Haldeman, M. Babeş-Vroman, A. Tjang and T. D. Nguyen, "CSF: Formative feedback in autograding", ACM Trans. Comput. Educ., vol. 21, no. 3, pp. 1-30, Sep. 2021.
2. A. Yadav, C. Mayfield, S. K. Moudgalya, C. Kussmaul and H. H. Hu, "Collaborative learning self-efficacy and student performance in CS1 POGIL", Proc. 52nd ACM Tech. Symp. Comput. Sci. Educ., pp. 775-781, 2021.
3. M. Yağcı, "Educational data mining: Prediction of students' academic performance using machine learning algorithms", Smart Learn. Environ., vol. 9, no. 1, pp. 11, Dec. 2022.
4. S. Gaftandzhieva, A. Talukder, N. Gohain, S. Hussain, P. Theodorou, Y. K. Salal, et al., "Exploring online activities to predict the final grade of student", Mathematics, vol. 10, no. 20, pp. 3758, Oct. 2022.
5. X. Wang, X. Yu, L. Guo, F. Liu and L. Xu, "Student performance prediction with short-term sequential campus behaviors", Information, vol. 11, no. 4, pp. 201, Apr. 2020.