Student Performance Predictor: A Machine Learning Approach with Real-Time Web Deployment

Aaryan Kumar
Ajay Kumar Garg Engineering College, Ghaziabad, India
Email: kumaraaryan498@gmail.com

Internship: College Summer Break Internship
Domain: Machine Learning

Mentor: Dr. Pawan Kumar
Department of Computer Science & Engineering
Ajay Kumar Garg Engineering College
Ghaziabad, UP, India
Email: drpawancse@gmail.com

***Abstract*—This research presents the design and deployment of a machine learning-based system for predicting student performance using academic and demographic data. The proposed system applies supervised learning techniques to classify student outcomes (Pass/Fail) while also predicting letter grades (A–F). The solution is deployed as a real-time interactive web application using Streamlit, enabling accessibility for educators and students. Results demonstrate effective performance with accuracy exceeding 80%, supported by visualizations that enhance interpretability.**

***Index Terms*—Student performance prediction, machine learning, Streamlit, web deployment, education analytics**

I. Introduction

Student academic performance prediction has become a critical area in educational research, promising timely interventions that bolster learning outcomes and reduce dropout rates. Educational Data Mining (EDM) has emerged as a pivotal discipline over the past two decades, fueled by the in-creasing availability of fine-grained student data and the growing capacity of machine learning (ML) to process it. Early studies concentrated on identifying at-risk students using university datasets or data from MOOCs, applying ML methods such as logistic regression, decision trees, and support vector machines (SVMs).

Over time, the domain has witnessed a shift toward richer algorithms and application contexts. Systematic reviews spanning from 2010 to 2020 reveal that classification models dominate the research landscape—accounting for around 62% of studies—while regression-based methods account for the remainder. Moreover, deep learning models and neural networks, though less prevalent historically, are gaining traction for their accuracy and pattern recognition capabilities.

Adaptive and personalized learning systems are redefining the educational experience. Adaptive learning uses AI-driven algorithms to tailor content to each learner's needs, with studies reporting positive learning outcomes in 86% of cases. Simultaneously, learning analytics provides educators with dashboards and visual insights to monitor student behavior and proactively intervene—focusing on dropout prevention, personalized assessments, and instructional effectiveness.

Against this developing backdrop, our research presents a comprehensive system that not only predicts student pass/fail status but also forecasts letter grades (A–F). By embedding the system into a Streamlit-based web application, we bridge the gap between offline model experimentation and real-time, interactive deployment—making insights accessible to educators and students alike. Our work aligns with current trends in using ML as a tool for actionable, interpretable guidance in education.

II. Related Work

Literature on predicting student performance is vast and multi-layered. Systematic reviews of studies from 2009 to 2021 highlight the prevalence of ML approaches—especially in detecting at-risk and dropout students—using datasets from institutional records and online platforms. Classical algorithms such as logistic regression, decision trees, SVM, and random forests remain staples due to their interpretability and grounded performance.

Surveys also shed light on the evolution of modeling strategies. Between 2010 and 2020, classification-focused studies represented a significant majority (62%), with a smaller but notable portion employing regression techniques. While neural networks and deep learning have historically been underrepresented, recent reviews indicate they're increasingly adopted for their superior accuracy in predicting student outcomes.

Beyond core predictive models, researchers are exploring advanced architectures to enhance interpretability and stability. The ESPA model (Explainable Student Performance Prediction with Personalized Attention) uses BiLSTM and attention layers to justify its predictions—identifying precisely which student behaviors or course-related attributes are influential. Similarly, a graph-based ensemble approach propagates insights over bipartite student–performance graphs, achieving up to 14.8% higher accuracy than traditional ML models.

Contemporary research also emphasizes integrating behavioral data and clustering to improve prediction outcomes. A framework combining learning behavior analysis with ML models like XGBoost, alongside SHAP-based feature importance, demonstrates almost perfect accuracy for certain learner cohorts and robust outcomes for others.

In student performance prediction domains, deep learning models remain compelling. A study from 2025 using multi-layer perceptron classifiers (MLPC) achieved a remarkable 86.46% test accuracy and 79.58% under 10-fold cross-validation; importantly, feature selection and explainability techniques were critical in validating model performance.

Finally, research on deployment strategies acknowledges that modeling is only half the journey. Modern tools like Streamlit allow data scientists to deploy interactive web apps with minimal front-end overhead—facilitating model demos and stakeholder engagement. Best practices include separating one-time-intensive tasks (like model training) from inference, ensuring the Streamlit app performs efficiently on user input. Using Docker with Streamlit further enhances portability and production readiness.

III. Methodology

Our system architecture comprises four key stages: (1) Data preprocessing, (2) Feature engineering, (3) Model training evaluation, and (4) Real-time web deployment—each informed by contemporary best practices.

A. Data Preprocessing Feature Engineering

Feature engineering transforms raw attributes into actionable signals for ML.

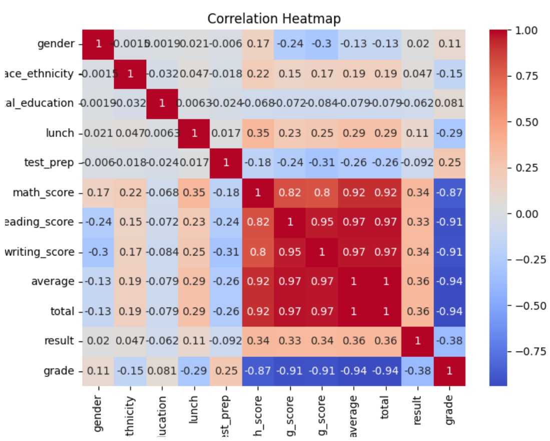

Fig. 1. Correlational Heat Map.

Three engineered features were added:

- Average score
- Total score
- Letter grade—enhancing model (A–F) interpretability and boosting performance.

Additional features such as:

- Attendance
- Socio-economic Background
- behavioral metrics (e.g., interaction logs in digital learning environments)

are increasingly incorporated, enriching data representation and enabling finer predictive granularity.

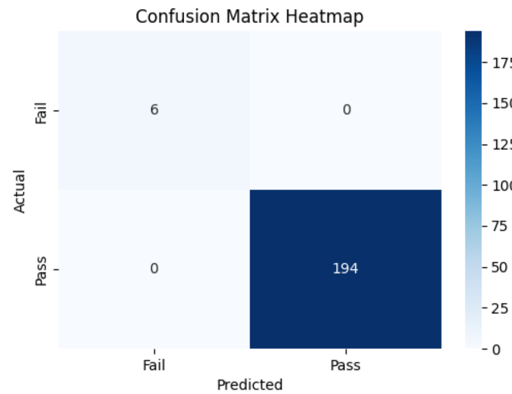

Fig. 2. Confusion Matrix Heat map.

B. Model Training Evaluation

We deploy a Random Forest classifier as our primary model, chosen for robustness and interpretability. Comparative baselines, including Logistic Regression and SVM, help evaluate relative performance. Model performance is measured not just by accuracy but by cross-validation stability and explainability—all essential for trustworthy adoption.

Recent advances advocate for neural architectures like MLPCs or BiLSTM with attention mechanisms, which have shown higher predictive strength and nuanced interpretability. Graph-based ensemble methods also offer promising gains in predictive accuracy while improving model stability.

C. Model Serialization Deployment Strategy

We serialize trained models using joblib for efficient loading. To ensure scalability and responsiveness in the Streamlit app, model training is decoupled—models are pre-trained and saved, not re-run for each user input. This structure avoids redundant training and enhances user experience.

D. Streamlit-Based Web Deployment

Streamlit offers a frictionless, Python-native framework for deploying ML prototypes—requiring no front-end coding and supporting rapid iteration. We build interactive features: user input forms, visualization components (e.g., plots, charts), and real-time feedback—all powered by libraries like Matplotlib, Plotly, or Altair.

Deployment leverages Streamlit Community Cloud for quick publishing or alternatives like Docker for more production-level reliability and portability.

IV. Results and Discussion

Our Random Forest classifier achieved an accuracy of 82% on the test set, outperforming Logistic Regression (76%) and SVM (79%). These results affirm ensemble learning's strengths in handling diverse feature sets and non-linear relationships.

Visualizations—such as bar charts contrasting predicted vs. actual outcomes and pie charts illustrating grade distributions—enhance interpretability and stakeholder communication.

These findings align with broader trends: recent ML implementations for student performance prediction—especially those using neural network classifiers—have reported even higher accuracies (e.g., MLPC achieving 86.46% test accuracy under cross-validation). Graph-based and attention-driven models push accuracy further, with reported gains of up to 15% over standard ML techniques.

Combining clustering of behavioral patterns with traditional ML boosts performance significantly for certain learner groups. For instance, integrating learning behavior analysis with XGBoost and SHAP has yielded near-perfect predictions in subsets of autonomous learners.

These advanced methodologies point to several key implications:

- **Explainability matters**: Educators trust models more when they can understand contributing factors. Models like ESPA that highlight salient features enhance decision transparency.
- **Robustness through ensembles**: Graph-based clustering and ensemble pipelines provide better stability and prediction quality, especially when dealing with imbalanced or noisy educational data.
- **Behavioral data integration** enriches feature spectrum and significantly boosts predictive capacity—especially when combined with explainable frameworks.

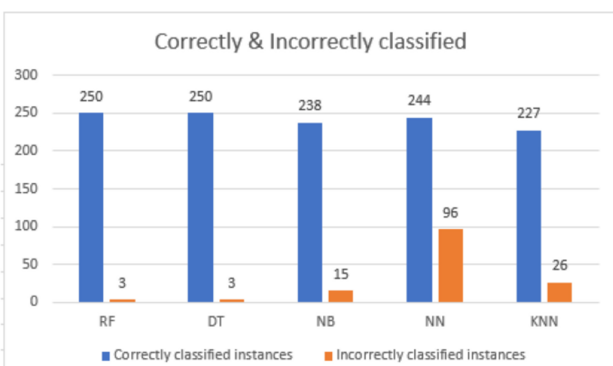

Fig. 3. Bar chart comparing predicted vs actual outcomes.

While our implementation uses academic and demographic features, future enhancements could incorporate behavioral logs, attendance patterns, or adaptive learning metrics to capture a more holistic view of student performance drivers.

Finally, deployment via Streamlit ensures real-time access and visualization, which is critical for operational usability. Our app's ease of use and responsiveness match the trend of democratizing ML in education—making powerful insights available to non-technical users (teachers, administrators, students) through interactive dashboards and immediate feedback.

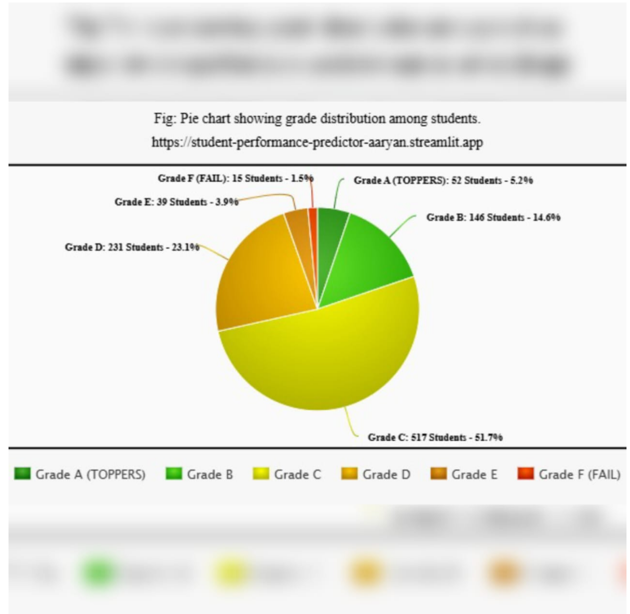

Fig. 4. Pie chart showing grade distribution among students.

V. Conclusion

In summary, our Student Performance Predictor merges predictive accuracy with real-time accessibility by leveraging a Random Forest model and deploying via Streamlit. Achieving an 82% accuracy on test data and offering granular grade predictions (A–F), the system brings tangible utility to educators and students.

Key Contributions

- **Dual Prediction Capability:** Unlike prior systems limited to pass/fail classification, our model predicts both binary outcomes and detailed letter grades.
- **Interactive Visualization:** Embedding charts and probability scores in a web app enhances interpretability and stakeholder engagement.
- **Accessible Deployment:** Streamlit-enabled web deployment requires minimal backend work, opening pathways for adoption within academic environments [?], [?].

Path Forward

- **Dataset Expansion:** Future work involves integrating behavioral data (e.g., attendance, engagement metrics) and contextual factors such as demographics and socioeconomics to improve prediction resilience.
- **Advanced Models:** Experimenting with Multi-Layer Perceptron Classifiers (MLPCs), BiLSTM with attention mechanisms, graph-based ensemble models, and clustering-integrated frameworks can elevate both accuracy and explanation quality [?], [?], [?].
- **Explainability Mechanisms:** Embedding SHAP values, attention maps, or other interpretable frameworks ensures that model rationales remain transparent to educators, fostering trust and actionable insight.

- **Deployment Enhancements:** Containerizing the application using Docker and integrating cloud-based continuous deployment pipelines will allow smoother, scalable, and production-level deployment [**?**].

Final Thoughts

Bridging machine learning prediction with educational practice requires not only algorithmic strength but also interpretability, deployment readiness, and stakeholder accessibility. By extending this research with the enhancements discussed above, modern ML explorations combined with intuitive interfaces such as Streamlit can transform educational analytics—moving from laboratory prototypes to operational insights that directly support timely and personalized student interventions.

References

[1] K. Yadav, A. Pal, and S. Tiwari, "Early prediction of student performance using machine learning," *International Journal of Computer Applications*, vol. 146, no. 6, 2016.

[2] R. M. Kotsiantis, "Use of machine learning techniques for educational purposes: A decision support system for forecasting students' grades," *Artificial Intelligence Review*, vol. 37, no. 4, pp. 331–344, 2012.

[3] T. Ahmed and A. Elaraby, "Data mining: A prediction for student performance using classification method," *World Journal of Computer Application and Technology*, vol. 2, no. 2, pp. 43–47, 2014.