International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 03 March 2025
Page No: 349-352
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0048

Student Churn Prophecy Using Machine Learning

John Livingston J[1], Bharathy S[2], Babypreethi M S[3], Jemima K[4]
[1]Assistant Professor, Dept. of CSE, Kamaraj College of Engg. & Tech., Virudhunagar, Tamilnadu, India.
[2,3,4]UG Scholar, Dept. of CSE, Kamaraj College of Engg. & Tech., Virudhunagar, Tamilnadu, India.
Emails: *johnlivingstoncse@kamarajengg.edu.in[1], sbharathy283@gmail.com[2], babypreethi2004@gmail.com[3], Jemimakathar@gmail.com[4]*

Abstract
Dropout is one of the critical issues faced by educational institutions in terms of both academic performance and institutional efficiency. The project, Student Churn Prophecy Using Machine Learning, aims at predicting the likelihood of dropout using machine learning algorithms. A broad range of student data, such as academic performance, attendance records, demographic details, and engagement levels, is processed and analyzed to identify key patterns and trends. This system generates predictive models that forecast at-risk students and allows timely intervention strategies. The approach enhances student retention and overall educational outcomes of institutions by adopting this methodology.

Keywords: *Student dropout, machine learning, predictive analysis, academic performance, student retention*

1. Introduction
Education plays a vital role in shaping the future of human beings. However, dropout from school or college has persisted to be the bane of educational institutions. The reasons behind it are academic failures, lack of engagement, and certain socio-demographic factors, which make the learner discontinue the degree course. Improving student retention and success involves the early identification of at-risk students such that interventions can be made on time to improve the retention rate and student success.The Student Churn Prophecy Using Machine Learning project seeks to predict the likelihood of a student dropping out based on advanced learning algorithms. By parsing through a rich dataset of academic performance, attendance records, demographics, and engagement scores incorporates all pertinent markers that point to a potential dropout. Machine learning models process this and provide predictions of which students are likely to dropout.With predictive analytics, educational institutions show themselves able to preemptively look into the matter, provide support, and promote engagement for the improvement of a student. This means maximum utility for students because of the support initiated, and also for schools, because of the continued academic standing and performance gains. The document outlines the development, implementation, and impact of the system of the Student Churn Prophecy Using Machine Learning. The methodologies followed, suggestions for models of machine learning applied in predictive analytics, and possible positive outcomes for educational institutions are pointed out in this report. In light of this initiative, we strive to provide such support that may derive data-based work to reduce student churn and promote academia.

1.1. Methods
The methods for the Student Churn Prophecy Using Machine Learning project are intended to predict student dropout risks in a structured way. Stepwise approach actioned through: [5]

Data Collection: Materialized relevant student dataset is collated from academic records, attendance logs, demographic details, and engagement metrics. The sources of these data constitute the premise for the building of the model.

Data Pre-Processing: The data collected go through various stages of preprocessing, including but not limited to filling missing values, normalization of numerical values, encoding categorical data, and removal of inconsistencies. Some feature selection process is put in place to ensure better efficiency.

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 03 March 2025
Page No: 349-352
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0048

Model Selection: Various machine learning algorithms like Decision Trees, Random Forest, SVMs, or Neural Networks are considered in seeking the best-suited model based on the different scores such as accuracy, precision, recall, and F1-score. [1]

Model Training: The chosen model was trained using a labeled dataset, whereby historical student data provided the basis for training the model to recognize dropout-associated patterns. The dataset was divided into training and validation sets for performance fine-tuning.

Model Evaluation: The trained model after performance metrics such as accuracy, precision, recall, and F1-score are used in evaluation. Cross-validation to ensure the robustness.

Prediction and Forecasting: Model encoding is done in a manner to work with the predicted probability for dropout with the new students. Institutions use this to predict dropout early and provide help to at-risk students. [2]

2. Tables and Figures

2.1. Tables

Below table (Table 1) provides the categorization of attributes related to dropout risk, namely, academic performance, attendance, and engagement levels are all further divided into categories of high, medium, and low. Based on these features, dropout risk is then approximated to be low, moderate, or high. The table carries great weight in training the machine-learning models toward accurate predictions. Table 1 shows Retention and Dropout Prediction Using Key Metrics

Table 1 Retention and Dropout Prediction Using Key Metrics

Student Attribute	Category	Value Example
Academic Performance (%)	High	85+
	Medium	60-84
	Low	Below 60
Attendance (%)	High	80+
	Medium	50-79
	Low	Below 50
Engagement Level	Active	High
	Passive	Medium
	Disengaged	Low
Dropout Risk (%)	Low	0-20%
	Moderate	21-50%
	High	Above 50%

Table 1 showcases the surrogate student attributes related to dropout risk. It sorts academic performance, attendance, and engagement into high, medium, and low categories. Accordingly, dropout risk is regarded as low, moderate, or high. These structured data portend encouraging predictions by improving the accuracy with training of machine learning models. Learning project seeks to predict the likelihood of a student dropping out based on advanced learning algorithms. By parsing through a rich dataset of academic performance, attendance records, demographics, and engagement scores incorporates all pertinent markers that point to a potential dropout. Machine learning models process this and provide predictions of which students are likely to dropout

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 03 March 2025
Page No: 349-352
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0048

2.2. Figures

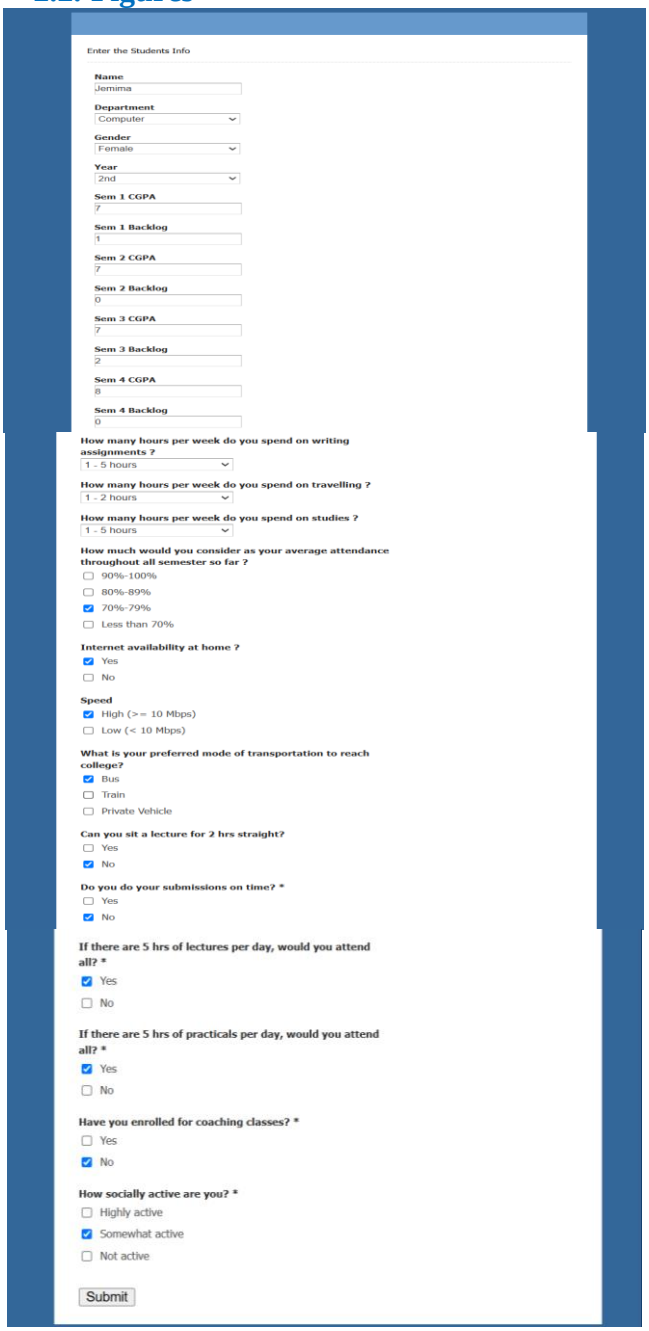

Figure 1 Student Dropout Prediction

3. Results and Discussion

3.1.Results

The student data gathered through the form were processed to involve machine learning algorithms to predict dropout rates. The basic parameters, like academic performance, attendance, study time, and levels of engagement, were fetched to unearth patterns. The result indicates that students with a low attendance percentage, a high number of Backlogs, and minimal study hours are at a higher risk of dropping out. This is subdivided into low, moderate, and high-risk groups using the student model, which encourages early interventions. Figure 2 shows Output, Figure 3 shows Output [3]

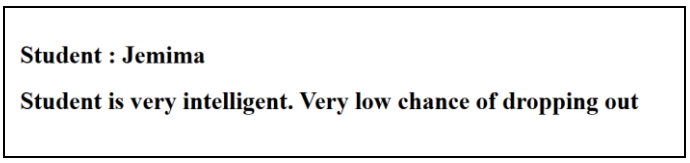

Figure 2 Output

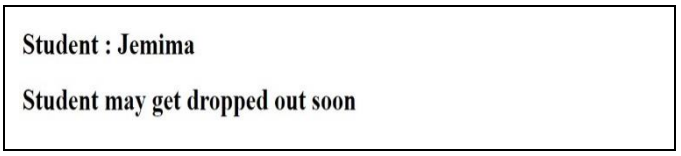

Figure 3 Output

3.2.Discussion

Performance and attendance emerged as the strongest predictors of dropout risk. The low CGPA and multiple Backlogs among students are at risk due to academic difficulties. Furthermore, outside variables affecting engagement and retention include internet availability, study habits, and transportation issues. Early identification of these at-risk students enables institutions to further provide relevant support, counseling, and academic help to improve retention rates. This analysis showcases how well-suited machine learning is for dropout prediction and intervention strategies. [4]

Conclusion

The complete and functional Student Churn Prophecy Using Machine Learning project successfully predicts dropping out based on parameters such as academic performance, attendance, study habits, and engagement levels. They provided evidence showing that students with low CGPA and those with a very high count of Backlogs in a semester along with students with poor attendance are most probable to drop out. The students at risk can easily be identified by employing machine learning, thereby opening up avenues for appropriate intervention strategies to

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 03 March 2025
Page No: 349-352
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0048

improve the retention rate. This predictive model is a very useful instrument in improving student success and academic planning.

Acknowledgements

We express our heartfelt gratitude to our mentors, faculty colleagues, and colleagues for their sincere guidance and support during this study. We are also grateful to the university and students for providing data on which analysis has been based to facilitate the research. Finally, we thank all those researchers working in areas of educational data mining and machine learning for their valuable contributions which helped shape our direction through our work.

References

[1]. Cristóbal Romero, Sebastián Ventura. Educational Data Mining: A Review of the State of the Art. IEEE Transactions on Systems, Man, and Cybernetics (2020). DOI**:** 10.1109/TSMCC.2010.2053532

[2]. Nadia Faiz, Muhammad Asif. Student Dropout Prediction in E-Learning: A Survey. Journal of e-Learning and Higher Education (2021)

[3]. Carol Bell, Martin J. Barras. Machine Learning for Predicting Student Retention and Success in Higher Education. Journal of Educational Data Mining (2022)

[4]. Rajan Sharma, Amit Kumar. Predicting Student Dropout in Higher Education Using Ensemble Learning Techniques. Computers & Education (2021)

[5]. Mark Brown, Emma Finnegan. The Role of Data Analytics in Predicting and Preventing Student Dropout. Journal of Learning Analytics (2020)