Excel International Journal of
Technology, Engineering & Management
Journal homepage: exceljournals.org.in
ISSN ONLINE: https://portal.issn.org/resource/ISSN/2349-8455
ISSN PRINT: https://portal.issn.org/resource/ISSN/2348-0378

Predicting Student Performance Analysis for Campus Recruitment Using Machine Learning Techniques

***1 Dr. M M Kavitha** ***2 Dr. K M Padmapriya** ***3 Dr. T A Sangeetha**
*1 Assistant Professor, *3 Associate Professor & Head, Kongu Arts and Science College, Erode
*2 Assistant Professor & Head, Hindusthan College of Science and Commerce, Erode

Published online: February 2026
DOI Link: https://doi.org/10.64971/j.cph.eijtem.v13.i1.10.2026

Article published link: https://exceljournals.org.in/detail.php?id=853

Abstract
Campus recruitment screening is traditionally based on static eligibility criteria such as CGPA and arrears, which may not accurately reflect students' employability. This paper presents a Machine Learning (ML) based screening model that analyzes student performance data to predict campus drive eligibility. Various supervised ML algorithms including Logistic Regression, Decision Tree, Random Forest, Support Vector Machine, and K-Nearest Neighbors are applied to a student performance dataset. The experimental results show that ensemble-based ML models provide higher prediction accuracy and better decision support for placement officers. The proposed system ensures fair, efficient, and data-driven student screening for campus recruitment.
Keywords: Machine Learning, Classification Algorithms, Random Forest, Decision Tree, Support Vector Machine, k-Nearest Neighbor.

1. Introduction
Campus recruitment is a crucial phase in higher education institutions. Current screening methods rely on limited academic indicators, often ignoring multidimensional performance factors. With the availability of institutional datasets, ML algorithms can identify hidden patterns in student performance and predict recruitment readiness more effectively. This study proposes an ML driven student screening system to assist placement cells in short listing eligible candidates for campus drives.

2. Related Work
The various studies analyze student performance to predict their employability using screening resume and their academic records. Recent studies highlight the effectiveness of ML in educational analytics: Logistic Regression has been widely used for employability prediction; Decision Trees offer interpretability for academic decision-making, Random Forest improves accuracy by reducing overfitting and SVM has shown robustness in classification of student outcomes. However, limited research addresses pre-campus drive screening, which is the focus of this work.

3. Proposed System
The proposed system aims to automate the student screening process for campus recruitment by leveraging machine learning techniques on academic and skill-based performance data. The system predicts whether a student is Eligible or Not Eligible for campus placements, thereby reducing manual effort, bias, and time consumption in the screening process. The dataset consists of academic and skill-based attributes collected from student records. The key features are Attendance Percentage, Internal Assessment Marks, Semester CGPA, Number of Arrears, Technical Skill Score, Soft Skill Evaluation, Internship Participation, Final Screening Status (Eligible / Not Eligible). The dataset is divided into training (80%) and testing (20%) sets.

4. Machine Learning Algorithms

4.2.1 Logistic Regression

Logistic Regression is used as a baseline classification model to estimate the probability of a student being eligible for campus recruitment. It analyzes the relationship between academic, technical, and skill-based features and the final eligibility outcome.

4.2.2 Decision Tree

The Decision Tree algorithm builds a tree-structured model by learning decision rules from student performance attributes. It provides clear interpretability by showing how different factors such as CGPA, attendance, and skills influence recruitment eligibility.

4.2.3 Random Forest

Random Forest is an ensemble learning technique that combines multiple decision trees to improve prediction accuracy. By aggregating results from several trees, it reduces overfitting and enhances the robustness of student eligibility predictions.

4.2.4 Support Vector Machine (SVM)

Support Vector Machine identifies an optimal hyperplane that best separates eligible and non-eligible students in the feature space. It is effective in handling high-dimensional student performance data and complex decision boundaries.

4.2.5 K-Nearest Neighbors (KNN)

K-Nearest Neighbors classifies a student based on the performance similarity with the nearest students in the dataset. The eligibility decision is determined by the majority class among the closest neighbors in the feature space.

5. Methodology

The architecture (Fig-1) represents an automated student screening system designed to support campus recruitment using machine learning. Student academic and skill-based data are collected from the institutional database and passed through a data preprocessing stage where cleaning, normalization, and feature selection are performed to ensure quality input. The processed data is then fed into multiple machine learning algorithms such as Logistic Regression, Decision Tree, Random Forest, SVM, and KNN to train predictive models. These models are evaluated using performance metrics like accuracy, precision, recall, and F1-score to select the best-performing model. Finally, the system predicts whether a student is Eligible or Not Eligible for campus placements, and the results are provided to the campus placement system to assist recruiters in efficient and unbiased student screening.

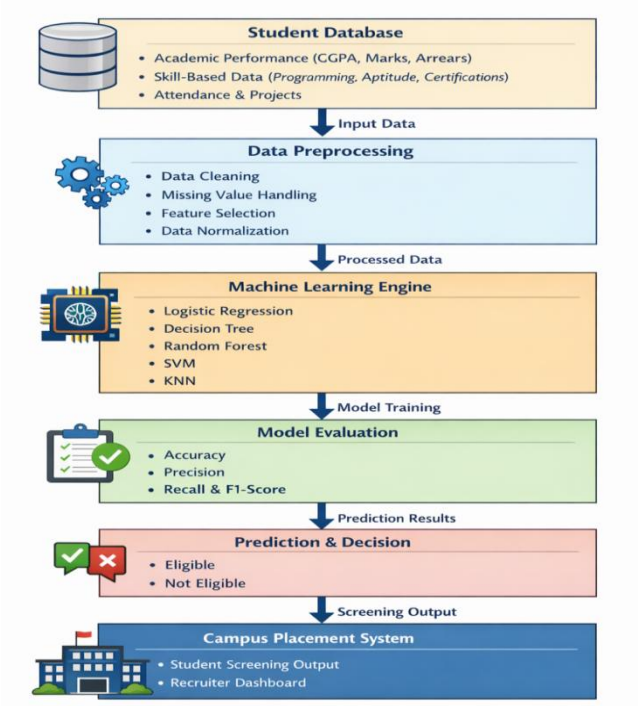


Fig-1: Automated Student Screening System for Campus Recruitment

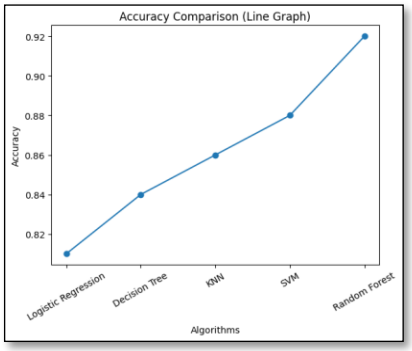
5. Performance Evaluation

To assess the effectiveness of ML models, the following metrics are used such as Accuracy, Precision, Recall, F1-Score. The proposed ML-based screening model are Student performance data collection, Data preprocessing and feature scaling, ML model training, Eligibility prediction and Automated shortlist generation. This system supports placement officers in making objective and scalable screening decisions.

6. Implementation

The machine learning models were implemented using Python 3.x with essential libraries such as NumPy and Pandas for data processing, Scikit-learn for model development and evaluation, and Matplotlib for data visualization, all executed within the Jupyter Notebook environment.

Algorithm	Accuracy	Precision	Recall	F1-Score
Logistic Regression	0.81	0.80	0.78	0.79
Decision Tree	0.84	0.83	0.82	0.82
KNN	0.86	0.85	0.84	0.84
SVM	0.88	0.87	0.86	0.86
Random Forest	**0.92**	**0.91**	**0.90**	**0.90**



The experimental results observed that Random Forest achieves the highest accuracy (92%) due to ensemble learning and better handling of feature interactions. The experimental results clearly indicate that ensemble-based models outperform individual classifiers.

Random Forest achieved the highest accuracy of 92%, demonstrating its effectiveness in screening students for campus recruitment. The model successfully captures complex relationships between academic performance, skill assessments, and employability status. This confirms that machine learning based screening provides a more reliable alternative to traditional eligibility criteria.

7. Conclusion

This study demonstrates that machine learning based screening provides a more accurate and objective approach to campus recruitment compared to traditional static criteria. The experimental result shows that ensemble models, particularly Random Forest, achieve superior prediction performance in identifying eligible students. The proposed system enables efficient, fair, and data-driven screening, assisting placement officers in making reliable recruitment decisions and improving overall campus placement outcomes.

References

1. Jiaze Chen, Liangcai Gao, and Zhi Tang, (2016). "Information extraction from resume documents in pdf format", pp. 1-8
2. Jie Chen, Chunxia Zhang and Zhendong Niu, (2018). "A two-step resume information extraction algorithm", pp. 1-10.
3. Saket Maheshwary and Hemant Misra, (2018). "Matching resumes to jobs via deep siamese network", In: Proceedings of The Web Conference, International World Wide Web Conferences Steering Committee, pp. 87-88.
4. Romero, C., & Ventura, S. (2020). Educational data mining. IEEE Transactions on Learning Technologies.
5. Lokesh. S, Mano Balaje. S, Prathish. E and B. Bharathi (2022). Resume screening and recommendation system using machine learning approaches, Computer Science & Engineering: An International Journal (CSEIJ), Vol 12, No 1,
6. Karthikeyan, T. et al. (2021). Placement prediction using ML algorithms. International Journal of Computer Applications.
7. Sharma, R., & Gupta, S. (2022). Employability prediction using machine learning. Education and Information Technologies.

How do I cite this article?

Dr. M M Kavitha et.al, Predicting Student Performance Analysis for Campus Recruitment Using Machine Learning Techniques, Excel International Journal of Technology, Engineering and Management, 2026; Volume -13, Issue-1_Page_57-59. DOI Link: https://doi.org/10.64971/j.cph.eijtem.v13.i1.10.2026