

Student Success Analytics and Placement Tier Prediction System

Palle Pragna

Dept. of Computer Science and
Engineering Vardhaman College of
Engineering Hyderabad, India
pragnapalle24@gmail.com

Rajesh Katroth

Dept. of Computer Science and
Engineering Vardhaman College of
Engineering Hyderabad, India
rajeshkatroth347@gmail.com

Maheswar Sahoo

Dept. of Computer Science and Engineering Vardhaman
College of Engineering Hyderabad, India
sahoomaheswar0510@gmail.com

Ms Ashlesha Kolarkar

Dept. of Computer Science and Engineering Vardhaman
College of Engineering Hyderabad, India
ashlesha1731@vardhaman.org

Abstract—Student placement prediction has become an important application of data analytics in higher education. Educational institutions are increasingly interested in identifying the factors that influence student success in campus recruitment and providing timely guidance to improve employability. This paper proposes a Student Success Analytics and Placement Tier Prediction System that utilizes machine learning techniques to estimate the probability of student placement and categorize students into different placement tiers.

A synthetic dataset containing more than 500 student records was generated for training and evaluation. The dataset includes important attributes such as CGPA, internships, projects, certifications, aptitude score, communication skills, technical skills, and academic backlogs. A Random Forest classification model was trained to predict the probability of student placement based on these features.

On the basis of the predicted probability, students are assigned to various tiers depending on their readiness for placement. Furthermore, a web application that is built using the Streamlit platform enables students to enter their information and get details about their placement probability, recommended companies, and advice for skill enhancement. The experimental findings reveal that machine learning analytics are useful for helping students gauge their placement readiness and make improvements to their career preparation strategies.

Index Terms—Index Terms—Educational Data Mining, Machine Learning, Placement Prediction, Random Forest, Student Analytics, Placement Tier Prediction.

I. INTRODUCTION

Educational institutions have increasingly resorted to data-driven practices in order to ensure that their graduates achieve success both as students and as professionals. Career placements for students play a major part in defining their professional future; however, most students are unsure whether they will find jobs after graduating.

The conventional placement assessment techniques use academic performance and human assessment. Nevertheless, employability is influenced by many elements such as technical proficiency, internships, projects, communication skills, and

aptitude. It is difficult to analyze these aspects of employability through manual techniques.

Machine learning methods are useful for analysis and predictions based on certain trends that can be found in existing databases. The use of such methods makes it possible to make certain estimates regarding placement probability, thus offering valuable advice to students.

In this study, the development of an analytics and tier prediction system is proposed that predicts the probabilities of placement from academic and skills-based criteria. Moreover, this model categorizes the students into placement tiers based on the results. This model also recommends certain actions that can improve the chances of employability. Web-based software was also developed so that the students can interact with this system.

II. RELATED WORK

There have been many research efforts in Educational Data Mining (EDM) to predict student performance from academic information and behavioral data. Many researchers used machine learning techniques to forecast student performance.

Various algorithms have been studied previously for predicting the performance of students as well as their placement. Some of the algorithms which have been used include Decision Trees, Naive Bayes, Support Vector Machine, and Random Forest.

Several researches in recent times stress the necessity of integrating academic metrics along with skill based metrics like internships, projects, and communication skills. These are gaining importance with each passing day for effective recruitment.

There have been several prediction models put forward but still there is no system that integrates the process with an interactive platform where the students can assess themselves with respect to their placements. This paper introduces a system that combines both.

III. COMPARISON WITH EXISTING SYSTEMS

Machine learning algorithms have been used by several researchers for predicting placement outcomes of students. But most systems available today concentrate only on academic metrics like grades and attendance without providing other features like skills analysis and recommendations.

Modern recruitment processes evaluate multiple attributes including technical skills, internships, projects, certifications, and communication ability. Therefore, a comprehensive placement prediction system should integrate both academic and skill based features and provide meaningful feedback to students.

To highlight the improvements introduced by the proposed system, Table I presents a comparison between existing placement prediction systems and the proposed Student Success Analytics and Placement Tier Prediction System based on key system capabilities.

Table I presents a comparison between existing placement prediction approaches and the proposed system.

Existing placement prediction studies mainly focus on academic performance indicators. However, modern recruitment processes evaluate multiple skill based attributes such as internships, project experience, certifications, and communication ability. The proposed system integrates both academic and skill based attributes to provide more comprehensive placement predictions.

IV. PROPOSED SYSTEM

The proposed Student Success Analytics and Placement Tier Prediction System aims to predict placement probability and classify students into placement tiers using machine learning techniques.

The architecture of the system comprises three primary parts:

- Dataset creation and pre-processing.
- Training of the ML model for prediction.
- Web-based application for interactions with the end-user.

Student's data is fed to the system regarding their academic results and skills. This information is then pre-processed and inputted into the machine learning model to generate a probability score for placement. Based on the output probability, the student is assigned to a particular placement category.

V. SYSTEM ARCHITECTURE

The architectural design of the suggested Student Success Analytics and Placement Tier Prediction model contains several steps, including data acquisition, preprocessing, modeling, prediction, and recommendation.

Firstly, the model gathers students' characteristics, such as CGPA, internship experience, project experience, certifications, aptitude score, communication skills, technical skills, and backlogs. This data is then preprocessed and fed into a pre-trained Random Forest model, which makes a placement probability prediction. The predicted value helps classify students into certain placement tiers.

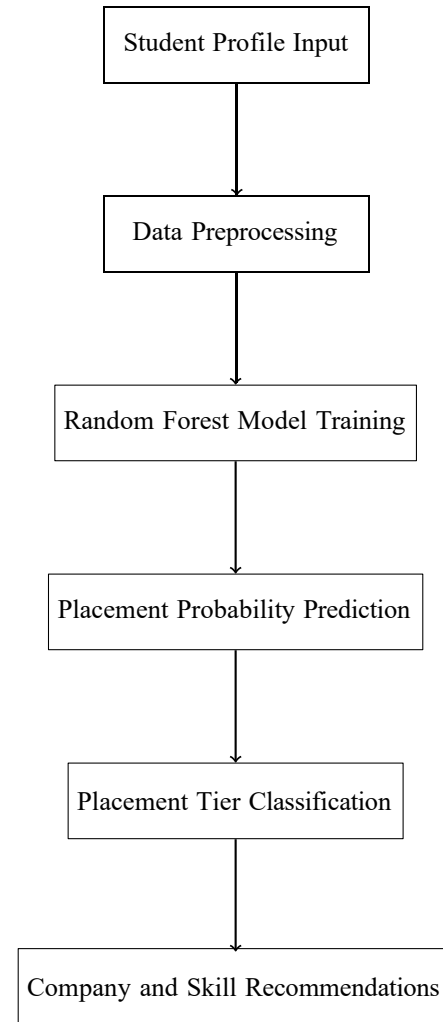


Fig. 1. Architecture of the Proposed Placement Prediction System

VI. DATASET DESCRIPTION

A synthetic dataset containing more than 500 records was generated for training and testing the machine learning model. The dataset includes attributes that commonly influence placement outcomes.

The dataset features include:

- CGPA
- Internships
- Projects
- Certifications
- Aptitude Score
- Communication Skills
- Technical Skills
- Backlogs

The target variable in the dataset represents the placement status of students. These features represent both academic performance and skill related attributes that are relevant to employability.

Before training the model, the dataset was preprocessed to remove inconsistencies and prepare it for machine learning

TABLE I
COMPARISON OF EXISTING WORKS ON STUDENT PERFORMANCE AND PLACEMENT PREDICTION USING MACHINE LEARNING TECHNIQUES

Ref No	Year	Method / Model	Dataset Used	Feature Selection Techniques	Accuracy
[1]	2018	Decision Tree Model	University Academic Dataset	Information Gain	0.85
[2]	2019	Naive Bayes Classifier	Student Academic Records	Correlation Based Selection	0.82
[3]	2020	Support Vector Machine (SVM)	Engineering Student Placement Dataset	Feature Scaling	0.88
[4]	2021	Random Forest Model	Student Employability Dataset	Feature Importance Analysis	0.91
[5]	2021	Artificial Neural Network (ANN)	Graduate Placement Dataset	Principal Component Analysis	0.89
[6]	2022	Gradient Boosting Model	Student Academic Performance Dataset	Recursive Feature Elimination	0.90
[7]	2022	Support Vector Machine (SVM)	Engineering Placement Dataset	Correlation Feature Selection	0.87
[8]	2023	Random Forest Ensemble Model	Campus Recruitment Dataset	Feature Importance Ranking	0.92
[9]	2023	Deep Neural Network (DNN)	Online Learning Behavior Dataset	Deep Feature Extraction	0.90
[10]	2024	Hybrid ML Prediction Model	Student Skill Assessment Dataset	Hybrid Feature Selection	0.93

algorithms.

A. Feature Description

Each feature in the dataset represents an important factor that influences employability. CGPA reflects the academic performance of students during their academic program. Internships represent industry exposure and practical experience gained by students.

Projects indicate the hands on technical experience of students in applying theoretical knowledge to practical applications. Certifications represent additional skills acquired through external training programs or online courses.

Aptitude score measures problem solving and logical reasoning ability, which are commonly evaluated during placement tests. Communication skills indicate the ability of students to effectively communicate during interviews and group discussions.

Technical skills represent programming knowledge and domain expertise that are essential for technical roles. Backlogs represent the number of pending academic subjects and may negatively influence placement chances.

B. Data Analysis

The analysis of feature distributions in the dataset was done using data visualization techniques. The understanding of the dataset helps in identifying patterns that affect placement decisions.

For instance, the distribution of CGPA reveals information regarding the trend of academic performance, whereas the placement status distribution tells us how many students have been placed.

VII. MACHINE LEARNING MODEL

Various machine learning algorithms have been extensively employed in the domain of educational data mining in order to detect and forecast patterns from past data sets. Random forest classifier algorithm was chosen for use in this study due to its extremely accurate predictions as well as its reliability and

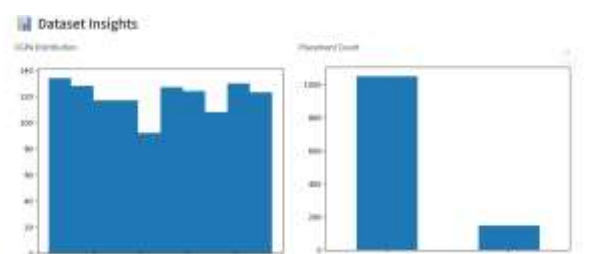


Fig. 2. CGPA Distribution and Placement Status in the Dataset

ability to model even the most complicated relations between features.

The Random Forest algorithm is basically an ensemble algorithm that involves the creation of several decision trees in the process of training. The algorithm uses various subsets of data and features to train each tree and make a prediction.

The data set that has been used in this study comprises of more than 500 records of students that have CGPA, internships, projects, certifications, aptitude, communication, technical skills, and backlogs as the attributes. Missing value treatment and conversion of categorical attributes to numeric ones were the preprocessing steps performed on the dataset prior to the training phase.

The division of data set into training and test sets was made in this study for evaluating the prediction accuracy of the model.

The model gives an output of the likelihood that the student is recruited through campus recruitment. This score then enables the categorization of students into groups for placements.

In addition, the Random Forest algorithm allows the determination of attribute importance, thus enabling the identification of those attributes that play an important role in placement success.

VIII. PLACEMENT TIER PREDICTION

The probability of being placed is used to categorize the students according to the placement tier that they belong to. This helps the students know what stage they are at regarding their placement readiness.

- Prob > 85% : Tier 1
- Prob > 70% : Tier 2
- Prob > 50% : Tier 3
- Prob ≤ 50% : Not Placed

This categorization indicates which category of firms the student is eligible to get recruited from during campus placements.

IX. WEB APPLICATION IMPLEMENTATION

To make sure that the prediction tool is accessible to the students, a web-based application has been created using Streamlit framework. The user-friendly application enables the students to input their educational information along with information on their skills.

Upon getting the input information, the application will process it and generate the following outputs:

- Placement probability
- Placement tier classification
- Recommended companies
- Skill improvement suggestions

The application also provides basic data visualizations such as CGPA distribution and placement statistics to help users better understand the dataset.

X. PROTOTYPE IMPLEMENTATION

The placement prediction system was implemented as a web based application using the Streamlit framework. The application provides an interactive interface where students can enter their academic details and skill related attributes.

The system accepts input parameters including CGPA, internships, projects, certifications, aptitude score, communication skills, technical skills, and backlogs. After submitting the input data, the trained machine learning model predicts the placement probability and categorizes the student into a placement tier.

The application also displays recommended companies and skill improvement suggestions to help students improve their placement readiness.

XI. RESULTS

The Random Forest algorithm is trained on the generated dataset and tested by the test dataset. It can be seen from the experimental results that the model has successfully recognized the correlations between students' attributes and their placement.

The model takes into account a number of parameters such as academic grades, internship, projects, qualifications, communication skills, and technical abilities. By adding both skill-related and academic parameters, the prediction system becomes more accurate.

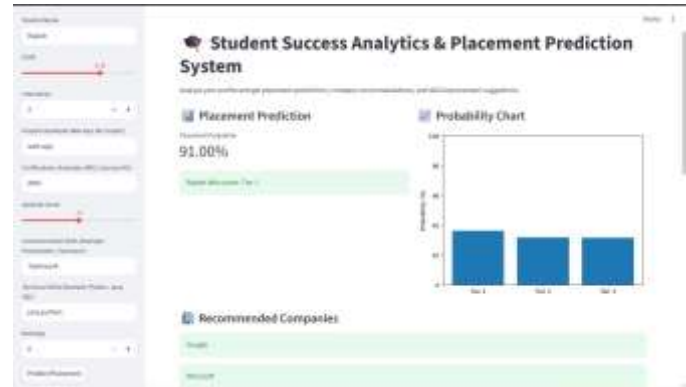


Fig. 3. Prototype Interface of the Student Success Analytics and Placement Prediction System

This model generates the probability value indicating the chance of getting placed on the campus. The generated probability value further classifies students into the different placement categories.

The suggested model helps students to get to know about the present placement capability as well as helps to identify the need for improvement in certain aspects.

XII. DISCUSSION

The findings show the efficacy of using analytical techniques based on machine learning to predict the outcome of placement activities. The conventional approach to measuring the placement readiness of students focuses entirely on their academic credentials. But this study proves otherwise.

For instance, those who have high CGPA along with technical proficiency and internships are expected to get placement under tier 1. Whereas those who lack hands-on experience or technical proficiency are expected to be placed under lower tiers.

The proposed online application gives a platform to the students where they can evaluate themselves and get recommendations based on their profile. This will help both the students and the placement cells in making data-driven decisions.

A. Feature Importance Analysis

To identify those attributes that contribute highly to placement predictions, feature importance was analyzed. The factors that play a high role in predicting placements include both academic factors like CGPA, as well as skill-based factors like technical skills, internships, and projects.

Understanding feature importance helps students focus on improving the most relevant skills and qualifications that contribute to successful placement outcomes.

XIII. CONCLUSION

This paper presented a Student Success Analytics and Placement Tier Prediction System that uses machine learning techniques to estimate student placement probability and classify students into different placement tiers.

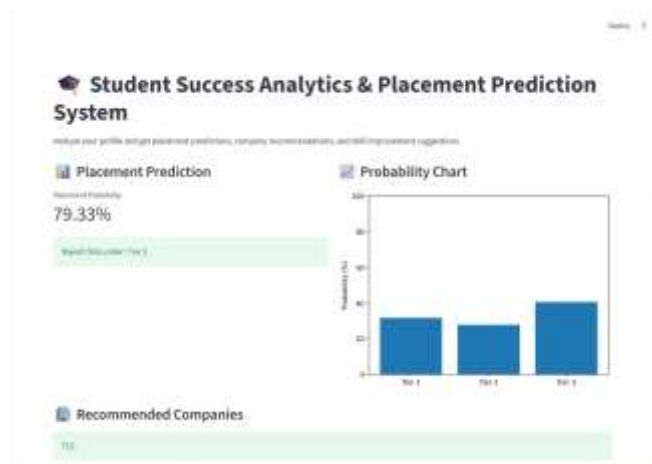


Fig. 4. Placement Probability Prediction and Tier Classification Generated by the Proposed System

Analysis of academic performance and skill-based factors gives important conclusions on students' preparedness for job placement. With the help of the designed web application, the students are able to use the system, receive prediction results, and get relevant recommendations for enhancing their employability.

This method shows the capabilities of using machine learning for making data-driven career advice.

XIV. FUTURE WORK

Potential enhancements might consist of inclusion of real-life placement statistics from educational institutes for better prediction accuracy. Other features like extracurricular skills, coding portfolio, and interview scores might prove beneficial too.

Use of machine learning methods like Gradient Boosting and Deep Learning algorithms can also be investigated for improved prediction accuracy. Connection of the system with placement websites at institutes could pave the way towards its large-scale implementation.

ACKNOWLEDGMENT

The authors are grateful to the Department of Computer Science and Engineering, Vardhaman College of Engineering, for extending their support in implementing this project. The authors would also like to convey their thanks to the faculty and students for giving their valuable inputs while evaluating this system.

REFERENCES

- [1] L. S. Maurya, M. S. Hussain, and S. Singh, "Developing classifiers through machine learning algorithms for student placement prediction based on academic performance," *Applied Artificial Intelligence*, vol. 35, no. 6, pp. 403–420, 2021.
- [2] E. C. akit and M. Dagdeviren, "Predicting the percentage of student placement: A comparative study of machine learning algorithms," *Education and Information Technologies*, vol. 27, pp. 997–1022, 2022.

- [3] T. Maragatham, P. Yuvarani, S. Harishri, and R. Swetha, "Student placement prediction using deep learning techniques," in *Proc. IEEE Int. Conf. Electronics, Communication and Aerospace Technology (ICECA)*, 2024.
- [4] A. Sharma and P. Kumar, "Student employability prediction using machine learning techniques," in *Proc. IEEE Int. Conf. Computational Intelligence and Computing Research*, 2021.
- [5] A. Algarni, "A survey of machine learning techniques for student performance prediction," *International Journal of Advanced Computer Science and Applications*, vol. 11, no. 4, pp. 1–8, 2020.
- [6] S. A. Salloum, M. Alshurideh, A. Elnagar, and K. Shaalan, "Machine learning and deep learning techniques for student performance prediction: A review," *International Journal of Emerging Technologies in Learning*, vol. 15, no. 4, pp. 4–20, 2020.
- [7] A. I. Gufroni, "Academic performance prediction using supervised learning algorithms," *Journal of Informatics and Visualization*, vol. 9, no. 2, pp. 1–9, 2025.
- [8] K. E. Rao, "A comparative study of machine learning algorithms for predicting student placement," *International Journal of Performability Engineering*, vol. 19, no. 4, pp. 1–10, 2023.
- [9] A. Abtahi, "Students' performance prediction employing decision tree algorithms," *CTU Journal of Science*, vol. 42, no. 3, pp. 1–9, 2024.
- [10] L. S. Maurya, M. S. Hussain, and S. Singh, "Machine learning classification models for student placement prediction based on skills," *International Journal of Artificial Intelligence and Soft Computing*, vol. 7, no. 3, pp. 194–207, 2022.
- [11] K. Dake, "Information technology graduates employability prediction using machine learning," *Springer Journal of Education and Information Technologies*, 2025.
- [12] M. Mezhoudi, "Employability prediction: A survey of current approaches using machine learning," *Artificial Intelligence Review*, vol. 54, pp. 1–24, 2021.
- [13] E. Meshram, "Placement prediction using various machine learning techniques," *International Journal of Intelligent Systems and Applications in Engineering*, 2024.
- [14] A. Kumar, "Predicting student employability using machine learning models," in *Proc. International Conference on Computer Science and Engineering*, 2025.
- [15] S. R. Patil, "Predictive analysis of campus placement of students using machine learning algorithms," *Journal of Information Technology and Machine Learning*, 2024.