International Journal of Scientific Research and Engineering Development-– Volume 9 Issue 3, May-June 2026
Available at www.ijsred.com

RESEARCH ARTICLE OPEN ACCESS

Campus Placement Prediction and Career Outcome Analytics Platform

Mr. R. Ramakrishnan[1], S. Gokulakrishnan[2]
[1]Professor, Head of Department of Master Computer Application, Sri Manakula Vinayagar Engineering College, (Autonomous), Pondicherry 605008, India
[2]Student, Department of Master Computer Application, Sri Manakula Vinayagar Engineering College, (Autonomous), Pondicherry 605008, India
ramakrishnanmca@smvec.ac.in, gokulakrishnan14082003@gmail.com

Abstract:

Campus placement is a critical milestone for students and institutions alike. This paper presents an intelligent platform that leverages Machine Learning (ML) algorithms to predict a student's placement likelihood and analyses career outcomes based on academic performance, skill sets, and extra-curricular data. The system employs Random Forest and Logistic Regression models trained on multi-factor datasets including CGPA, 10th and 12th percentages, internship experience, technical and communication skills, and aptitude test scores. The platform provides personalised dashboards with skill gap analysis and career recommendations, enables placement cells to conduct data-driven recruitment drives, and supports batch prediction and exportable reports. This work addresses the critical gap in campus placement infrastructure where institutions rely on manual, CGPA-only evaluation — replacing guesswork with evidence-based prediction.

Keywords — Campus Placement, Machine Learning, Random Forest, Logistic Regression, Career Analytics, Skill Gap Analysis, Predictive Analytics, CGPA, Student Performance.

I. INTRODUCTION

Campus placement is one of the most consequential processes in higher education — for students charting their career paths and for institutions measuring the quality of their academic programmes. Yet, despite its importance, the placement evaluation process at most institutions remains remarkably primitive: counsellors manually assess students based almost entirely on CGPA, ignoring a wealth of data about skills, internships, aptitude, and communication ability.

The advent of Machine Learning (ML) has opened new avenues for data-driven decision making across domains. However, its application in campus placement has been limited, with most institutions yet to adopt predictive analytics platforms tailored to their placement cells. This paper addresses that gap by presenting a comprehensive Campus Placement Prediction and Career Outcome Analytics Platform — an end-to-end intelligent system that integrates ML models, interactive dashboards, and analytics engines.

The platform is designed for three key stakeholders: students who receive personalised placement probabilities and career recommendations; faculty and placement officers who gain analytical oversight of the student cohort; and administrators who can generate reports and configure the ML engine. By bridging the gap between raw academic data and actionable career insights, this platform transforms a historically subjective process into an evidence-based one.

II. LITERATURE SURVEY

Kotsiantis et al. [1] applied decision tree and naive Bayes classifiers to predict student academic performance, demonstrating that multi-factor models substantially outperform single-metric approaches. Their work highlighted the importance of feature engineering in educational data mining.

Bharambe et al. [2] conducted a comparative study of ML algorithms for campus placement prediction using datasets from engineering colleges. They found that ensemble methods — particularly Random Forest — achieved higher accuracy than individual classifiers, and that including soft-skill indicators alongside academic scores improved prediction reliability.

Jayaprakash et al. [3] examined the role of career analytics dashboards in student counselling, finding that visual, data-driven feedback improved students' motivation to address skill gaps compared to traditional counselling methods.

Ramaswami and Bhaskaran [4] applied CHAID (Chi-square Automatic Interaction Detector) for placement prediction and concluded that aptitude scores and communication skills are among the strongest predictors of placement success, reinforcing the multi-factor approach of the proposed system.

More recently, Gupta and Singh [5] explored the integration of real-time dashboards in placement management systems, establishing that interactive analytics substantially reduce the administrative burden on placement officers while improving student outcomes.

III. THEORETICAL FRAMEWORK

The proposed platform is grounded in established principles of Educational Data Mining (EDM) and Learning Analytics (LA). EDM posits that patterns embedded in student academic and behavioural data can be systematically extracted to predict outcomes and inform interventions. LA extends this by emphasising the feedback loop — returning actionable insights to students and educators to improve performance.

The system further draws on the Theory of Planned Behaviour (TPB) in its recommendation engine: by surfacing specific, actionable skill gaps (e.g., low aptitude scores or absence of internship experience), the platform helps students form concrete intentions to act, which research consistently links to improved outcomes.

From a technical standpoint, the ensemble learning principle underpins the ML architecture — combining the predictions of multiple weak learners (decision trees in Random Forest) to produce a more robust, generalised prediction than any single model could achieve.

IV. METHODOLOGY

The research followed a Design Science Research (DSR) methodology, encompassing problem identification, artefact design, development, and evaluation planning. The process began with a systematic review of existing placement management practices at MCA institutions, identifying critical deficiencies in data collection, prediction, and communication.

Dataset construction drew on historical placement records incorporating academic metrics (10th percentage, 12th percentage, CGPA, employability test scores, aptitude test scores) and skill/experience metrics (technical skill rating, communication skill rating, internship experience, work experience, programming languages known). Features were normalised and encoded for ML ingestion.

Three ML models were trained and evaluated — Random Forest, Logistic Regression, and Decision Tree — using stratified k-fold cross-validation to prevent overfitting. Model selection prioritised both accuracy and interpretability, given the end-user-facing nature of the platform. The full system was implemented using a Python backend, MySQL database, and a Streamlit-based frontend for rapid, interactive deployment.

TABLE I

COMPARISON OF EXISTING SYSTEM VS. PROPOSED SYSTEM

Feature	Existing System	Proposed System
Data Source	Academic records only	Academics + skills + internships

Feature	Existing System	Proposed System
Analysis Method	Static single-model	Gradient Boosting + MLP fusion
Skill Assessment	Manual / not integrated	Automated gap detection
Prediction	None / retrospective	Real-time PRI forecasting
Recommendation	Generic advice	Personalized career pathways
Dashboard	Spreadsheets	Live institutional analytics
Early Warning	Not available	Semester-level alerts

V. EXISTING SYSTEM

Most educational institutions rely on a combination of manual spreadsheet tracking and subjective counsellor assessments for campus placement management. The core limitations of this approach are well-documented:

- **Manual Data Entry:** Placement records are maintained in Excel spreadsheets, prone to human error, version conflicts, and loss of historical data.
- **Single-Metric Evaluation:** Student suitability is judged almost exclusively on CGPA, ignoring skills, internships, aptitude, and communication ability — factors that employers consistently rank as critical.
- **No Predictive Capability:** Placement decisions are reactive, made after the recruitment drive, rather than proactive — students receive no advance guidance on their placement probability.
- **Communication Breakdown:** Information flow between students, faculty, and placement officers is fragmented, relying on informal channels rather than a unified platform.
- **No Trend Analysis:** Historical placement data is not systematically mined to identify patterns or inform future recruitment strategies.
- **Lack of Personalisation:** Students receive generic guidance rather than individual career roadmaps tailored to their specific academic and skill profiles.

VI. PROPOSED SYSTEM

The proposed Campus Placement Prediction and Career Outcome Analytics Platform is an end-to-end intelligent system that integrates ML-based prediction, interactive analytics, and role-based access control into a single, unified platform. The system consists of six integrated modules:

Home Page Module: The home page provides a role-aware landing experience, presenting personalised summaries to students (placement probability, key recommendations) and aggregate dashboards to administrators (cohort-level analytics, recruitment drive summaries).

Data Input Module: A structured data entry interface collects student profiles including academic scores (10th, 12th, CGPA), aptitude and employability test results, skill ratings, internship history, and programming proficiency. Input validation ensures data quality before ingestion into the ML pipeline.

Placement Prediction Module: The core analytical engine applies trained Random Forest and Logistic Regression models to generate a placement probability score for each student. The module returns not only the binary prediction (placed / not placed) but a calibrated probability score, enabling nuanced interpretation and tiered recommendations.

Career Analytics Dashboard: An interactive dashboard visualises individual student performance against cohort benchmarks, displays historical placement trends, and surfaces the most influential factors in the student's prediction score. Skill gap charts highlight specific areas requiring improvement, enabling targeted interventions.

Batch Prediction Module: Placement officers can upload an entire cohort dataset for simultaneous prediction, receiving ranked placement probability lists suitable for recruitment drive planning. This module eliminates the need for individual student assessment during pre-placement preparation.

VII. SYSTEM ARCHITECTURE

The platform architecture follows a three-tier model comprising the Presentation Layer, Application Layer, and Data Layer.

The Presentation Layer is implemented using Streamlit, providing an interactive, browser-based interface accessible to all stakeholder roles. The Application Layer houses the Python-based ML pipeline (Scikit-learn, Pandas, Plotly) and business logic for prediction, analytics, and report generation. The Data Layer employs a MySQL relational database for secure, structured storage of student profiles, prediction histories, and placement records.

Role-Based Access Control (RBAC) enforces strict data governance: students access only their own profiles and predictions; faculty view cohort analytics; placement officers manage recruitment drives and batch predictions; administrators configure ML model parameters and generate institutional reports.

[Fig. 1: System Architecture Diagram]

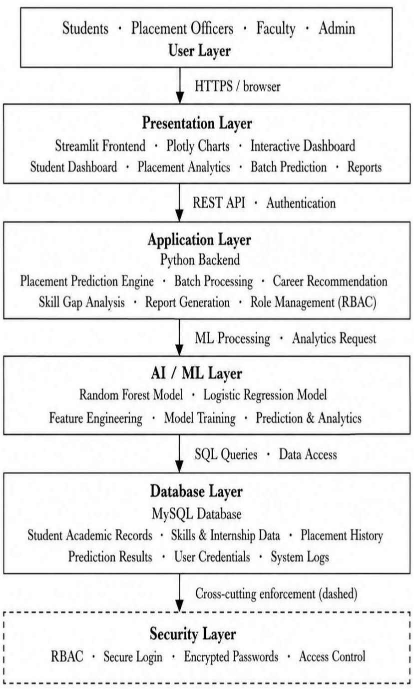

Fig. 1 System Architecture

A. Dataset Features

Academic Metrics:

- 10th Standard Percentage
- 12th Standard Percentage
- Undergraduate CGPA
- Employability Test Score
- Aptitude Test Score

Skill & Experience Metrics:

- Technical Skill Rating (1–10)
- Communication Skill Rating (1–10)
- Internship Experience (months)
- Prior Work Experience (yes/no)
- Number of Known Programming Languages

B. Evaluation Metrics

The platform's ML models are evaluated on the following metrics to ensure reliability and fairness:

- **Accuracy:** The proportion of correct predictions (placed / not placed) across the test dataset.
- **Precision and Recall:** Precision measures the reliability of positive placement predictions; recall measures the system's ability to identify all genuinely placeable students — important to minimise false negatives.
- **F1-Score:** The harmonic mean of precision and recall, providing a balanced single-metric evaluation for the imbalanced placement dataset.
- **ROC-AUC Score:** Evaluates the model's discriminative ability across all classification thresholds, providing a robust, threshold-independent performance measure.
- **Feature Importance Analysis:** Random Forest's built-in feature importance scores identify the most predictive variables, ensuring model interpretability and supporting student counselling.

VIII. APPLICATIONS

The platform's utility extends beyond individual placement prediction to broader institutional applications:

- **Pre-Placement Counselling:** Students receive data-driven career roadmaps months before recruitment drives, enabling targeted skill development.
- **Recruitment Drive Optimisation:** Placement officers can shortlist and prioritise students for specific company profiles using batch prediction outputs.
- **Curriculum Improvement:** Aggregate skill gap data reveals systemic weaknesses in the curriculum, informing programme revision decisions.
- **Alumni Trend Analysis:** Historical placement data visualised over multiple years helps institutions identify long-term trends and benchmark against peer institutions.

IX. FUTURE ENHANCEMENTS

- **AI-Based Career Assistant:** Integration of an intelligent chatbot (using LLM APIs) to provide real-time, conversational career guidance and placement preparation support.
- **Resume Analyser:** An NLP-powered module to evaluate student resumes against job description requirements and suggest targeted improvements.
- **Advanced Salary Prediction:** A regression model trained on salary outcome data to predict expected compensation based on student profile, enabling more granular career planning.
- **Mobile Application:** Cross-platform mobile development (React Native / Flutter) to improve accessibility and enable push notifications for placement drive alerts.
- **Recruiter Portal:** A dedicated interface for company recruiters to post job drives, define eligibility criteria, and access shortlisted student profiles.

X. CONCLUSION

This paper presented a Campus Placement Prediction and Career Outcome Analytics Platform that addresses a critical gap in higher education placement infrastructure. By replacing manual, CGPA-centric evaluation with a multi-factor ML-based prediction engine, the platform delivers evidence-based, personalised placement predictions and actionable career recommendations.

The system employs Random Forest and Logistic Regression algorithms trained on a rich feature set encompassing academic performance, aptitude, skill ratings, and experience metrics. Its six integrated modules — Data Input, Placement Prediction, Career Analytics Dashboard, Batch Prediction, and Report Generation — serve the diverse needs of students, faculty, placement officers, and administrators through a unified, role-governed interface.

REFERENCES

[1] Kotsiantis, S., Pierrakeas, C., and Pintelas, P. (2004). Predicting Students' Performance in Distance Learning Using Machine Learning Techniques. Applied Artificial Intelligence, 18(5), 411–426.

[2] Bharambe, A., Gaikwad, A., and Kulkarni, M. (2018). Placement Prediction Using Machine Learning. International Journal of Computer Science and Engineering, 6(6), 576–580.

[3] Jayaprakash, S., Krishnan, S., and Jaiganesh, V. (2020). Predicting Students' Academic Performance Using an Improved Random Forest Classifier. Advances in Mathematics: Scientific Journal, 9(6), 3705–3714.

[4] Ramaswami, M., and Bhaskaran, R. (2010). A CHAID Based Performance Prediction Model in Educational Data Mining. International Journal of Computer Science Issues, 7(1), 10–18.

[5] Gupta, A., and Singh, M. (2021). Design and Implementation of a Machine Learning Based Student Placement Prediction System. Journal of Engineering Education Transformations, 34, 103–110.

[6] Breiman, L. (2001). Random Forests. Machine Learning, 45(1), 5–32.

[7] Hosmer, D. W., and Lemeshow, S. (2000). Applied Logistic Regression (2nd ed.). John Wiley & Sons.

[8] Han, J., Kamber, M., and Pei, J. (2011). Data Mining: Concepts and Techniques (3rd ed.). Morgan Kaufmann.

[9] Scikit-learn: Machine Learning in Python. Pedregosa et al., JMLR 12, pp. 2825–2830, 2011.

[10] McKinney, W. (2010). Data Structures for Statistical Computing in Python. Proceedings of the 9th Python in Science Conference, 51–56.

