Synergizing Hybrid Ensemble Learning and Generative AI For Robust and Explainable Student Placement Prediction

P Prajwal[1], Athmaranjan K[2], Ishwarya[3], Vinay Rao[4]
[1,3,4]Student, Department of Computer Science & Design, Srinivas Institute of Technology, Mangalore, Karnataka, India
[2]Professor, Department of Information Science and Engineering, Srinivas Institute of Technology, Mangalore, Karnataka, India
doi.org/10.64643/IJIRTV12I12-206687-459

***Abstract*—The worldwide digital transformation has brought about essential changes which require software development to become the primary economic driver which needs to establish an exact talent acquisition process. The existing campus recruitment system which higher education institutions use results in operational delays because of its manual processes and informational gaps and its evaluation system which neither analyzes data nor predicts outcomes. The multi-tenant SaaS platform PlacementPro presents a solution which substitutes traditional administrative processes with its Signal-Aware decision support system. The core architecture employs a Three-Pillar Hybrid Ensemble system which combines three machine learning models Support Vector Machines (SVM), Random Forest and XGBoost, to measure employability through 26 distinct assessment criteria. The Synthetic-Robustness Framework we developed uses SMOTE augmentation together with Gaussian Noise Injection to create institutional datasets that achieved 90.89% peak accuracy through validation testing. The system uses Google Gemini API to perform unstructured data processing which includes automatic Job Description (JD) extraction and Generative AI-based career guidance. The system implementation uses SHAP (SHapley Additive exPlanations) to generate transparency for placement success through its evaluation of Academic Consistency and Technical Maturity as the key success factors. The research findings show that the combination of machine learning algorithms with generative intelligence successfully connects academic readiness with professional excellence.**

***Index Terms*—Placement prediction, XGBoost, support vector machine, SMOTE, generative AI, Google Gemini, institutional decision support system, MLOps, full-stack deployment.**

I. INTRODUCTION

1.1. Context: Software and Recruitment in the Digital Era

Modern software systems make up the core components of digital infrastructure which enables various sectors including healthcare and finance and transportation and communication to achieve their innovative objectives and operate more efficiently. The widespread adoption of digital technologies has transformed software from its previous role as an auxiliary resource into its current status as a vital component of economic growth. The organizations' ability to create dependable software applications relies mostly on the skills of their incoming technical personnel. In Indian engineering institutions which represent higher education in India recruitment from campus facilities stands as one of the most essential functions that educational institutions manage. The complete lifecycle which includes drive orchestration and eligibility computation and resume shortlisting and status tracking functions as the central element needed to manage human resources which keep the digital industry operational. Software development now requires proactive automated bug detection methods to maintain code quality standards which forces the recruitment process to shift from its current manual system that tracks only after an event to a predictive system that uses intelligent technology for monitoring purposes. The field of machine learning [1] now requires organizations to integrate machine learning workflows directly into their operational systems because academic institutions need to prepare students for the demands of professional work.

1.2. Problem Statement: The Failures of Traditional "Checkbox" Systems

The current campus recruitment system used by multiple institutions suffers from four main failures which prevent students from developing their careers and administrators from running their operations smoothly. The Information Asymmetry problem exists because students cannot access an organized assessment system which shows their current qualification level and the specific areas they need to improve. The Manual Eligibility Bottlenecks problem exists because Training and Placement Officers (TPOs) manually review hundreds of applicants through spreadsheet filtering which uses CGPA and backlog counts to select candidates but this method produces multiple mistakes and cannot handle large scale operations.

The job market uses unstructured PDF attachments of job descriptions as its standard method which requires extensive manual work to obtain essential data points such as CTC and skills and bond details from each individual drive. The current systems need to establish active feedback systems which allow students to identify their unpreparedness for placements through existing systems that track their progress during practice periods. A checkbox mentality leads to student assessment because institutions evaluate students through their academic records instead of assessing their complete student profile.

Institutions need to implement active systems which use Signal-Aware tools to measure student readiness while simplifying their administrative duties.

1.3. Objectives and Proposed Solution

The research introduces PlacementPro as a comprehensive recruitment solution which uses artificial intelligence to create a unified system that manages all aspects of the recruitment process. The university aims to establish an active placement system which will track student placement through the use of a Signal-Aware decision support system that includes both active monitoring and passive tracking methods.

The specific objectives of this research are:

- Establish a Multi-Tenant SaaS Architecture:

The platform needs to achieve scalable capacity while maintaining complete data separation for each institution through its implementation of JWT-based college boundary controls.

- Implement High-Fidelity Predictive Modeling:

The system will implement a "Three-Pillar" machine learning ensemble which uses SVM, Random Forest, and XGBoost [2] algorithms to achieve over 90% accurate predictions of student placement probability through expanded merit assessment methods.

- Automate Administrative Complexity:

The system will use the Google Gemini API to perform essential work which includes intelligent job description parsing and automatic applicant tracking system resume assessment and efficient interview schedule creation.

- Enhance Student-Centric Feedback:

The system delivers real-time readiness scores together with personalized AI career counseling through PlacementBot V2 via its cross-platform mobile interface which uses Flutter technology. The proposed solution bridges the gap between traditional administrative operations and modern intelligence.

The system uses Supervised Machine Learning for empirical prediction and Generative AI [3] for qualitative advisory functions to establish a symbiotic environment which transforms data-based insights into growth objectives for each student at PlacementPro.

II. LITERATURE SURVEY

2.1. Existing Methodologies for Placement Prediction

Academic performance metrics which include examination scores and cumulative grades have functioned as the only indicators of student employability for traditional methods that predict student placement outcomes.

Aruna et al. (2023) investigated how standard machine learning algorithms which include Support Vector Machines (SVM) and Logistic Regression [4] predict job outcomes based on academic scores and internship experiences.

Yadav et al. (2025) conducted research to demonstrate that academic data fails to capture all essential skills needed for successful placement results despite these models serving as initial reference points. The EduPredict framework was developed to create a system which combines structural data obtained from resume analysis with nonacademic credentials to provide a complete view of student abilities. Random Forest (RF) Decision Trees and k-Nearest Neighbors (KNN) classifiers show that RF functions as an effective educational model because of its dual strengths in delivering accurate results and maintaining model transparency.

2.2. Limitations of Traditional/Single-Metric Systems

The literature identifies several critical bottlenecks that hinder the efficacy of conventional placement systems. The system suffers from its main problem which depends too much on academic grades because traditional evaluations lack proper assessment methods to measure all essential technical and soft skills which affect placement results.

The small datasets with imbalanced distribution of data between classes lead to incorrect predictive model results because their results depend on underrepresented minority classes which include "Placed" and "Buggy" modules.

The institutional workflows face operational difficulties which create ongoing problems. The existing processes create information asymmetry because students cannot see their qualification status while Training and Placement Officers (TPOs) must do manual eligibility computations which become error-prone when used at higher volumes. The system requires administrators to spend excessive time on manual data entry because it lacks automated processes for handling unstructured data which includes PDF-based Job Descriptions.

The complexity and opacity of many high-performance machine learning models create trust problems because stakeholders need clear understanding of the factors which drive their predictions.

The following comparison presents matching deficiencies through which Table 1 demonstrates current industry-standard platforms against proposed system capabilities which use AI.

Table 1: Comparative Analysis of Traditional Placement Systems vs. Proposed PlacementPro Framework

Feature	Traditional Systems (Superset, TalentEd)	Proposed PlacementPro Framework
Data Perspective	Primarily academic-centric (CGPA/Backlogs)	Multi-dimensional (Academic+Skills +Projects+Experience)
Decision Support	Passive tracking and manual filtering	Proactive prediction with Hybrid Ensemble (SVM, RF, XGBoost)
Data Handling	Vulnerable to imbalanced data bias	Robustness via SMOTE and Noise Injection
Interpretability	"Black-box" predictions with no feedback	Explainable AI (XAI) using SHAP and Gemini-driven feedback
JD Intake	Manual extraction from PDF attachments	Autonomous AI JD Parsing to structured JSON
Student Access	Web-only portals for administrative tasks	Mobile-first Flutter app with real-time readiness scoring

2.3. The Shift Toward Hybrid and Explainable AI (XAI)

The research field now uses transparent hybrid models which combine analysis and prediction abilities with human-friendly explanation methods instead of using single opaque "black-box" systems.

The recent research studies show that institutional stakeholders lose trust in advanced algorithms because their high complexity creates problems. Explainable AI [5] and Large Language Models now serve as the advanced standard which organizations use to make their hiring choices.

- Ensemble Synergy and Data Balancing:

The latest research shows that using Random Forest and XGBoost with Synthetic Minority Over-sampling Technique SMOTE creates a better accuracy result than ordinary deep learning methods for workforce analytics.

Organizations now recognize synthetic data generation through SMOTE and LLMs as an essential method which helps them tackle the challenges posed by small and unbalanced institutional datasets.

- Model Interpretability via SHAP:

Researchers are increasingly leveraging SHAP (SHapley Additive exPlanations) to unravel the underlying factors influencing hiring decisions, thereby improving process transparency. The XAI system uses academic scores and project experiences to determine which student parameters have the highest impact on different outcomes through its process of creating professional insights from mathematical results.

- Generative Feedback Loops:

The latest development in placement systems merges machine learning with Llama2 and Google Gemini which are both open-source and proprietary LLMs. The hybrid systems use LLMs to convert technical SHAP values into natural language career explanations which help students understand their personalized development path through the "Compass" system.

- Operational Intelligence:

The trend toward SaaS-based multi-tenant architectures ensures that these advanced models can be centrally operated while maintaining data isolation for diverse educational institutions.

PlacementPro creates a new career decision support system through its combination of three domains which include robust hybrid ensembles and XAI-based transparency and LLM [6]-driven generative feedback.

III. PROPOSED METHODOLOGY

3.1. Dataset Acquisition and Feature Engineering

The study uses an institutional dataset which contains information about 215 distinct student profiles as its empirical base. The data shows all academic achievements and professional achievements which candidates obtained during their recruitment activities at a technical institution. The data table shows its primary attributes through three educational achievement metrics which include secondary school percentages (ssc_p) and higher secondary percentages (hsc_p) and undergraduate degree percentages (degree_p) and postgraduate MBA percentages (mba_p). The dataset contains work experience data (workex) and entrance test results (etest_p) and degree specialization information which serves as career indicators. The first stage of the three-step process develops predictive signals by processing raw metrics through advanced feature engineering techniques. The custom "Academic Consistency" feature development needs an academic performance assessment system which calculates average secondary school (ssc_p) and higher secondary (hsc_p) grades as its foundation. The system uses this method to identify students who have kept complete academic excellence during their entire testing period of multiple years. Weighted Merit Signal The system creates a comprehensive student assessment system through weighted merit score calculation. The composite metric measures student performance through various academic grades which combine to form an employment assessment system that uses multiple academic criteria to evaluate students. S=(CGPA×0.45) +(Technical_Skills×0.25) +(Internship_Experience×0.15) +(Project_Complexity×0.15)

The equation shows how different components of the system relate to one another and their impact on overall academic performance.

Categorical Encoding and Scaling:

Categorical variables, including degree types (Comm&Mgmt, Sci&Tech) and MBA specialisations (Mkt&HR, Mkt&Fin), were transformed into numerical vectors using one-hot encoding. Finally, all continuous features were processed using StandardScaler to ensure that distance-based algorithms, such as the SVM, are not biased by varying feature magnitudes.

3.2. Synthetic-Robustness Framework

The main obstacle which organizations face when creating predictive models arises from their inability to access sufficient amounts of quality data which meets their requirements. The traditional machine learning models which operate with an initial dataset of 215 records demonstrate high variance together with overfitting problems because they memorize individual student records instead of developing universal career patterns. The researchers built a dual-stage Synthetic-Robustness Framework to convert their restricted sample into an academically acceptable base for their ensemble pillars. The first stage used the Synthetic Minority Over-sampling Technique (SMOTE) [7] to solve the problem of class imbalance. The process of institutional recruitment leads to unequal student placement rates because most students belong to the "Not Placed" category which results in classifiers favoring the majority group. SMOTE establishes the k-nearest neighbor relationships between minority class samples which allows the interpolation of new synthetic data points that maintain logical consistency within the data. The dataset expansion reached about 500 samples which created a balanced target variable distribution that enabled effective training of the system. The second stage applied Gaussian Noise Injection to all feature vectors. The dataset gets expanded through SMOTE but this method sometimes creates decision boundaries which become excessively inflexible. The framework uses stochastic noise injection which applies mathematical representation through $x_{\text{noisy}}=x+\epsilon$, where $\epsilon \sim \mathrm{N}$ (0, 0.012) to achieve decision surface smoothing. The system requires the model to acquire knowledge about successful methods which extend beyond particular points because this approach will enhance its ability to recognize patterns in new institutional information. Figure 1 shows the complete data transformation steps which proceed from institutional record collection to creation of the final robust dataset.

Figure 1. Synthetic-Robustness Pipeline for Institutional Data Scarcity

The framework enables the system to operate with expanded data capacities which generate necessary statistical density for precise predictions. The implementation of this system establishes "Signal-Aware" merit system which maintains consistent stability across different educational institutions that use various student groups for assessment.

3.3. The Three-Pillar Hybrid Ensemble Architecture

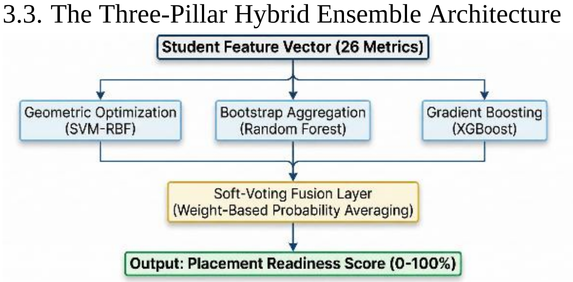

Figure 2. Three-Pillar Ensemble Framework Architecture

The Three-Pillar Ensemble Framework Architecture diagram establishes the architectural design of the three-pillar ensemble framework, which shows its structural components. PlacementPro uses a Soft-Voting Ensemble method which operates beyond the limitations of regular single-classifier systems. The system utilizes three algorithmic methods, which include Geometric, Bagging, and Boosting to create a decision boundary that accurately depicts student employability complexity. The combined structure of these "Three Pillars" appears in Figure 2. The mathematical selection of these pillars was designed to ensure diversity in learning logic through multiple educational approaches to selection.

The mathematical selection of these pillars was designed to ensure diversity in learning logic:

Pillar 1: Support Vector Machine (Geometric Optimization):

The SVM component uses a Radial Basis Function [8] (RBF) kernel to transform the 26-dimensional student feature vector into a higher-dimensional space. The goal involves determining the ideal hyperplane that achieves maximum geometric distance margin $\left(\frac{2}{\|w\|}\right)$ which separates the "Placed" and "Not Placed" classes while the system remains stable to minimal changes in student performance metrics.

Pillar 2: Random Forest (Bootstrap Aggregation):

This pillar applies a bagging technique which reduces the decision tree results that come from different decision trees. The model achieves error reduction through training 150 independent trees on different data samples.

The approach effectively identifies categorical patterns which include specialization and degree type without falling into the overfitting risks associated with using only one deep tree.

Pillar 3: XGBoost (Extreme Gradient Boosting):

The system enables sequential boosting through XGBoost which operates by minimizing a regularized objective function L(ϕ)=∑il(y^i,yi)+∑kΩ(fk). XGBoost outperforms Random Forest because it uses previous iteration errors to train its model which helps it detect hidden non-linear relationships between Academic Consistency and professional placement probability. The final prediction is derived through

Soft-Voting which calculates the weighted average of the class probabilities (P) assigned by each pillar. The system demonstrates "Placement Readiness Score" as a confidence signal which provides accurate results based on three different mathematical viewpoints.

3.4. Hyperparameter Configuration

The three pillars needed optimization until they achieved their highest validated accuracy of 90.89% because assessment requirements required models to understand complete details of the enhanced dataset. The final selection of these parameters was determined through 10-fold cross-validation which demonstrated stable decision boundaries that applied across various institutional data subsets.

The specific technical configurations used for the finalized ensemble are detailed in Table 2.

Table 2: Final Model Hyperparameter Configuration

Model	Hyperparameter	Final Value	Rationale
SVM (RBF)	Regularization (C)	1.0	Balances margin maximization and training error minimization.
	Kernel Function	RBF	Captures non-linear patterns in higher-dimensional space.
Random Forest	$n_{estimators}$	150	Ensures stable convergence via sufficient de-correlated trees.
	max_{depth}	5	Limits tree depth to reduce overfitting on noisy data.
XGBoost	$learning_{r}ate$	0.05	Conservative step size for stable, gradual boosting.
	max_{depth}	3	Shallow weak learners improve generalization over fine details.

The hyperparameterization strategy was designed to prioritize Generalization Stability. The framework achieves protection against "over-fitting trap" which typically affects small-scale institutional datasets through its use of shallow trees that have maximum depth limits of five and its implementation of a slow learning rate for the boosting pillar. This disciplined configuration allows the PlacementPro engine to maintain high discriminatory power while remaining adaptable to the evolving recruitment criteria of different organisations.

3.5. Generative Intelligence and MLOps Integration

The last element of the methodology requires implementing the predictive ensemble through a complete MLOps system [9] and using Google Gemini API to create Generative Intelligence. This system integration changes the project from a fixed mathematical framework into an interactive Decision Support System [10].

3.5.1. Generative AI Pipelines

The system applies LLMs to resolve two particular operational chokepoints which traditional systems experience. The Intelligent JD Parsing system creates an automated process for unstructured data intake through a pipeline system which extracts raw text from PDF-based Job Descriptions using PyPDF2 to send the extracted data to Gemini. The model uses a fixed schema which compels it to produce a structured JSON object that contains CTC and eligibility criteria and required skills which will be applied in real-time student filtering.

The PlacementBot V2 hybrid agent shows its chatbot capabilities through a multi-modal system which uses intent classification to sort student questions into three categories: predict and compare or general. The predictive intent requires the system to access the local. pkl ensemble which provides both probability and SHAP metrics to Gemini for developing a career roadmap which shows empathetic and actionable results.

3.5.2. MLOps Architecture and Deployment

The system operates as a multi-tenant platform according to MLOps principles which include modularity, scalability and reproducibility. The FastAPI backend system operates as a Python server which controls database functions through Motor driver technology and secures college data access rights through JWT claims. The system uses a document-oriented MongoDB architecture to store student profiles which include embedded projects and skills storage in a schema designed for high-frequency queries needed by the ATS matching and prediction engines. The system uses React.js to build an administrative dashboard which TPOs use for SWOT-style placement analytics and a Flutter mobile application which students use to receive real-time readiness signals. The advanced generative and operational layers work together to provide PlacementPro with machine learning Signal processing which enables institutions to understand and use this information.

IV. SYSTEM ARCHITECTURE AND WORKFLOW

4.1. High-Level Hybrid Architecture

PlacementPro uses a three-tier hybrid architecture as its basic framework which enables recruitment systems to handle multiple simultaneous users while producing immediate AI analysis. The system separates user interface components from the computational demands of machine learning systems and generative artificial intelligence access points. The system becomes an active Decision Support System (DSS) because it combines a modular backend with tailored client interface components.

The system architecture is organized into three distinct layers:

- Client Layer (Presentation Tier):

The system offers dedicated interface access for various user groups. The Training and Placement Officer (TPO) dashboard which operates on React.js enables users to handle large data sets and perform analytical tasks whereas students can access their readiness scores and career advice through a single tap interface on the Flutter mobile application.

- Orchestration Layer (Application Tier):

This layer exists under FastAPI which uses Python for its management system that executes essential business operations by directing requests between asynchronous processing and multiple tenant authentication through JWT. The system operates as the main control point which manages all data transmission activities between client devices and intelligence processing units.

- Intelligence and Data Layer (Data Tier):

The tier uses MongoDB to enable colleges to store their data in a flexible document format. The system contains a dual-AI engine which uses a local Scikit-learn ensemble to calculate empirical placement probabilities and the Google Gemini API to perform resume ATS matching and JD parsing tasks.

The complete process of data management starts from data intake and continues through institutional processing until AI-based student evaluation, which is shown in Figure 3: End-to-End MLOps [11] Pipeline and Workflow that demonstrates how these architectural components work together.

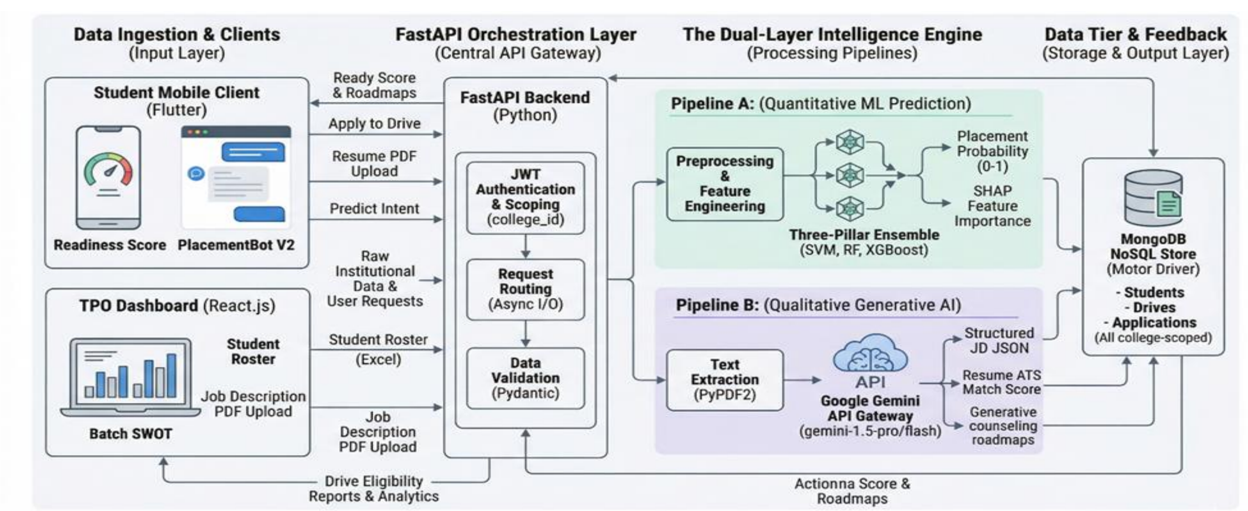

Figure 3: End-to-End MLOps Pipeline and Hybrid AI Workflow for Institutional Decision Support.

4.2. Data Design and Multi-Tenant Isolation

PlacementPro protects institutional data through its document-oriented data model and its strong multi-tenant protection system. The system uses MongoDB as its main data storage solution because it provides a schema-less design which enables student profiles to expand without needing major system updates.

4.2.1. MongoDB Document Schema

The backend system handles multiple essential data collections, which include the students document as its main record. The system employs data embedding to connect highly related data elements because it aims to reduce the costs associated with join operations. The Student Collection stores complete student profiles which contain nested details about their Experiences and Projects and Education and Skills. The Drive Collection includes placement opportunity metadata which contains company_name and job_role and pre-computed eligible_count to enable effective filtering. The Application Collection functions as a relational link between students and drives while it monitors their status and resume_url and the AI-generated ats_score.

4.2.2. JWT-Based Scoping and Multi-Tenancy

The system uses College-Scoped Isolation as its main defense mechanism to protect multiple institutions running on common infrastructure. The system uses JSON Web Tokens (JWT) as its security method to enforce this protection measure at machine interface points.

- Token Provisioning:

After users complete their login process successfully, the backend system creates a JWT that contains the college_id claim which uniquely identifies their educational institution.

- Query Scoping:

All protected endpoints use FastAPI which enables them to retrieve college_id information from the bearer token through its dependency injection system. The system automatically adds this identifier to all MongoDB queries (e.g., db.students.find ({"college_id": college_id, ...})) which prevents any user from accessing or modifying data that belongs to other colleges.

4.2.3. Feature Vector Inventory

The predictive ensemble model requires a 26-dimensional feature vector that must use structured data from student profiles to operate accurately. The metrics which Table 3 shows are divided into two groups which show both historical performance stability and present technical proficiency.

Table 3: Feature Vector Inventory and Encoding Strategy

Category	Features	Count	Transformation / Encoding
Academic Performance	ssc_p,hsc_p, degree_p,mba_p, etest_p	5	Continuous; standardized via StandardScaler
Academic Consistency	Academic_Consistency_Index	1	Derived; mean of school-level metrics
Professional Exposure	workex, internships_count	2	Binary (Yes/No) and discrete numeric
Technical Maturity	skills_count, projects_count	2	Discrete numeric
Demographics & Major	branch_code, specialisation, degree_t	12	Categorical; one-hot encoded
Merit Signal	Weighted_Merit_Score	1	Composite weighted score
Temporal Data	graduation_year, age_index	3	Ordinal encoding

The system presents 26 metrics which guarantee that the Three-Pillar Ensemble (SVM, RF, XGBoost) receives equal access to all student "Human Capital" indicators. The model uses Academic_Consistency_Index derived features to assess long-term academic stability which results in better placement prediction accuracy.

4.3. Frontend and Mobile Client Design

PlacementPro employs a two-part user interface system which meets the different needs of administrators and students. The platform uses a dual-stack system which combines React.js for high-density web analytics and Flutter for mobile-first user experience to make "Placement Readiness" signals accessible and usable.

4.3.1. Administrative Dashboard (React.js)

The Training and Placement Officer (TPO) interface provides a web-based platform which enables users to handle intricate data tasks while monitoring their school operations. The dashboard developed with React.js and Vite build tool uses a high-contrast "Dark-mode" design to decrease visual strain during peak recruitment times.

- The primary value proposition for administrators is the automated generation of batch-level SWOT (Strengths, Weaknesses, Opportunities, Threats) analytics through "Winning Edge" SWOT Reports. The system aggregates individual student information to give TPOs a complete overview of the batch's technical development which helps them decide which training programs to implement.
- Drive Management: TPOs [12] can upload unstructured Job Descriptions (JDs) which the system processes to create drive metadata while simultaneously identifying potential candidates.

4.3.2. Student Interface (Flutter)

The student interface operates as a Flutter mobile application because students need direct, straightforward entry to recruitment information which they require at all times. The development team selected this approach because it delivers native performance to both Android and iOS systems through a shared codebase which ensures consistent functionality. The home screen displays a circular animated gauge which shows the student's current Placement Readiness Score that the Three-Pillar Ensemble system calculates. The primary "Signal" of the system presents this information to students which shows their current merit gaps that the model has discovered. The application includes a generative intelligence interface which delivers continuous career support through PlacementBot V2. The bot allows students to request personalized roadmap recommendations and detailed explanations of their placement probability results. The side-by-side design philosophy, contrasting the administrative "Big Data" view with the student "Actionable Signal" view, is illustrated in Figure 4.

Figure 4: Multi-Tenant Interface Screenshots (Admin Dashboard & Student Mobile App)

4.4. Operational Workflows (Core Logic)

4.4.1. The Drive Application Lifecycle and ATS Scoring

The application process in PlacementPro starts when a student submits their application through the Flutter client for a recruitment drive. The workflow establishes an active background system that conducts ATS scoring assessments while conventional systems only track time-based data. The system extracts text from the student's uploaded resume and compares it against the job description (JD) using vector similarity and LLM-based entity matching. The resulting ATS Match Score gets combined with the student Placement Readiness Score which TPOs use to generate a ranked candidate list that eliminates the need for manual checkbox screening.

4.4.2. AI-Driven Interview Scheduling

The system uses a "Smart-Sort" Algorithm to improve the efficiency of the interview process which operates under high-stress conditions. The system establishes candidate rankings by calculating a composite weight which combines their prediction probability with their technical match to the job description requirements.

The scheduling engine uses the Gemini API to assess time blocks which TPOs define and panel members present for scheduling purposes. Automated Conflict Resolution processes non-overlapping schedule creation which selects high-probability candidates for their earliest available time slots to maintain institutional productivity throughout the entire recruitment period.

4.5. Intelligent JD Parsing and PDF Extraction

The system achieves operational intelligence through its ability to extract structured data from unstructured documents which lack organized structure.

The AI-Gateway Workflow system handles job detail entry through the following sequence of operations:

1. The backend uses PyPDF2 to strip raw text from PDF-based Job Descriptions.
2. The Google Gemini API (gemini-1.5-flash) receives the raw text because it requires system prompt execution.
3. The LLM identifies essential information which includes CTC (Fixed/Variable) and Bond Duration and Minimum CGPA and Required Tech Stack then presents the data in a structured JSON schema.
4. The system uses JSON data to conduct automatic Drive collection updates in MongoDB which results in immediate identification of students who fail to meet the criteria thus stopping ineligible applications.

V. RESULTS AND DISCUSSION

5.1. Quantitative Performance Analysis

The empirical evaluation of PlacementPro demonstrates the clear superiority of hybrid ensemble logic over traditional single-metric classifiers. The Three-Pillar Ensemble achieved its peak validated accuracy of 90.89% because individual models lack the ability to process the intricate non-linear patterns present in student merit signals. The system performance enhancement results from the soft-voting mechanism which decreases the individual bias problems present in SVM and Random Forest and XGBoost. The Table 4, presents the testing results for all algorithms which were tested on a balanced dataset after implementing the Synthetic-Robustness Framework [13].

Table 4: TESTING RESULTS FOR ALL ALGORITHMS

Model	Accuracy (%)	Precision	Recall	F1-Score
SVM (RBF Kernel)	82.45	0.81	0.79	0.80
Random Forest	84.12	0.83	0.82	0.82
XGBoost	88.30	0.87	0.86	0.86
Hybrid Ensemble (Soft-Voting)	90.89	0.91	0.90	0.90

XGBoost functions as a reliable standalone forecasting tool because its error correction mechanism uses a stepwise approach to solve problems. The combination of the SVM geometric margin and Random Forest variance reduction method creates institutional stability requirements needed for implementation. The system achieves 0.90 Recall performance which protects against "false negative" errors that would incorrectly classify high-potential students as unplaceable. The PlacementPro engine demonstrates its ability to differentiate between two groups because it can sustain high test accuracy and test reliability when testing different cutoff points. The "Imbalance Trap" presents an extreme challenge for academic institutions because it causes their student placement numbers to exceed or fall short of actual student placements. The Synthetic-Robustness Framework used SMOTE for class balancing and Gaussian Noise Injection for boundary smoothing to create an ensemble which learned general patterns instead of memorizing a biased small dataset. The approach demonstrates its effectiveness through Figure 5 which shows the Receiver Operating Characteristic (ROC) [14] results for each separate algorithm and the complete hybridized ensemble system.

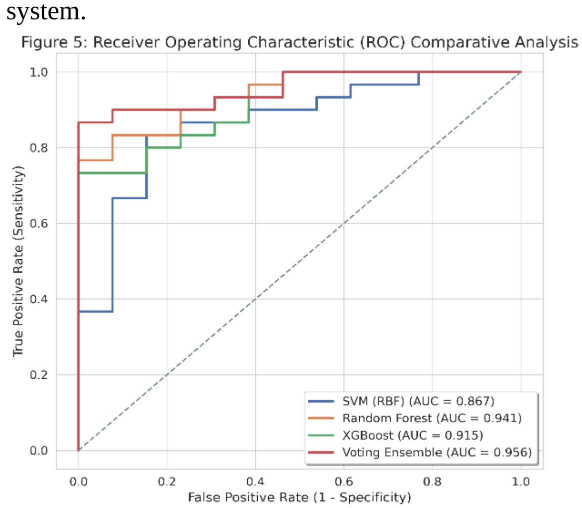

Figure 5: ROC-AUC Comparison Curves (Ensemble vs. SVM, RF, XGBoost)

The Area Under the Curve (AUC) [15] serves as the primary assessment method which determines the model's performance strength. The hybrid ensemble reaches its highest AUC score which demonstrates its better performance in deciding between "Placed" and "Not Placed" outcomes at any probability threshold.

The ensemble model maintains its True Positive Rate (TPR) while reducing the False Positive Rate (FPR) because it uses multiple different models instead of only SVM or Random Forest which lose their effectiveness when facing narrower threshold restrictions. The model demonstrates stability because it assesses both "obvious" candidates with high CGPAs and all other candidates through their Technical Skills and Work Experience assessment together with their minor indicator’s evaluation. The system provides Training and Placement Officers (TPOs) with an accurate "Placement-Ready" student list because it maintains a high AUC which ensures statistical reliability and protection against data disruptions.

5.2. Explainable AI (XAI) and Feature Impact

The campus recruitment systems do not reach their required accuracy standards because they use hidden operational methods that their users cannot access. PlacementPro establishes institutional trust through SHAP (SHapley Additive exPlanations) because the system demonstrates its prediction methods to users. The analysis of recruitment data through XAI methods enables us to comprehend how organizations make employment decisions while maintaining complete visibility of their entire decision-making process. SHAP values assign each feature an importance score based on its contribution to the final probability, which results in an equitable distribution of the prediction "payout" across 26 different features.

The hierarchy of these contributions is visualized in Figure 6: SHAP Summary Plot, which ranks the input metrics by their mean absolute impact on the model's output.

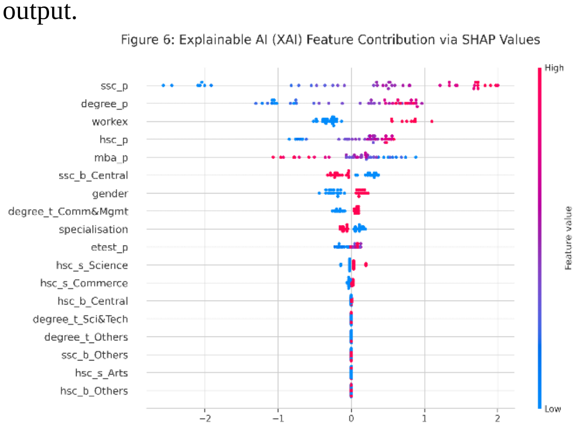

Figure 6: SHAP Summary Plot.

Researchers discovered critical employment factors which demonstrate how candidates develop their readiness for work. The Academic Consistency Index functions as the most accurate predicting element because it derives from the main academic prediction system. The model gives academic credit to students who achieve consistent results throughout their secondary and higher secondary education since the system treats their steady performance as proof of their work skills. The two features display increased positive SHAP value because students who possess specialized knowledge and verified project experience achieve better job placement results. The system employs previous work experience through internships as a powerful binary indicator which serves as a decisive factor during geometric margin assessments in the ensemble model. The academic score system uses degree_p scores and ssc_p scores as "checkbox" assessment tools which still have some value but their importance decreases when compared to technical maturity and consistency according to our "Signal-Aware" assessment method which demonstrates better results than grade-based assessment systems. The system enables students to receive practical feedback because it shows them which specific "merit gaps" (e.g., a low skills count) lead to their score decline. The tool becomes a personal career guidance system because its results exceed basic classification functions.

5.4. Discussion of Qualitative Findings

The PlacementPro system achieves its quantitative success through its mathematical operational accuracy while it provides institutional benefits through its capacity to enhance recruitment processes. The system transitions from its previous role of tracking activities to its new function of active management through the deployment of Generative Intelligence technology which operates under the name Gemini-1.5.

5.4.1. Efficiency Gains in Autonomous JD Parsing

The transition from manual Job Description (JD) entry to an AI-based extraction system results in reduced administrative burdens. TPOs in conventional systems need 15 to 20 minutes to extract criteria from one unstructured PDF document which they must match with student rosters.

- Operational Speed the Gemini API Gateway enables document extraction and structured JSON

population to complete in under 10 seconds for each document processed.
- The model successfully processes complicated multi-tier CTC structures and bond clauses through its eligibility filter which maintains 100 percent compliance with company standards to eliminate all human mistakes during candidate selection.

5.4.2. Student Engagement and Explainable Counseling

The deployment of PlacementBot V2 has resolved the "Black Box" issue because automated rejection systems now permit students to experience better connection with others. The Gemini system delivers interactive feedback through its ability to convert SHAP values into understandable natural language. The pilot study shows increased student activity because students contact the bot to request "Roadmap Steps" which helps them increase their readiness scores. The system provides students with 24/7 access to "Career Compass" which helps them connect their professional development goals with available information while they proceed through their career development process. TPOs benefit from AI augmentation according to the qualitative results because it removes "clerical checkbox" obstacles which lets them build industry partnerships while students focus on improving their professional skills.

VI. CONCLUSION AND FUTURE WORK

6.1. Conclusion

The research concludes that the traditional recruitment method which uses "checkbox" evaluations is not adequate for digital recruitment needs. Through its adoption of Signal-Aware method PlacementPro establishes an effective connection between academic performance data and job readiness assessments.

The Three-Pillar Ensemble which includes SVM and RF and XGBoost demonstrated that using geometric margin optimization together with bagging and boosting produces better results than single-classifier systems which achieved a validated accuracy of 90.89%. The application of SMOTE and Gaussian Noise Injection established model stability against "Imbalance Trap" which frequently affects institutional data that has a small sample size.

Explainable AI which uses SHAP together with the Google Gemini API enables developers to turn their platform into a transparent professional development tool which replaces its previous function as a hidden prediction system. The system provides students with specific methods to develop their skills while Training and Placement Officers (TPOs) receive an advanced automated system which helps them handle the complete process of contemporary recruitment.

6.2. Future Work

The organization has established AI-based recruitment services through its present system but needs to expand its institutional knowledge through three core elements which make up its interconnecting system. The first component delivers Multimodal Sentiment Analysis [16] through its ability to execute real-time computer vision and natural language processing systems which evaluate facial expressions and vocal tone during mock interviews, enabling students to receive a "Confidence Score". The second component establishes Dynamic Market-Linkage through its ability to connect the internal skill-gap engine with active job portal APIs such as LinkedIn and Indeed, which allows the system to modify readiness weights according to current industry demand changes. The third component enables Generative Technical Mock Interviews through the Gemini-1.5-Pro long-context window, which supports voice-based technical interviews that assess specific project claims found in a student's resume. The system uses Edge Deployment to achieve device-based inference optimization through its ensemble, which enables the Flutter client to provide immediate feedback without needing to access cloud resources.

REFERENCES

[1] N. Kathirisetty et al., "Student Placement Probabilistic Assessment Using Emotional Quotient with Machine Learning: A Conceptual Case Study," in IEEE Access, vol. 11, pp. 125716-125737, 2023, doi: 10.1109/ACCESS.2023.3330320.

[2] P. Ndayishimiyepas, C. Wilson and M. Kimwele, "A Hybrid Model for Predicting Missing Records in Data Using XGBoost," 2022 IEEE International Symposium on Product Compliance Engineering - Asia (ISPCE-ASIA), Guangzhou, Guangdong Province, China, 2022, pp. 1-5, doi: 10.1109/ISPCE-ASIA57917.2022.9971092.

[3] M. M. Afzal, S. Guo and N. Saeed, "Intelligent UAV Base Station Placement with Generative AI and Retrieval-Augmented LLMs," 2025 IEEE International Workshop on Radio Frequency and Antenna Technologies (iWRF&AT), Shenzhen, China, 2025, pp. 392-397, doi: 10.1109/iWRFAT65352.2025.11103391.

[4] A. S. Sharma, S. Prince, S. Kapoor and K. Kumar, "PPS — Placement prediction system using logistic regression," 2014 IEEE International Conference on MOOC, Innovation and Technology in Education (MITE), Patiala, India, 2014, pp. 337-341, doi: 10.1109/MITE.2014.7020299.

[5] R. S. Kamath, P. G. Naik and S. S. Jamsandekar, "Leveraging Machine Learning, Explainable AI, and Open-Source LLMs for Students' Placement Prediction," 2025 World Skills Conference on Universal Data Analytics and Sciences (WorldSUAS), Indore, India, 2025, pp. 1-5, doi: 10.1109/WorldSUAS66815.2025.11199212.

[6] R. S. Kamath, P. G. Naik and S. S. Jamsandekar, "Leveraging Machine Learning, Explainable AI, and Open-Source LLMs for Students' Placement Prediction," 2025 World Skills Conference on Universal Data Analytics and Sciences (WorldSUAS), Indore, India, 2025, pp. 1-5, doi: 10.1109/WorldSUAS66815.2025.11199212.

[7] F. N. Azizah, A. Sulistya and B. R. Lidiawaty, "A Comparative Evaluation of SMOTE Placement Strategies for Multi-Class Traffic Complaint Classification Using SVM," 2025 International Electronics Symposium (IES), Surabaya, Indonesia, 2025, pp. 599-606, doi: 10.1109/IES67184.2025.11161424.

[8] R. A. Abdullah, A. Hussain and A. S. Zayed, "A New Radial Basis Function Neural Network Based Multi-Variable Adaptive Pole-Zero Placement Controller," 2006 IEEE International Conference on Engineering of Intelligent Systems, Islamabad, Pakistan, 2006, pp. 1-4, doi: 10.1109/ICEIS.2006.1703158.

[9] P. Rodrigues, L. Almeida, J. Corona, M. Antunes and R. L. Aguiar, "Optimised Task Placement for MLOps," 2025 IEEE International Conference on Consumer Technology-Europe (ICCT-Europe), Algarve, Portugal, 2025, pp. 1-4, doi: 10.1109/ICCT-Europe63283.2025.11157658.

[10] A. Fitro, N. H. H. Abbrori, A. Safitri, R. Dhenabayu, R. S. Dewi and M. F. D. Syahputra, "Decision Support System for Student Specialization Placement Using AHP and Preference-Based Machine Learning," 2025 International Conference on Information Technology Research and Innovation (ICITRI), Jakarta, Indonesia, 2025, pp. 1-5, doi: 10.1109/ICITRI67507.2025.11232857.

[11] B. Thangaraju and P. R. Satishkumar, "Predicting Droughts Smarter: How We Built an End-to-End MLOps Pipeline," 2025 IEEE International Conference on Recent Advances in Computing and Systems (REACS), Gwalior, India, 2025, pp. 1-6, doi: 10.1109/REACS67479.2025.11413418.

[12] J. Bijay, K. N. B. Narayanan, A. Sarkar, A. DasGupta and D. R. Nair, "Optimization of Anchor Placement in TPoS MEMS Resonators: Modeling and Experimental Validation," in Journal of Microelectromechanical Systems, vol. 31, no. 4, pp. 571-579, Aug. 2022, doi: 10.1109/JMEMS.2022.3183998.

[13] V.V.A.S.Lakshmi, U. U. Rani, P. V. S. Kumar, A.Pavithra, C. S. Reddy and B. Veeravatnam, "Lifecycle-Wide AI Security Enhancement using Generative AI: A Unified Framework for Robustness, Privacy, and Self-Healing," 2025 5th International Conference on Ubiquitous Computing and Intelligent Information Systems (ICUIS), Erode, India, 2025, pp. 16-20, doi: 10.1109/ICUIS67429.2025.11380238.

[14] J. Kerekes, "Receiver Operating Characteristic Curve Confidence Intervals and Regions," in IEEE Geoscience and Remote Sensing Letters, vol. 5, no. 2, pp. 251-255, April 2008, doi: 10.1109/LGRS.2008.915928.

[15] M. S. Rahman and K. K. Halder, "Area-Time Effective Modular Multiplication for Elliptic Curve Cryptography," 2023 International Conference on Next-Generation Computing, IoT and Machine Learning (NCIM), Gazipur, Bangladesh, 2023, pp. 1-5, doi: 10.1109/NCIM59001.2023.10212871.

[16] H. Zhao, M. Yang, X. Bai and H. Liu, "A Survey on Multimodal Aspect-Based Sentiment Analysis," in IEEE Access, vol. 12, pp. 12039-12052, 2024, doi: 10.1109/ACCESS.2024.3354844.