# Pathnix: Machine Learning-Based Placement Prediction and Career Recommendation for Engineering Students

**Praveen M. Hallur,**
Department of Computer Science and Business Systems
Dr. Ambedkar Institute of Technology, Bengaluru, India

## Abstract

Placement preparation is an important challenge for engineering students because academic performance alone does not completely represent their employability. Students may have different levels of technical skills, projects, internships, certifications, and preparation experience. Traditional placement systems mainly store student information and placement records but provide limited predictive analysis and personalized career guidance.

**Pathnix** is a machine learning-based placement prediction and career recommendation system developed to address this limitation. The system analyzes student information such as academic performance, technical skills, projects, internships, certifications, and other placement-related attributes. Logistic Regression is used as the primary machine learning algorithm to predict placement probability. Based on the prediction and student profile, the system provides readiness analysis, identifies weak areas, and generates career recommendations such as suitable career paths, job roles, company types, salary ranges, and improvement suggestions. The system also provides separate dashboards for students, faculty, and administrators.

The current evaluation achieved approximately **78% accuracy, 78% precision, 76% recall, and 77% F1-score**. The results demonstrate that machine learning can provide useful insights from structured student data and support data-driven placement preparation. Pathnix can therefore serve as a decision-support platform for students and faculty rather than simply functioning as a conventional student record system.

**Keywords:** Machine Learning, Placement Prediction, Logistic Regression, Career Recommendation, Student Analytics, Educational Data Mining.

# I. INTRODUCTION

Engineering students face increasing competition during campus placements. Academic performance is important, but placement outcomes are also influenced by technical skills, internships, projects, certifications, communication abilities, and preparation practices.

Traditional placement management systems primarily focus on storing student information and maintaining placement records. Although such systems are useful for administration, they provide limited information about a student's future placement readiness.

A student with a high CGPA may still have limited practical skills, while another student with an average CGPA may have strong programming skills, internships, and projects. Therefore, evaluating placement readiness using academic performance alone may not provide a complete picture.

Machine Learning can be used to analyze multiple student attributes and identify patterns associated with placement outcomes. Educational Data Mining and machine learning techniques have been increasingly applied to student performance and employability prediction.

Pathnix addresses this problem by combining student data, machine learning prediction, readiness analysis, and career recommendation in a single web-based system.

The basic workflow of Pathnix is:

**Student Data → Prediction → Readiness Analysis → Recommendation → Action**

The objective is not to guarantee placement but to provide students and faculty with useful information for improving placement preparation.

# II. RELATED WORK

Several studies have explored the use of machine learning for student performance and employability prediction.

Educational Data Mining techniques have been used to analyze student information and identify patterns related to academic and career outcomes. Different machine learning algorithms, including Logistic Regression, Random Forest, Support Vector Machine, K-Nearest Neighbors, and Naïve Bayes, have been investigated for employability prediction.

Existing research demonstrates that student attributes can be used to develop predictive models for placement and employability analysis. However, many systems primarily focus on predicting placement status and provide limited personalized recommendations.

Pathnix extends this approach by combining placement prediction with readiness analysis and career recommendations. The system attempts to provide not only a prediction but also practical guidance regarding areas that students can improve.

# III. PROPOSED SYSTEM

## A. System Overview

Pathnix is a web-based placement prediction and career recommendation platform.

The system accepts student information and processes the data through a machine learning pipeline. Logistic Regression is used to estimate the student's placement probability. The prediction is then combined with additional profile information to generate readiness and career recommendations.

The major components are:

1. Student Management
2. Machine Learning Prediction
3. Placement Readiness Analysis
4. Career Recommendation
5. Faculty Dashboard
6. Student Dashboard
7. Report Generation

## B. Input Features

The system considers multiple student-related attributes.

| Feature | Description |
|---|---|
| CGPA | Academic performance |
| Attendance | Academic participation |
| Technical Skills | Technical and programming abilities |
| Projects | Practical development experience |
| Internship | Industry exposure |
| Certifications | Additional technical qualifications |
| Backlogs | Academic performance indicator |
| Preparation | Placement preparation activities |

These features provide a broader representation of student placement readiness than academic performance alone.

---

## C. Machine Learning Model

Pathnix uses **Logistic Regression** as the primary classification algorithm.

The placement prediction problem is represented as a binary classification problem where the model predicts whether the student belongs to the positive or negative placement class.

The logistic function is:

[
P(Y=1|X)=\frac{1}{1+e^{-z}}
]

where (X) represents the input student features and (z) represents the weighted combination of those features.

The model produces a probability value between 0 and 1, which can be presented to the user as a percentage.

For example:

**Placement Probability = 78%**

The probability is then used as one of the inputs for the system's readiness analysis.

---

## D. Career Recommendation

The prediction is not presented as the final output.

Pathnix follows the process:

**Placement Prediction**

↓

**Readiness Analysis**

↓

**Weak Area Detection**

↓

**Career Recommendation**

↓

**Improvement Suggestions**

The system can provide:

- Placement probability
- Readiness level
- Weak areas
- Career path
- Predicted role
- Company type

- Expected salary range
- Improvement suggestions

These recommendations are intended to assist students in planning their preparation.

---

# IV. SYSTEM IMPLEMENTATION

## A. System Architecture

Pathnix follows a client-server architecture.

The main workflow is:

**User**

↓

**Web Interface**

↓

**Flask Backend**

↓

**Database / ML Model**

↓

**Prediction Engine**

↓

**Recommendation Engine**

↓

**Dashboard / Report**

The frontend provides the user interface, while the backend handles authentication, student data, prediction requests, database operations, and report generation.

---

## B. Student Module

The student module provides:

- Student login
- Profile access
- Placement prediction
- Readiness analysis
- Career recommendations
- Weak-area identification
- Report download

The student can view the generated prediction and understand the areas that require improvement.

---

## C. Faculty Module

The faculty module provides:

- Faculty login
- Student search
- Student profile viewing
- Student data management
- Placement prediction
- Readiness monitoring
- Placement-related analysis
- Report generation

This allows faculty members to support students using data-driven information.

---

## D. Administrator Module

The administrator manages the overall system.

The administrator can manage:

- User accounts
- Student information
- Faculty information
- Institutional data
- System settings

Role-based access ensures that different users receive access to appropriate system functionality.

# V. RESULTS AND DISCUSSION

## A. Model Performance

The current evaluation of the Pathnix prediction model produced the following results:

| Metric | Result |
|---|---|
| Accuracy | 78% |
| Precision | 78% |
| Recall | 76% |
| F1-Score | 77% |

These results indicate that the selected student attributes provide useful information for predicting placement outcomes.

## B. Performance Analysis

**Accuracy** represents the overall proportion of correctly classified samples.

[
Accuracy=\frac{TP+TN}{TP+TN+FP+FN}
]

**Precision** represents the proportion of predicted positive cases that are actually positive.

[
Precision=\frac{TP}{TP+FP}
]

**Recall** represents the proportion of actual positive cases correctly identified.

[
Recall=\frac{TP}{TP+FN}
]

**F1-score** provides a combined measure of precision and recall.

[
F1=2\times\frac{Precision\times Recall}{Precision+Recall}
]

The obtained results show reasonably balanced performance across precision, recall, and F1-score.

---

## C. Discussion

The model achieved approximately 78% accuracy, indicating that the selected academic and employability-related features contain useful predictive information.

Logistic Regression was selected because it is suitable for binary classification, computationally efficient, and capable of producing probability estimates. It is also relatively simple to interpret compared with more complex machine learning models.

The system extends the prediction beyond a simple classification result. For example, instead of displaying only a placement prediction, Pathnix can provide:

**Placement Probability:** 72%

**Readiness:** Moderate

**Weak Area:** Technical Skills

**Recommended Action:** Improve programming and DSA

**Career Path:** Software Development

This makes the prediction more useful for student preparation.

---

## D. Logistic Regression and Random Forest

Random Forest can provide stronger performance for some datasets because it can model nonlinear relationships. However, the two models must be evaluated using the same dataset and evaluation methodology before making a numerical comparison.

| Parameter | Logistic Regression | Random Forest |
|---|---|---|
| Model Type | Linear classifier | Tree-based ensemble |
| Interpretability | High | Moderate |
| Training | Fast | Moderate |
| Non-linear Patterns | Limited | Strong |
| Probability Output | Yes | Yes |
| Pathnix | Primary | Future Comparison |

Random Forest can therefore be considered as a future experimental comparison rather than claiming an unverified accuracy value.

---

# VI. APPLICATIONS AND LIMITATIONS.

## A. Applications

| **Pathnix can be used in:** | |
|---|---|
| Engineering<br>**To monitor student placement readiness.** | Colleges: |
| Placement<br>**To identify students who may require additional training.** | Cells: |
| Faculty<br>**To support personalized student guidance.** | Counseling: |
| Student<br>**To help students understand their current strengths and weaknesses.** | Self-Assessment: |
| Placement<br>**To identify common skill gaps and plan training programs.** | Training: |
| Career<br>**To provide possible career paths and improvement suggestions.** | Guidance: |

## B. Limitations

The current system has several limitations.

1. The dataset size may limit generalization.
2. Student data may reflect the characteristics of a particular institution.
3. Some employability factors cannot be completely represented through structured data.
4. Placement probability is not a guarantee of actual placement.
5. Company hiring policies and market conditions can change.
6. The recommendation component requires further validation with larger datasets.
7. The model requires testing on data from multiple institutions before large-scale deployment.

---

# VII. FUTURE ENHANCEMENTS

The Pathnix system can be improved in several ways.

## 1. Larger Dataset

Student data from multiple colleges can be collected to improve model generalization.

## 2. Advanced Machine Learning

Future versions can compare Logistic Regression with Random Forest, XGBoost, SVM, and other algorithms.

## 3. Explainable AI

Explainable AI techniques can be integrated to show why a particular prediction was generated.

## 4. Job Market Integration

Real-time job and skill-demand information can be integrated into career recommendations.

## 5. Mobile Application

A mobile application can allow students to access their predictions and recommendations conveniently.

# VIII. CONCLUSION

Pathnix demonstrates the application of machine learning to placement prediction and career guidance for engineering students. The system combines student information with Logistic Regression to estimate placement probability and extends the prediction through readiness analysis and personalized recommendations.

The current evaluation achieved approximately **78% accuracy, 78% precision, 76% recall, and 77% F1-score**. These results indicate that structured academic and employability-related student attributes can provide useful information for placement prediction.

The main contribution of Pathnix is the integration of prediction and recommendation into a single platform. Instead of simply informing students whether they are likely to be placed, the system provides information about readiness, weak areas, career paths, and possible improvement strategies.

Therefore, Pathnix can serve as a decision-support system for students, faculty, and placement cells. With larger datasets, additional machine learning models, explainable AI, and external validation, the system can be further developed into a more comprehensive career-readiness platform.

# REFERENCES


*[1] R. Liu and A. Tan, "Towards Interpretable Automated Machine Learning for STEM Career Prediction,"* **Journal of Educational Data Mining**, *vol. 12, no. 2, 2020.*

*[2] "Employability Prediction: A Survey of Current Approaches, Research Challenges and Applications,"* **Journal of Ambient Intelligence and Humanized Computing**, *2021.*

*[3] K. Sreenivasa Rao, N. Swapna, and P. Praveen Kumar, "Educational Data Mining for Student Placement Prediction Using Machine Learning Algorithms,"* **International Journal of Engineering and Technology**, *vol. 7, no. 1.2, 2018.*

*[4] K. C. Sekhar and K. Santhosh Kumar, "Undergraduate Student's Campus Placement Determination Using Logistic Regression Analysis for Predicted Probabilities on Uncertain Dataset,"* **International Journal of Intelligent Systems and Applications in Engineering**, *2022.*

*[5] M. Ruparel and P. Swaminarayan, "Student Placement Prediction Using Various Machine Learning Techniques,"* **International Journal of Intelligent Systems and Applications in Engineering**, *2023.*

*[6] M. Ruparel and P. Swaminarayan, "Enhancing Student Placement Predictions with Advanced Machine Learning Techniques,"* **Journal of Information Systems Engineering and Management**, *2024.*

*[7] "Prediction of Campus Placement Using Hybrid Logistic Regression Model,"* **IEEE Conference Publication**, *2025.*

*[8] R. Tariq, D. G. Vargas, F. Ali, M. Gonzalez-Mendoza, and C. S. Torres-Castillo, "What Determines Student Employability? Educational Data Mining Through Machine and Deep Learning Approach,"* **IEEE-RITA**, *vol. 20, pp. 271–289, 2025.*

*[9] Scikit-learn Developers, "Logistic Regression,"* **Scikit-learn Documentation.**

*[10] Scikit-learn Developers, "Machine Learning in Python,"* **Scikit-learn Documentation.**