International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 05 Issue: 03 | March – 2026 DOI:10.55041/ISJEM05832
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

A Comparative Machine Learning Method for Predicting Campus Placement using Academic and Skill-Based Features

Dr. K. Satyam[1], Konda Swetha[2]

[1]Associate Professor, Department of MCA, Annamacharya Institute of Technology & Sciences, Tirupati, Andhra Pradesh, India.

[2] Post Graduate, Department of MCA, Annamacharya Institute of Technology & Sciences, Tirupati, Andhra Pradesh, India.

Abstract

Campus placement has a big impact on students' careers, yet predicting placement outcomes is still challenging due to the influence of many academic and skill-based factors. This work provides a machine learning-based approach to predict student placement status using structured data. The proposed system makes use of features like academic performance (SSC, HSC, MCA), programming skills, internships, projects, certificates, and participation in hackathons. Three supervised learning algorithms—Random Forest, Decision Tree, and Logistic Regression—are put into practice and evaluated. Experimental results show that Random Forest and Decision Tree models are more accurate than Logistic Regression. The technology is further integrated into an internet application to provide forecasts in real time. This tactic can improve employability and give institutions and students a better understanding of placement trends.

Keywords

Campus Placement Prediction, Machine Learning, Decision Tree, Random Forest, Logistic Regression, Classification

I.Introduction

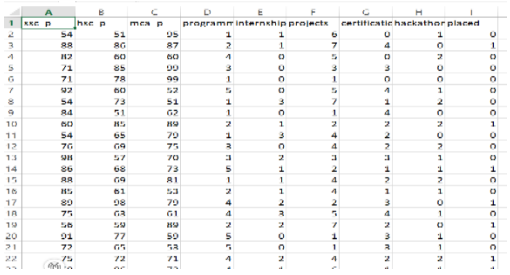
A crucial stage in a student's academic career is campus placement, which acts as a link between learning and professional work. Securing a placement in today's fiercely competitive labour market depends not just on academic achievement but also on a variety of skill-based and experiential elements, including involvement in technical activities, internships, certifications, and programming knowledge. Finding the critical elements that affect placement success has grown in significance for educational institutions and students alike. Conventional approaches to assessing students for placement prospects sometimes rely on manual evaluation or narrow criteria, which may not adequately reflect the intricate interactions between many traits.

Intelligent systems that can evaluate several variables at once and make precise predictions are therefore becoming more and more necessary. In this field, machine learning has become a potent tool that makes data-driven decision-making possible by revealing hidden patterns in massive datasets. In educational data mining, machine learning algorithms—especially classification models—have been utilised extensively to forecast student performance and results. These models can produce accurate predictions by learning the association between input variables and placement outcomes by utilising past data. These prediction algorithms can support universities in creating focused training programs and assist students in identifying areas for improvement. This study proposes a machine learning-based method that combines academic and skill-based variables to predict campus placement outcomes. To assess their efficacy, several supervised learning algorithms—such as Random Forest, Decision Tree, and Logistic Regression—are put into practice and contrasted. Furthermore, a web-based application that lets users enter their information and get real-time placement forecasts incorporates the suggested system.

II. Dataset Description

Students' academic and skill-based characteristics, which are thought to have a significant impact on campus placement outcomes, are included in the structured dataset used in this study. The dataset is arranged in tabular fashion, with each column denoting a particular trait and each row denoting a single student.

International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 05 Issue: 03 | March – 2026 DOI:10.55041/ISJEM05832
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata



	A	B	C	D	E	F	G	H	I
1	ssc_p	hsc_p	mca_p	programm	internship	projects	certificatic	hackathor	placed
2	54	51	65	1	1	0	0	1	0
3	88	86	87	2	1	7	4	0	1
4	82	80	80	4	0	5	0	2	0
5	71	85	99	3	0	3	3	0	0
6	71	78	99	1	0	1	0	0	0
7	92	80	52	5	0	5	4	1	0
8	54	73	51	1	3	7	1	2	0
9	84	51	62	1	0	1	4	0	0
10	80	85	89	2	1	2	2	2	1
11	54	65	70	1	3	4	2	0	0
12	76	69	75	3	0	4	2	2	0
13	98	57	70	1	2	1	1	1	0
14	86	68	73	5	1	2	1	1	1
15	88	69	81	1	1	4	2	2	0
16	85	61	53	2	1	4	1	1	0
17	89	98	79	4	2	2	3	0	1
18	75	68	61	4	3	5	4	1	0
19	56	59	89	2	2	7	2	0	1
20	91	77	59	5	0	1	3	1	0
21	72	65	53	5	0	1	3	1	0
22	75	72	71	4	2	4	2	2	1

Fig: Dataset

The dataset includes the following features:

• SSC Percentage (ssc_p): Indicates how well a student performed in secondary school. It shows the student's fundamental academic strength.

• HSC Percentage (hsc_p): Indicates a student's success in higher secondary education and is a key factor in determining consistency and topic understanding.

• MCA Percentage (mca_p): A crucial component in assessing a student's technical proficiency, it shows postgraduate academic success.

• Programming Skills: Indicates the student's degree of coding expertise. Stronger programming expertise, which is necessary for technical placements, is indicated by higher values.

• Internship Experience (internship): Indicates how many internships the student has finished. Internships improve employability and offer real-world experience.

• Projects (projects): Shows how many personal or academic projects have been finished. Projects show practical experience and problem-solving skills.

• Certifications: Indicates how many certifications the student has earned in pertinent fields, demonstrating extracurricular abilities.

• Hackathon Participation: Reflects involvement in hackathons or coding competitions, demonstrating collaboration and practical skills.

• Placement Status (placed): This is the goal variable, where

- 1 denotes the student's placement.
- 0 denotes that the pupil is not assigned.

III. Results and Discussion

A. Model Training Results

Accuracy is the main parameter used to assess the performance of the installed machine learning models. A comparison table is used to display the outcomes of the training phase.

With an accuracy of 82.7%, Logistic Regression outperformed the other two models in predicting placement outcomes. On the other hand, both Random Forest and Decision Tree models performed better on the provided dataset, achieving 100% accuracy. Because Decision Tree and Random Forest models can capture intricate and non-linear correlations between input characteristics and the target variable, their accuracy is better. As an ensemble technique, Random Forest reduces variance by mixing several decision trees to further improve performance.

Perfect accuracy, however, could also be a sign of overfitting, particularly if the dataset size is small. This implies that even while the model does remarkably well on the training and testing data, more validation with larger datasets is required to guarantee generalisation.

International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 05 Issue: 03 | March – 2026 DOI:10.55041/ISJEM05832
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

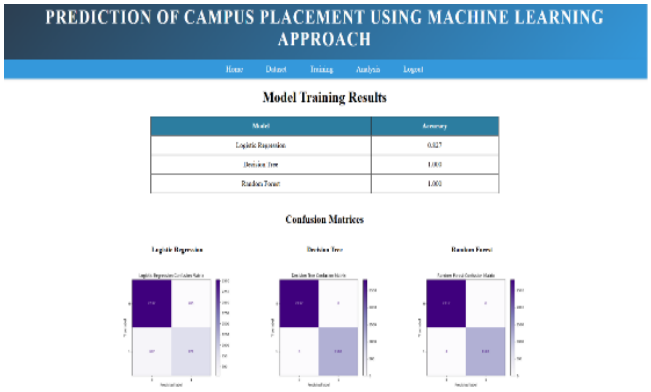


Fig: Model Results

By contrasting real and anticipated values, the confusion matrix offers a thorough assessment of the model's performance.

• Logistic Regression: A few misclassifications, including false positives and false negatives, are displayed in the confusion matrix. This explains why it is less accurate than other models.

• Decision Tree: Almost flawless classification with little to no misclassification is shown by the confusion matrix. Its excellent accuracy is a result of this.

• Random Forest: The Random Forest model's resilience is confirmed by its extremely accurate forecasts with small errors, much like the Decision Tree model.

B. Interpretation of Metrics:

- True Positives (TP): Correctly predicted placed students
- True Negatives (TN): Correctly predicted non-placed students
- False Positives (FP): Incorrectly predicted as placed
- False Negatives (FN): Incorrectly predicted as not placed

IV. System Implementation and User Interface

Fig: Web Interface

The suggested approach is put into practice as a web application that enables real-time user interaction with the machine learning model. Students can enter their academic and skill-related information with ease thanks to the interface's straightforward, accessible, and user-friendly design.

Key attributes like SSC percentage, HSC percentage, MCA percentage, programming skills, number of internships, completed projects, certifications earned, and hackathon participation are all required to be entered into the application's input form. The attributes utilised to train the machine learning models match these inputs.

The system processes the input data and sends it to the trained machine learning model on the backend after the user enters the necessary information. The placement outcome is then predicted by the model and shown on the same interface. Whether or not the student is likely to be placed is indicated by the prediction result. The system's frontend is constructed using HTML and CSS, while the backend is created using Python and a machine learning framework. A web framework manages the frontend and backend interaction, guaranteeing seamless data flow and real-time prediction.

V. Conclusion

This study offers a machine learning-based method that uses academic and skill-oriented variables to predict campus placement outcomes. Among the models assessed, Random Forest and Decision Tree outperformed Logistic Regression. Real-time and user-friendly engagement is made possible by the prediction model's incorporation into a web-based system. The outcomes demonstrate how data-driven methods can increase placement forecast accuracy. The potential for overfitting, however, suggests that more validation with bigger datasets is required. All things considered, the suggested system offers institutions and students a workable way to improve decision-making and employability.

References

[1] J. Han, M. Kamber, and J. Pei, Data Mining: Concepts and Techniques, 3rd ed. San Francisco, CA, USA: Morgan Kaufmann, 2011.

[2] F. Pedregosa et al., “Scikit-learn: Machine Learning in Python,” Journal of Machine Learning Research, vol. 12, pp. 2825–2830, 2011.

[3] L. Breiman, “Random Forests,” Machine Learning, vol. 45, no. 1, pp. 5–32, 2001.

[4] J. R. Quinlan, “Induction of Decision Trees,” Machine Learning, vol. 1, no. 1, pp. 81–106, 1986.

[5] C. M. Bishop, Pattern Recognition and Machine Learning. New York, NY, USA: Springer, 2006.

[6] T. Hastie, R. Tibshirani, and J. Friedman, The Elements of Statistical Learning, 2nd ed. New York, NY, USA: Springer, 2009.