International Journal of

DATA SCIENCE AND IOT MANAGEMENT SYSTEM

Peer Reviewed, Referred & Indexed Journal

ISSN: 3068-272X ww.ijdim.com Original Research Paper

# COLLEGE PLACEMENT ANALYSIS

[1]N.Bhargavi,[2]B.Sai Vamshi,[3]S.Ramya sri,[4]K.Naveen

[1]Assistant Professor, [234]Students

Department of CSE(Data Science)

Siddhartha institute of technology & sciences,narapally

bhargavi.cse@siddhartha.co.in, 23TQ1A6714@siddhartha.co.in,
23TQ1A6721@siddhartha.co.in, 23TQ1A6734@siddhartha.co.in

## ABSTRACT

Campus placement is one of the most important aspects of higher education institutions as it directly reflects the employability of graduating students and the effectiveness of academic training provided by colleges. Placement analysis helps institutions understand placement trends, identify factors affecting recruitment success, and improve student preparation strategies for future placement drives.This project, Collage Placement Analysis, focuses on analyzing previous-year and current-year student placement datasets using data analytics and machine learning techniques. The datasets contain academic and skill-related variables such as Student ID, Branch, Gender, CGPA, 10th Percentage, 12th Percentage, BTech Percentage, Communication Skill, Programming Skill, Internship Experience, Aptitude Score, Placement Status, and Package Offered.The project uses data preprocessing, exploratory data analysis, visualization, and predictive modeling to understand placement patterns and estimate future placement outcomes. Historical placement records are first analyzed to identify trends such as placed versus non-placed percentages, branch-wise package distribution, the impact of CGPA on placement, and the role of internships in employability. Machine learning models such as Logistic Regression are then used to predict placement status for current-year students, while Random Forest Regression is applied to estimate expected salary packages. The project also provides visual outputs such as pie charts, bar graphs, and scatter plots, making placement trends easy to interpret for academic institutions and placement officers. The final results help colleges improve placement training, identify students requiring additional support, and strengthen institutional placement planning.

International Journal of

DATA SCIENCE AND IOT MANAGEMENT SYSTEM

Peer Reviewed, Referred & Indexed Journal

ISSN: 3068-272X ww.ijdim.com Original Research Paper

## I INTRODUCTION

Campus placement plays a crucial role in shaping students' careers after graduation. Colleges and universities invest significant effort in organizing placement drives, collaborating with companies, and preparing students for recruitment processes.However, analyzing placement trends manually can be difficult because placement datasets contain multiple variables such as:

- Academic performance
- Degree specialization
- Work experience
- Skills and internships
- Salary offers

Placement data analysis helps institutions identify patterns that influence employment outcomes. By analyzing historical placement data, colleges can understand which factors contribute to successful placements and higher salary packages.Data analytics techniques allow educational institutions to examine placement trends across departments, companies, and student profiles. These insights help improve training programs and enhance employability among students.The analysis helps identify branch-wise placement performance, academic impact on placement, internship influence, and expected placement percentages, supporting institutions in improving placement training and student employability

## II LITERATURE SURVEY

Campus placement analysis has become an important research area in educational data mining and machine learning. Researchers analyze student data to understand the factors affecting job placements.Several studies have used machine learning algorithms to predict placement outcomes using historical student data such as academic performance, skills, and work experience. Algorithms like Decision Tree and Random Forest are commonly used to analyze placement datasets and predict the probability of students getting placed.Another study on campus placement prediction applied algorithms such as Logistic Regression, Support Vector Machine (SVM), Random Forest, and AdaBoost to forecast whether a student will get placed based on academic and training factors. These approaches help institutions understand placement trends and improve training programs.Research has also shown that academic attributes such as grades, test scores, and internships significantly influence placement predictions. Comparative studies indicate that algorithms like Random Forest and K-Nearest Neighbors often achieve high prediction accuracy in placement analysis.Other

International Journal of

DATA SCIENCE AND IOT MANAGEMENT SYSTEM

Peer Reviewed, Referred & Indexed Journal

ISSN: 3068-272X ww.ijdim.com Original Research Paper

research has focused on developing placement management systems that analyze student data and assist colleges in improving placement rates. These systems use data preprocessing, classification algorithms, and visualization techniques to generate insights about student employability.Overall, the literature shows that data analytics and machine learning techniques are effective tools for analyzing placement data and identifying factors influencing student employment outcomes..

## III SYSTEM ANALYSIS

College placement analysis in Data Science involves studying student data, company participation, and hiring trends to understand and improve placement outcomes. The system collects data such as student academic performance, skills, attendance, and previous placement records, along with company requirements and job roles offered. Using data science techniques like data cleaning, visualization, and predictive modeling, the system identifies patterns such as which skills lead to higher placement chances, which companies visit frequently, and what salary trends look like. It helps colleges make better decisions by predicting future placement performance, guiding students on required skills, and improving training programs. Overall, the system transforms raw placement data into meaningful insights that support students, faculty, and placement cells in achieving better results**.**

### Existing system

The existing system for college placement analysis in Data Science is mostly manual and lacks advanced analytical capabilities. Placement cells typically maintain student and company data using spreadsheets or basic databases, where information is stored but not deeply analyzed. The process mainly depends on past records, simple sorting, and human observation to understand placement trends. There is no proper use of predictive models or intelligent systems to forecast student success or company requirements. As a result, it becomes difficult to identify skill gaps, track performance effectively, or provide personalized guidance to students. This traditional approach is time-consuming, less accurate, and does not fully utilize the potential of data for improving placement outcomes.

### Disadvantages of existing system

- Time-consuming due to manual data handling and updates
- Prone to human errors in data entry and analysis
- Lack of predictive analysis for future placement trends
- No proper identification of student skill gaps

International Journal of

DATA SCIENCE AND IOT MANAGEMENT SYSTEM

Peer Reviewed, Referred & Indexed Journal

ISSN: 3068-272X ww.ijdim.com Original Research Paper

- Difficult to manage and analyze large volumes of data

**Proposed system**

The proposed system for college placement analysis in Data Science uses advanced data-driven techniques to improve the overall placement process. It collects and integrates data from multiple sources such as student profiles, academic records, skill sets, and company requirements into a centralized system. By applying machine learning algorithms, data visualization, and predictive analytics, the system can identify patterns and trends that help predict student placement chances and suggest suitable career paths. It also provides insights into in-demand skills, helping students focus on areas that increase their employability. The system supports real-time data updates, reduces manual effort, and enhances accuracy in decision-making. Overall, it offers an intelligent and efficient approach to improve placement outcomes for both students and institutions.

**Advantages of proposed system**

- Provides accurate and data-driven insights
- Reduces manual work and saves time
- Predicts student placement chances using machine learning
- Helps identify skill gaps and improvement areas
- Supports better decision-making for placement cells

## IV METHODOLOGY

The methodology of the Collage Placement Analysis project explains the steps followed to collect, process, analyze and predict student placement outcomes using academic and skill-based datasets. The main goal of this methodology is to transform raw placement data into meaningful insights that help institutions understand placement trends and improve student preparation

.The first step in the methodology is data collection. Two datasets are used in this project: a previous-year placement dataset and a current-year student dataset, both stored in CSV format. The previous-year dataset contains student academic details, placement status, and package offered, while the current-year dataset contains student academic records for prediction

.The second step is data preprocessing. In this stage, datasets are loaded into Python using the Pandas library. Categorical columns such as Branch and Placement Status are converted into numerical form using Label Encoding for model training

International Journal of

DATA SCIENCE AND IOT MANAGEMENT SYSTEM

Peer Reviewed, Referred & Indexed Journal

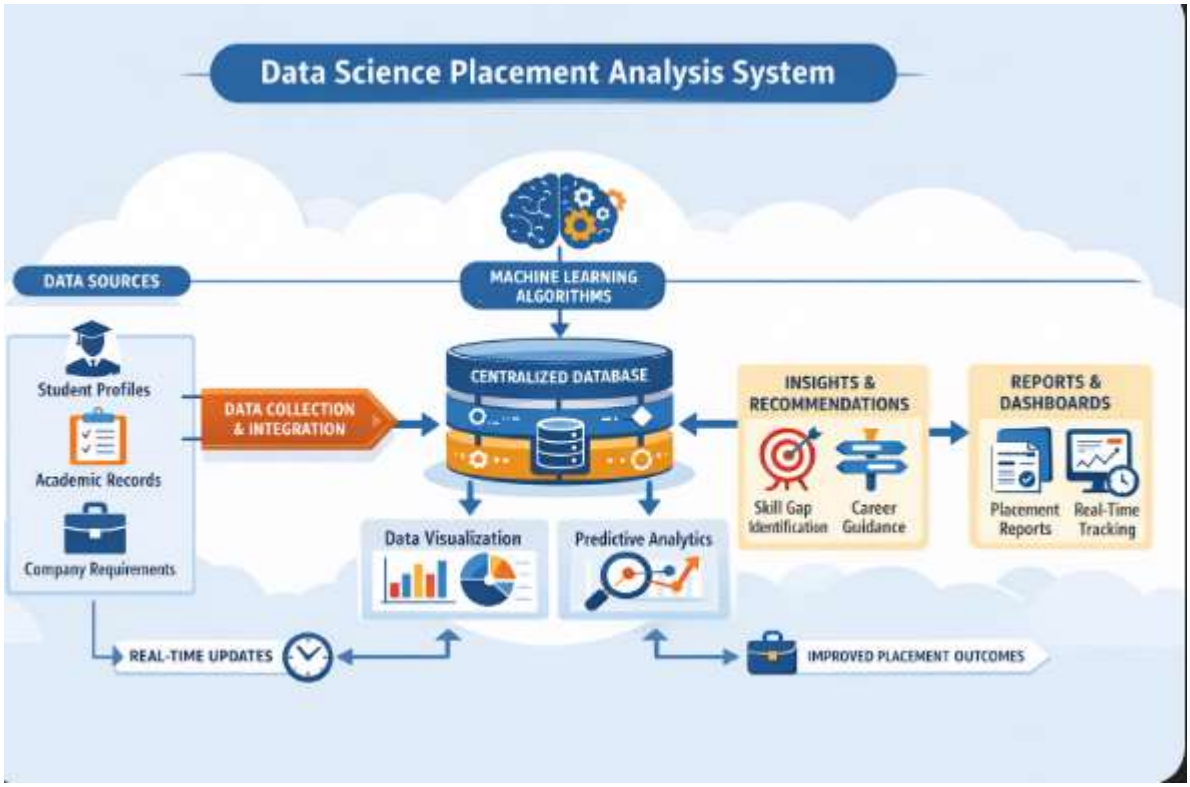
ISSN: 3068-272X ww.ijdim.com Original Research Paper

.The third step is data analysis. Previous-year placement data is analyzed to identify placement percentage, highest package by branch, CGPA impact on placement, and internship influence.

The fourth step is data visualization. Matplotlib is used to generate pie charts, bar graphs, and scatter plots for clear understanding of placement trends.

The fifth step is prediction and result generation. Logistic Regression predicts current-year placement outcomes, while Random Forest Regression estimates expected salary packages

**System Architecture**




The system architecture of the college placement analysis system is designed to efficiently collect, process, and analyze data to improve placement outcomes. It follows a layered approach consisting of data collection, data processing, analysis, and presentation layers. In the first layer, data is gathered from multiple sources such as student profiles, academic records, and company requirements. This raw data is then stored in a centralized database after undergoing data cleaning and integration processes.

International Journal of

DATA SCIENCE AND IOT MANAGEMENT SYSTEM

Peer Reviewed, Referred & Indexed Journal

ISSN: 3068-272X ww.ijdim.com Original Research Paper

In the processing layer, the system applies data science techniques such as data preprocessing, feature selection, and transformation to make the data suitable for analysis. The analysis layer uses machine learning algorithms and predictive models to identify patterns, such as student performance trends and skill demands in the job market. These models help in predicting placement chances and recommending suitable career paths.

Finally, the presentation layer displays the results through dashboards, reports, and visualizations, making it easy for placement officers and students to understand insights. The system also supports real-time updates, ensuring that the information remains current and useful.

## V RESULTS&OUTPUT

```
Current Year Prediction Output:

   Student ID Branch  CGPA  BTech Percentage  Internship Experience  \
0        S101    CSE  8.60              86.5                      0
1        S102     DS  6.13              60.3                      0
2        S103   AIML  5.85              61.3                      0
3        S104    ECE  6.86              67.7                      0
4        S105  CIVIL  8.26              80.3                      1
..        ...    ...   ...               ...                    ...
95       S196    CSE  8.00              81.5                      1
96       S197     DS  7.38              73.7                      0
97       S198   AIML  7.87              79.5                      0
98       S199    ECE  7.60              78.9                      0
99       S200  CIVIL  5.85              57.3                      0

   Predicted Placement  Expected Package
0              Placed          7.300800
1          Not Placed          0.000000
2          Not Placed          0.000000
3          Not Placed          0.000000
4              Placed          5.933800
..                ...               ...
95             Placed          7.086300
96             Placed          2.572967
97             Placed          8.980533
98             Placed          9.070200
99         Not Placed          0.000000

[100 rows x 7 columns]

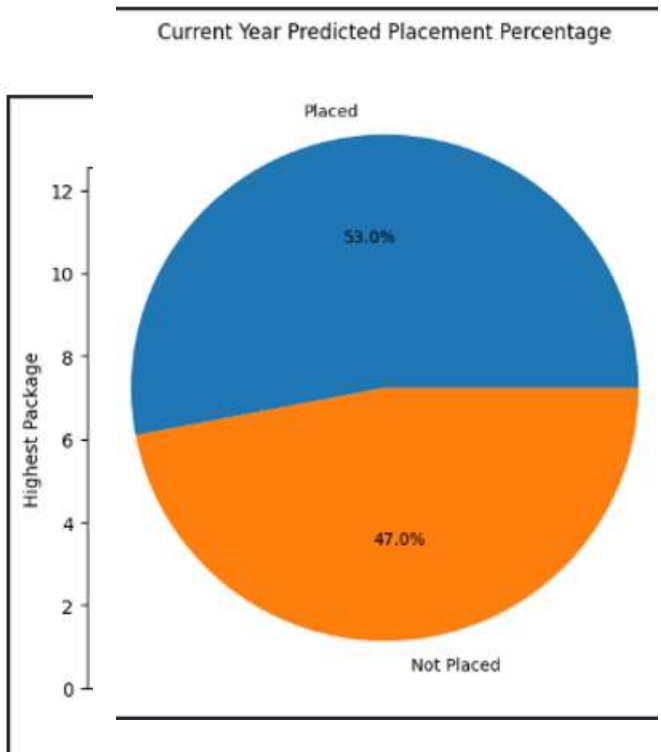
Prediction file saved as placement_prediction_output.csv
```

International Journal of

DATA SCIENCE AND IOT MANAGEMENT SYSTEM

Peer Reviewed, Referred & Indexed Journal

ISSN: 3068-272X ww.ijdim.com Original Research Paper




## VI CONCLUSION

The Collage Placement Analysis project provides useful analysis of student placement data using Python, Pandas, and machine learning techniques. However, the system can be further improved by adding advanced features and expanding its capabilities.In the future, the system can be integrated with real-time institutional databases so that student academic records and placement information are automatically updated and analyzed continuously. This would help placement officers monitor placement performance more efficiently and maintain updated student records.Another improvement could be the use of advanced machine learning algorithms to improve placement prediction accuracy. Models such as Decision Tree, Random Forest Classification, Support Vector Machine, and Neural Networks can be applied to capture more complex placement patterns.The system can also be enhanced by developing an interactive dashboard that displays placement insights using advanced visualization tools. Such a dashboard would allow users to easily explore branch-wise placement trends, salary distributions, and student performance reports.Additionally, the project can be extended to include more student-related factors such as certifications, coding skills, project experience, interview performance, and extracurricular activities for deeper analysis.The system may also

International Journal of

DATA SCIENCE AND IOT MANAGEMENT SYSTEM

Peer Reviewed, Referred & Indexed Journal

ISSN: 3068-272X ww.ijdim.com Original Research Paper

be improved by integrating cloud-based storage, which would allow handling larger datasets and easier access from different departments.

## REFERENCE

[1] Kumar, R. D., Prudhviraj, G., Vijay, K., Kumar, P. S., & Plugmann, P. (2024). Exploring COVID-19 through intensive investigation with supervised machine learning algorithm. In Handbook of Artificial Intelligence and Wearables (pp. 145-158). CRC Press.

[2] Swathi, B., Vijay, K., Sushanth Babu, M., & Dinesh Kumar, R. (2024, November). Machine Learning Techniques in Cloud Based Intrusion Detection. In The International Conference on Artificial Intelligence and Smart Environment (pp. 557-564). Cham: Springer Nature Switzerland.

[3] Sv satyakrishna, shirisha rangu ,bhargavi nalacheruve.(2024) Prospective investigation on colorectal cancer with SMOTE on machine learning Algorithm

[4] Dr.G.Vishnu Murthy, BhargaviNalacheruve 1Professor, Department of computer Science & engineering, Anurag University, TS, India. 2Student, Department of computer Science & engineering, Anurag University, TS, India.

[5] V. N. S. Manaswini, K. K, C. Nigam, S. S. Ali, R. Niranjana, and Suman, “Real-Time Object Detection in Drone Surveillance Using YOLOv5,” in Proc. 2025 3rd Int. Conf. IoT, Communication and Automation Technology (ICICAT), Gorakhpur, India, 2025, pp. 1–6, doi: 10.1109/ICICAT68430.2025.11414670.

[6] B. Soundarya, V. N. S. Manaswini, M. Ayyakrishnan, R. D. Kumar, “Contextual Analysis of Big Data Analytics in Intelligent Transportation Frameworks,” in Intersection of Artificial Intelligence, Data Science, and Cutting-Edge Technologies: From Concepts to Applications in Smart Environment, Lecture Notes in Networks and Systems, vol. 1353, Cham: Springer, 2025, doi: 10.1007/978-3-031-88304-0_79.

[7] R. D. Kumar, V. N. S.Manaswini, “Applications of blockchain in smart cities: detecting fake documents from land records using blockchain technology,” in Blockchain for Smart Cities, Elsevier, 2021, pp. 105–117, doi: 10.1016/B978-0-12-824446-3.00017-X.

[8] Tejavath Veeramma, Badarla Anil, Guguloth Ravinder, “An advanced movie recommender using collaborative filtering and sentiment analysis,” International Research Journal of Modernization in Engineering Technology and Science, vol. 7, no. 7, July 2025, doi: 10.56726/IRJMETS81618.

[9] Ravi Kumar Banoth, Ramana Murthy B V, “Automatic crop recommendation system using LightGBM and decision tree machine learning models,” Journal of

International Journal of

DATA SCIENCE AND IOT MANAGEMENT SYSTEM

Peer Reviewed, Referred & Indexed Journal

ISSN: 3068-272X ww.ijdim.com Original Research Paper

Machine and Computing, vol. 5, no. 1, pp. 343, Jan. 2025, doi: 10.53759/7669/jmc202505026.

[10] Ravi Kumar Banoth, Dr. B.V. Ramana Murthy, "Smart agriculture through IoT and machine learning for analyzing carbon footprints," in Proc. Int. Conf. Computer Science and Communication Engineering (ICCSCE), Apr. 2025.

[11] Ravi Kumar Banoth, B. V. Ramana Murthy, "Soil image classification using transfer learning approach: MobileNetV2 with CNN," SN Computer Science, vol. 5, art. no. 199, 2024, doi: 10.1007/s42979-023-02500-x.