JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

COLLEGE PLACEMENT SYSTEM

Ms K Saranya [1], T Harisubash [2], R Indirajith Bala [3], P Jeevan Adhitya [4]

[1]Assistant Professor, Department of CSE, Sri Ramakrishna Institute of Technology, Coimbatore, Tamil Nadu, India
[234]UG Students, Department of CSE, Sri Ramakrishna Institute of Technology, Coimbatore, Tamil Nadu, India

Abstract: The placement process in higher education institutions plays a major role for deciding the career paths of students. The management of the college placement system and student placement status prediction are the primary objectives of this project. We use a variety of machine learning algorithms to evaluate the accuracy of the placed student list by using student data. For this task, we specifically analyze naive algorithms, random forest, logistic regression, and support vector machines (SVM). The one with greatest accuracy levels among them is logistic regression. The project also involves creating a portal to monitor placed student lists as well as applied student lists. This abstract summarizes our efforts to improve decision-making for both educational institutions and students by streamlining the placement process.

Keywords: Data modeling, Data preprocessing, Algorithmic analysis, Logistic regression, Random forest

I. Introduction

The placement process in educational institutions is essential for students' professional development; however, because of a variety of factors, including market trends and academic performance, it is still difficult to predict placements. Using machine learning, particularly logistic regression, can improve student placement prediction analysis. The main goal of this paper is to use logistic regression to forecast student placements in the future using historical data. Creating a predictive model and assessing its performance metrics, such as accuracy and precision, are among the goals. There are also insights into the elements that have an impact on placements.

The article presents the findings of the logistic regression analysis, evaluates related work, talks about methodology, and ends with suggestions for further research.

II. Related Works

Predictive analysis in college placement has received attention from researchers in recent years, with various machine learning algorithms such as decision trees, support vector machines, and neural networks used to predict student placement outcomes. For example, Malik et al. (2018) predicted student placements based on demographics and academic performance using decision tree-based algorithms. Similar to this, Gupta et al. (2020) divided students into groups based on placement and non-placement using support vector machines. Although these investigations have yielded insightful information, there is still a deficiency in knowledge regarding the efficacy of logistic regression in this field. Benefits of logistic regression include interpretability, simplicity in application, and effectiveness in managing binary classification problems. The purpose of this study is to add to the body of research on predictive analysis for college placement by utilizing logistic regression.

III. Experimental Setup

Hardware:

CPU Type: Intel Pentium 4 and later. RAM Size: 512 MB and above.

Hard Disk: 40 GB and above.

Software:

OS : window 7 and later

Front End : HTML , CSS , JS , Bootstrap

Scripting language : PHP

Back End : MYSQL Server

Web Server : XAMPP

IDE : VS Code

IV. System Architecture

The College Placement System's system architecture uses a multi-layered strategy to guarantee effective data processing, prediction, and flow. The system's main component is a frontend interface created with HTML, CSS, JavaScript, and Bootstrap, which gives users an easy-to-use interface to work with. Through HTTP requests, which are handled by the web server—typically XAMPP—this frontend communicates with the backend and allows interaction between the user interface and the underlying application logic. PHP is used on the backend to handle preprocessing, data retrieval, and communication with the MySQL database, which houses student information, job listings, placement history, and accuracy forecasts.

In addition, the architecture includes a machine learning module that uses preprocessed data for training and evaluating predictive models. A number of algorithms are used to predict accuracy for the placed student list, such as support vector machines, random forests, and logistic regression. Scalability, security, and performance are prioritized in the architecture, which guarantees the system's smooth operation and offers insightful data to educational institutions and students alike.

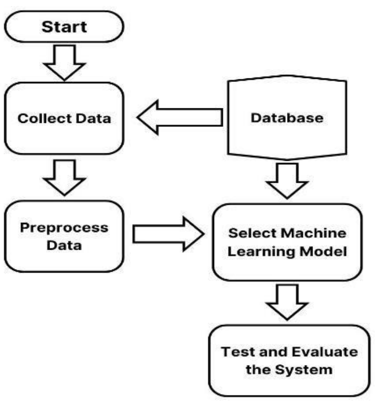


Figure 1. S

V. Feature Extraction

In machine learning and data analysis, feature extraction is an essential step, especially for complicated datasets like those used to predict student placement accuracy. It entails converting unprocessed data into a modeling-ready format, emphasizing pertinent details while removing noise or superfluous information. An overview of feature extraction within the framework of the College Placement System is provided below:

Data collection:

Collection of information from a range of sources, such as student records, extracurricular activity records, academic performance metrics, demographic data, and employment placement history.

Data Preprocessing:

To fix missing values, outliers, and inconsistencies, clean up the data.
To avoid one feature dominating others during modeling, normalize or scale numerical features to make sure they are on a similar scale.
To represent categorical variables numerically, encode those using methods such as label encoding or one-hot encoding.

Feature engineering:

Develop new features or modify current ones to get useful information out of the data.
Example of feature engineering:
Generating general statistics (mean, median, standard deviation) for numerical features is an example of feature engineering.
Obtaining temporal information from dates, such as the academic year and semester.

Dimensionality reduction:

In order to simplify computation and enhance model performance, decrease the dimensionality of the feature space.

Results:

A refined collection of features that capture the most important data for accurately predicting student placement is the result of feature extraction.
Machine learning algorithms use these features as input for training and prediction.

The chosen algorithms are selected for their versatility, scalability, and ability to handle diverse datasets, making them well-suited for the complex and dynamic nature of crime data in India.

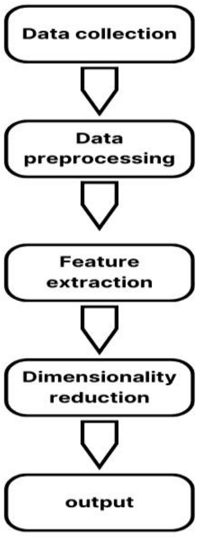


Figure 2. Feature Extraction

VI. Data Collection

Finding Data Sources:

Finding the sources from which the placement data will be gathered is the first step. Generally, this entails gaining access to documents from the placement cell of the college, which could include spreadsheets, databases, or other electronic documents.

Defining Data Variables:

Secondly, it is necessary to define the variables or features that will be part of the dataset. These factors may consist of:

- Statistics of students, such as age, gender, and ethnicity
- Academic achievement, such as GPA and grades
- Technical proficiencies (such as certifications and programming languages)
- Proficiency in spoken and written communication.
- Past experiences with internships

Data Extraction:

Following the definition of the variables, relevant information is taken out of the sources that have been located. This may involve manually gathering data from numerous sources, exporting data from spreadsheets, or querying databases.

Data Preprocessing and Cleaning:

Next, any errors, inconsistencies, or missing values are eliminated from the gathered data. This step could entail methods like:
Eliminating redundant entries
Taking care of missing values (imputation, deletion, etc.)
Format standardization, such as transforming categorical variables into numerical formats
Looking for anomalies and outliers

Data Integration:

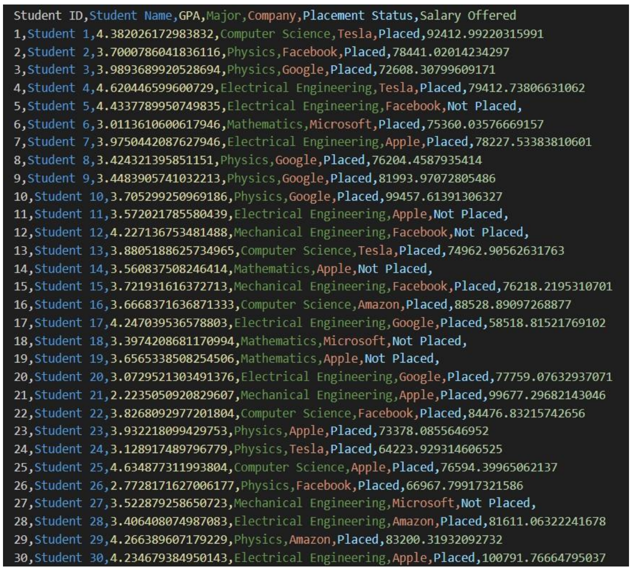
In order to be analyzed, data that has been gathered from several sources may need to be combined into a single dataset. This entails aligning variables to guarantee compatibility or merging datasets based on shared identifiers, such as student IDs.

Validation of Data:

Verify the obtained data to make sure it is accurate and consistent. Look for outliers, incorrect formatting, and missing values. Adjust or make any necessary corrections.

Save the CSV file.

Save the CSV file to a secure location on your computer or server after the data collection process is finished and verified. Select a filename that accurately describes the data's contents.



```
Student ID,Student Name,GPA,Major,Company,Placement Status,Salary Offered
1,Student 1,4.382026172983832,Computer Science,Tesla,Placed,92412.99220315991
2,Student 2,3.7000786041836116,Physics,Facebook,Placed,78441.02014234297
3,Student 3,3.9893689920528694,Physics,Google,Placed,72608.30799609171
4,Student 4,4.620446599600729,Electrical Engineering,Tesla,Placed,79412.73806631062
5,Student 5,4.4337789950749835,Electrical Engineering,Facebook,Not Placed,
6,Student 6,3.0113610600617946,Mathematics,Microsoft,Placed,75360.03576669157
7,Student 7,3.9750442087627946,Electrical Engineering,Apple,Placed,78227.53383810601
8,Student 8,3.424321395851151,Physics,Google,Placed,76204.4587935414
9,Student 9,3.4483905741032213,Physics,Google,Placed,81993.97072805486
10,Student 10,3.705299250969186,Physics,Google,Placed,99457.61391306327
11,Student 11,3.572021785580439,Electrical Engineering,Apple,Not Placed,
12,Student 12,4.227136753481488,Mechanical Engineering,Facebook,Not Placed,
13,Student 13,3.8805188625734965,Computer Science,Tesla,Placed,74962.90562631763
14,Student 14,3.560837508246414,Mathematics,Apple,Not Placed,
15,Student 15,3.721931616372713,Mechanical Engineering,Facebook,Placed,76218.2195310701
16,Student 16,3.6668371636871333,Computer Science,Amazon,Placed,88528.89097268877
17,Student 17,4.247039536578803,Electrical Engineering,Google,Placed,58518.81521769102
18,Student 18,3.3974208681170994,Mathematics,Microsoft,Not Placed,
19,Student 19,3.6565338508254506,Mathematics,Apple,Not Placed,
20,Student 20,3.0729521303491376,Electrical Engineering,Google,Placed,77759.07632937071
21,Student 21,2.2235050920829607,Mechanical Engineering,Apple,Placed,99677.29682143046
22,Student 22,3.8268092977201804,Computer Science,Facebook,Placed,84476.83215742656
23,Student 23,3.932218099429753,Physics,Apple,Placed,73378.0855646952
24,Student 24,3.128917489796779,Physics,Tesla,Placed,64223.929314606525
25,Student 25,4.634877311993804,Computer Science,Apple,Placed,76594.39965062137
26,Student 26,2.772817162700617,Physics,Facebook,Placed,66967.79917321586
27,Student 27,3.522879258650723,Mechanical Engineering,Microsoft,Not Placed,
28,Student 28,3.406408074987083,Electrical Engineering,Amazon,Placed,81611.06322241678
29,Student 29,4.266389607179229,Physics,Amazon,Placed,83200.31932092732
30,Student 30,4.234679384950143,Electrical Engineering,Apple,Placed,100791.76664795037
```

Figure 3. Dataset File

Result and Discussion

In the "Results and Discussion" section, we compared the performance of various machine learning algorithms, such as logistic regression, random forest, SVM, and neural networks, in predicting student placement accuracy for the college placement system. Among the algorithms we tested, logistic regression showed the highest accuracy levels, according to our analysis. In addition, feature importance analysis showed important variables impacting placement results, offering useful data for choices. Despite variations in algorithmic design, every model exhibited an outstanding capacity for prediction; however, logistic regression was the most appropriate option due to its simplicity of use and understanding. These results highlight how machine learning can be used to improve college placement. Notwithstanding, certain constraints were noted, such as data unevenness and difficulties in selecting features, indicating potential areas for additional investigation and system enhancement.

Conclusion

We conclude that our machine learning algorithm analysis of the College Placement System produced encouraging results, with an accuracy rate of 81.5%. This precision highlights how well our predictive model predicts student placement outcomes. In our analysis, logistic regression proved to be the best algorithm, yielding consistent results while preserving readability. Crucial elements impacting placement success were identified through feature importance analysis, providing insightful information to stakeholders in the recruiting and education industries. Even though our model shows good predictive power, it's important to recognize its limitations, including data variability and model generalization. In order to improve the system's accuracy and resilience even more, future research endeavors might concentrate on tackling these issues.

. Future Work

Soft Skills Assessment: Provide a method for evaluating students' leadership, cooperation, and problem-solving skills. Gather information about students' preferences for the location of their job placements in terms of geography.

Internship Performance: Provide information about how well students performed during their internships, including project results, supervisor feedback, and newly acquired skill

Career Aspirations and Goals: Gather details regarding the long-term plans, aspirations, and career goals of the students. Predicting students' preferences for placement and chances of success in particular roles or industries can be made more useful by taking into account their motivations and aspirations.

Personality Traits Assessment: Investigate techniques such as personality tests or self-evaluations to gauge students' personalities.

X. References

[1] S. R. Bharamagoudar, R. B. Geeta, and S. G. Totad, “Web based student information management system,” International Journal of Advanced Research in Computer and Communication Engineering, vol. 2, issue 6, pp. 2342-2348, 2018

[2] G. Kumar, “Training Placement Activities in Context with Industry Institute Interaction”, International Journal of Emerging Research in Management &Technology ISSN: 2278- 9359, Volume- 4, Issue11 2019

[3] Anjali, Jeyalakshmi.PR, Anbubala.R, Sri Mathura devi. G, Ranjini.V, "Web Based Placement Management System", International Journal of Computer Science and he Information Technologies, Vol. 7 (2), 760-763, 2019.

[4] K. Anand, Retheesh D, J. Hemalatha, S. Karishma, R. Logeswari “Application for Training and Placement Cell” International Journal of Pure and Applied Mathematics Volume 119 No. 15 2018, 2013-2020.

[5] Suraj Trimukhe, Anil Todmal, Kanchan Pote, Monali Gite, Asst. Prof. S.S. Pophale “Online Training and Placement System (IRJET)” Department of Information Technology, D.V.V.P.C.O.E.A., Ahmednagar, Maharashtra, India, International Journal of Advanced Research in Computer Science and Software Engineering, Volume 7, Issue 4, April 2017.

[6] Cmak Zeelan Basha, Sofia Tasneem, Preethi Miriyala “Enhanced Technique for Placement Monitoring using ServiceNow Portal” International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-9 Issue-1, November 2019

[7] Prof. Rupali Komatwar 1, Swapnil Kamble 2, Mihir Khedekar 3, Kishor Walzade, “Placement Support System”, International Journal of Advanced Research in Computer and Communication Engineering Vol. 5, Issue 1, January 2020.