

A
MINI PROJECT REPORT
ON
PLACEMENT PREDICTION

SUBMITTED TO THE SAVITRIBAI PHULE PUNE UNIVERSITY, PUNE
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE AWARD OF THE DEGREE OF

BACHELOR OF ENGINEERING INFORMATION TECHNOLOGY

BY

Akshay Koganur (33242)
Avinash Maharnavar (33248)

Under the guidance of
Prof. R. B. MurumKar



DEPARTMENT OF INFORMATION TECHNOLOGY
PUNE INSTITUTE OF COMPUTER TECHNOLOGY
SR. NO 27, PUNE-SATARA ROAD, DHANKAWADI PUNE - 411 043.
AY: 2022-2023

SCTR's PUNE INSTITUTE OF COMPUTER TECHNOLOGY
DEPARTMENT OF INFORMATION TECHNOLOGY



C E R T I F I C A T E

This is to certify that the Project work entitled
Placement Prediction

Submitted by

Akshay Koganur (33242)
Avinash Maharnavar (33248)

is a bonafide work carried out under the supervision of Prof. R. B. Murumkar and it is submitted towards the partial fulfillment of the requirements of Savitribai Phule Pune University, Pune for the award of the degree of Bachelor of Engineering (Information Technology).

Prof. R.B. Murumkar.
Project Guide

Dr. A. S. Ghotkar
HOD IT

Dr. S. T. Gandhe
Principal

Date: 11/07/2022
Place: Pune

Acknowledgement

We would like to express my thankful gratitude to Prof. R. B. Murumkar Sir who guided our machine learning mini project group in this mini project work. Sir gives us valuable guidance, suggestions and advice that helped us to complete this project work in short time span.

I would also like to sincerely thanks to Head of Information technology department Dr. A. S. Ghotkar who gives us this opportunity of doing machine learning mini project which gives us lot of learning experience. We would also like to thanks to the SPPU University for providing such course that gives important knowledge other than academics

Akshay Koganur (33242)

Avinash Maharnavar (33248)

Abstract

An educational institution has one of the most important objective that is placement of a student. Educational institutions are always working for the placement of student by introducing new courses and skills. The analytical study of the skills of student, both technical and soft skills give idea about the student will placed or not. This model will give idea about the skill to be prepared for the placement. The proposed model predicts whether the student will place or not. It uses technical and soft skills. In this project the previous year student's data is used to predict the chance of student placement. The bult model is based on the decision tree classification algorithm. This classification model classifies the student into the placed or not placed category. This classification model will help the student to check their progress easily from time to time. Parameters used for the prediction are the academic score, internship done or not, total number of backlogs etc. The proposed model is also compared with the other classification model with their accuracy. The accuracy obtained from this model is 85%. The model also helps in analysis of difference between skillset of placed and not placed student.

Keywords: Machine learning, Decision tree algorithm, python flask.

Contents

Certificate.....	i
Acknowledgement.....	ii
Abstract.....	iii
Contents.....	iv
List of Figures.....	v
1. Introduction	1
1.1 Introduction	1
1.2 Purpose	1
1.3 Scope.....	2
1.4 Objectives.....	2
1.5 Abbreviations	2
2. Literature Survey.....	3
3. System Architecture and Design	4
3.1 Detail Architecture.....	4
3.2 Dataset Description.....	4
3.3 Detail Phases	5
3.4 Algorithms.....	5
4. Experimentations and Result	6
4.1 Phase-wise Results.....	6
4.2. Explanation with example.....	6
5. Conclusion and Future Scope.....	8
5.1 Conclusion:.....	8
5.2 Future Scope:	8

List of Figures

Sr. No.	Figure Name	Page No.
1.		
2.		
3.		
4.		

1. Introduction

1.1 Introduction

The placements of students are considered to be the main objective of each institution. The basic success of college is measured by the placements of students. It is important for the student as well institutions to an early idea of the current state of skillset of students for the future placement. Educational institutions are always working for the placement of student by introducing new courses and skills. The will get to know that on which parameters they have to focus more. Sometimes it is difficult for the student to prepared for the placement with all the skills. This project will help student to focus on the skillset only required for the placement. The analytical study of the skills of student, both technical and soft skills give idea about the student will placed or not. This model will give idea about the skill to be prepared for the placement. The proposed model predicts whether the student will place or not. It uses technical and soft skills. In this project the previous year student's data is used to predict the chance of student placement.

This classification model is based on the decision tree classification algorithm. This classification model classifies the student into the placed or not placed category. This classification model will help the student to check their progress easily from time to time. Parameters used for the prediction are the academic score, internship done or not, total number of backlogs etc. The proposed model is also compared with the other classification model with their accuracy.

1.2 Purpose

To study the machine learning algorithms and implantation of same for the prediction of the student. The model built using decision tree algorithm able to predict the placement of student based on previous year student data. Help students to improve their skillset from time to time. Compare Built model's accuracy with already existing models for placement prediction to understand efficiency of the algorithm used.

1.3 Scope

The project is based on the decision tree classification algorithm in machine learning. It involves the study and comparison between the different machine learning algorithms. The model also able to study the analysis of student skillset require for placement.

1.4 Objectives

1. To study the decision tree algorithm.
2. To implement decision tree algorithm for placement prediction.
3. To compare the performance of decision tree algorithm with other machine learning algorithms to understand efficiency of algorithms.

1.5 Abbreviations

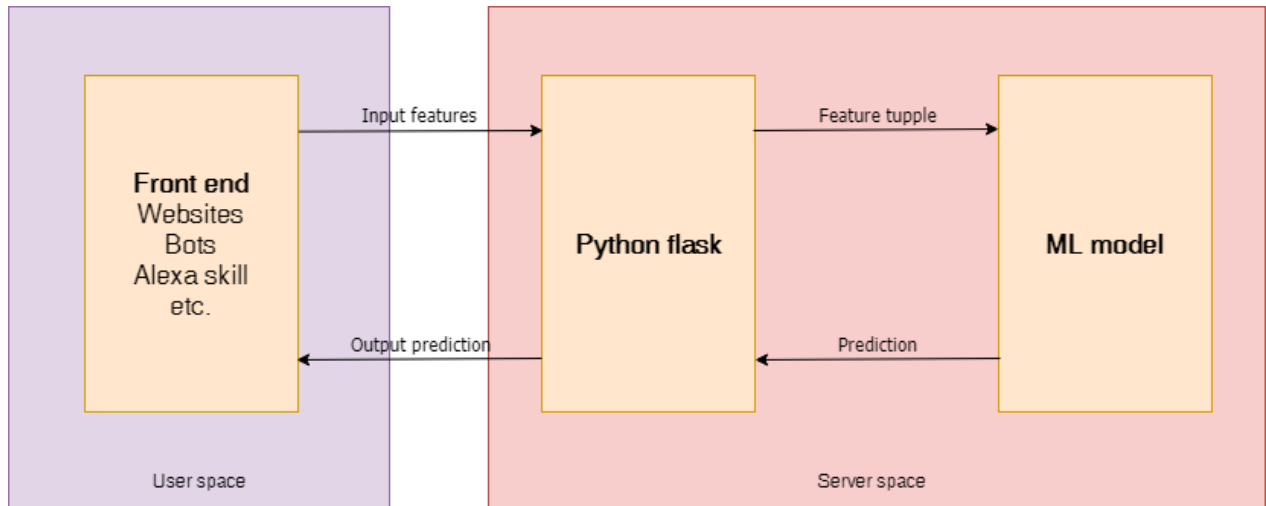
CSS – Cascading Style Sheets.
HTML- Hyper Text Markup Language.
ML- Machine Learning.
REST- Representational State Transfer.
API- Application Programming Interface.
SVM- Support Vector Machine.
KNN- K-Nearest Neighbors.
XGBoost- Extreme Gradient boosting.

2. Literature Survey

Publish year	Paper	Tech./Algorithms Used	Result (%)
2022	Placement Prediction and Analysis using Machine learning.	1.Random Forest. 2.Catboost.	87 88
2022	Campus Placement Prediction.	1.SVM 2.XGBoost.	87 90
2020	Placement Prediction using various machine learning Models and their Efficiency Comparison.	1.Random Forest. 2.KNN 3.Logistic Regression. 4.SVM	96 95 97 100
2019	Campus Placement Prediction Using Machine Learning Techniques.	1.Decision Tree 2.Random Forest.	86 84

3. System Architecture and Design

3.1 Detail Architecture



Frontend gives the input features from the user to the server. Frontend is designed using Simple HTML and CSS. HTML defines the meaning and structure of web content. HTML makes the structure of the website. HTML is often accompanied by CSS. CSS is used for styling the web pages. CSS describes how elements should be rendered on screen, on paper, in speech, or on other media. HTML and CSS together form the visual representation of the web page that is visible to the user. The input request will be given to the trained ML model which is deployed on Python's Flask framework. Flask is a lightweight web application framework. It is designed to make getting started quick and easy, with the ability to scale up to complex applications. Making RESTful APIs by Flask is easy.

3.2 Dataset Description

Engineering Placement Prediction Dataset was downloaded from Kaggle through this url : <https://www.kaggle.com/datasets/tejashvi14/engineering-placements-prediction>

Above dataset consists of 2967 datapoints. Dataset has following attributes : Age, Gender, Stream(i.e. I.T., CSE, EntC, etc.), number of internships done, CGPA, whether lives in hostel or not, Total backlogs(Active as well as Passive), Placed or not. Placed or not attribute has just two values 0 or 1 where 0 indicating that candidate is not placed while 1 indicating that the candidate is placed.

3.3 Detail Phases

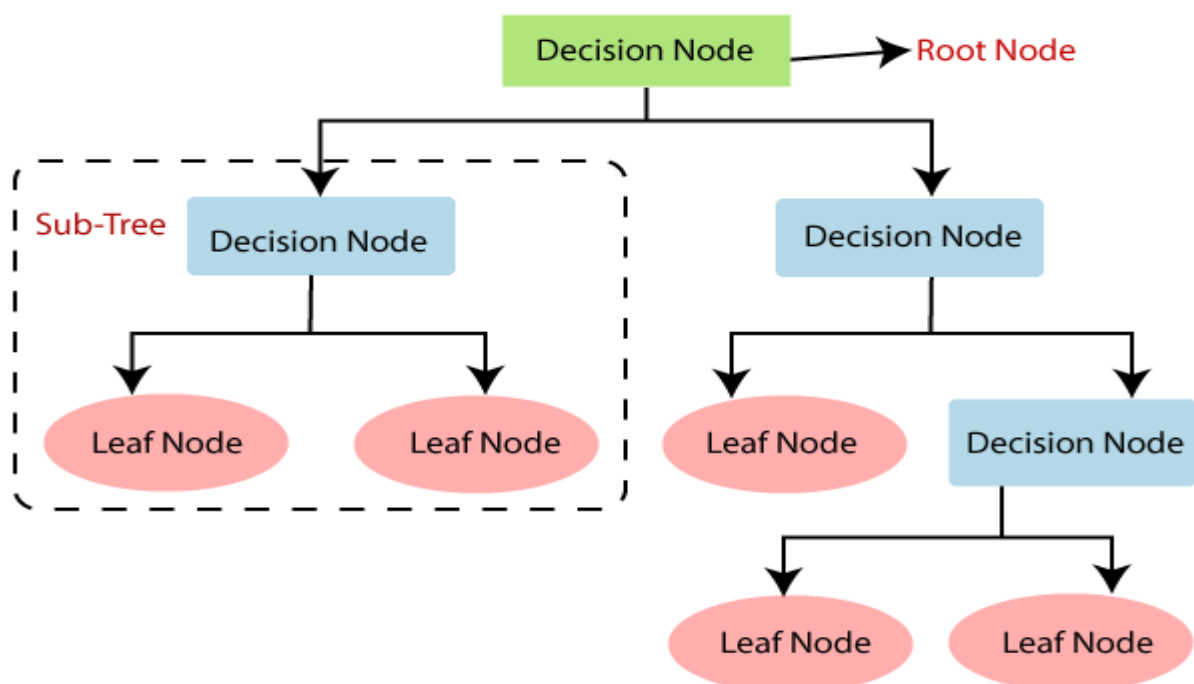
1)Data preprocessing: It includes finding null values in dataset, choosing important features for training and label encoding the dataset. In some cases, the dataset contains missing values. A common plan to handle the matter is to require a mean of all the values of the same column and have it to replace the missing data.

2)Training and testing dataset and checking its performance: Now the next step is to split our dataset into two. Training set and a Test set. We will train our machine learning models on our training set, i.e., our machine learning models will try to understand any correlations in our training set and then we will test the models on our test set to examine how accurately it will predict.

3)Deploying ML model on Flask: ML model which we have prepared will be deployed on Flask. Flask is a Python-based micro framework used for developing small-scale websites. Flask is used to handle API requests.

3.4 Algorithms

Decision Tree algorithm can be used to train the model on Placement prediction dataset. Decision Tree is a Supervised learning technique that can be used for both classification and Regression problems, but mostly it is preferred for solving Classification problems. It is a tree-structured classifier, where internal nodes represent the features of a dataset, branches represent the decision rules and each leaf node represents the outcome.



4. Experimentations and Result

4.1 Phase-wise Results

Phase 1: The dataset didn't contain any null or missing values. Also, the attributes 'Hostel' and 'Age' didn't were not important for training the dataset, hence these attributes were deleted from the dataset. Label encoding was applied on attributes 'Stream' and 'Gender' to convert values in these attributes to int.

Phase 2: The dataset was divided into training and testing dataset in the 70:30 proportion. 70% in training and 30% in testing. The model showed 85.7% accuracy, 96.6% precision, 76.9% recall and 85.7% F1 score.

Phase 3: Flask was installed and build the frontend part of this project was prepared as well as API requests were prepared. At the end, website is ready.

4.2. Explanation with example

The website works based on the information provided by the user. This input will be given to trained model to predict the output. For example, user provides these inputs: Gender is male, Stream is CSE, Previously done 1 internship, CGPA is 9 and no backlog. The trained model will predict based on input and it will be shown to user as :

Welcome to Placement Prediction App
Your Result is
Yes
This Website uses Decision Tree Model to predict results

4.3. Comparison of result with standard

4.4. Accuracy

The model showed 85.7% accuracy, 96.6% precision, 76.9% recall and 85.7% F1 score.

```
acc = metrics.accuracy_score(y_test,y_pred)
acc
```

```
0.8573033707865169
```

```
pre = metrics.precision_score(y_test,y_pred)
pre
```

```
0.9669211195928753
```

```
re = metrics.recall_score(y_test,y_pred)
re
```

```
0.7692307692307693
```

```
f1 = metrics.f1_score(y_test,y_pred)
f1
```

```
0.8568207440811726
```

4.5. Visualization

4.6. Tools used

Linux OS: Ubuntu/Windows, Jupiter notebook.

PC with the configuration as Pentium IV 1.7 GHz. 128M.B RAM, 40 G.B HDD, 15" Color Monitor, Keyboard, Mouse.

Languages: Python

Framework: Flask

Libraries: NumPy, Pandas, Matplotlib, Seaborn, sklearn.

5. Conclusion and Future Scope

5.1 Conclusion:

The campus placement activity is incredibly a lot of vital as institution point of view as well as student point of view. In this regard to improve the student's performance, a work has been analyzed and predicted using the classification algorithms Decision Tree and then this model is deployed on Flask. Making website of this model will help to increase the reach of this model to as many students as possible.

5.2 Future Scope:

More instances can be added to the dataset to get more accurate results. Dataset can be trained by different machine learning algorithms to check which performs the best and deploy the website on that model for better results.

Savitribai Phule Pune University, Pune Third Year Information Technology (2019 Course) 314449 : Seminar		
Teaching Scheme:	Credit Scheme:	Examination Scheme:
Practical (PR) : 01 hrs/week	01 Credits	TW : 50 Marks
Prerequisites: <ol style="list-style-type: none"> 1. Project Based Learning 2. Software Engineering 		
Course Objectives: Seminar should make the student attain skills like: <ol style="list-style-type: none"> 1. To gather the literature of specific area in a focused manner. 2. To summarize the literature to find state-of-the-art in proposed area. 3. To identify scope for future work. 4. To present the case for the intended work to be done as project. 5. To report literature review and proposed work in scientific way. 		
Course Outcomes: On completion of the course, students will be able to– <p>CO1: Understand, interpret and summarize technical literature.</p> <p>CO2: Demonstrate the techniques used in the paper.</p> <p>CO3: Distinguish the various techniques required to accomplish the task. CO4: Identify intended future work based on the technical review.</p> <p>CO5: Prepare and present the content through various presentation tools and techniques in effective manner.</p> <p>CO6: Keep audience engaged through improved interpersonal skills.</p>		
Guidelines for Seminar Selection and Presentation		
<ol style="list-style-type: none"> 1) Student shall identify the area or topics in Information Technology referring to recent trends and developments in consultation with industry (for their requirement) and institute guide. 2) Student must review sufficient literature (reference books, journal articles, conference papers, white papers, magazines, web resources etc.) in relevant area on their topic as decided. 3) Seminar topics should be based on recent trends and developments. Guide should approve the topic by thoughtfully observing different techniques, comparative analysis of the earlier algorithms used or specific tools used by various researchers in the domain. 4) Research articles could be referred from IEEE, ACM, Science direct, Springer, Elsevier, IETE, CSI or from freely available digital libraries like Digital Library of India (dli.ernet.in), National Science Digital Library, JRD Tata Memorial Library, citeseerx.ist.psu.edu, getcited.org, arizona.openrepository.com, Open J-Gate, Research Gate, worldwidescience.org etc. 5) Student shall present the study as individual seminars in 20 – 25 minutes in English which is followed by Question Answer session. 6) Guide should ensure that students are doing literature survey and review in proper manner. 7) Guide should give appropriate instructions for effective presentation. 8) Attendance of all other students in the class for presentation is mandatory. 		



Timeline is suggested to follow throughout the semester:

- 1) **Week– 01:** Discussion to understand what is technical paper, how to search, where to search?
- 2) **Week– 02:** Download technical papers (minimum four), getting approved from Guide and Prepare abstract summary of all papers downloaded.
- 3) **Week– 03 & 04:** Read and understand in detail the decided research papers about the problem statement, techniques used, experimental details and results with conclusion from identified papers.
- 4) **Week– 05:** Review of the studied papers by Guide / Panel.
- 5) **Week – 06 & 07:** Search / Find equivalent techniques (other than the one proposed in technical paper) so performance / complexities can be improved (by amortized analysis, not actual implementation).
- 6) **Week – 08 & 09:** Prepare presentation with outline as The topic, its significance, The research problem, Studied solutions (through research papers) with strengths and weaknesses of each solution, comparison of the solutions to research problem, future directions of work, probable problem statement of project, tentative plan of project work
- 7) **Week – 10:** Write Seminar report.
- 8) **Week – 11:** Deliver Presentation to Guide/ Panel.
- 9) **Week–12:** Verification of Seminar report and Submission.

Guidelines for Seminar report

1. Each student shall submit two copies of the seminar report in appropriate text editing tool/software as per prescribed format duly signed by the guide and Head of the department/Principal.
2. Broad contents of review report (20-25 pages) shall be
 - a) Title Page with Title of the topic, Name of the candidate with Exam Seat Number / Roll Number, Name of the Guide, Name of the Department, Institution, Year & University.
 - b) Seminar Approval Sheet/Certificate.
 - c) Abstract and Keywords.
 - d) Acknowledgments.
 - e) Table of Contents, List of Figures, List of Tables and Nomenclature.
 - f) Chapters need to cover topic of discussion-
 - i. Introduction with section including organization of the report,
 - ii. Literature Survey
 - iii. Motivation, purpose and scope and objective of seminar
 - iv. Details of design/technology/Analytical and/or experimental work, if any/
 - v. Discussions and Conclusions,
 - vi. Bibliography/References (in IEEE Format),
 - vii. Plagiarism Check report,
3. Students are expected to use open source tools for writing seminar report, citing the references and plagiarism detection.

Guidelines for Lab /TW Assessment:	
<ol style="list-style-type: none">1. A panel of reviewers constituted by seminar coordinator (where guide is one of the member of the panel) will assess the seminar during the presentation.2. Student's attendance for all seminars is advisable.3. Rubric for evaluation of seminar activity:<ol style="list-style-type: none">i. Relevance of topic - 05 Marksii. Relevance + depth of literature reviewed - 10 Marksiii. Seminar report (Technical Content) - 10 Marksiv. Seminar report (Language) - 05 Marksv. Presentation Slides - 05 Marksvi. Presentation & Communication Skills - 05 Marksvii. Question and Answers - 10 Marks <p style="text-align: right;">TOTAL: 50 Marks</p>	
Reference Book:	
<ol style="list-style-type: none">1. Andrea J. Rutherford, Basic Communication Skills for Technology, Pearson Education Asia, 2nd Edition.2. Lesikar, Lesikar's Basic Business Communication, Tata McGraw, ISBN: 256083274, 1st Edition.	
Text Book :	
<ol style="list-style-type: none">1. Sharon J. Gerson, Steven M. Gerson, Technical Writing: Process and Product, Pearson Education Asia, ISBN: 130981745, 4th Edition.	