

# IDENTIFY PATTERNS AND TRENDS IN CAMPUS PLACEMENT DATA USING MACHINE LEARNING

Submitted in partial fulfillment of requirement for the award of the Degree

Bachelor of Computer Science

In the faculty of Computer Science of Bharathiar University, Coimbatore.

Submitted by

TEAM ID: NM2023TMID18424

**NIRMALA .V**

**SANTHIYA .A**

**RASIKADEVI .S**

**JANANI .S**

Under the guidance of lecturer ,

**Dr. V.PUNITHAVATHI MCA.,M.Phil., DOA., Ph.D.**

Department of Computer Science



**DEPARTMENT OF COMPUTER SCIENCE**

**LRG GOVERNMENT ARTS COLLEGE FOR WOMEN**

**(Affiliated To Bharathiar University)**

**TIRUPUR-4**

**APRIL-2023**

# **LRG GOVERNMENT ARTS COLLEGE**

## **NAAN MUDHALVAN PROJECT WORK**

**(AFFILIATED TO BHARATHIAR UNIVERSITY)**

**TIRUPUR-641602**

**TITLE :Identifying patterns and Trends in campus placement data  
using machine learning**

This is to certify that this is a bonafide record of work done by the above students  
of III B.Sc (CS) Degree**NAAN MUDHALVAN PROJECT** during the year  
2022-2023

Submitted for the Naan Mudhalvan project work held

On .....20

**CLASS TUTOR**

[Dr.V.punithavathi]

**HEAD OF THE DEPARTMENT**

[Dr.R.Parimala]

# **INDEX**

<b>CHAPTER NO</b>	<b>CONTENTS</b>
<b>1</b>	INTRODUCTION
<b>2</b>	PROBLEM DEFINITION & DESIGN THINKING
<b>3</b>	RESULT
<b>4</b>	ADVANTAGES & DISADVANTAGES
<b>5</b>	APPLICATIONS
<b>6</b>	CONCLUSION
<b>7</b>	FUTURE SCOPE
<b>8</b>	APPENDIX

## **INTRODUCTION**

### **1.1 OVERVIEW**

Identifying patterns and trends in campus placement data is an important task for universities, colleges, and recruitment agencies. Machine learning algorithms can be used to analyze large amounts of placement data and identify patterns and trends that can inform recruitment strategies and decision-making processes.

In this overview, we will discuss how machine learning can be used to identify patterns and trends in campus placement data. The first step in this process is to collect and preprocess the data. This may involve gathering data from multiple sources, cleaning and transforming the data into a format that can be used by machine learning algorithms.

Once the data is prepared, various machine learning algorithms can be applied to identify patterns and trends. For example, clustering algorithms can be used to group similar job profiles or companies together based on certain attributes, such as salary, location, or industry sector. Decision trees and regression models can be used to identify key factors that influence job placement success, such as academic performance or previous work experience.

## **1.2 PURPOSE**

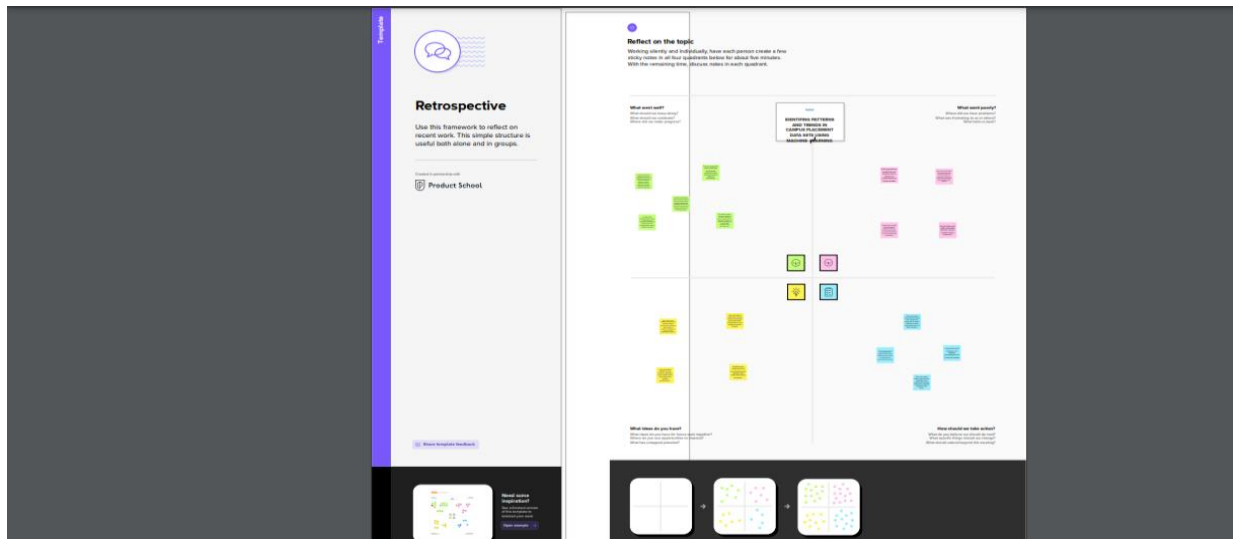
can be used to inform recruitment strategies, improve the effectiveness of the campus placement process, and better support students seeking job placements. By identifying which industries and skillsets are in high demand, universities and colleges can tailor their programs and support services to better prepare students for job opportunities.

Additionally, companies can use this information to identify potential candidates that have the skills and experience they are looking for. This can help to reduce the time and effort needed to find the right candidates for job openings, improving the overall efficiency of the recruitment process.

## **PROBLEM DEFINITION&DESIGN THINKING**

## 2.1 EMPATHY MAP

By using empathy mapping to better understand the needs and concerns of different stakeholders involved in campus placement, if recruiters are struggling to identify potential candidates, insights gained from analyzing placement data may help to better target potential candidates.



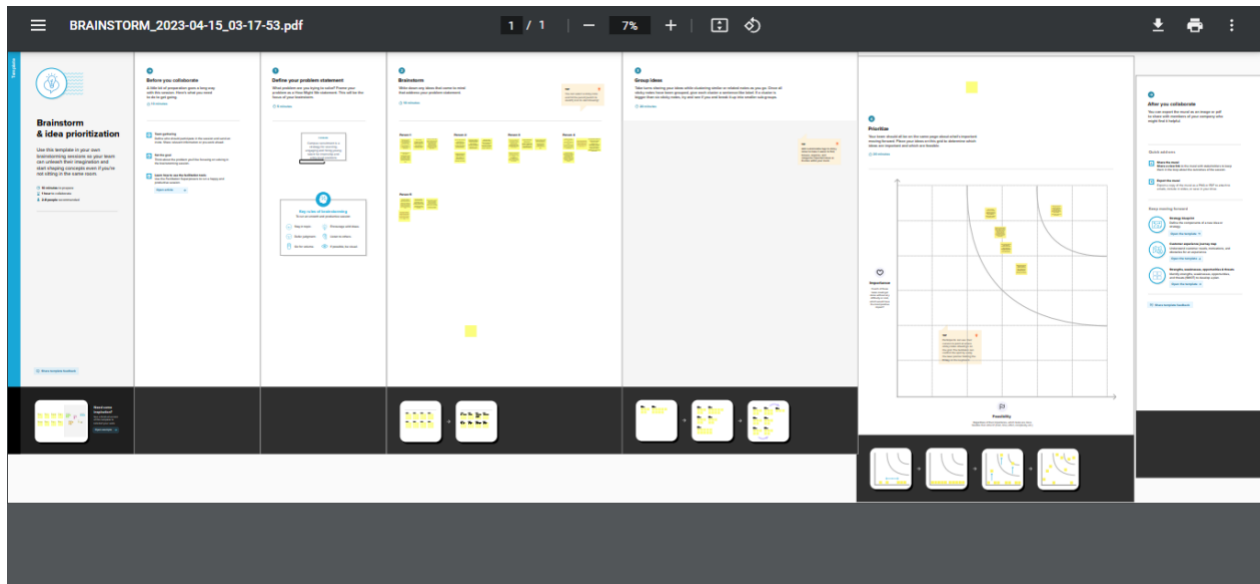
## 2.2 IDEATION AND BRAINSTORMING

ideation and brainstorming for identifying patterns and trends in campus placement is to be open-minded and collaborative. By

working with stakeholders and experimenting with different approaches, it's possible to gain valuable insights that can inform recruitment strategies and support services for students seeking job placements.

**Collaborate with stakeholders:** Bring together stakeholders involved in the campus placement process, such as recruiters, students, and university officials. Encourage them to share their ideas and insights about the placement process, and use this as a starting point for brainstorming.

**Use visualization tools:** Consider using visualization tools such as graphs, charts, and heat maps to help identify patterns and trends in the data. This can help to make complex data more accessible and easier to understand.



# RESULT

## Campus Placement Prediction

Enter Details for forecast.

SSC PR%:[45-100]

mba\_p:[1-99]

hsc\_s

HSC PR%:[45-100]

gender

degree\_t

Degree PR%:[35-100]

ssc\_b

workex

etest\_p:[1-99]

hsc\_b

specialisation

Predict

Student will be

**Placed**

Return

Project by

**Devansh Mistry**

Jr. Data Scientist

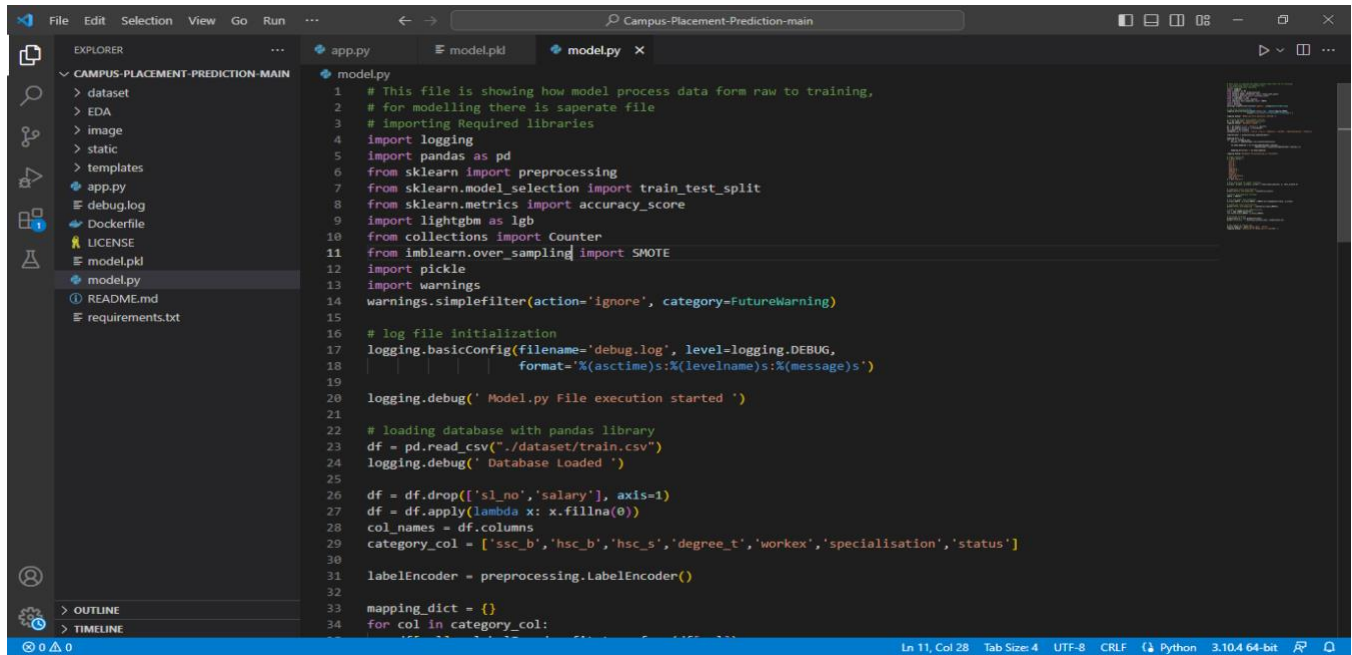
Connect



*Shivam MISTRY*



# SAMPLE CODING



The screenshot shows a code editor with a dark theme. The Explorer panel on the left shows a project named 'CAMPUS-PLACEMENT-PREDICTION-MAIN' with files like 'dataset', 'EDA', 'image', 'static', 'templates', 'app.py', 'debug.log', 'Dockerfile', 'LICENSE', 'model.pkl', 'model.py', 'README.md', and 'requirements.txt'. The main editor area shows the code for 'model.py'. The code includes imports for logging, pandas, sklearn, lightgbm, collections, imblearn, pickle, and warnings. It sets up logging, loads a CSV dataset, preprocesses it (dropping 'sl\_no' and 'salary', filling NaNs), and defines category columns. It also initializes a LabelEncoder and a mapping dictionary.

```
1 # This file is showing how model process data form raw to training.
2 # for modelling there is saperate file
3 # importing Required libraries
4 import logging
5 import pandas as pd
6 from sklearn import preprocessing
7 from sklearn.model_selection import train_test_split
8 from sklearn.metrics import accuracy_score
9 import lightgbm as lgb
10 from collections import Counter
11 from imblearn.over_sampling import SMOTE
12 import pickle
13 import warnings
14 warnings.simplefilter(action='ignore', category=FutureWarning)
15
16 # log file initialization
17 logging.basicConfig(filename='debug.log', level=logging.DEBUG,
18                     format='%(asctime)s: %(levelname)s: %(message)s')
19
20 logging.debug(' Model.py File execution started ')
21
22 # loading database with pandas library
23 df = pd.read_csv("../dataset/train.csv")
24 logging.debug(' Database Loaded ')
25
26 df = df.drop(['sl_no', 'salary'], axis=1)
27 df = df.apply(lambda x: x.fillna(0))
28 col_names = df.columns
29 category_col = ['ssc_b', 'hsc_b', 'hsc_s', 'degree_t', 'workex', 'specialisation', 'status']
30
31 labelEncoder = preprocessing.LabelEncoder()
32
33 mapping_dict = {}
34 for col in category_col:
```

## ADVANTAGES

**Efficiency:** Machine learning algorithms can process and analyze large amounts of data quickly and accurately, saving time and reducing errors compared to manual analysis.

**Data-driven decision-making:** By identifying patterns and trends in campus placement data, institutions can make data-driven decisions to improve their academic programs and placement rates, leading to better outcomes for their graduates.

**Customization:** Machine learning algorithms can be customized to the specific needs of each institution, allowing them to focus on the most relevant factors affecting their placement rates.

**Improved accuracy:** Machine learning algorithms can provide more accurate predictions and recommendations based on historical placement data and current trends in the job market.

Scalability: Machine learning algorithms can be easily scaled to handle larger datasets, allowing institutions to analyze placement data from multiple years and make more informed decisions.

## **DISADVANTAGES**

Data quality: Machine learning algorithms rely on high-quality and accurate data to provide reliable predictions and recommendations. If the placement data is incomplete or inaccurate, it can affect the accuracy of the results.

Bias: Machine learning algorithms can be biased based on the data they are trained on, which can result in inaccurate predictions or recommendations. Institutions need to ensure that the data used to train the algorithms is diverse and representative.

Technical expertise: Implementing machine learning algorithms requires technical expertise and specialized skills that may not be available within the institution. Institutions may need to invest in training or hiring personnel with the necessary skills to implement and maintain the algorithms.

Cost: Implementing machine learning algorithms can be expensive, requiring investments in hardware, software, and personnel. Institutions need to carefully evaluate the cost-benefit of implementing these algorithms before making a decision.

Interpretability: Machine learning algorithms can provide accurate predictions and recommendations, but they can be difficult to interpret. Institutions need to ensure that they can understand and explain the results of the algorithms to stakeholders.

## **APPLICATIONS**

**Program evaluation:** Institutions can use machine learning algorithms to evaluate the effectiveness of their academic programs by analyzing placement data for different majors and identifying trends and patterns.

**Career counseling:** By analyzing placement data, institutions can provide more accurate and personalized career counseling to their students, helping them choose the right major and prepare for the job market.

**Employer outreach:** Institutions can use machine learning algorithms to identify which industries and companies are hiring the most graduates, and target their employer outreach efforts accordingly.

**Curriculum development:** By identifying patterns and trends in placement data, institutions can adjust their curriculum to better align with the needs of the job market, ensuring that their graduates have the skills and knowledge employers are looking for.

**Student recruitment:** Institutions can use machine learning algorithms to analyze placement data and identify which majors and programs have the highest placement rates, using this information to attract prospective students.

**Institutional benchmarking:** Institutions can compare their placement rates and patterns to those of peer institutions, identifying areas where they may need to improve and learning from successful strategies used by other institutions.

## **CONCLUSION**

Identifying patterns and trends in campus placement data using machine learning can provide valuable insights for colleges and universities. By analyzing placement data, institutions can identify which majors, industries, and locations offer the best job opportunities for their graduates, and make data-driven decisions to improve their academic programs and placement rates.

While there are some potential disadvantages to implementing machine learning algorithms, such as the need for high-quality data and technical expertise, the benefits outweigh the drawbacks for most institutions. Implementing machine learning algorithms can lead to more efficient and accurate analysis of placement data, better career counseling for students, improved curriculum development, and more effective employer outreach efforts.

Overall, identifying patterns and trends in campus placement data using machine learning can lead to better outcomes for graduates, improved institutional performance, and a competitive advantage for colleges and universities.

## **FUTURE SCOPE**

Identifying patterns and trends in campus placement data using machine learning has a lot of potential for future applications. Here are some potential areas of future scope:

**Improving student employability:** By analyzing the campus placement data, machine learning models can identify the skills and traits that are most sought-after by employers. This information can be used by educational institutions to design their curriculum and training programs to better prepare students for the job market.

**Predicting job market trends:** By analyzing the campus placement data over time, machine learning models can identify trends in the job market, such as which industries are growing, which skills are in demand, and which regions have the most job opportunities. This information can help students and job seekers make more informed decisions about their career paths.

Identifying bias in hiring: Machine learning models can be trained to analyze the campus placement data for signs of bias in the hiring process. For example, they can detect whether certain groups of students are consistently overlooked or underrepresented in the hiring process. This information can be used to improve diversity and inclusion in the workplace.

Personalized career guidance: By analyzing individual students' performance in the campus placement process, machine learning models can provide personalized career guidance and job recommendations based on their skills, interests, and strengths.

Improving recruitment processes: Employers can use machine learning models to analyze their campus placement data to identify the most effective recruitment strategies and optimize their hiring processes.

## **APPENDIX**

**Identifying patterns and Trends in campus placement data using machine learning**

### **VIDEO LINK ABOUT THIS PROJECT**

<https://youtu.be/XgaBy7aO4cM>

### **GITHUB LINK**

<https://github.com/saran858/Identify-Patterns-and-trends-campus-placement-data-using-machine-learning>