# An Internship Report
On

# CAMPUS RECRUITMENT USING MACHINE LEARNING

*"know your chances of getting placed"*

*Submitted in partial fulfilment of the requirement for the award of the degree of*

**BACHELOR OF ENGINEERING**
in
**COMPUTER SCIENCE AND ENGINEERING**

*Submitted by*

Mohit Notani
Sanjay K
Mohammed Shiyaz
Ananya Y K
Meghana K
Podaralla Revathi
R Yogetha

*Under the Guidance of*

<table>
<tr><td>NEXT UP</td><td>MENAHI SHAYAN</td></tr>
<tr><td>ORGANISATION</td><td>INSTRUCTOR</td></tr>
</table>

**DEVELOPED UNDER NEXT UP INDIA AS PART OF**

A NEXT UP
CURATED PROGRAM

*https://nextup.co.in*

# 2020-2021

# ACKNOWLEDGEMENT

It gives us immense pleasure to present before you our project titled "**CAMPUS RECRUITMENT IN MACHINE LEARNING**". The joy and satisfaction that accompany the successful completion of any task would be incomplete without the mention of those who made it possible. We are glad to express our gratitude towards our prestigious ogranisation **NEXT-UP** for providing us with utmost knowledge, encouragement and the maximum facilities in undertaking this project.

We wish to express sincere thanks to Menahi Shayan and Nishank Swamy for all their support.

We express our deepest gratitude and special thanks to our Resource Persons **Mr. PARAM MIRANI** and **Mr. RAHUL SANDHU** for all their guidance and encouragement.

# ABSTRACT

Education in Machine Learning is very useful, the field of education to predict campus placement of students. Campus placement is a very important concern for any educational organization and wants to improve placement. Educational in Machine Learning uses the new technique and tools for discovering the knowledge by analysing the database to support for prediction process in educational institution. The datasets analyse the different machine learning techniques and implement machine learning technique to enhance prediction for campus placement in any higher education institute. In this datasets takes student academic data to predict the standing of placement as an input. Flask software is used for design and implements to making clusters of complete database which is classify the students according to their performance and qualification. The parameters to calculating performance of student including academic performance, communication skills, technical skills and work-experiences are measured to ability of a student for placement. The educational institution can predict the campus placement of each student and improve the placement of the organization.

# TABLE OF CONTENTS

# 1.INTRODUCTION

Campus Placements are considered to be very important for each and every college. The basic success of the college is measured by the campus placement of the students. Every student takes admission to the colleges by seeing the percentage of placements in the college. Hence, in this regard the approach is about the prediction and analyses for the placement necessity in the colleges that helps to build the colleges as well as students to improve their placements.

In Campus Placement Prediction system predicts the probability of a undergrade students getting placed in a company by applying supervised-learning classification algorithms such as Logistic Regression. The main objective of this model is to predict whether the student he/she gets placed or not in campus recruitment. For this the data consider is the academic history of student like overall percentage in SSLC, PU, Degree and work experience. The algorithms are applied on the previous years data of the students.

# 2.TASK DESCRIPTION

## 2.1 PROBLEM STATEMENT & DESCRIPTION

1.  **Context**

• With the number of millennials and freshers increasing at an exponential rate, conducting campus recruitment and hiring fresh talent has become an inexplicable part of most company's hiring procedure. For companies hiring freshers is one of the most advantageous decision to get fresh, energetic and committed individuals brimming with new ideas and excited for new experiences out to fulfil their dreams.

2.  **Task Details**

• To identify the factors affecting campus placement based on the data. For example, degree marks, high school marks, gender, work experience, etc. and to predict the chances of getting placed based on the identified factors as input.

• Additionally, to identify:

    i.    The importance of percentage in getting placed.

    ii.    The impact of switching fields in the chances of getting placed.

    iii.    How gender affects the likelihood of placement for the same parameters.

3.  **Datasets**

    • Campus Recruitment Dataset

    • Additional Campus Recruitment Data

## 2.2 TEAM

| TEAM MEMBERS | ROLE |
|---|---|
| 1.Mohit Notani | Team leader & Backend Team |
| 2. Sanjay K | Backend Head |
| 3.Mohammed Shiyaz | Backend Team |
| 4.R Yogetha | Documentation & Frontend Team |
| 5. Ananya Y K | Frontend Head |
| 6. Meghana K | Frontend Team |
| 7.Podaralla Revathi | Frontend Team |

## 2.3 DATASETS & OTHER DATA:

These datasets are used for machine-learning research and have been cited in peer-reviewed academic journals. Datasets are an integral part of the field of machine learning.

## I.    PRIMARY DATASETS

| sl_no | gender | ssc_p | hsc_p | hsc_s | degree_ | degree_t | worke | status |
|---|---|---|---|---|---|---|---|---|
| 1 | M | 67 | 91 | Comme | 58 | Sci&Tech | No | Placed |
| 2 | M | 79.33 | 78.33 | Science | 77.48 | Sci&Tech | Yes | Placed |
| 3 | M | 65 | 68 | Arts | 64 | Comm&Mg | No | Placed |
| 4 | M | 56 | 52 | Science | 52 | Sci&Tech | No | Not Placed |
| 5 | M | 85.8 | 73.6 | Comme | 73.3 | Comm&Mg | No | Placed |

In the above mentioned primary datasets following will be Independent variable and Dependent variable:

i.    Independent Variable: gender ,ssc_p, hsc_p, hsc_s, degree_p, degree_t, workex.

ii.    Dependent Variable: status.

## II.    CUSTOMISED DATASETS

| sl_no | gender | tenth_ma | pu_mark | pu_stream | deg_ma | deg_stream | worke | status |
|-------|--------|----------|---------|-----------|--------|------------|-------|--------|
| 1 | M | 58 | 70 | Commerce | 61 | Comm&Mgmt | No | Not Placed |
| 2 | M | 58 | 61 | Commerce | 60 | Comm&Mgmt | Yes | Placed |
| 3 | M | 69.6 | 68.4 | Commerce | 78.3 | Comm&Mgmt | Yes | Placed |
| 4 | F | 47 | 55 | Science | 65 | Comm&Mgmt | No | Not Placed |
| 5 | F | 77 | 87 | Commerce | 59 | Comm&Mgmt | No | Placed |

In the above mentioned Customised datasets following will be Independent variable and Dependent variable:

**i.    Independent Variable:** gender, tenth_marks, pu_marks, pu_stream, deg_marks, deg_stream, work_ex.

**ii.    Dependent Variable:** status.

Where, **Independent Variable** is the variable the experimenter changes or controls and is assumed to have a direct effect on the dependent variable.

The **Dependent variable** is the variable being tested and measured in an experiment, and is 'dependent' on the independent variable.

# 3.WORKLOG

| Timestamp | Task completed | Members Involved | GitHub Commit ID |
|---|---|---|---|
| Jul 14, 2020 | Initial Commit | Mohit | 509cf45 |
| Jul 17, 2020 | Initial Training Model | Sanjay | 0695eb8 |
| Jul 17, 2020 | Files Uploaded | Mohit | f1dc97f |
| Jul 20, 2020 | Files Uploaded | Sanjay, Shiyaz | b6ea44a |
| Jul 20, 2020 | Files Uploaded | Sanjay | 9cfed52 |
| Jul 25, 2020 | Files Uploaded | Shiyaz | 1bfb3c7 |

# 4.SYSTEM REQUIREMENTS

The Software Requirements Specification is priced at the culmination of the analysis . The function and performance allocated to software as put of system are refined by establishing in complete information description ,a detailed functional and behavioural description, an indication of performance, requirements and design constraints appropriate validation criteria other to requirements

The Method was of using various new tech to make website as responsive and as efficient possible

## 4.1 SOFTWARE REQUIREMENTS

1. Operating System: Windows 7 or higher  version

2. Browser: Google Chrome/Internet Explorer

3. Technology Used: HTML5,CSS, JS,PYTHON

4. Editor-Sublime Text3,Python 3(Flask)

## 4.2 HARDWARE REQUIREMENTS

1   Computer with a 1.1 GHz or faster processor

2   Minimum 4GB of DDR3 RAM or more

3   2.5 GB of available hard-disk space

4   5400 RPM hard drive

5   600*800 or higher-resolution display

# 5.IMPLEMENTATION AND CODING

## FRONT-END

For front-end we have used the frame-work HTML,CSS,JAVA-SCRIPT.
The meta which we used for designing the home page:

```html
<meta charset="UTF-8">
    <meta http-equiv="X-UA-Compatible" content="IE=edge">
    <meta name="viewport" content="width=device-width, initial-scale=1">
    <meta name="description" content="APPTON HTML5 Template is a simple Smooth
 Personal App Landing Template" />
    <meta name="keywords" content="App, Landing, Business, Onepage, Html, Busi
ness" />
```

The style-sheets which we used for designing the web page:

```html
    <link href="assest/css/plugins.css" rel="stylesheet">
    <link href="assest/css/theme.css" rel="stylesheet">
    <link href="assest/css/icons.css" rel="stylesheet">

    <!--======= MAIN STYLESHEETS =======-->
    <link href="style.css" rel="stylesheet">
    <link href="assest/css/responsive.css" rel="stylesheet">
```

The java-scripts which we used for designing the web page:

```html
<!--======= SCRIPTS JS =======-->
    <script src="assest/js/vendor/jquery-1.12.4.min.js"></script>
    <script src="assest/js/vendor/bootstrap.min.js"></script>

    <!--======= PLUGINS JS =======-->
    <script src="assest/js/vendor/jquery.easing.1.3.js"></script>
    <script src="assest/js/vendor/jquery-migrate-1.2.1.min.js"></script>
    <script src="assest/js/vendor/jquery.appear.js"></script>
    <script src="assest/js/owl.carousel.min.js"></script>
    <script src="assest/js/stellar.js"></script>
    <script src="assest/js/waypoints.min.js"></script>
    <script src="assest/js/jquery.counterup.min.js"></script>
    <script src="assest/js/wow.min.js"></script>
    <script src="assest/js/jquery-modal-video.min.js"></script>
    <script src="assest/js/stellarnav.min.js"></script>
    <script src="assest/js/placeholdem.min.js"></script>
    <script src="assest/js/contact-form.js"></script>
    <script src="assest/js/jquery.ajaxchimp.js"></script>
    <script src="assest/js/jquery.sticky.js"></script>
```

The slides used for the web-page:

```
<div class="welcome-slider-area white font16">
<div class="welcome-single-slide">
<div class="slide-bg-one slide-bg-overlay"></div>
<div class="welcome-area">
<div class="container">
<div class="row flex-v-center">
<div class="col-md-8 col-lg-7 col-sm-12 col-xs-12">
<div class="welcome-text">
<h1>CAMPUS PLACEMENT</h1>
<p>Was certainty remaining engrossed applauded sir how discovery. Settl
ed opinion how enjoyed greater joy adapted too shy. Now properly surprise expe
nses.</p>
<div class="home-button">
</div>
</div>
</div>
</div>
</div>
</div>
</div>
```

Synopsis Area:

```
<section class="section-padding" id="synopsis">
<div class="container">
<div class="row">
<div class="col-md-8 col-lg-8 col-md-offset-2 col-lg-offset-2 col-sm-
12 col-xs-12">
<div class="text-center wow fadeIn">
<h2 class="xs-font20">Synopsis</h2><p></p>
</div></div>
</div>
</div></section>
```

For the Developer team:

```
<section class="course-area padding-top" id="developerTeam">
<div class="container-fluid">
<div class="row">
<div class="col-md-8 col-lg-8 col-md-offset-2 col-lg-offset-2 col-sm-
12 col-xs-12">
<div class="area-title text-center wow fadeIn">
<h2 class="xs-font26">Our Development Team</h2>
</div></div>
</div>
```

For About us:

```html
<section class="section-padding" id="aboutus">
<div class="container">
<div class="row">
<div class="col-md-8 col-lg-8 col-md-offset-2 col-lg-offset-2 col-sm-
12 col-xs-12">
<div class="text-center wow fadeIn">
<h2 class="xs-font20">About Us</h2>
<p></p></div>
</div></div>
</div></section>
```

For the Prediction page:

```html
!DOCTYPE html>
<html>
    <head>
        <title>Prediction</title>
        <link rel="stylesheet" href="css/style1.css" type="text/css">
    </head>
    <body class="text-center">

     <div class="simple-form">
         <h1>PREDICTION FORM</h1>
```

The input fields for the prediction page:

```html
<form id="registration">
<input type="text" name="fname" id="button" placeholder="Enter Your Fir
st Name" required> <br><br>
<input type="text" name="lname" id="button" placeholder="Enter Your Las
t Name" required> <br><br>
<select id="ph">
<option disabled="disabled" selected="selected" id="ph">Choose Your Gen
der....</option>
```

For designing the placed page:



For designing the not placed page:

## BACK-END

For front-end we have used the frame-work Python.
Importing the Libraries:

```python
In [2]: import pandas as pd
        import numpy as np
        import sidetable
        import joblib
```

Reading the datasets:

```python
In [3]: d1 = pd.read_csv("Data.csv")

In [4]: d1
```

Out[4]:

| | gender | tenth_marks | pu_marks | pu_stream | deg_marks | deg_stream | workex | status |
|---|---|---|---|---|---|---|---|---|
| 0 | M | 58.0 | 70.00 | Commerce | 61.00 | Comm&Mgmt | No | Not Placed |
| 1 | M | 58.0 | 61.00 | Commerce | 60.00 | Comm&Mgmt | Yes | Placed |
| 2 | M | 69.6 | 68.40 | Commerce | 78.30 | Comm&Mgmt | Yes | Placed |
| 3 | F | 47.0 | 55.00 | Science | 65.00 | Comm&Mgmt | No | Not Placed |
| 4 | F | 77.0 | 87.00 | Commerce | 59.00 | Comm&Mgmt | No | Placed |
| ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 327 | M | 81.1 | 78.60 | Arts | 35.84 | Sci&Tech | No | Not Placed |
| 328 | M | 81.3 | 85.12 | Arts | 68.50 | Sci&Tech | No | Not Placed |
| 329 | M | 83.5 | 89.90 | Science | 76.30 | Sci&Tech | No | Not Placed |
| 330 | M | 85.5 | 90.63 | Science | 74.50 | Sci&Tech | No | Not Placed |
| 331 | M | 80.0 | 83.33 | Science | 65.53 | Sci&Tech | No | Not Placed |

332 rows × 8 columns

The model which we used for the campus recrfuitment is Logistics Regression:

```python
In [15]: from sklearn.metrics import mean_absolute_error
         from sklearn.linear_model import LogisticRegression
         lr_model = LogisticRegression(max_iter=100)
         from sklearn.model_selection import train_test_split
         X = d1.drop('status', axis='columns')
         y = d1['status']

In [16]: X_train, X_test, y_train, y_test = train_test_split(X,y,test_size = 0.2, random_state=10)
```

```
In [17]: lr_model.fit(X_train,y_train)

C:\ProgramData\Anaconda3\lib\site-packages\sklearn\linear_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge
(status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regression
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Out[17]: LogisticRegression(C=1.0, class_weight=None, dual=False, fit_intercept=True,
                   intercept_scaling=1, l1_ratio=None, max_iter=100,
                   multi_class='auto', n_jobs=None, penalty='l2',
                   random_state=None, solver='lbfgs', tol=0.0001, verbose=0,
                   warm_start=False)
```

The training accuracy in Logistic Regression:

```
In [18]: y_pred = lr_model.predict(X_train)
         from sklearn.metrics import accuracy_score
         accuracy_score(y_train, y_pred)

Out[18]: 0.8339622641509434
```

The testining accuracy in Logistic Regression:

```
In [19]: y_pred = lr_model.predict(X_test)
         accuracy_score(y_test, y_pred)

Out[19]: 0.7910447761194029
```

## FLASK CONNECTIVITY

```
import numpy as np
from flask import Flask, request, render_template
import joblib

app = Flask(__name__)

@app.route('/')
def home():
```

```python
    return render_template('prediction.html')

@app.route('/predict')
def predict():
    model = joblib.load('C:\\Users\\Sanjay\\OneDrive\\Documents\\app\\env\\model.joblib')
    if request.args.get('ph') == 'Male':
        temp=1
    else:
        temp=0

    sslc=float(request.args.get('sslc'))

    if request.args.get('stream') == 'Commerce':
        temp1=1
    elif request.args.get('stream') == 'Science':
        temp1=2
    else:
        temp1=0
    pu=float(request.args.get('pu'))
    if request.args.get('deg') == 'Comm and Mgmt':
        temp2=0
    elif request.args.get('deg') == 'science and tech':
        temp2=2
    else:
        temp2=1
    degree=float(request.args.get('degree'))
    if request.args.get('wex') == 'Yes':
        temp3=1
    else:
        temp3=0

    data = [temp,sslc,pu,temp1,degree,temp3,temp2]
    print(data)
    result = model.predict(np.array(data).reshape(1,-1))
    if result[0] == 1:
        resFile = open('C:\\Users\\Sanjay\\OneDrive\\Documents\\app\\env\\congrats.html').read()
    else:
        resFile = open('C:\\Users\\Sanjay\\OneDrive\\Documents\\app\\env\\sad.html').read()

    return resFile
```
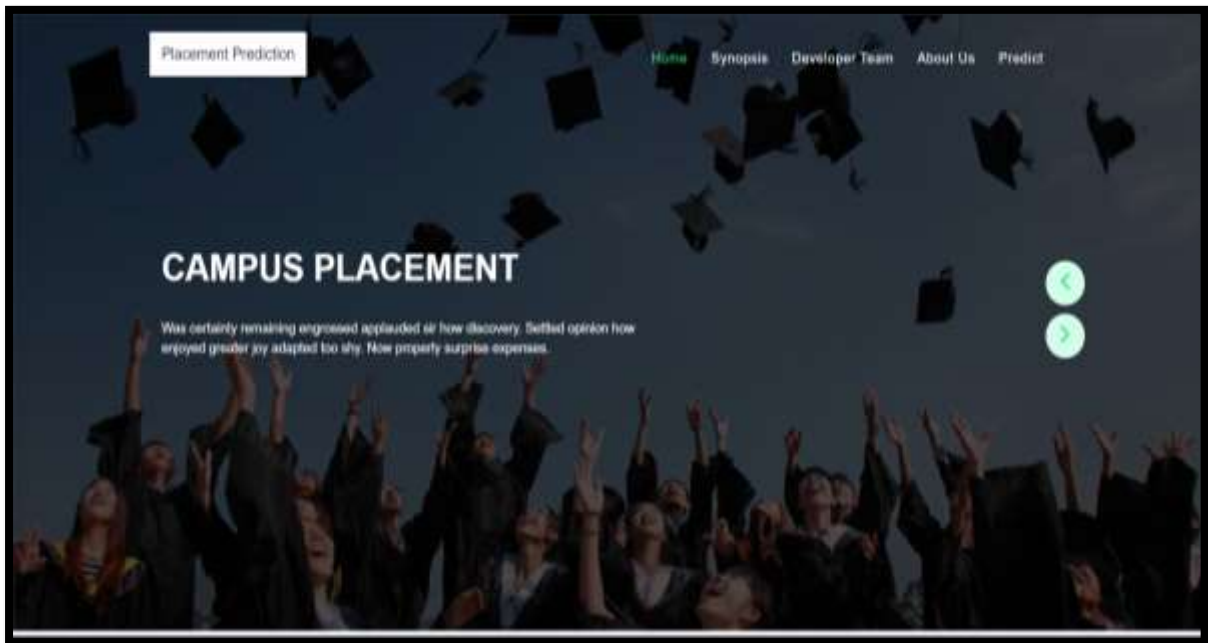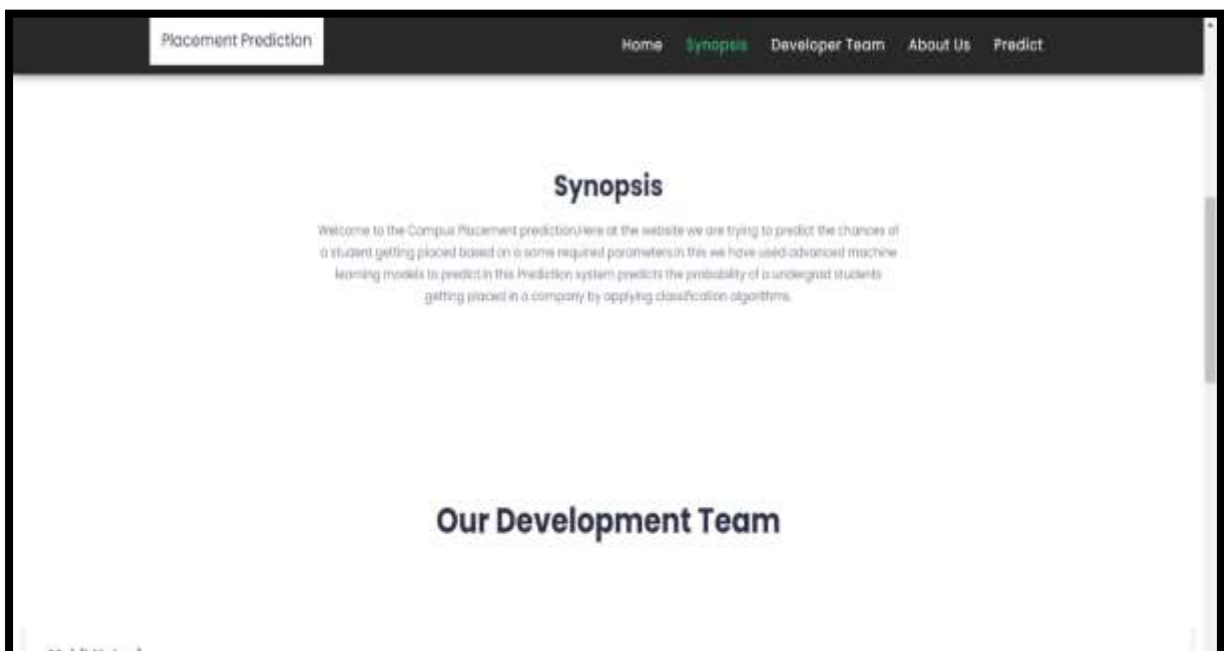
# 6.WEB DESIGN SNAPSHOTS
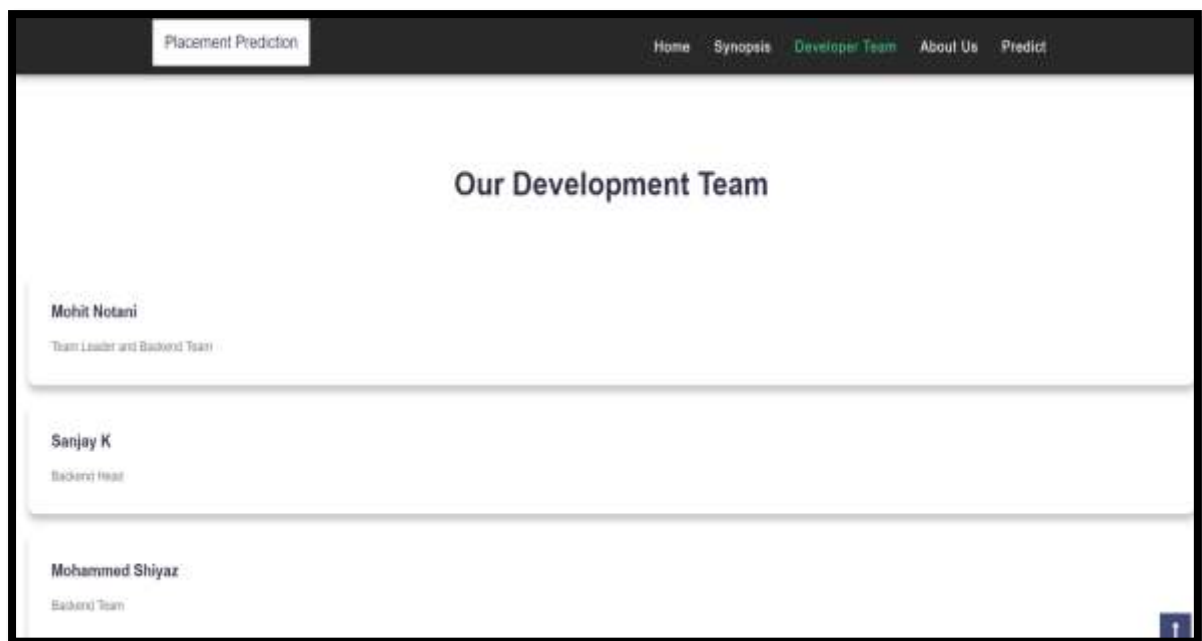


**Fig 6.1 HOME PAGE**
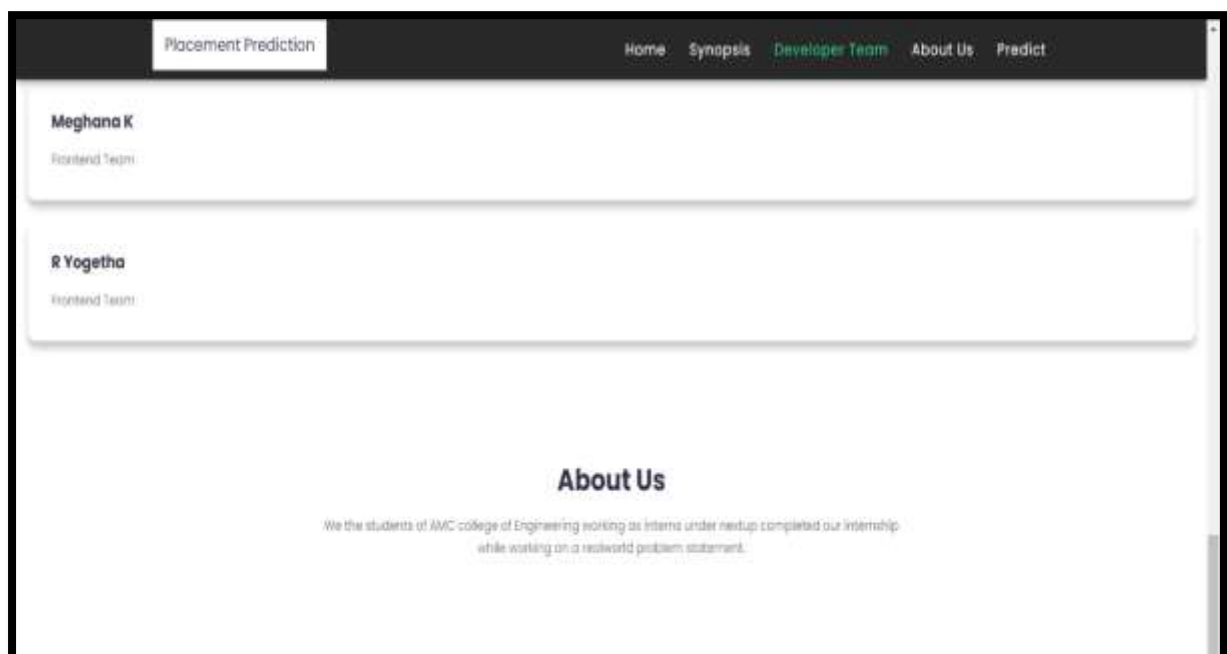


**Fig 6.2 SYNOPSIS**

**Fig 6.3 DEVELOPER TEAM**



**Fig 6.4 ABOUT US**

**Fig 6.5 PREDICTION FORM**

# 7.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 algorithm logistic regression algorithm to validate the approaches. The algorithms are applied on the data set and attributes used to build the model. The accuracy obtained after analysis for logistic regression is 80%.

# 8.REFERENCES

1. https://www.kaggle.com/benroshan/factors-affecting-campus-placement

2. https://www.researchgate.net/publication/328629392_student's_campus_placement_data

3. https://guides.github.com

4. https://guides.github.com/introduction/flow/

5. https://guides.github.com/features/pages/

6. https://scikit-learn.org/stable/