

# **IDENTIFYING PATTERNS AND TRENDS IN CAMPUS PLACEMENT DATA USING MACHINE LEARNING**

**TEAM ID : NM2023TMID35515**

**TEAM SIZE : 4**

**TEAM LEADER : S. PRSANTH**

**TEAM MEMBERS : RAJA .K**

**SANTHOSH . S**

**SARAVANAN . R**

# 1. INTRODUCTION

## 1.1. OVERVIEW

The goal of this project is to use machine learning techniques to analyze campus placement data and identify patterns and trends. Campus placements are an important part of a student's education, and this project aims to identify key factors that can impact placement outcomes.

To achieve this goal, the project involves several steps. The first step is to collect the necessary data from different sources, such as the placement cell of the institute, alumni surveys, or publicly available data. The data is then preprocessed, which includes cleaning, transforming, and normalizing it.

Once the data is preprocessed, the next step is to explore it to gain insights. Data exploration involves using descriptive statistics and data visualization techniques to understand the distribution of the data, identify any outliers, and discover any relationships between the variables.

Based on the insights gained from exploring the data, a suitable machine learning algorithm is selected. The selection of the algorithm is dependent on the nature of the data and the research question. For example, classification algorithms can be used to predict whether a student will get placed or not, whereas regression algorithms can be used to predict the salary offered to the placed students.

The selected algorithm is then trained and tested using the data. The performance of the model is evaluated, and any necessary adjustments are made to improve its accuracy.

Finally, the results of the model are interpreted, and conclusions are drawn based on the patterns and trends identified in the campus placement data. These insights can be used to inform decision-making and improve the outcomes of campus placements.

Some of the factors that can be analyzed in this project include academic performance, work experience, location preferences, salary expectations, and industry trends. By identifying the factors that impact placement outcomes, the project can help students, colleges, and companies make better decisions about placements.

In addition, the project can provide insights into the effectiveness of the placement process, and identify areas where improvements can be made. For example, if the model identifies that students with internships are more likely to get placed, the institute can focus on providing more internship opportunities to its students.

Overall, this project is a valuable exercise in using machine learning to gain insights into campus placement data. By identifying patterns and trends, the project can help improve the outcomes of campus placements

and provide valuable information to stakeholders in the education and employment sectors.

## **1.2. PURPOSE**

The purpose of this project is to gain insights from campus placement data using machine learning techniques. By identifying patterns and trends in the data, the project can provide valuable information to stakeholders in the education and employment sectors.

One potential application of this project is to improve the outcomes of campus placements. By understanding the factors that impact placement outcomes, colleges and companies can make better decisions about placements, which can ultimately benefit the students. For example, if the model identifies that certain skills or experiences are more likely to lead to successful placements, colleges can focus on developing those skills in their students.

In addition, the insights gained from this project can help colleges and companies better understand the job market and the needs of employers. This can help them tailor their curriculums and programs to better align with the demands of the job market, which can benefit both the students and the companies.

Another potential application of this project is to identify areas where improvements can be made in the placement process. By analyzing the data, the project can identify any inefficiencies or areas where the placement

process can be improved. For example, if the model identifies that certain locations or industries have low placement rates, the placement cell can focus on increasing the number of opportunities in those areas.

Overall, the use of this project is to gain insights from campus placement data that can inform decision-making and lead to better outcomes for students, colleges, and companies.

## **2. PROBLEM DEFINITION AND DESIGN THINKING**

Problem definition and design thinking are critical components of any project, including the project on identifying patterns and trends in campus placement data using machine learning.

The problem definition phase involves clearly defining the problem that the project aims to solve. In the case of this project, the problem is the lack of understanding of the factors that impact campus placement outcomes. This lack of understanding can lead to inefficiencies in the placement process and suboptimal outcomes for students.

Design thinking, on the other hand, is a problem-solving approach that involves empathizing with the end-users, defining the problem, ideating possible solutions, prototyping, and testing. It is a user-centered approach that prioritizes the needs and experiences of the end-users.

In the context of this project, design thinking can be applied to gain a better understanding of the needs and experiences of the end-users, such as the students, colleges, and companies involved in the placement process. For example, design thinking can be used to understand the challenges faced by students in securing placements, the criteria used by companies to select candidates, and the processes followed by colleges to facilitate placements.

Once the problem is clearly defined and the needs of the end-users are understood, design thinking can be used to ideate and prototype possible

solutions. For example, the project can explore different machine learning algorithms and techniques to identify the factors that impact campus placement outcomes. Prototypes can be built to test and refine the solutions, and feedback can be gathered from the end-users to further improve the project.

In summary, problem definition and design thinking are critical components of the project on identifying patterns and trends in campus placement data using machine learning. By clearly defining the problem and using design thinking to empathize with the end-users, ideate and prototype solutions, the project can be more effective in addressing the needs of the stakeholders and achieving its goals.

### 3. ADVANTAGES AND DISADVANTAGES

#### 3.1. Advantages:

**Accurate predictions:** Machine learning models can analyze large volumes of placement data to identify patterns and trends accurately. This can help students, colleges, and companies to make informed decisions about the placement process.

**Personalization:** Machine learning models can provide personalized recommendations to students based on their skills, interests, and placement history. This can improve their chances of securing a suitable job.

**Efficient:** Machine learning models can automate various aspects of the placement process, such as resume screening and candidate selection, making the process faster and more efficient.

**Cost-effective:** Machine learning models can reduce the cost of recruitment for companies by streamlining the hiring process and identifying the best candidates quickly.

**Data-driven:** Machine learning models can use data to identify the most effective strategies for improving the placement process, such as skill enhancement programs and outreach programs.



### **3.2. Disadvantages:**

Data quality: The accuracy of machine learning models depends on the quality of data used to train them. If the data is incomplete or biased, the models may not produce accurate predictions.

Technical expertise: Developing and implementing machine learning models requires technical expertise and resources, which may be a challenge for some organizations.

Privacy concerns: Collecting and analyzing personal data can raise privacy concerns among students and may require strict data protection measures.

Lack of transparency: Machine learning models can be complex and opaque, making it difficult to understand how they arrive at their predictions. This can reduce trust and confidence in the model's predictions.

Overall, the advantages of using machine learning to analyze campus placement data outweigh the disadvantages. However, it is essential to address the potential challenges and concerns through appropriate data governance, transparency, and accountability measures.

## 4. APPLICATIONS

The solution of identifying patterns and trends in campus placement data using machine learning can be applied in various areas, including:

### **Higher education institutions:**

Colleges and universities can use this solution to optimize their placement process and improve the employability of their students. By analyzing placement data, colleges can identify the skills and qualifications that employers are looking for and develop targeted skill enhancement programs for their students.

### **Recruitment agencies:**

Recruitment agencies can use this solution to streamline their candidate selection process and identify the best candidates for specific job roles. By analyzing placement data, agencies can identify the characteristics and qualifications that lead to successful placements and use this knowledge to improve their recruitment process.

### **Companies:**

Companies can use this solution to improve their hiring process and reduce the cost of recruitment. By analyzing placement data, companies can identify the characteristics and qualifications that lead to successful placements

and use this knowledge to screen and select the most suitable candidates for their job roles.

**Government agencies:**

Government agencies can use this solution to monitor and improve the employability of their citizens. By analyzing placement data, governments can identify the skills and qualifications that are in demand and develop targeted skill enhancement programs to improve the employability of their citizens.

**Career counseling services:**

Career counseling services can use this solution to provide personalized guidance to their clients based on their placement history and qualifications. By analyzing placement data, career counselors can identify the most suitable job roles and industries for their clients and provide them with targeted advice on how to improve their employability.

In summary, the solution of identifying patterns and trends in campus placement data using machine learning has a wide range of applications in various industries and sectors. It can be used to optimize the placement process, improve employability, reduce recruitment costs, and provide personalized guidance to job seekers.

## 5. APPENDIX

### SOURCE CODE

```
from flask import Flask, render_template, request
import pickle
import numpy as np
app = Flask(__name__)
model = pickle.load(open("rdf.pkl", 'rb'))
@app.route('/')
def hello():
    return render_template("index.html")
@app.route('/guest', methods = ["POST"])
def Guest():
    return render_template("secondpage.html")
@app.route('/y_predict', methods=["POST"])
def y_predict():
    if request.method == "POST":
        sen1 = request.form["sen1"]
        sen2 = request.form["sen2"]
        sen3 = request.form["sen3"]
        sen4 = request.form["sen4"]
        sen5 = request.form["sen5"]
        sen6 = request.form["sen6"]
        X_test = np.array([[sen1, sen2, sen3, sen4, sen5, sen6]], dtype=float)
        prediction = model.predict(X_test)
        prediction = prediction[0]
        return render_template("secondpage.html", y=prediction)
    else:
        return "Invalid request method"
if __name__ == '__main__':
    app.run(debug=True)
```

Naamudhavan - Smartintenz x Identifying-Patternsand-Trends x +

github.com/iamharisharath/Identifying-Patternsand-Trends-in-Campus-Placement-Data-using-Machine-Learning/blob/main/collegePlace.csv

Code

main

Go to file

Search this file

iamharisharath Add files via upload

2967 lines (2967 loc) · 384 KB

Raw

	Age	Gender	Stream	Internships	CGPA	Hostel	HistoryOfBacklogs	PlacedOrNot
1	22	Male	Electronics And Communication	1	8	1	1	1
2	21	Female	Computer Science	0	7	1	1	1
3	22	Female	Information Technology	1	6	0	0	1
4	21	Male	Information Technology	0	8	0	1	1
5	22	Male	Mechanical	0	8	1	0	1
6	22	Male	Electronics And Communication	0	6	0	0	0
7	21	Male	Computer Science	0	7	0	1	0
8	21	Male	Information Technology	1	7	0	0	0
9	21	Male	Computer Science	2	6	0	0	1
10	21	Female	Computer Science	1	6	1	0	0
11	22	Male	Computer Science	1	7	0	0	0
12	22	Female	Electrical	1	8	0	1	1
13	21	Female	Computer Science	2	6	1	1	0
14	21	Male	Computer Science	1	8	0	1	1
15	21	Female	Electronics And Communication	2	8	0	0	1
16	22	Male	Mechanical	0	8	1	0	1
17	22	Female	Computer Science	1	7	1	0	0

download (1).png

download.png

app (1).py

app.py

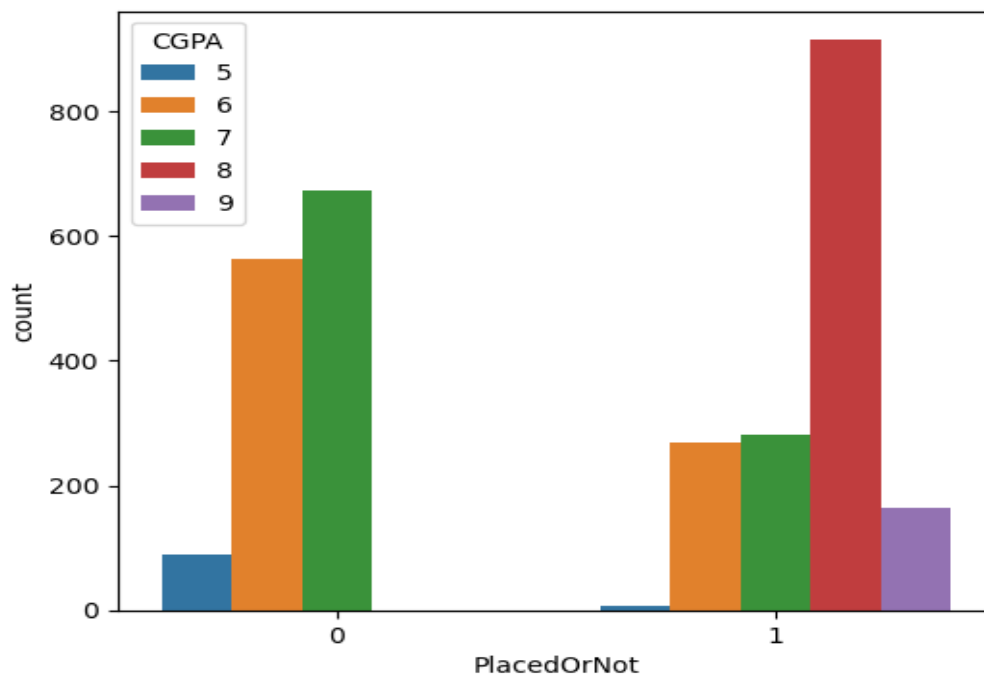
Show all

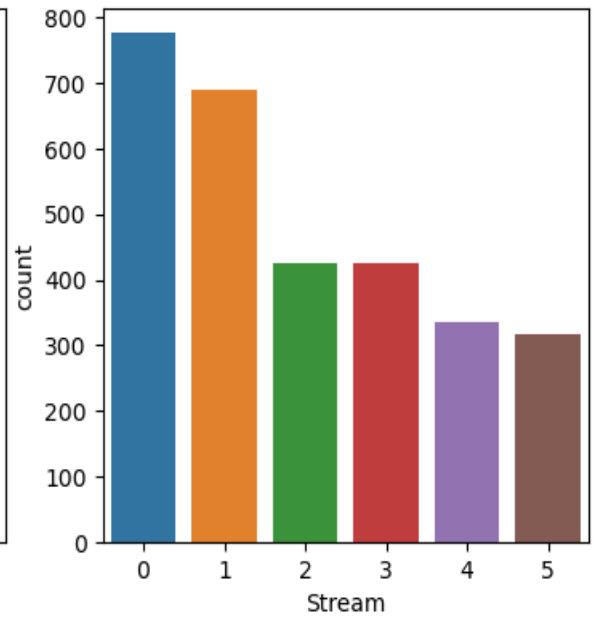
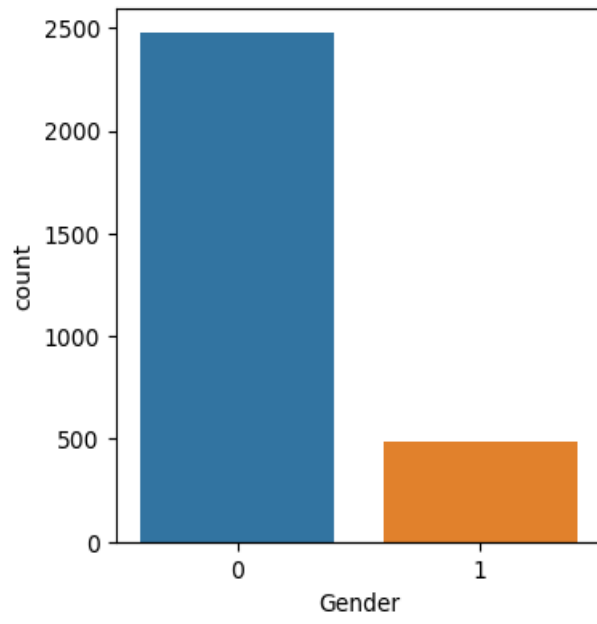
Type here to search

12:01

ENG

29-04-2023





## 6. RESULT

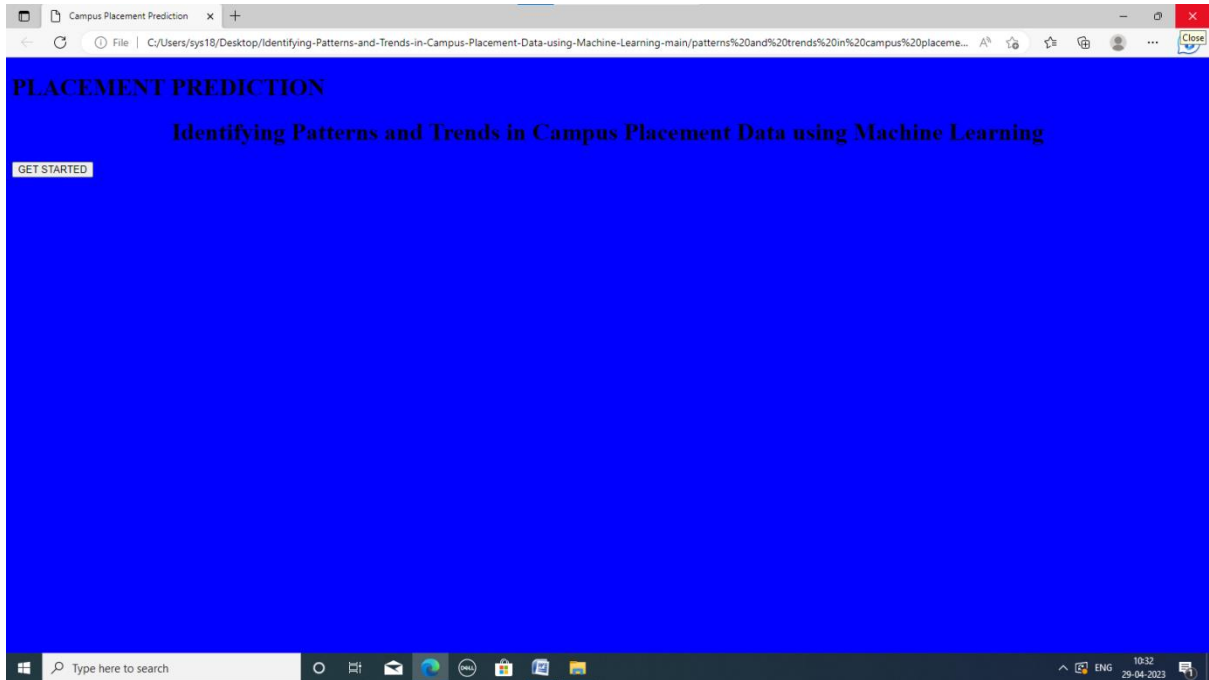
The data set used for is further spitted into two sets consisting of two third as training set and one third as testing set. Among the two algorithms applied random forest shown the best results. The efficiency of the two approaches is compared in terms of the accuracy.

The accuracy of the prediction model/classifier is defined as the total number of correctly predicted/classified instances. Accuracy is given by using following formula:

$$\text{Accuracy} = (\text{TP} + \text{TN} / \text{TP} + \text{FN} + \text{FP} + \text{TN}) * 100$$

where TP, TN, FN, FP represents the number of true positives, true negative, false negative and false positive cases.

# 1. CAMPUS PLACEMENT PREDICTION





## 2. FILE THE DETAILS:

Fill the details

Cy/Users/sys18/Desktop/Identifying-Patterns-and-Trends-in-Campus-Placement-Data-using-Machine-Learning-main/patterns%20and%20trends%20in%20campus%20placeme...

### FILL THE DETAILS

4

4

4

4

4

4

4

Submit

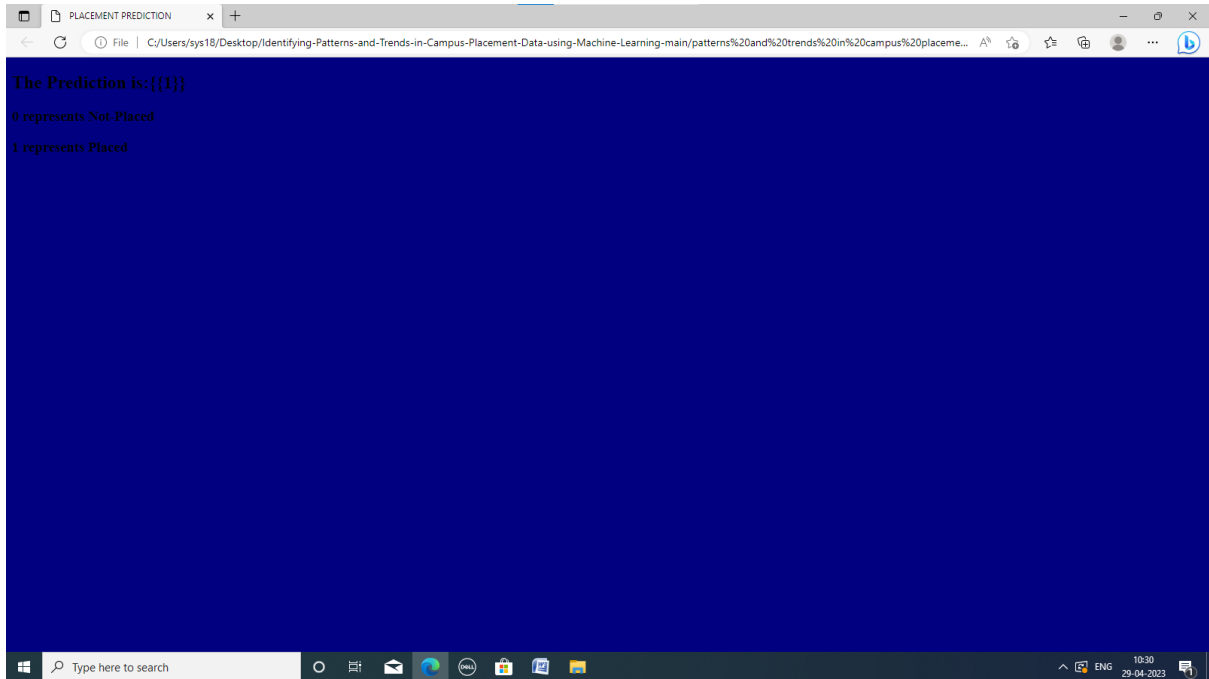
### 3. IDENTIFY THE DETAILS:

The screenshot shows a web browser window with a form titled "Fill the Details in Form". The form is centered on a green background. The form fields are as follows:

Field Label	Value
Age	18
Gender	1
Weight	6
Height	4
Blood Pressure	8
Number of Smoking	1

At the bottom of the form is a green button labeled "Submit".

## 4. PLACEMENT PREDICTION



## 7. CONCLUSION

In conclusion, identifying patterns and trends in campus placement data using machine learning is an effective solution that can provide valuable insights into the placement process. By analyzing placement data, colleges, universities, recruitment agencies, companies, government agencies, and career counseling services can identify the skills, qualifications, and characteristics that lead to successful placements and improve the employability of job seekers.

While there may be some challenges and concerns, such as data quality, technical expertise, privacy concerns, and lack of transparency, these can be addressed through appropriate data governance, transparency, and accountability measures.

Overall, the solution of identifying patterns and trends in campus placement data using machine learning has numerous advantages, such as accurate predictions, personalization, efficiency, cost-effectiveness, and data-driven decision making. Therefore, it is a promising approach that can transform the placement process and improve the employability of job seekers in various industries and sectors.

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 the Random forest 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 Decision tree is 84% and for the Random Forest is 86%. Hence, from the above said analysis and prediction it's better if the Random Forest algorithm is used to predict the placement results.

## **8. FUTURE SCOPE**

The future scope of the project of identifying patterns and trends in campus placement data using machine learning is significant. Some of the potential areas of growth and development include:

### **Integration with other systems:**

The project can be integrated with other systems such as learning management systems, student information systems, and applicant tracking systems. This integration can improve the accuracy of predictions and provide a more comprehensive understanding of the placement process.

### **Adoption by more institutions:**

As the benefits of the project become more apparent, more institutions, both public and private, are likely to adopt it. This adoption will lead to a larger pool of data and more accurate predictions.

### **Advancements in machine learning:**

The advancements in machine learning techniques, such as deep learning and natural language processing, can further enhance the accuracy and efficiency of the project.

### **Integration with job portals:**

The project can be integrated with job portals to provide personalized job recommendations to students based on their skills and qualifications. This

integration can improve the employability of students and reduce the time and cost of job searching.

### **Application in other industries:**

The project can be applied in other industries, such as healthcare, finance, and retail, to analyze employment data and provide insights into the job market.

In conclusion, the project of identifying patterns and trends in campus placement data using machine learning has a significant future scope. With the advancements in technology and the growing demand for data-driven decision making, this project is likely to become more prevalent and influential in the years to come.

## 9. REFERENCES

- [1]. Mangasuli Sheetal B, Prof. Savita Bakare “Prediction of Campus Placement Using Data Mining Algorithm Fuzzy logic and K nearest neighbour” International Journal of Advanced Research in Computer and Communication Engineering Vol. 5, Issue 6, June 2016 .
- [2]. Ajay Shiv Sharma, Swaraj Prince, Shubham Kapoor, Keshav Kumar “PPS-Placement prediction system using logistic regression” IEEE international conference on MOOC, innovation and Technology in Education (MITE), December 2014.
- [3]. Jai Ruby, Dr. K. David “Predicting the Performance of Students in Higher Education Using Data Mining Classification Algorithms - A Case Study” International Journal for Research in Applied Science & Engineering Technology (IJRASET) Vol. 2, Issue 11, November 2014.
- [4]. Ankita A Nichat, Dr. Anjali B Raut “Predicting and Analysis of Student Performance Using Decision Tree Technique” International Journal of Innovative Research in Computer and Communication Engineering Vol. 5, Issue 4, April 2017.
- [5]. Oktariani Nurul Pratiwi “Predicting Student Placement Class using Data Mining” IEEE International Conference 2013.
- [6]. Ajay Kumar Pal and Saurabh Pal, “Classification Model of Prediction for Placement of Students”, I. J. Modern Education and Computer Science, 2013, 11, 49-56
- [7]. Ravi Tiwari and Awadhesh Kumar Sharma, “A Data Mining Model to Improve Placement”, International Journal of Computer Applications