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

1.INDRODUCTION :

Campus recruitment is a strategy for sourcing, engaging and hiring young talent for internship and entry-level positions. Where it has various factors on candidates getting hired such as work experience,exam percentage etc., Finally it contains the status of recruitment and remuneration details.

1.1 OVERVIEW :

Identifying Patterns and Trends in Campus Placement Data using Machine Learning Project Description

Campus recruitment is a strategy for sourcing, engaging and hiring young talent for internship and entry-level positions. College recruiting is typically a tactic for medium- to large-sized companies with high-volume recruiting needs, but can range from small efforts (like working with university career centers to source potential candidates) to large-scale operations (like visiting a wide array of colleges and attending recruiting events throughout the spring and fall semester).Campus recruitment often involves working with university career services centers and attending career fairs to meet in-person with college students and recent graduates.Our solution revolves around the placement season of a Business School in India.

● Entered input is analyzed by the model which is integrated.

● Once model analyzes the input the prediction is showcased on the UI To accomplish this, we have to complete all the activities listed below,

● Data collection o Collect the dataset or create the dataset

● Visualizing and analyzing data o Univariate analysis o Bivariate analysis o Multivariate analysis o Descriptive analysis

● Data pre-processing o Checking for null values o Handling outlier o Handling categorical data o Splitting data into train and test

● Model building o Import the model building libraries o Initializing the model o Training and testing the model o Evaluating performance of model o Save the model

● Application Building o Create an HTML file o Build python code Project Flow: Create the Project folder which contains files as shown below

● We are building a flask application which needs HTML pages stored in the templates folder and a python script app.py for scripting ● rdf.pkl is our saved model.

● Training folder contains a model training file Milestone 1: Define Problem / Problem Understanding

● Activity 1: Specify the business problem Refer Project Description

● Activity 2: Business requirements The business requirements for a project aimed at "Identifying Patterns and Trends in Campus Placement Data using Machine Learning" would likely include the following: ❖Access to campus placement data: The project would require access to data on student performance, qualifications, and job placement outcomes.

❖Machine learning expertise: The project would require individuals with expertise in machine learning, data science and statistical analysis to develop and implement the algorithms and models needed to analyze the data.

❖Data storage and management: The project would require a robust and secure data storage and management system to store and organize the large amounts of data used in the analysis.

❖Infrastructure for model deployment: The project would require infrastructure for deploying the models and algorithms developed, including hardware, software, and cloud-based resources.

● Activity 3: Literature Survey (Student Will Write)

❖There have been several studies that have used machine learning techniques to identify patterns and trends in campus placement data. R. Suresh, published in the International Journal of Computer Science and Mobile Computing in 2015, used k-means clustering and decision trees to analyze campus placement data and identify patterns that could be used to predict placement outcomes.

❖Another study by authors V.V. Kulkarni and K.S. Patil, published in the International Journal of Engineering Research and Technology in 2012, used decision tree and neural network algorithms to analyze campus placement data and identify factors that influence student placement. ❖A study by authors S.S. Bhosale, S.S. Raut, and D.S. Kulkarni, published in the International Journal of Emerging Research in Management & Technology in 2013, used decision tree and Naive Bayes algorithms to analyze campus placement data and predict student placement outcomes.

❖In general, these studies found that machine learning techniques were effective at identifying patterns and trends in campus placement data, and could be used to predict student placement outcomes with high accuracy.

❖The business impact of a project that uses machine learning to identify patterns and trends in campus placement data could be significant. By analyzing data on factors such as student performance, qualifications, and job placement outcomes, the project could help organizations make more informed decisions about recruiting and hiring new graduates.

● Activity 1: Collect the dataset There are many popular open sources for collecting the data. This data is downloaded from kaggle.com. Link: https://www.kaggle.com/code/neesham/prediction-of-placements/data

● Activity 1.1: Importing the libraries

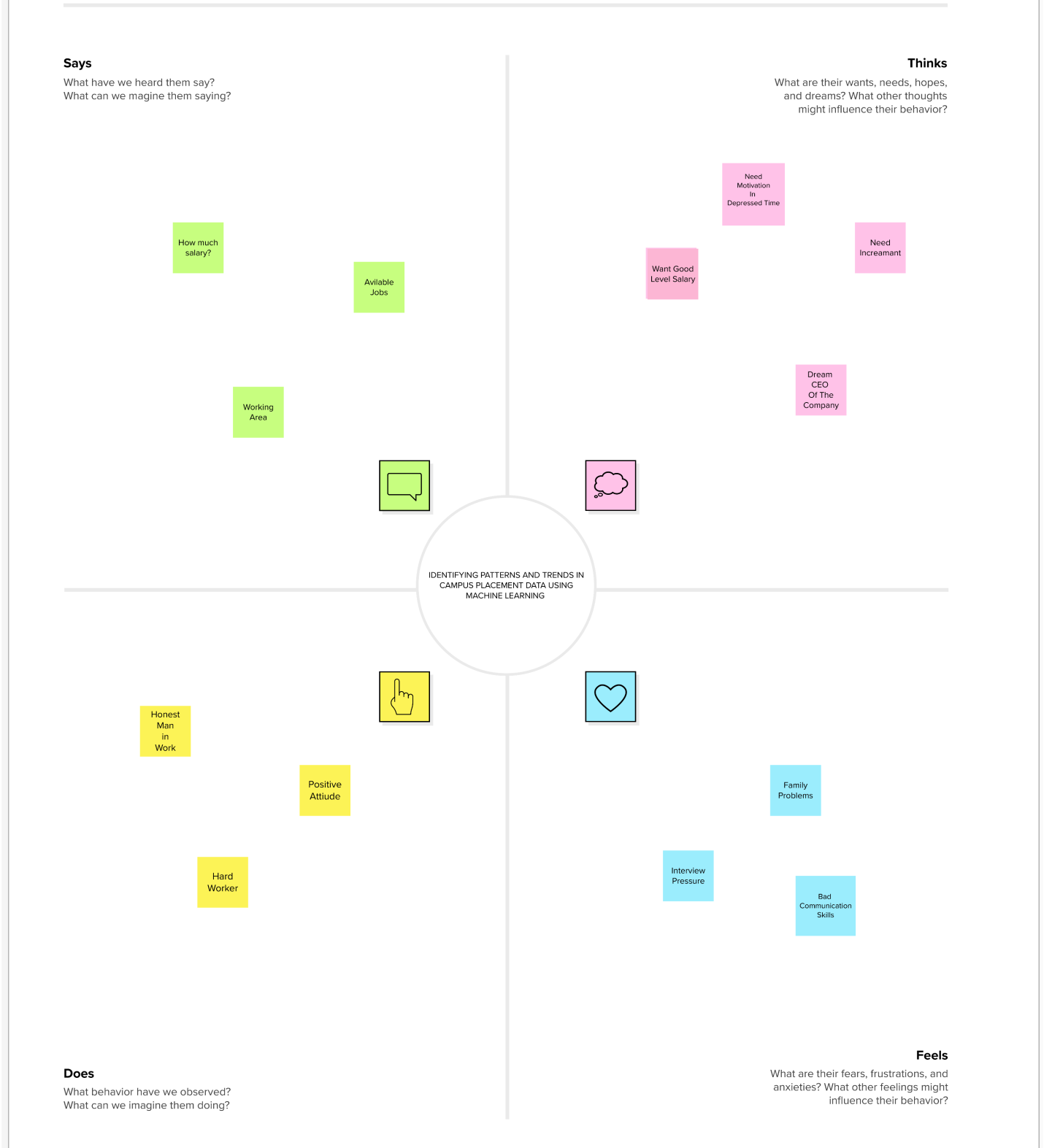
● Activity 1.2: Read the Dataset Our dataset format might be in .csv, excel files, .txt, .json, etc.

1.2 PURPOSE :

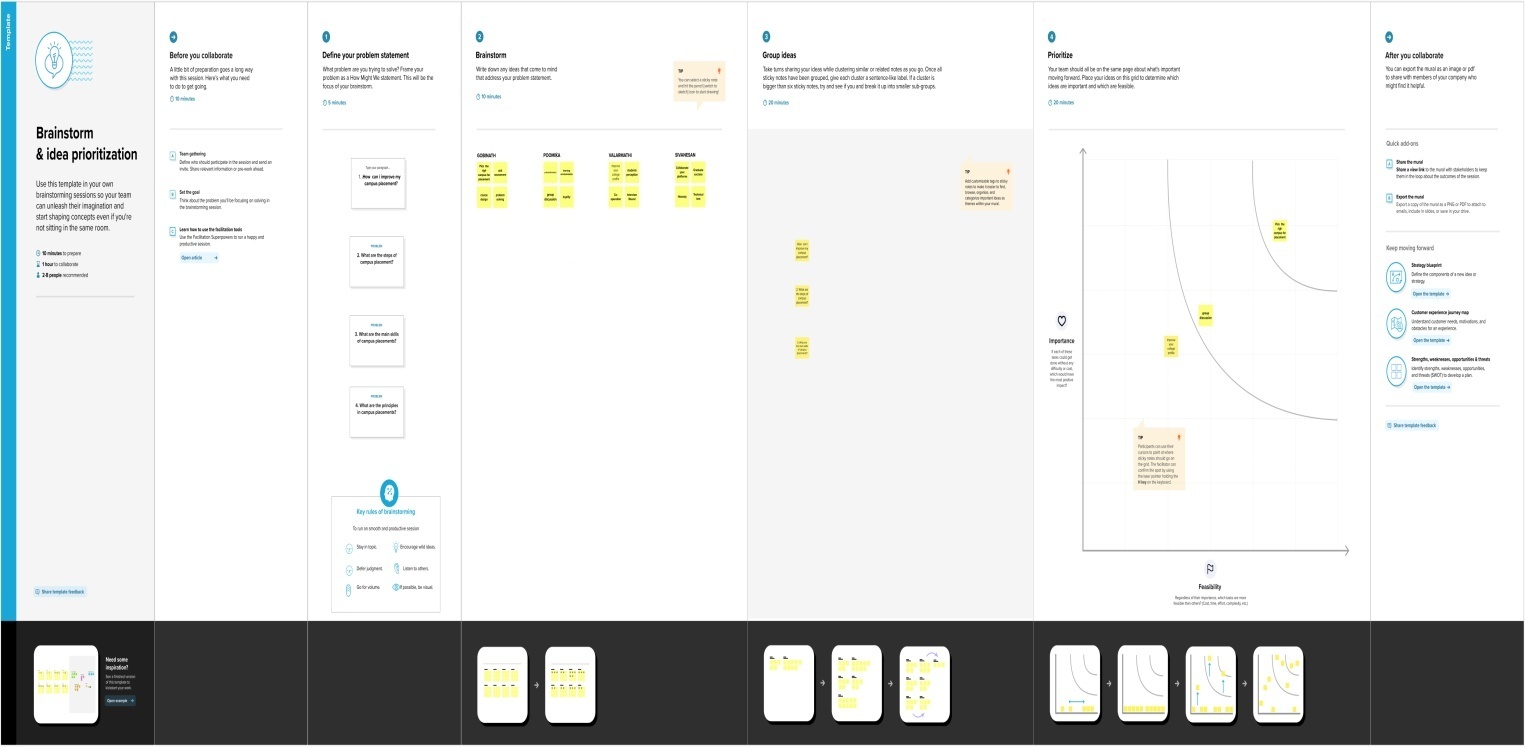
* The purpose of campus placement is so simple to provide you a start of your career and a safe future. There are two types of placements one is on campus placement and the other one is off campus placement.
* If you are going to take an on campus placement that means you will get a job so easily and you can think more then wherever you want to go. The recruiter companies which come to your college know that they are going to recruit fresher students and they don't expect so much from that students.
* On the other hand , if you are going for off campus recruitment then you can feel that there is an increment in the expectations of the recruiting company.And it will be a little bit difficult or you can say you have to put some extra efforts in getting yourself placed in a reputed organization.
* At the end I would suggest you to start working on yourself as soon as possible to make yourself more capable and skillful so that there is no need to think about the outer world’s competition.Once you have a bunch of skills ,know one can replace you.

2.PROBLEM DEFINITION & DESIGN THINKING :

2.1 EMPATHY MAP

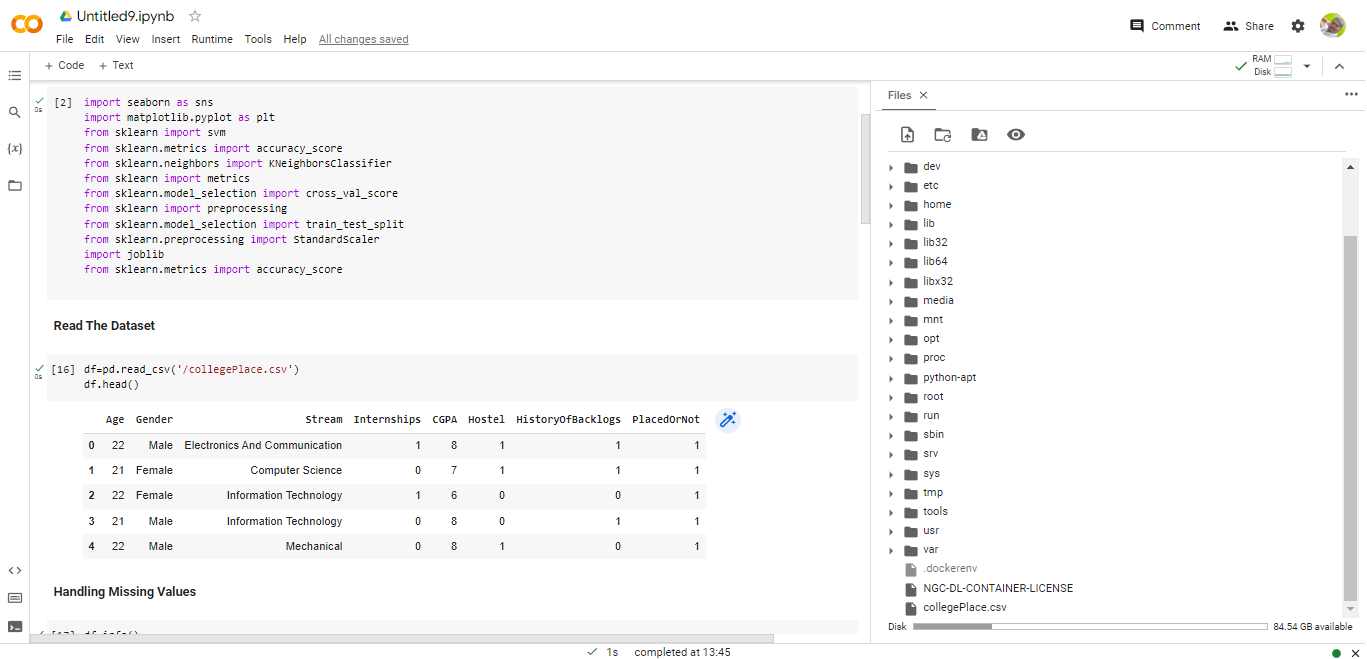


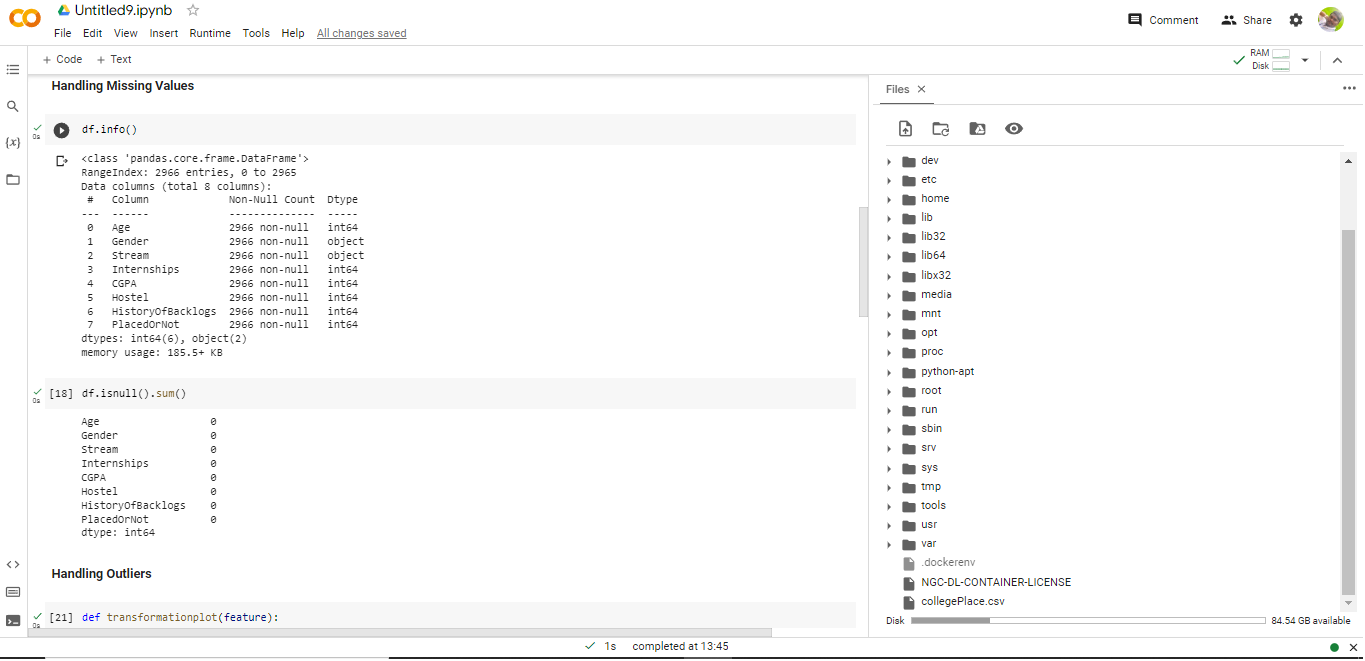
2.2 IDEATION & BRAINSTROMING MAP :

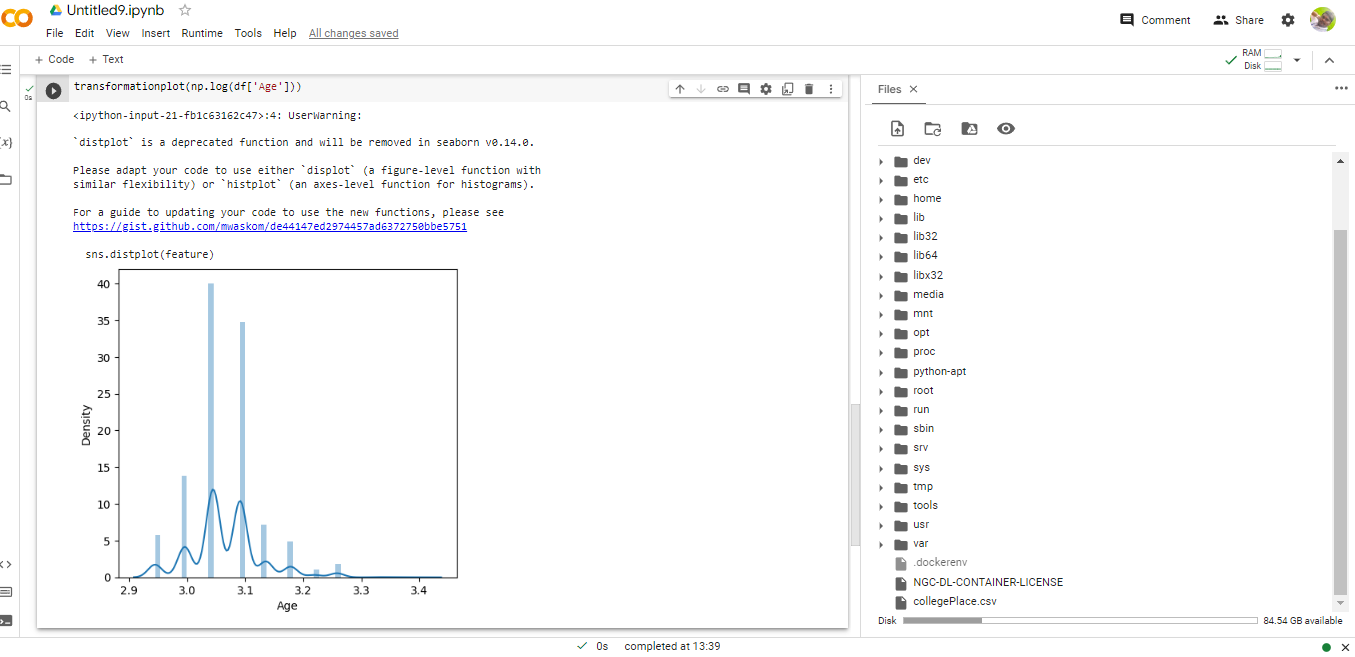


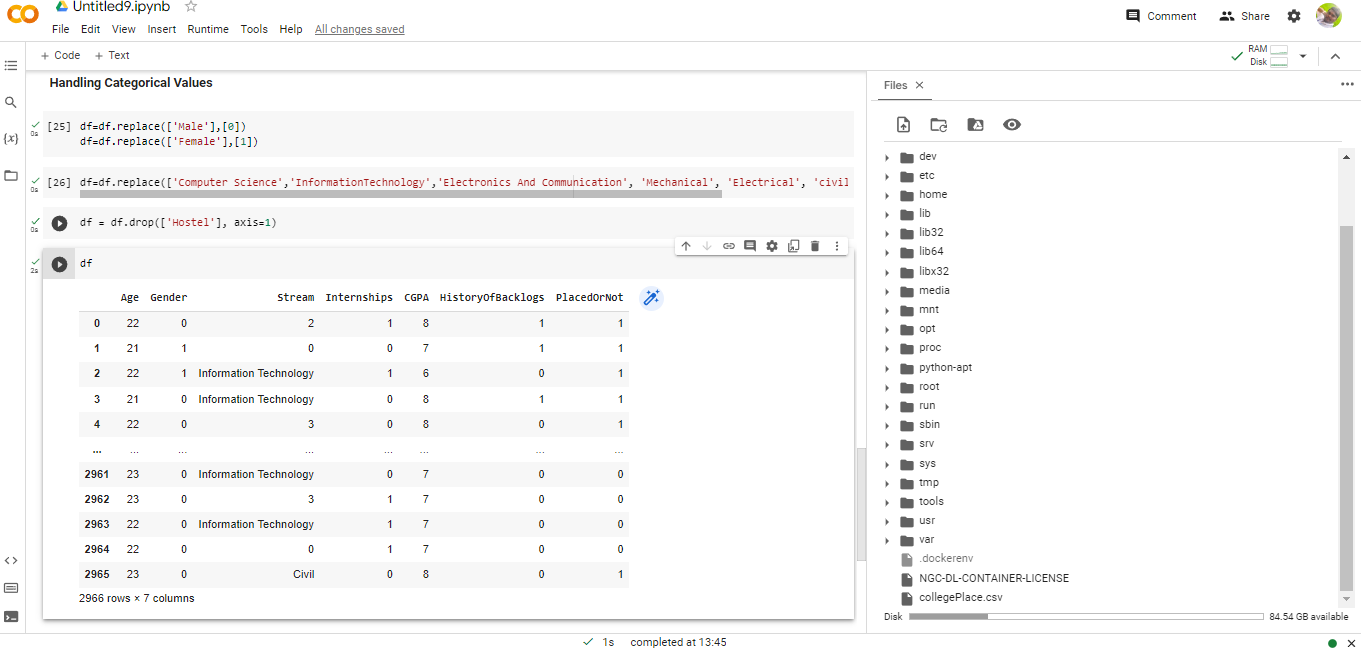
3.RESULT :

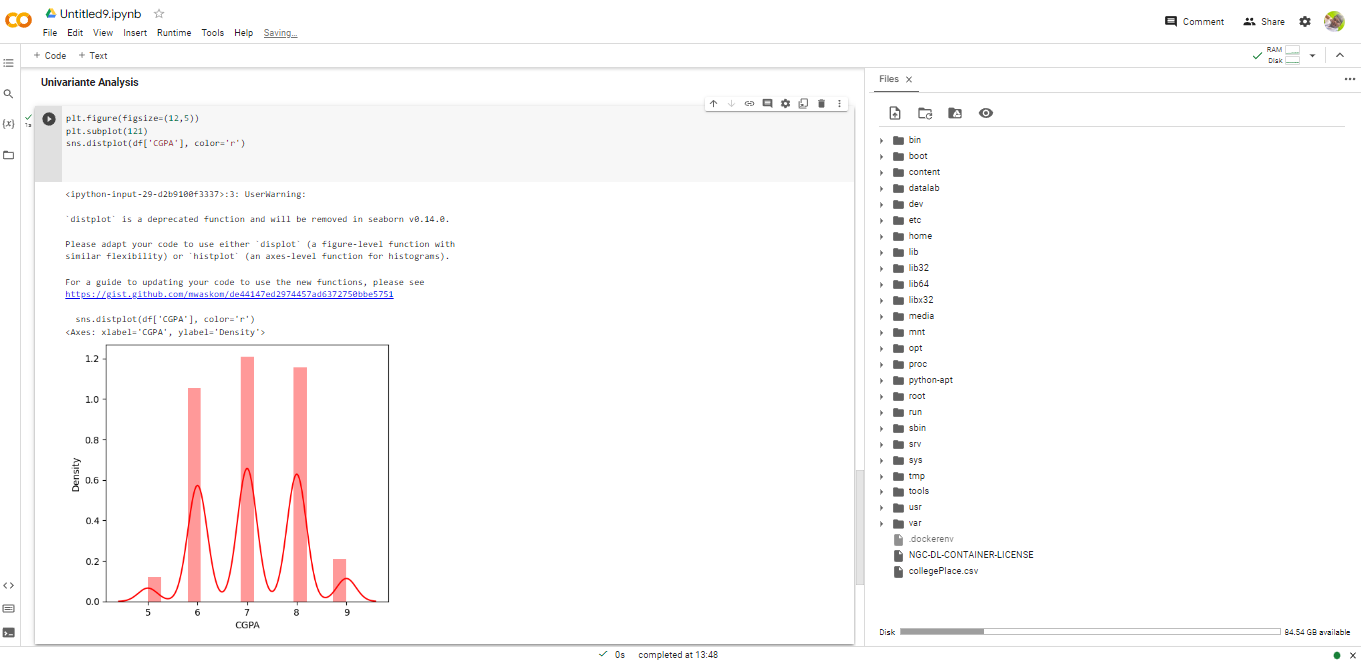
3.1 COLAB RESULT:

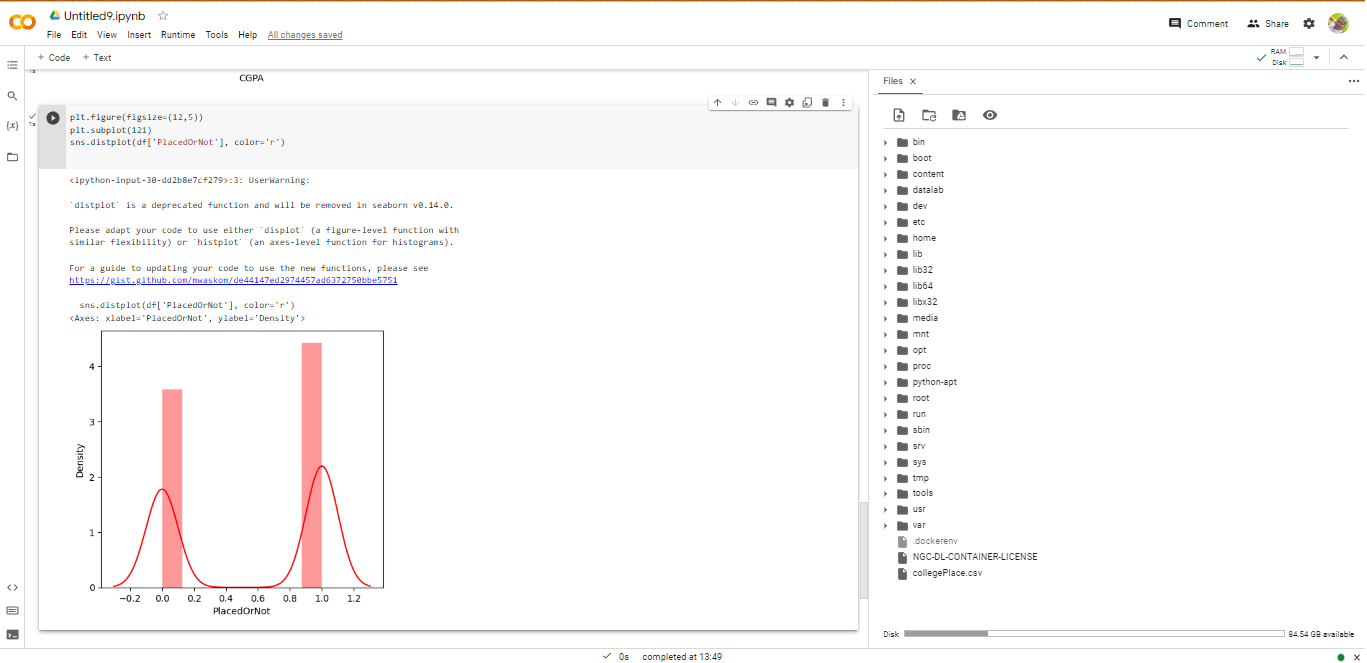




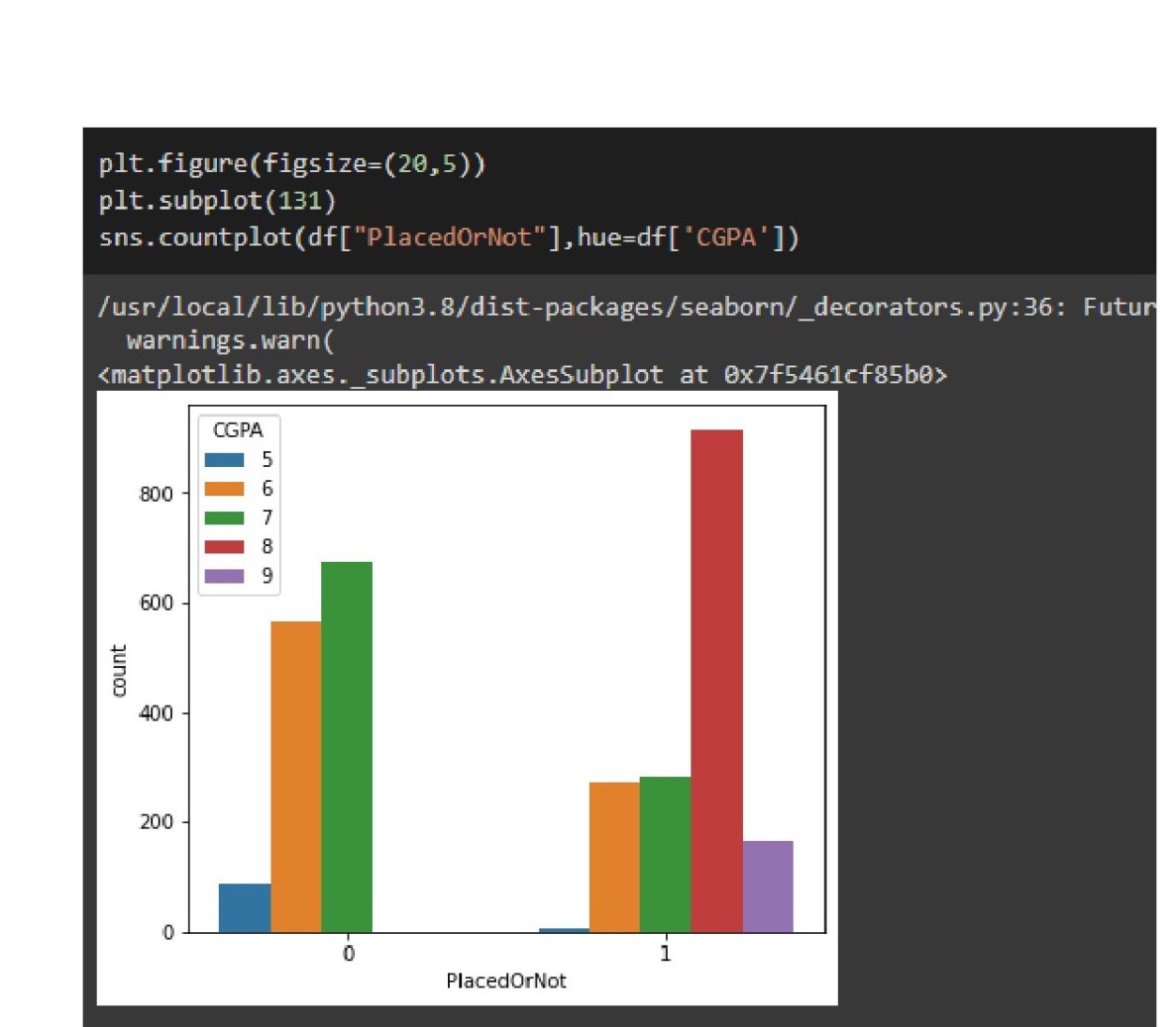


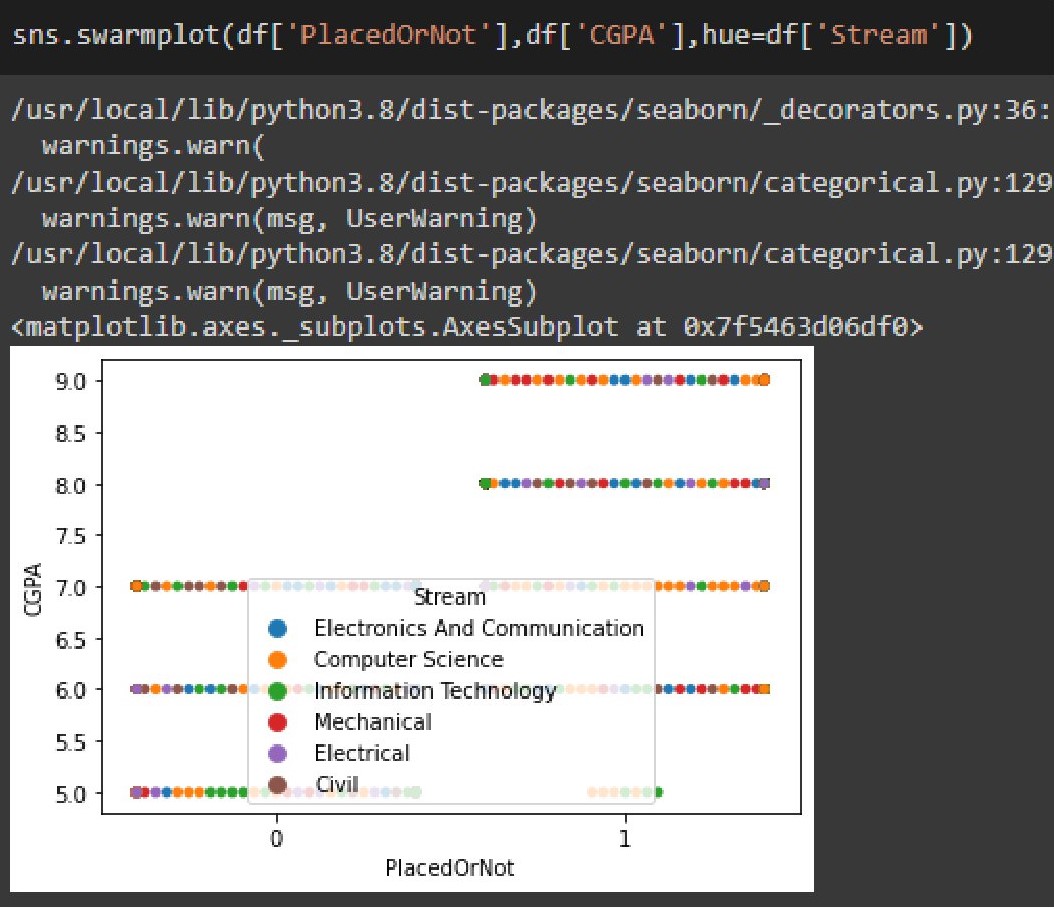












3.2 WEBSITE RESULT :



4.ADVANTAGES & DISADVANTAGES :

* Advantages of campus placement
* Saves Time & Efforts.
* Improved Retention Rates.
* Getting New Knowledge & Skills.
* Quick Learners & Multi-tasking candidates.
* Good relationship between Organization & Campus.
* High Volume of Talent Pool.
* Resumes are the only way to select a candidate.
* Limited Staff & Time.
* Disadvantage of campus placement
* High Volume of Talent Pool
* Resumes are the only way to select a candidate
* Limited Staff & Time
* Employer Branding to attract the younger generation
* Outdated recruiting practices

5.APPLICATIONS:

Campus placement is a process in which companies visit educational institutions such as colleges and universities to hire students for employment. The main objective of campus placement is to provide students with job opportunities and to fulfill the staffing needs of companies.

Here are some applications of campus placement:

* Job Opportunities for Students: Campus placement provides job opportunities for students, who are about to graduate or have recently graduated from college. By participating in campus placement drives, students can explore various career options and secure jobs in companies that match their skills and interests.
* Hiring Skilled Workforce: Campus placement allows companies to hire fresh talent from educational institutions. This helps companies to recruit skilled employees who have the necessary academic knowledge and technical skills required for the job.
* Cost-Effective Recruitment: Campus placement can be a cost-effective way for companies to recruit new employees as it eliminates the need for expensive job postings, recruiting agencies, and advertising. Also, companies can save time and resources by conducting interviews and recruitment processes on the campus itself.
* Building Relationships with Educational Institutions: Campus placement can help companies build a strong relationship with educational institutions. By recruiting students from these institutions, companies can establish a positive image and create goodwill among students and faculty.
* Improving Employer Branding: Campus placement can also be used as a tool for employer branding. By participating in campus placement drives, companies can showcase their work culture, values, and growth opportunities, and attract top talent to their organization.

6.CONCLUSION:

● Training folder contains a model training file­­­

Milestone 1: Define Problem / Problem Understanding

● Activity 1: Specify the business problem Refer Project Description

● Activity 2: Business requirements The business requirements for a project aimed at "Identifying Patterns and Trends in Campus Placement Data using Machine Learning" would likely include the following:

❖Access to campus placement data: The project would require access to data on student performance, qualifications, and job placement outcomes.

● Activity 3: Literature Survey (Student Will Write)

❖There have been several studies that have used machine learning techniques to identify patterns and trends in campus placement data.

­

7.FUTURE SCOPE :

The future scope for campus placement is quite promising as organizations continue to rely on universities and colleges to source their entry-level talent. Here are some potential trends and developments to watch for in the future of campus placement:

* Virtual Recruitment: With the pandemic accelerating the adoption of virtual recruitment, it is likely that virtual campus placement will become a norm. This would mean that students from different geographical locations would have equal opportunities to participate in campus recruitment processes, making it easier for organizations to source talent from diverse backgrounds.
* Emphasis on Soft Skills: While technical skills will always be important, organizations are increasingly prioritizing soft skills such as communication, adaptability, and teamwork when hiring for entry-level positions. This trend is expected to continue, and students who have honed their soft skills through extracurricular activities or internships may have an advantage in the recruitment process.
* Increased Automation: As organizations adopt new technologies to streamline their recruitment processes, it is likely that automation will become more prevalent in campus placement. This could mean that resume screening, pre-employment assessments, and even initial interviews are conducted using artificial intelligence or other automated tools.
* Focus on Diversity and Inclusion: With more organizations recognizing the importance of diversity and inclusion in their workplaces, it is likely that campus placement processes will also be designed with this in mind. This could mean that organizations partner with universities and colleges that have a diverse student body, or that they implement targeted outreach efforts to students from underrepresented backgrounds.

Overall, campus placement is likely to continue to be an important source of entry-level talent for organizations in the future, and students who are proactive in developing their skills and building their networks are likely to be well-positioned for success.

8.APPENDIX :-

SOURCE CODE

**­­­IMPORTING THE LIBRARIES :-**

import numpy as np

import pandas as pd

import os

import seaborn as sns

import matplotlib.pyplot as plt

from sklearn import svm

from sklearn.metrics import accuracy\_score

from sklearn.neighbors import KNeighborsClassifier

from sklearn import metrics

from sklearn.model\_selection import cross\_val\_score

from sklearn import preprocessing

from sklearn.model\_selection Import train\_test\_split

from sklearn.preprocessing import StandardScaler

importjoblib

from sklearn.metrics import accuracy\_score

**Read The Dataset :-**

df = pd.read\_csv("/collegePlace.csv")

df.head()

Handling Missing Values:-

df.info()

df.isnull().sum()

**Handling Outliers:-**

def transformationplot(feature):

plt.figure(figsize=(12,5))

plt.subplot(1,2,1)

sns.distplot(feature)

transformationplot(np.log(df['Age']))

**Handling Categorical Values:-**

df=df.replace([‘Male’], [0])

df = df.replace([Female], [1])

df = df.replace([‘Computer Science’,’InformationTechnology','ElectronicsAnd Communication’, ‘Mechanical’, ‘Electrical’, ‘civil’],[0,1,2,3,4,5])

df = df.drop([‘Hostel’], axis=1)

df

**Univariante Analysis:-**

plt.figure(figsize=(12,5))

plt.subplot(121)

sns.distplot(df[‘CGPA’], color=‘r’)

plt.figure(figsize=(12,5))

plt.subplot(121)

sns.distplot(df[‘PlacedOrNot], color=’r’)

**Bivariate Analysis:-**

plt.figure(figsize=(18,4))

plt.subplot(1,4,1)

sns.countplot(data[‘Gender\*])

plt.subplot(1,4,2)

sns.countplot(data[‘Education’])

plt.show()

**Multivariate Analysis:-**

plt.figure(figsize=(20,5))

plt.subplot(131)

sns.countplot(df[“PlacedOrNot”],hue=df[‘CGPA’])

sns.swarmplot(df[‘PlacedOrNot’],df[‘CGPA’],hue=df[‘Stream’])

**Scaling The Data:-**

Sc=StandardScaler()

x\_bal=sc.fit\_transform(x\_bal)

x\_bal = pd.DataFrame(x\_bal,columns=names)

**Splitting The Data Into Train And Set :-**

X = standardized\_data

Y = df[‘PlacedOrNot’]

X\_train, X\_test, Y\_train, Y\_test = train\_test\_split(X,Y,test\_size = 0.2, stratify=Y, random\_state=2)

**SVM Model:-**

Classifier = svm.SVC (kernel=’linear’)

Classifier.fit(X\_train, Y\_train)

SVC (kernel=”linear”)

X\_train\_prediction = classifier.predict(x\_train)

training\_data\_accuracy = accuracy\_score (X\_train\_prediction, Y\_train)

print(‘Accuracy score of the training data: ‘, training\_data\_accuracy)

Accuracy score of the training data: 0.7685497470489839

**KNN model:-**

best\_k = {“Regular”:0}

best\_score (“Regular”:0}

for k in range(3, 50, 2):

##Using Regular training set

knn\_temp = KNeighborsClassifier(n\_neighbors=k) #Instantiate the model

knn\_temp.fit(X\_train, Y\_train) # Fit the model to the training set

knn\_temp\_pred = knn\_temp.predict(X\_test) # Predict on the test set

score = metrics.accuracy score(Y\_test, knn\_temp\_pred) \* 100 #Get accuracy

if score >= best score[“Regular”] and score < 100: #Store best params

best\_score[”Regular”] = score

best\_k[“Regular”] = k

print(“---Results---\nk: {}\nScore: {}”.format(best\_k, best\_score))

Knn = KNeighborsClassifier(n\_neighbors=best\_k[“Regular”])

Knn.fit(X\_train, Y\_train)

Knn\_pred = knn.predict(X\_test)

testd = accuracy\_score(knn\_pred, Y\_test)

---Results---

K: {‘Regular’: 7}

Score:{‘ Regular’: 86.19528619528619}

**Artificial Nerual Network Model :-**

import tensorflow as tf

from tensorflow import keras

from keras, models import sequential

from tensorflow,keras import layers

classifier = sequential()

#first layer

classifier.add(keras.layers.Dense(6,activation = ’relu’,input\_dim = 6))

classifier.add(keras.layers.Droupout(0,50))

#second layer

classifier.add(keras.layers.Dense(6,activation = ’relu’))

classifier.add(keras.layers.Droupout(0,50))

#final output layer

classifier.add(keras.layers.Dense(1,activation = ’sigmoid’))

#compiling model

Loss\_1 = tf.keras.losses.Binarycrossentropy()

Classifiers.compile(optimizer = ‘Adam’,loss = loss\_1,metrics = [‘accuracy’])

#fitting the model

Classifier.fit(X\_train,Y\_train,batch\_size = 20,epochs = 100)

**SAVE THE BEST MODEL:-‘**

Import pickle

Pickle.dumb(knn,open(“placement.pkl”,’wb’))

Model = pickle.load(open(‘placement.pkl’,’rb’))