**VISVESVARAYA TECHNOLOGICAL UNIVERSITY**

**“JnanaSangama”, Belgaum -590014, Karnataka.**



**LAB REPORT ON**

**MACHINE LEARNING**

***Submitted by:***

## MANJIL RAJ PANTA(1BM21CS103)

***in partial fulfillment for the award of the degree of***

**BACHELOR OF ENGINEERING**

***in***

## COMPUTER SCIENCE AND ENGINEERING



**B.M.S. COLLEGE OF ENGINEERING**

**(Autonomous Institution under VTU)**

## BENGALURU-560019

March 2024-June 2024



B. M. S. College of Engineering,

**Bull Temple Road, Bangalore 560019**

(Affiliated To Visvesvaraya Technological University, Belgaum)

**Department of Computer Science and Engineering**

**CERTIFICATE**

This is to certify that the Lab work entitled “**MACHINE LEARNING**” carried out by **MANJIL RAJ PANTA(1BM21CS103),** who is bonafide student of **B. M. S. College of Engineering.** It is in partial fulfillment for the award of **Bachelor of Engineering in Computer Science and Engineering** of the Visvesvaraya Technological University, Belgaum during the year 2023-24. The Lab report has been approved as it satisfies the academic requirements in respect of Machine Learning Lab **- (22CS3PCMAL)** work prescribed for the said degree.

**Sunayana S Dr. Jyothi S Nayak**

Assistant Professor Professor and Head

Department of CSE Department of CSE

BMSCE, Bengaluru BMSCE, Bengaluru

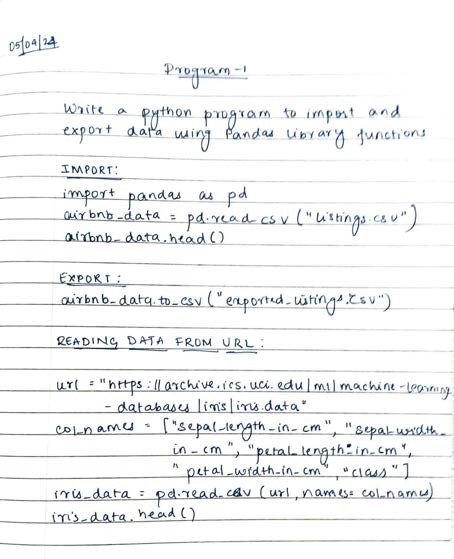
# INDEX

|  |  |  |
| --- | --- | --- |
| **Sl.**  **No.** | **Experiment Title** | **Page No.** |
| 1 | Write a python program to import and export data using Pandas library  functions | 3-5 |
| 2 | Demonstrate various data pre-processing techniques for a given dataset | 6-11 |
| 3 | Use an appropriate data set for building the decision tree (ID3) and apply this knowledge to classify a new sample. | 12-18 |
| 4 | Implement Linear and Multi-Linear Regression algorithm using  appropriate dataset | 19-25 |
| 5 | Build Logistic Regression Model for a given dataset | 26-30 |
| 6 | Build KNN Classification model for a given dataset. | 31-34 |
| 7 | Build Support vector machine model for a given dataset | 35-38 |
| 8 | Build Artificial Neural Network model with back propagation on a given  dataset | 39-41 |
| 9 | 1. Implement Random forest ensemble method on a given dataset. 2. Implement Boosting ensemble method on a given dataset. | 42-51 |
| 10 | Build k-Means algorithm to cluster a set of data stored in a .CSV file. | 52-55 |
| 11 | Implement Dimensionality reduction using Principle Component Analysis  (PCA) method. | 56-59 |

Date:05-04-2024

# PROGRAM 1

Write a python program to import and export data using Pandas library functions



**Import:**

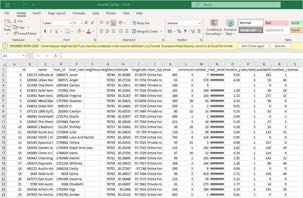
import pandas as pd # Read the CSV file

airbnb\_data = pd.read\_csv("listings.csv") # View the first 5 rows airbnb\_data.head()



**Export:**

airbnb\_data.to\_csv("exported\_listings.csv")



**Reading data from URL:**

url = "https://archive.ics.uci.edu/ml/machine-learning-databases/iris/iris.data" # Define the column names

col\_names = ["sepal\_length\_in\_cm", "sepal\_width\_in\_cm", "petal\_length\_in\_cm", "petal\_width\_in\_cm", "class"]

# Read data from URL

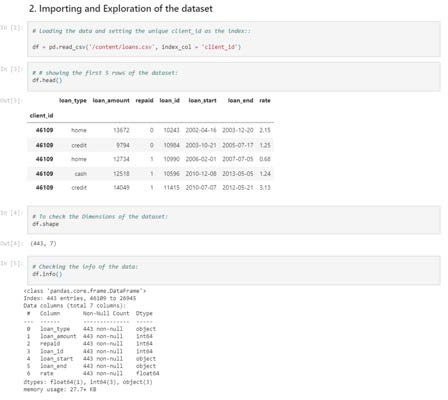
iris\_data = pd.read\_csv(url, names=col\_names) iris\_data.head()

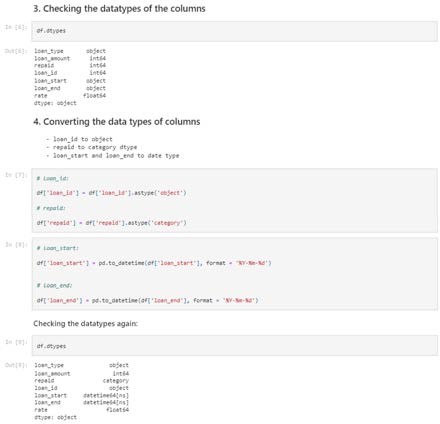


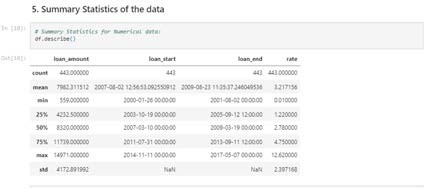
Date:05-04-2024

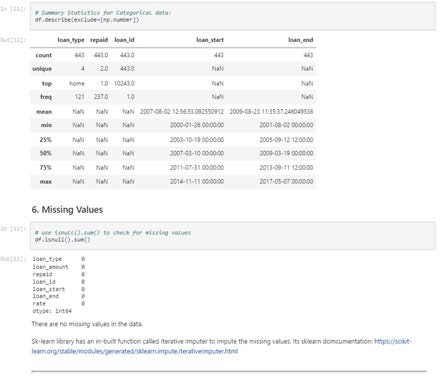
# PROGRAM 2

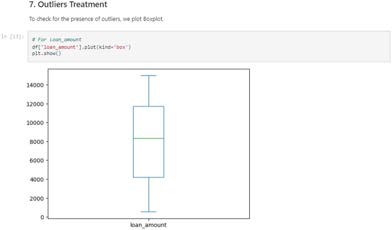
Demonstrate various data pre-processing techniques for a given dataset Code and Output

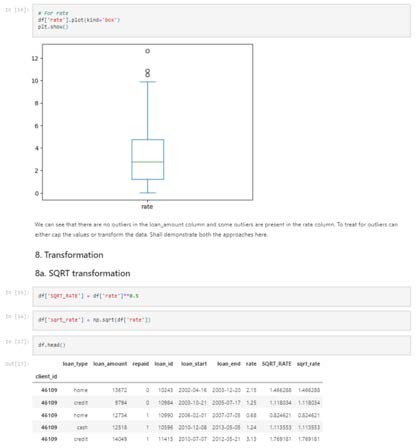


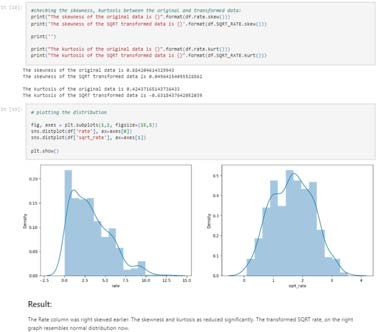


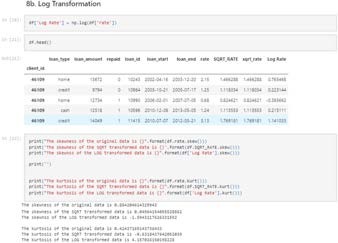


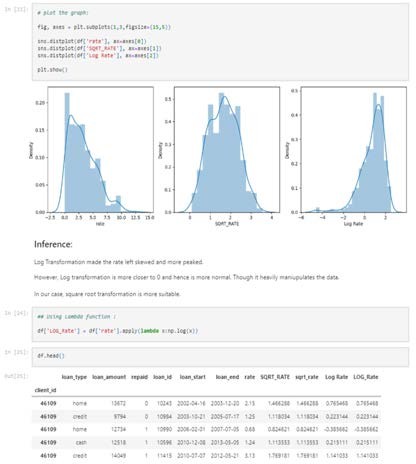








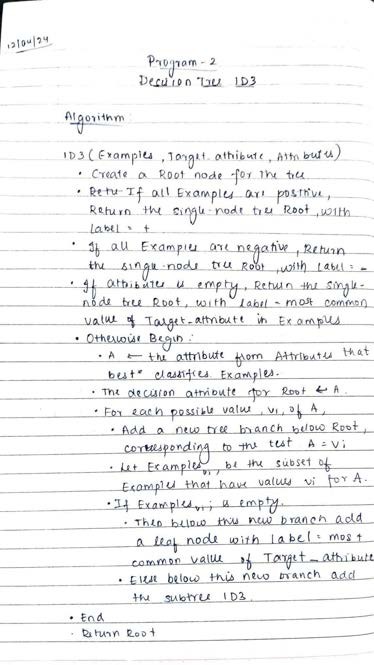




Date:12-04-2024

# PROGRAM 3

Use an appropriate data set for building the decision tree (ID3) and apply this knowledge to classify a new sample

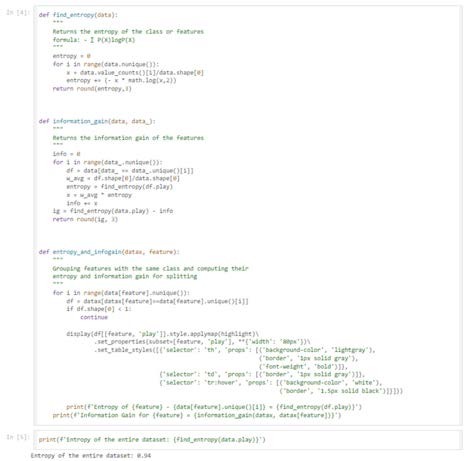
Algorithm:

Code:

**Importing Database**



**Entropy of the dataset**



**Entropy and Information Gain of temperature**



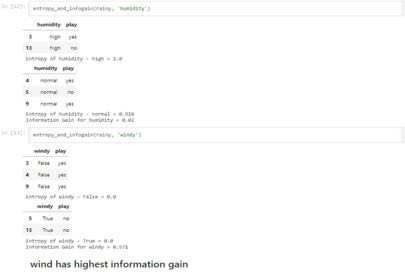
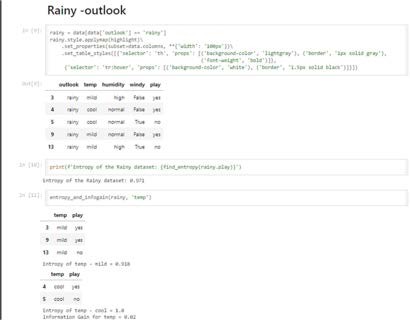
**Entropy and Information Gain of humidity**



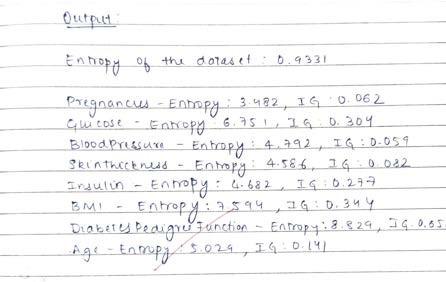
**Entropy and Information Gain of windy**



**Rainy Outlook**



Output



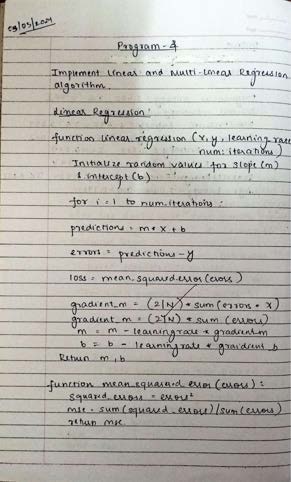
Date:19-04-2024

# PROGRAM 4

Implement Linear and Multi-Linear Regression algorithm using appropriate dataset

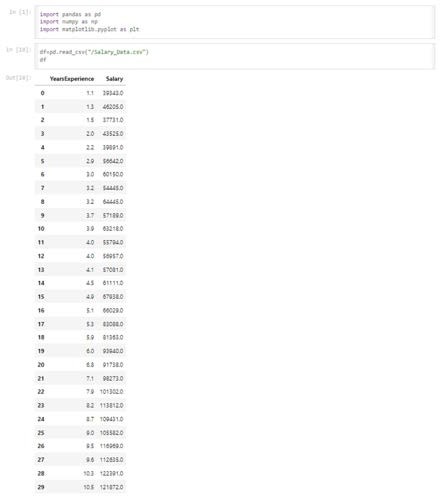
## LINEAR REGRESSION:

Algorithm



Code

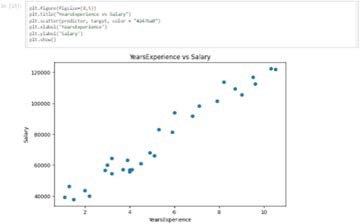
**Importing Dataset**



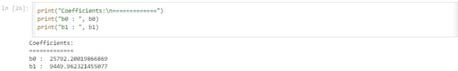
**Slope and Intercept calculation**

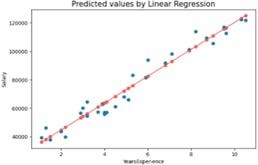


**Predicted Values Graph**



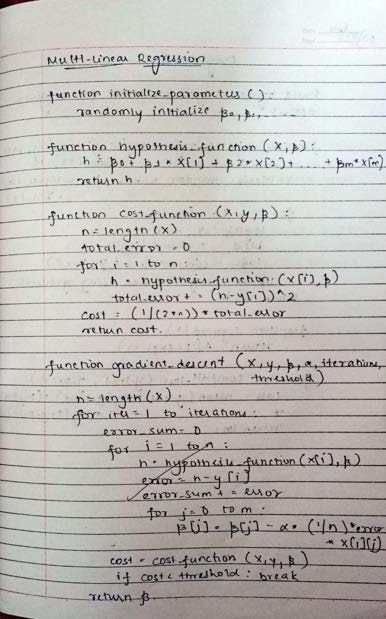
Output



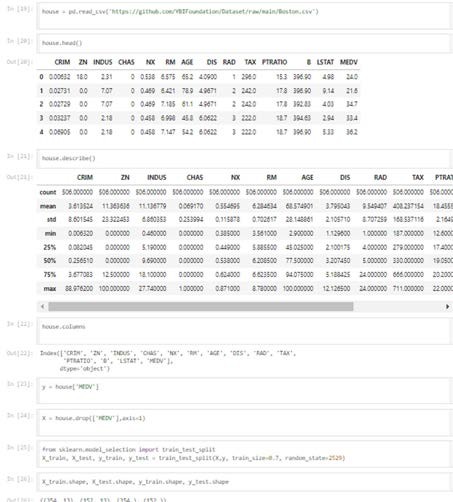


## MULTIPLE LINEAR REGRESSION:

Algorithm



Code





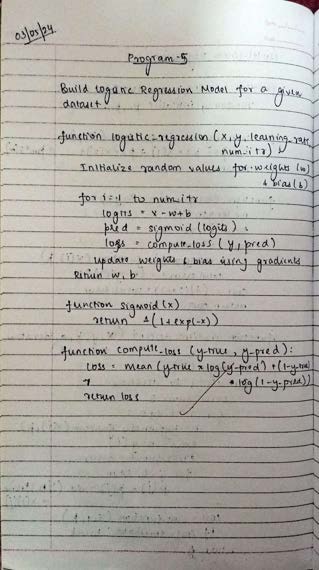
Output



Date:03-05-2024

# PROGRAM 5

Build Logistic Regression Model for a given dataset Algorithm

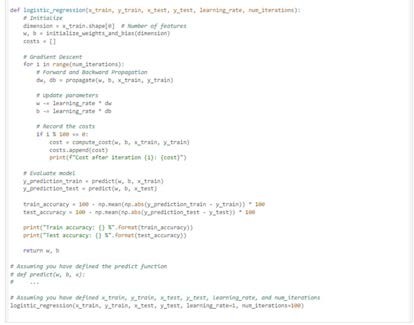


Code

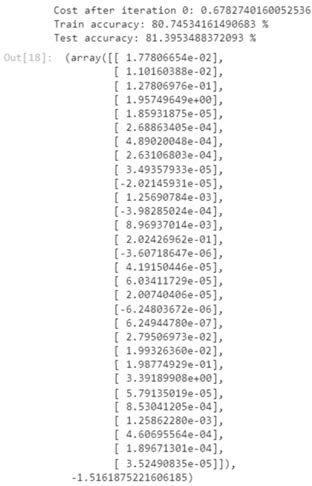








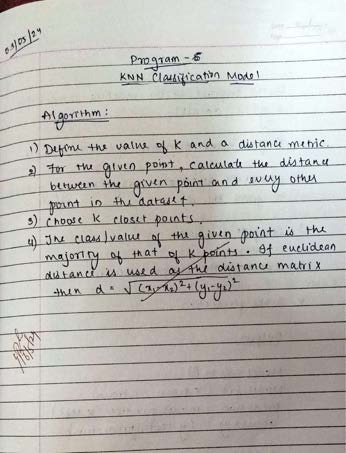
Output



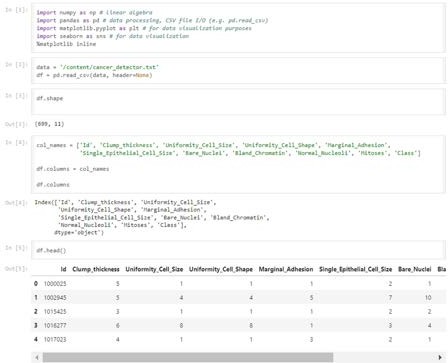
Date:19-04-2024

# PROGRAM 6

Build KNN Classification model for a given dataset. Algorithm



Code







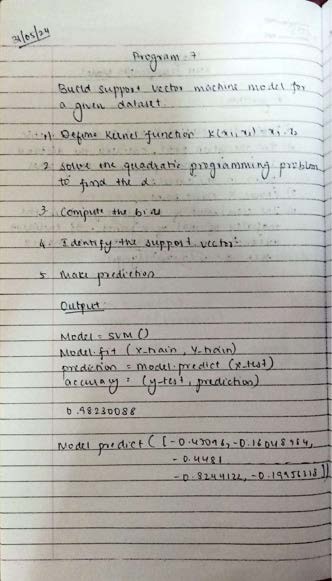
Output



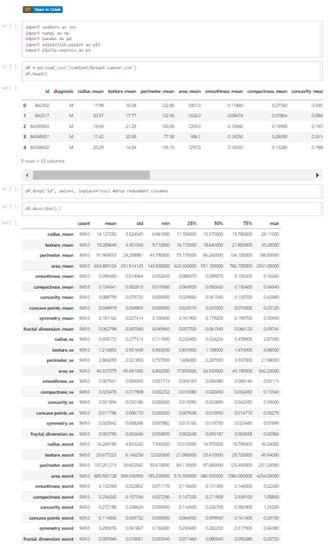
Date:24-05-2024

# PROGRAM 7

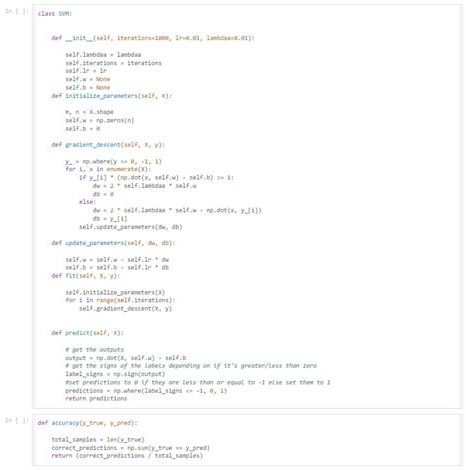
Build Support vector machine model for a given dataset Algorithm



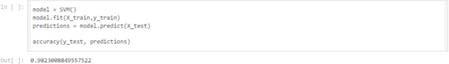
Code







Output

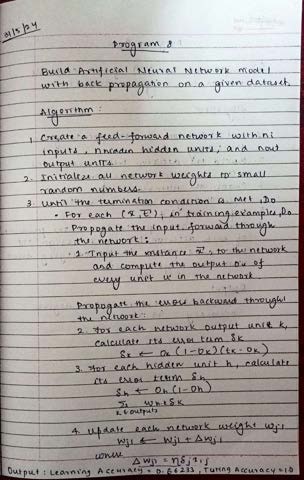


Date: 31-05-2024

# PROGRAM 8

Build Artificial Neural Network model with back propagation on a given dataset

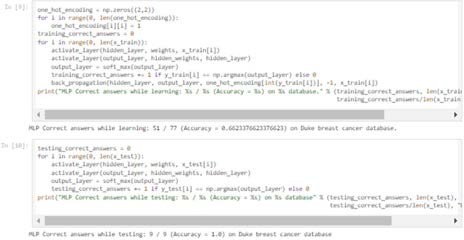
Algorithm



Code



Output

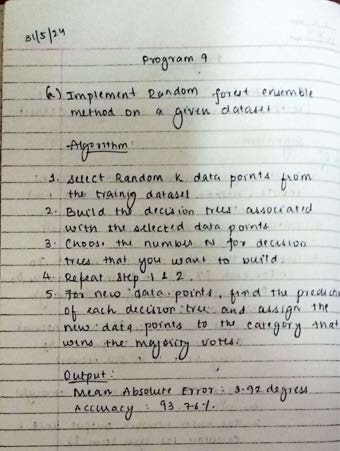


Date: 31-05-2024

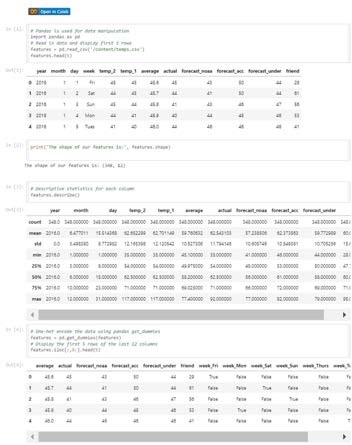
# PROGRAM 9

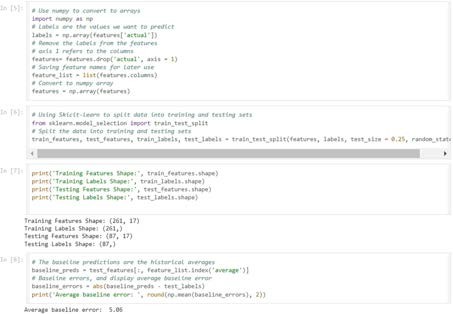
1. **Implement Random forest ensemble method on a given dataset.**

Algorithm



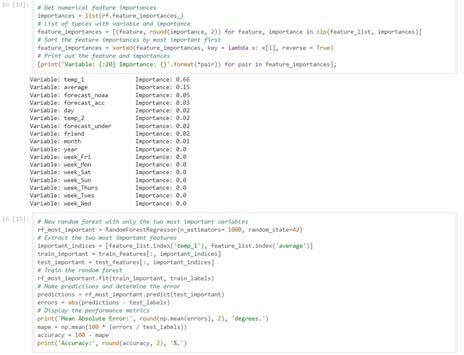
Code



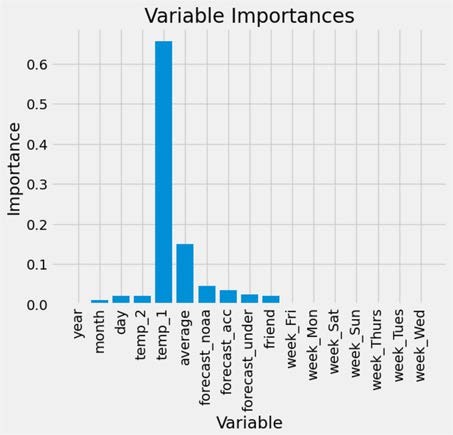






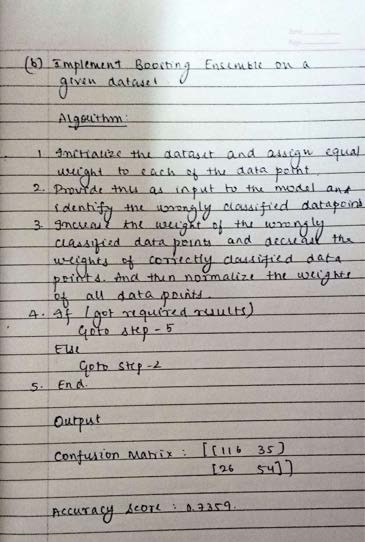


Output

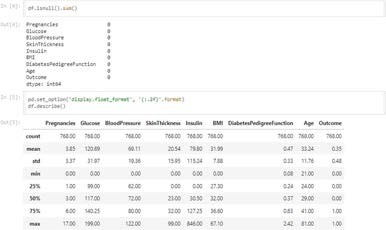
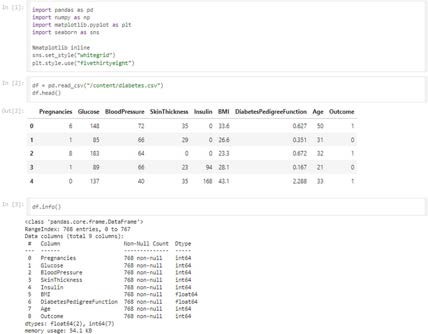


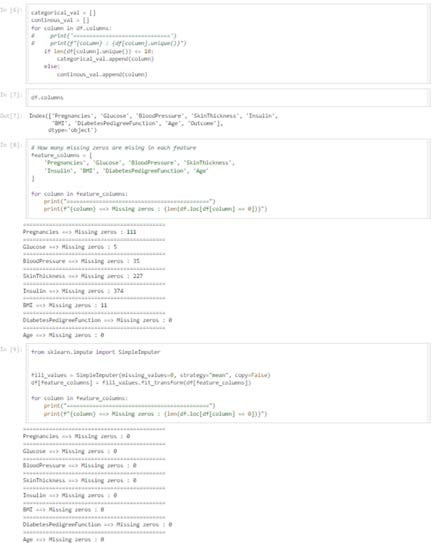
1. **Implement Boosting ensemble method on a given dataset.**

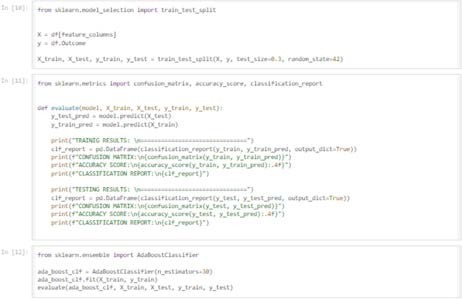
Algorithm



Code



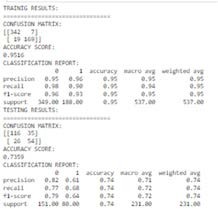




Output-AdaBoost



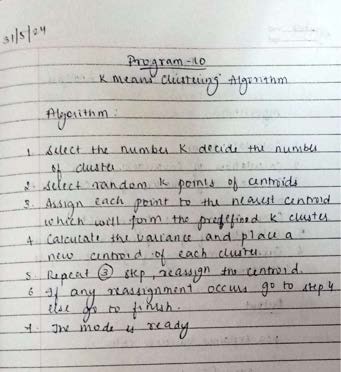
Output- GradientBoost



Date: 24-05-2024

# PROGRAM 10

Build k-Means algorithm to cluster a set of data stored in a .CSV file. Algorithm

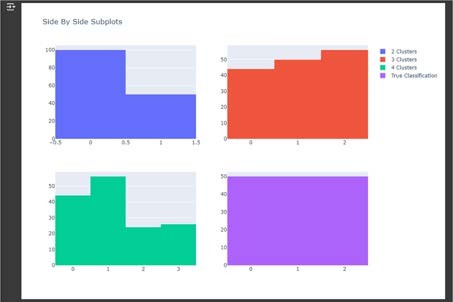


Code





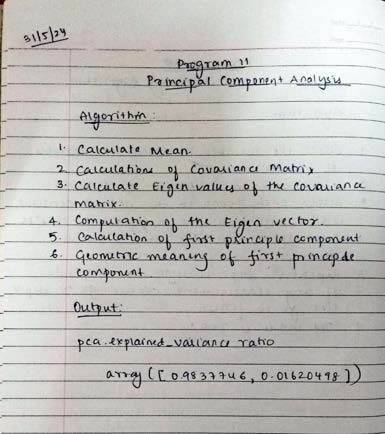
Output



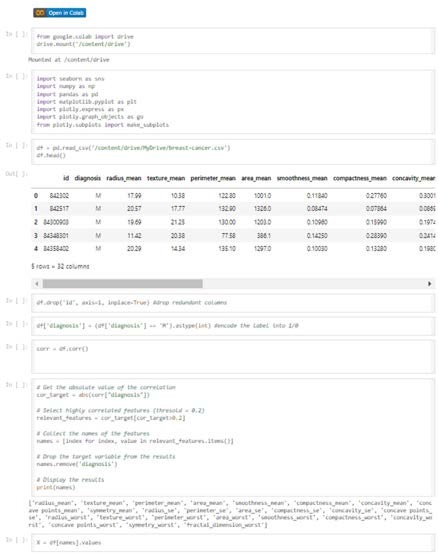
Date: 24-05-2024

# PROGRAM 11

Implement Dimensionality reduction using Principle Component Analysis (PCA) method.



Code







Output

