Machine Learning Assignment – 2

**Name: Akhil Patlori**

**Student Id: 700744211**

1. Pandas

1. Read the provided CSV file ‘data.csv’. <https://drive.google.com/drive/folders/1h8C3mLsso-R-sIOLsvoYwPLzy2fJ4IOF?usp=sharing>

**A screenshot of a computer program

Description automatically generated with medium confidence**

* Read the provided CSV file

A screenshot of a computer

Description automatically generated with medium confidence

**Github Link: https://github.com/AkhilPatlori/MachineLearningAssignment2**

2. Show the basic statistical description about the data.

A screenshot of a computer

Description automatically generated with medium confidence

3. Check if the data has null values. a. Replace the null values with the mean

A screenshot of a computer code

Description automatically generated with low confidence

A screenshot of a computer

Description automatically generated with medium confidence

4. Select at least two columns and aggregate the data using: min, max, count, mean.

A screenshot of a computer program

Description automatically generated with low confidence

5. Filter the dataframe to select the rows with calories values between 500 and1000.

A screenshot of a computer

Description automatically generated with medium confidence

6. Filter the dataframe to select the rows with calories values > 500 and pulse <100

A screenshot of a computer

Description automatically generated with medium confidence

7. Create a new “df\_modified” dataframe that contains all the columns from df except for “Maxpulse”.

A screenshot of a computer

Description automatically generated

8. Delete the “Maxpulse” column from the main df dataframe

A screenshot of a computer

Description automatically generated with medium confidence

9. Convert the datatype of Calories column to int datatype.

A screenshot of a computer

Description automatically generated with medium confidence

10. Using pandas create a scatter plot for the two columns (Duration and Calories).

A screen shot of a graph

Description automatically generated with medium confidence

2. Scikit-learn

1. Implement Naïve Bayes method using scikit-learnlibrary.

a. Use the glass dataset available in Link also provided in your assignment.

b. Use train\_test\_split to create training and testing part.

2. Evaluate the model on testing part using score and classification\_report(y\_true, y\_pred)

A screenshot of a computer

Description automatically generated with medium confidence

A screenshot of a computer

Description automatically generated with low confidence

A screenshot of a computer

Description automatically generated with low confidence

A screenshot of a computer

Description automatically generated

1. Implement linear SVM method using scikit library

a. Use the glass dataset available in Link also provided in your assignment.

b. Use train\_test\_split to create training and testing part.

2. Evaluate the model on testing part using score and classification\_report(y\_true, y\_pred)

A screenshot of a computer program

Description automatically generated with medium confidence

A screenshot of a computer

Description automatically generated with low confidence

A screenshot of a computer

Description automatically generated with medium confidence