Assignment-4

Name: Kamsani Vijay Kumar

Student Id:700737910

Github link:

https://github.com/Vijaykamsani/ML\_Assignment\_-70037910

Video :

https://drive.google.com/file/d/1SAEg5AlbjXa9xfl2MIadZ7fSc876m\_4b/view?usp=share\_link

Programming:

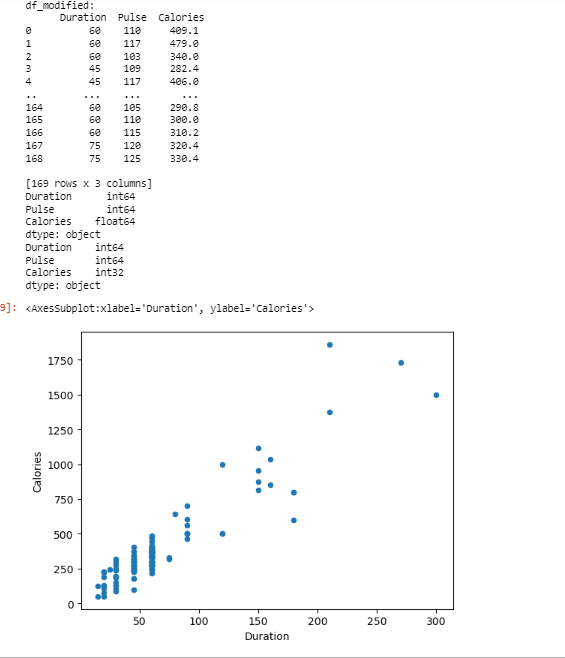
1.Pandas



Output:

Table

Description automatically generated



Comment:

The basic statistical description of the data is printed using the describe () function of pandas.The code checks if the data contains any null values. If there are any, the null values are replaced with the mean value of the column using the fillna() function. Two columns, Duration and Calories, are selected from the DataFrame and the aggregation functions min, max, count, and mean are applied to them using the agg() function. The resulting DataFrame is stored in 'agg\_df'.A new DataFrame named 'filtered\_df1' is created by filtering the rows in 'df' where the Calories value is between 500 and 1000. Another DataFrame named 'filtered\_df2' is created by filtering the rows in 'df' where the Calories value is greater than 500 and the Pulse value is less than 100.A new DataFrame named 'df\_modified' is created by dropping the 'Maxpulse' column from the 'df' DataFrame using the drop() function. The 'Maxpulse' column is deleted from the 'df' DataFrame using the drop() function The 'Calories' column in the 'df' DataFrame is converted to integer data type using the astype() function.A scatter plot of the 'Duration' column on the x-axis and the 'Calories' column on the y-axis is created using the plot.scatter() function of pandas.

Question1:

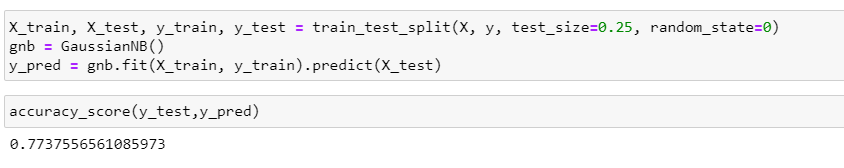


Graphical user interface, text, application

Description automatically generated

Chart, histogram

Description automatically generated



Comment:

Extracting data from the location and droping some necessary data fromthe sheet and describing the dataframe information, with the info() function.so, we have to compare the survived value and sex in the sheet so the because of the sex is in the string format we are converting into integer format and displaying as 0 as female and 1 as male. Droping some of the null values from the sheet and filling with the na values. creating x and y independent varibles. when we see the accuracy ratio Correlation Relation Between Survived and Sex:-0.548066 - A moderate negative (downhill sloping) relationship The correlation is clearly seen when the Sex (Male, Female) changes the survival chances changes because, When Sex is Female , there is a high chance of Survival and vice versa .Therefore, Sex is an Important Feature for the Survival chances. When we check with the naive bayes rule we got the accuracy rate. As shown in the output.

Question2:



Graphical user interface, text, application, email

Description automatically generated

Chart, bar chart, histogram

Description automatically generated

Comments:

Using the naive bayes algorithm, we are taking the training an testing the data of the glass file and here we training the x\_tain and y train and then we are predicting the x\_test data then we are comparing with the predicted value . the classification report for all the classes also we can see in the above output. later we are using the liner \_svc then we are predicted the values we got the accuracy.

svc got an accurate value. Compared with the all accuracy value we got On Glass Dataset, Linear SVC has better ­Accuracy than SVC and Naïve Bayes method. SVM are different from other Naive Bayes algorithm because of the way they choose the decision boundary that maximizes the distance from the nearest data points of all the classes. The decision boundary created by SVMs is called the maximum margin classifier or the maximum margin hyper plane.