

DS-312-Application of Data Science-A

Assignment 2 Report

Name:

Muhammad Nadeem 201980050

Bilal Butt

Uzba 201980049

Esha 201980022

Submitted to:

Sir Umar Ramzan

Report:

In this assignment the steps involved define below:

First we need some libraries so we import all those libraries that we need then we read csv file and after this we compute Mean, Median , Standard deviation and Five Number Summary Acc. To each Attribute. Then we find dataset shape then describe dataset like find count, unique, top and frequency. Then we find class attribute and find output like weather it is classification or multiclassification. Then we find relation between sepal and petal length. Then we find Percentage between each entity that consist in output class. Then we plot distance plot against each numerical attribute that give me the information on x-axis sepal length and y-axis density or probability. That give me the information define probability of occurrence of sepal length. Then we show count of output class entities. In which we have three entities we check these three entities against Four numeric parameters i.e Sepal

width, Sepal length, Petal width and Petal length. We Train 3 models on dataset i.e KNN, Logistic Regression and Decision Tree. And we find accuracy Acc. to 2 different ways 1st is .score it find accuracy from average/mean accuracy and 2nd is accuracy_score that find accuracy through confusion matrix. Finally we plot graph against knn ,Decision tree and Logistic Regression for compute the Accuracy through .score way.

The End