Muhammad Nadeem

201980050

DS-326-A-Machine Learning for Data Science

Proposal of Assignment 2

Submitted to:

Sir Usman Ali

Work That I do,

Data Pre-Processing:

In this Assignment we have dataset that consist of Many folders. From each folder we have many files. I collect just 100 watt files from each folder and drop [Current-B, Current-C, Time Stamp] columns from each file and collect first 100000 rows from each file. And make chunks of 1000 rows then take a transpose of every 1000 rows that I selected. Initially we 100000 rows and 1 column but when I set this completely we have 100 rows and 1000 columns of each files. And then we I os Library (The OS module in python provides functions for creating and removing a directory(folder), fetching its contents.) and then I can concat each file to another file with labels. After this we have 1400 rows and 1001 columns. Column number 1001 is a label column.

KNN:

I can implement knn from scratch as previous assignment. In which we have two methods first is euclidean_distance() in which we find Euclidean distance using some math functions and implement Euclidean formula. And second method is predict() and then drop a label column, then set data for Training and testing. Then give k=3 and predict and

then find the model accuracy, precision, recall, Confusion Matrix, Sensitivity, Specificity and F1 Score.

Logistic Regression for Binary Classification Problem:

In which I can implement using Built in Libraries. In which we drop label column and set testing size of dataset. Built in libraries work as self so then find the model Accuracy, precision, recall, Confusion Matrix, Sensitivity, Specificity and F1 Score.

Logistic Regression For Binary Class From Scratch:

In which we implement sigmoid function from scratch that learn formula in class. And then make class for logistic regression in which we initialize those things that used in this class e.g bias , weigths, learning rate (lr) etc. and then make method and take gradient descent in which then make predict() method that used sigmoid function and give me the prediction. After this we set test size of dataset using train_test_split() and then predict classifier of test dataset and then find accuracy from scratch.

Logistic Regression for Multi Class:

In which we read dataset and drop label column and then set testsize after this implement logistic regression model Built In and Find the model Accuracy, precision, recall, Confusion Matrix, Sensitivity, Specificity and F1 Score.

That's a Assignment...