CSE4/574 Fall 2022 Introduction to Machine Learning Programming Assignment 1
Classification and Regression
TA: Meng Ding

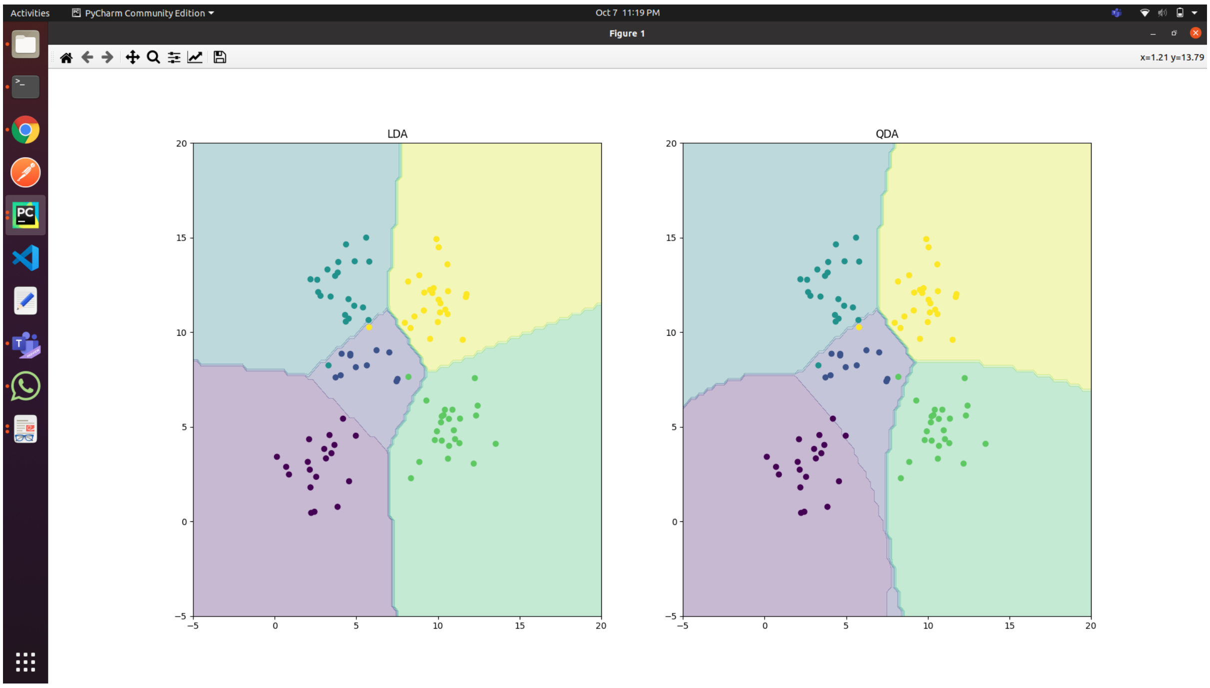
Team
NAME:Mallikharjuna Rao Annam
UBIT NAME: mallikha (50465177)
NAME: Sai Sridhar Reddy Palli
UBIT NAME: spalli (50468339)
NAME: Kajol
UBIT NAME: kajol (50478691)

Problem 1: Experiment with Gaussian Discriminator

After the experiment with gaussian discriminator of both Linear and Quadratic Discriminant Analysis which is also called (LDA) and (QDA). Below are the following accuracies with the test data.

LDA Accuracy : - **97.0 %**
QDA Accuracy :- **95.0 %**.

The plot below represents the graphs for both LDA and QDA for the given data.



The linear discriminant analysis is having the highest accuracy when compared to the quadratic discriminant analysis. By seeing the above graph, can say that the LDA will only be flexible for linear data while the QDA is flexible for quadratic data. In the real world QDA should give the highest accuracy because it has a covariance matrix for each and every class. As per the given data, it doesn't contain much data LDA gives us more accuracy. While QDA will work more efficiently for huge datasets.Because of linearity in LDA the lines are straight through the data but in QDA it’s a bit curved lines among the splits.

Problem 2: Experiment with Linear Regression

Below are the MSE details produced on the given data using the linear regression both with and without intercepts.

Train Data

MSE without intercept**[19099.44684457]**
MSE with intercept **[2187.16029493]**

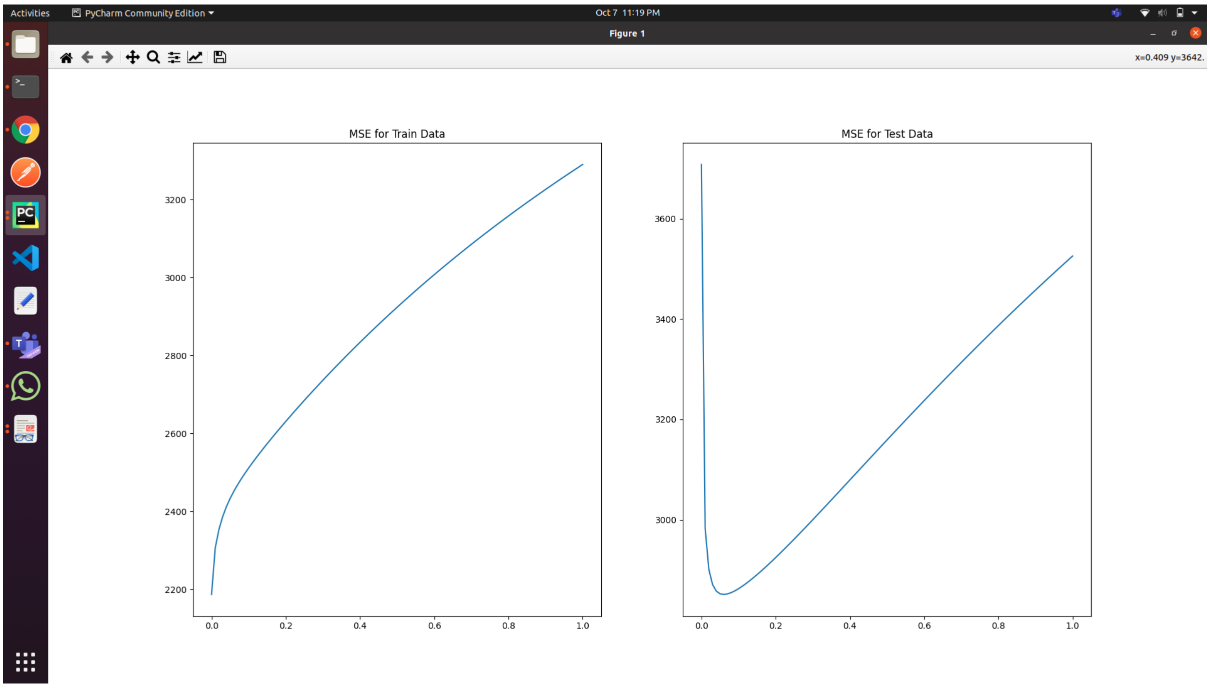
Test Data

MSE without intercept**[106775.36145122]**
MSE with intercept **[3707.84018096]**

From the above values, using the intercepts show the least error when compared to the MSE without intercepts for both test and train datas.

Problem 3: Experiment with Ridge Regression

a.) MSE for Ridge Regression on train data: **[2187.16029493]** , lambda: **0.0**
MSE for Ridge Regression on test data: **[2851.33021344]**, lambda: **0.06**



Since Ridge regression uses the regularisation, the MSE has the lesser values when compared to the OLE.

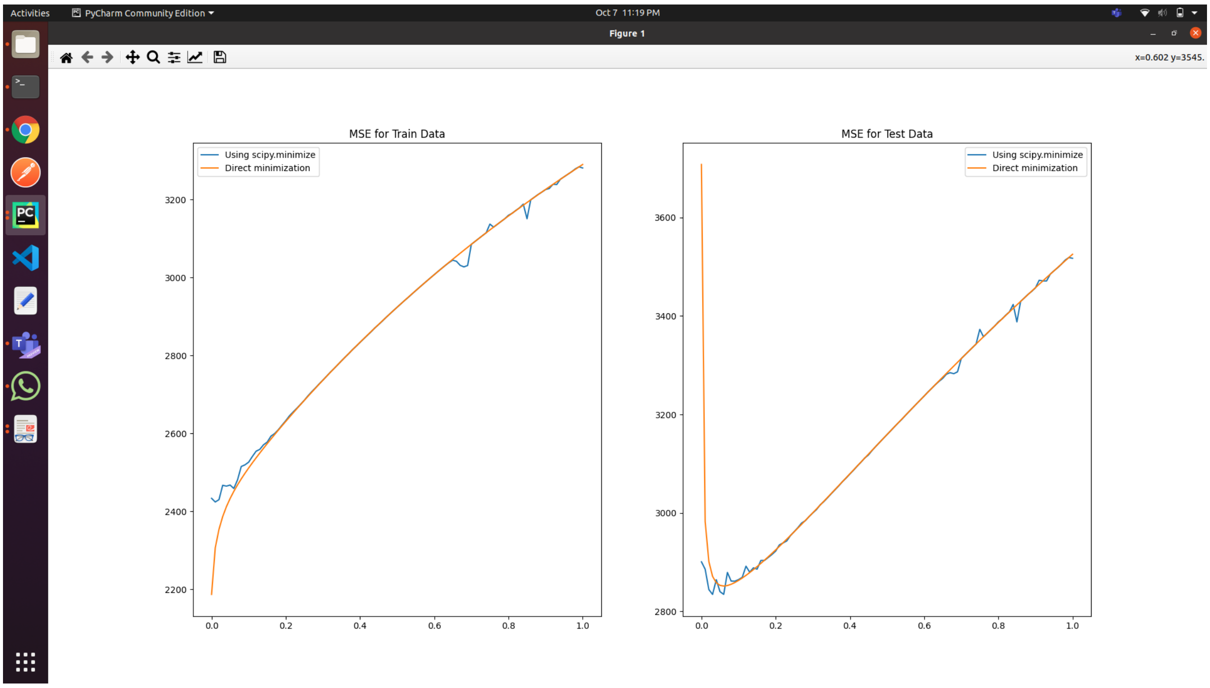
From Problem 2 and 3 :
Below are the MSE values for both OLE and Ridge Regression

MSE with intercept of OLE on train data **[2187.16029493]**
MSE with intercept of OLE on test data **[3707.84018096]**
MSE for Ridge Regression on train data with intercept: **[2187.16029493]**
MSE for Ridge Regression on test data with intercept : **[2851.33021344]**

From the above values, it's clear that MSE of ridge regression is lesser than OLE. Therefore Ridge regression works more efficiently than OLE.

Problem 4: Using Gradient Descent for Ridge Regression Learning

MSE for Gradient Descent on train data **2424.2539394878904**
MSE for Gradient Descent on test data **2834.2667881914676**



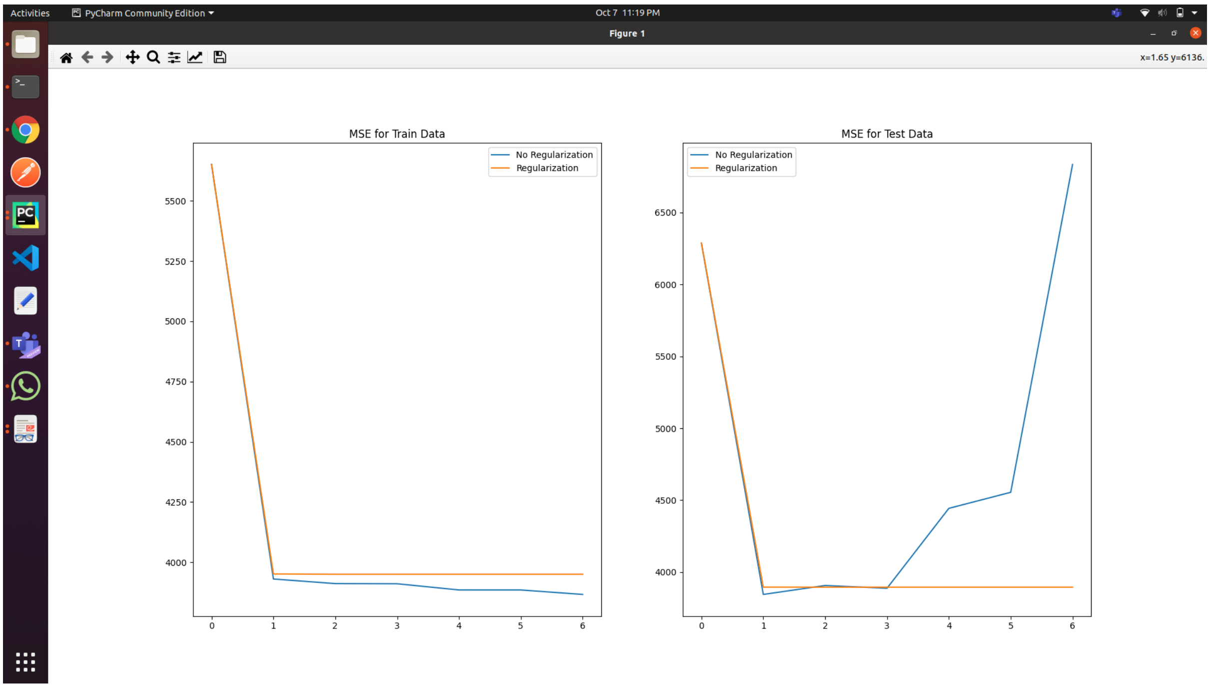
From the above graph, can say that apart from the lambda value either high or low, both ridge regression and gradient descent for ridge regression gives the same result and graph. Only the place where lambda value varies in a high value the graph shows some variations.

Problem 5: Non-linear Regression

MSE for train data **3845.034730173414**
MSE for test data **3866.88344944605**

Non-Linear Regression:

- To the given problem, we have shown the result of MSE values for both the Test and Train data by using regression and without regression as well.
- Test Data without regularisation, MSE rises dramatically after reaching a nadir at p=1. When we apply regularisation the values of MSE decreases infinitesimally starting at p = 1.
- Train Data with regularisation, Beginning with p = 1, MSE remains constant. MSE decreases linearly as the degree of polynomial increases.
- We can also see the training data. In both cases, with and without regularisation, the error gradually decreases.This is because higher order polynomials have more complex curves which can be better fit.
- In test data, the output of the regularised execution is much similar to that of the training data which was calculated. This is because lambda in regularisation technique will hold the weights under control and try to prevent overfitting.
- When it comes to unregularised data, There is no such check. As a result, some weights are much higher when compared to others, which results in overfitting.



PROBLEM 6 – INTERPRETING RESULTS

- Linear Regression without intercepts
  Train data - **19099.44684457**
  Test data - **106775.36155856**
- Linear Regression using intercepts
  Train data - **2187.16029493**
  Test data - **3707.84018153**
- Ridge Regression
  Train data - **2187.16029493**
  Test data - **2851.33021344**
- Ridge Regression Gradient Descent
  Train data - **2424.2539394878904**
  Test data - **2834.2667881914676**
- Non-linear Regression
  Train data - **3845.034730173414**
  Test data - **3866.88344944605**

By seeing the following types of regressions MSE values on the same data, Among all these Linear regression using intercepts and ridge regression gives the minimum error values for training data 2187.16029493 . As per the general concept whichever model is giving the least error we choose it as the best model. **Linear regression using intercepts and ridge regression** gives best only on training data. But by comparing the test data **Ridge Regression using Gradient Descent** gives the best solution. Overall Ridge regression without gradient-descent and Ridge regression using gradient descent produce similar error values. Sometimes running time may vary based on the size of the data but types of regularization will take care of producing the best results.