GROUP 18 Report

Department of Computer Science and Engineering

University at Buffalo, Buffalo, NY 14260

Hemanth Kumar Mutta

[hmutta@buffalo.edu](mailto:hmutta@buffalo.edu)

Desireddy Sai Sankeerthana

[saisanke@buffali.edu](mailto:saisanke@buffali.edu)

Sai Sandeep Lankisetty

[slankise@buffalo.edu](mailto:slankise@buffalo.edu)

**REPORT 1: Experiment with Gaussian Discriminators**

Accuracy on test data details as follows:

Accuracy of the test data for the implementation of two functions that are LDA Test and QDA Test where LDA Test is for Linear Discriminant Analysis and QDA Test is for Quadratic Discriminant Analysis

|  |  |  |
| --- | --- | --- |
|  | LDA | QDA |
| Accuracy | 97.00 | 96.00 |

Plotting the discriminating boundary for linear and quadratic discriminators as follows

**LDA QDA**

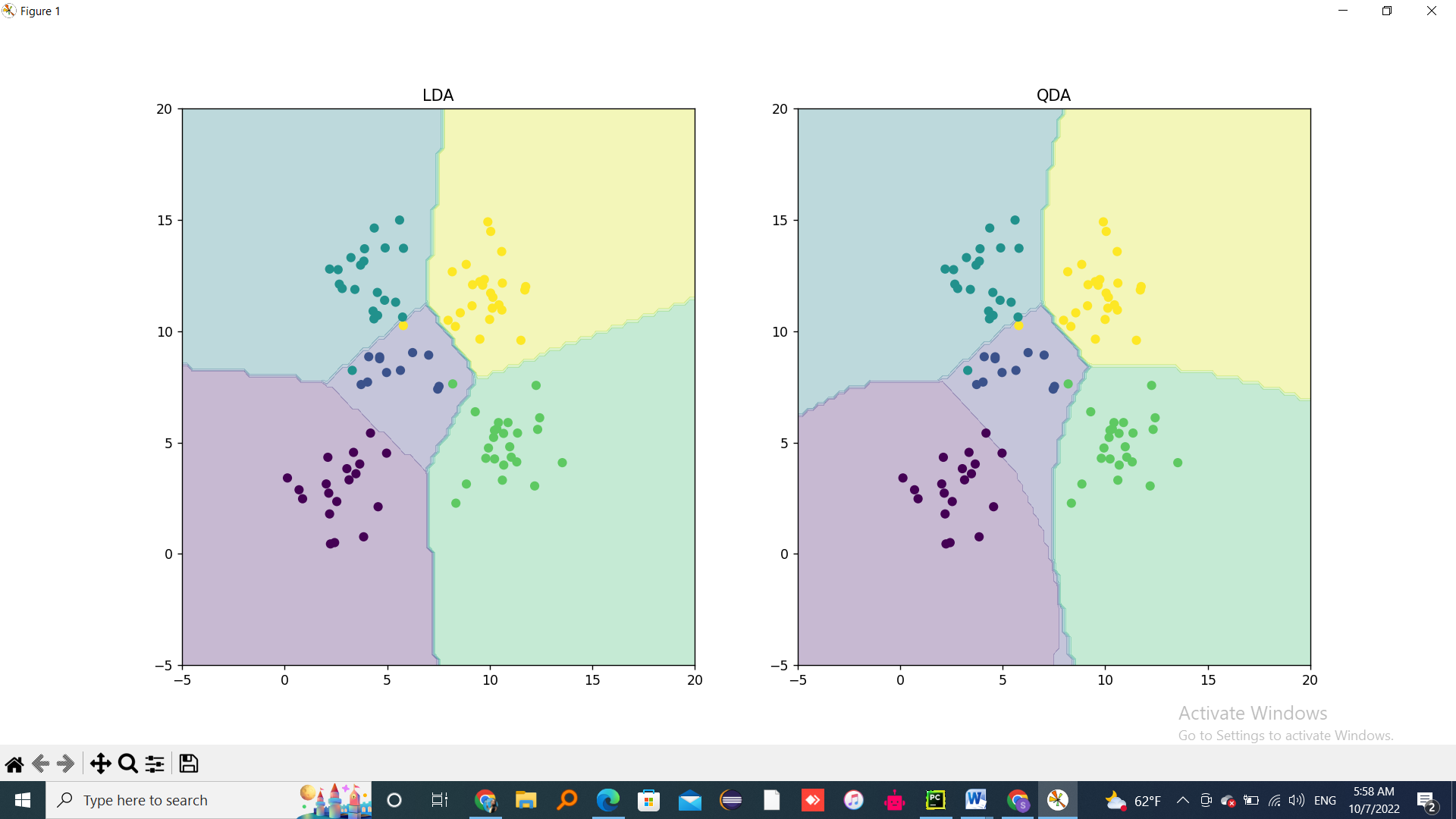
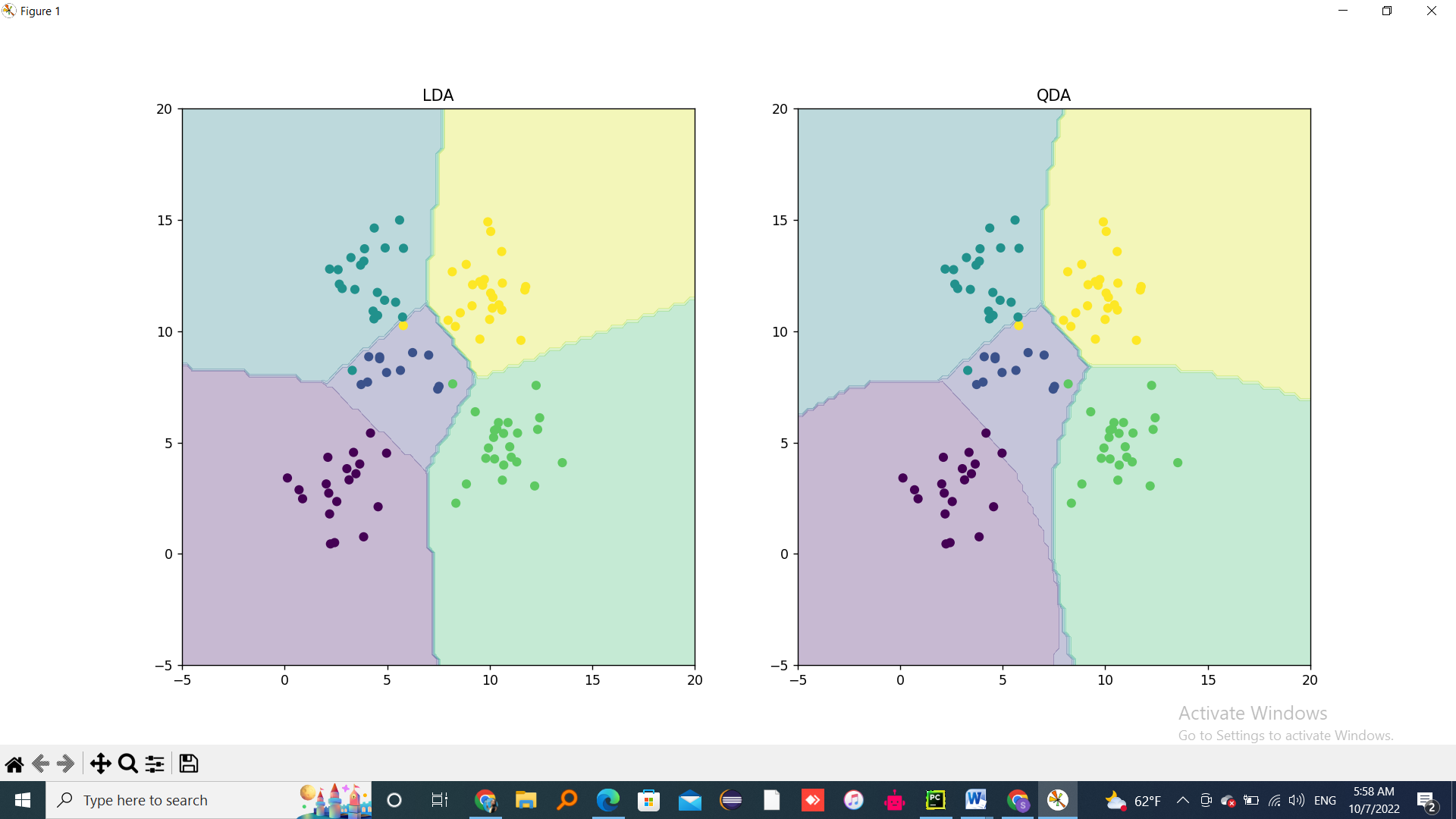
 

Figure 1:LDA and QDA

* LDA has a common covariance matrix. So, a covariance matrix that is common to all classes in a data set.
* As LDA has covariance matrix same for all classes that forces all the boundaries to be in straight lines
* QDA on other side, Observation of each class is drawn from a normal distribution.
* QDA says that each class has its own covariance matrix .hence, they doesn’t stop leading to non-linear boundaries

**REPORT 2: Experiment with Linear Regression**

linear regression, which is a very widely used method for predicting a real-valued output (also called the dependent variable or target) y ∈ R, given a vector of real-valued inputs (also called independent variables, explanatory variables, or covariates) x ∈ R D. The key property of the model is that the expected value of the output is assumed to be a linear function of the input,

E [y|x] = wTx,

which makes the model easy to interpret, and easy to fit to data.

Table values for Mean Squared Error

|  |  |  |
| --- | --- | --- |
| MSE Values | Test Value | Train Value |
| With intercept | 3707.8401817 | 2187.16029493 |
| Without intercept | 106775.36150046 | 19099.44684457 |

The Linear Regression model with intercept is the better one .Intercept allows the model to gives more accurately through the data and hence provides better results.

**REPORT 3: Experiment with Ridge Regression**

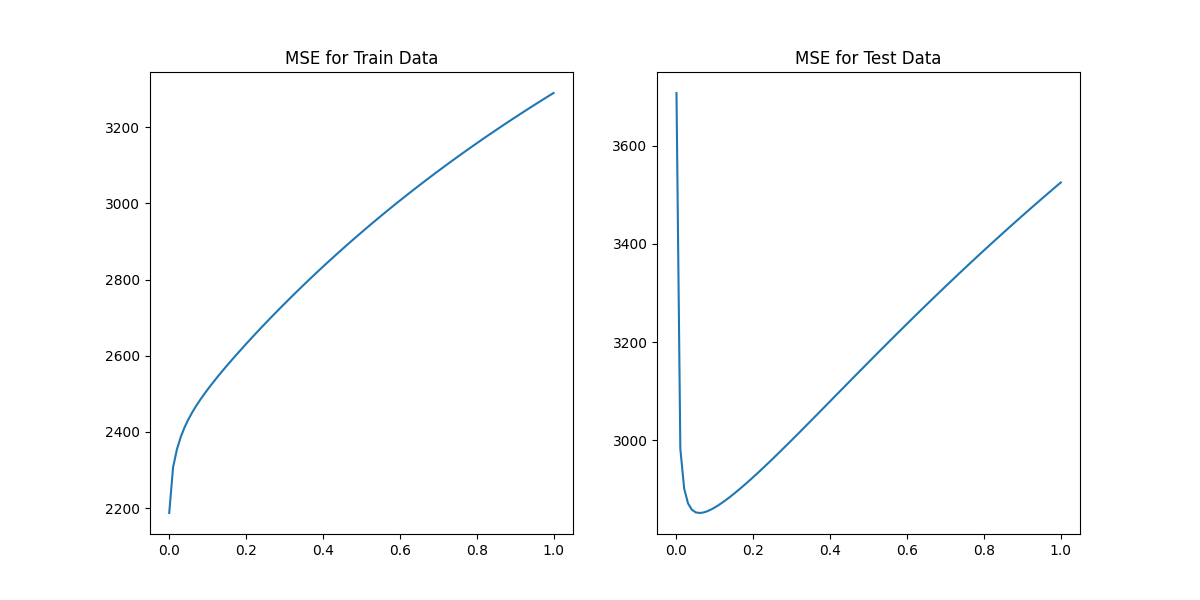


Table values for Mean Squared Error (MSE) at 0.06

|  |  |  |
| --- | --- | --- |
| Lambda | Test | Train |
| 0.0 | 3.708 \* 103 | 2.187 \* 103 |
| 0.01 | 2.982 \*103 | 2.307 \*103 |
| 0.02 | 2.901\*103 | 2.354\*103 |
| 0.03 | 2.871\*103 | 2.387\*103 |
| 0.04 | 2.858\*103 | 2.412\*103 |
| 0.05 | 2.853 \*103 | 2.433 \*103 |
| 0.06 | 2.851\*103 | 2.452\*103 |

Λ(lambda) = 0.06

3.1 Learned Weight Comparison

The weights compared with RR (zero less than lambda less than or equal to one) were found to be generally orders of magnitude smaller than weight observed from regular LR that is due to penalty of LR as it has larger weight magnitude

RR-Ridge Regression LR-Linear regression

|  |  |  |
| --- | --- | --- |
|  | Train Value | Test Value |
| OLE with intercept | 2187.16 | 3707.84 |

OLE intercept is similar with ridge regression. If we set optimal value at 0.06 we get

test value and training data at 0.0 since at that point the MSE is the least

**REPORT 4: Experiment with Gradient Descent for Ridge Regression Learning**

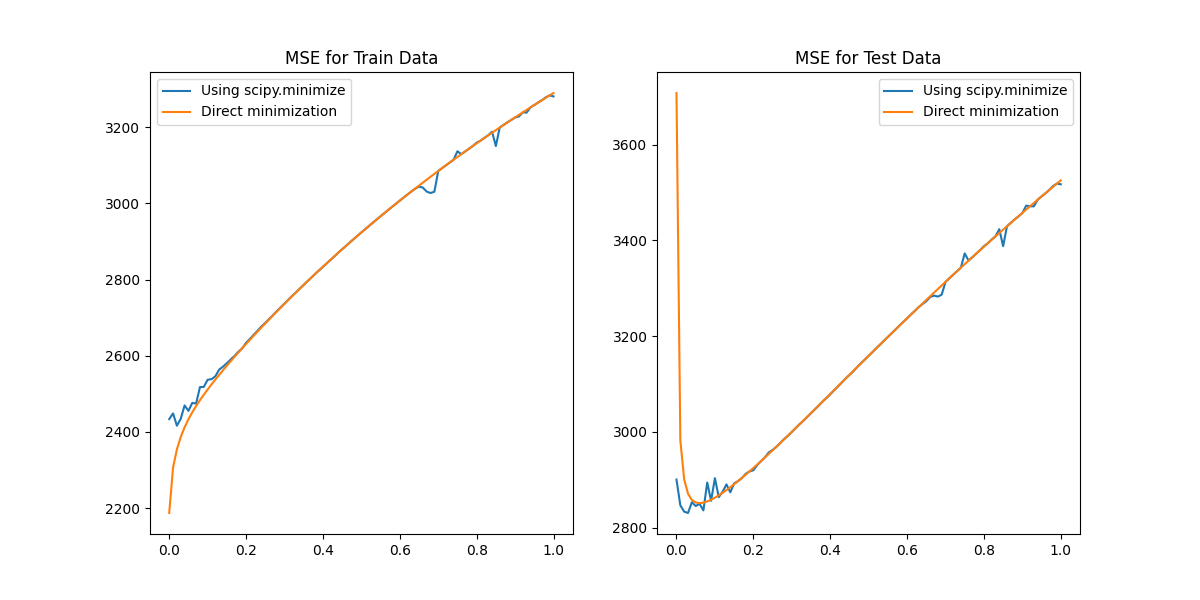
****

Figure 3:Gradient Descent for MSE

The above plotting for iteration 100.As we increase number of iterations the scipy.minimize gets more data and gives results that get very similar to the direct ones.

|  |  |  |
| --- | --- | --- |
|  | Train value | Test value |
| MSE | 2.452\*103 | 2.851\*103 |

**REPORT 5: Non-Linear Regression**

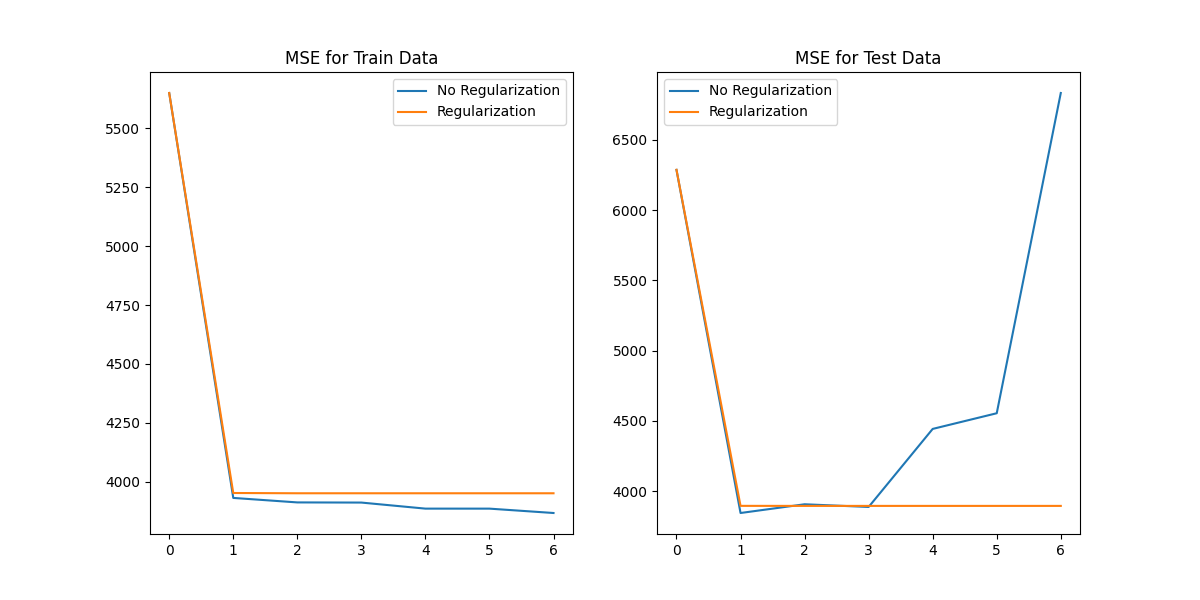
****

Figure 4:Non-linear regression

The optimal value of p 🡪 no regularization is 1

The optimal value of p 🡪 with regularization is 4

|  |  |  |
| --- | --- | --- |
|  | No regularization | With regularization |
| P 🡪1 | 3845.03 | 3895.85 |
| P 🡪4 | 4443.32 | 3895.58 |

**REPORT 6: Interpreting Results**

|  |  |  |
| --- | --- | --- |
| Model | Test | Train |
| Linear Regression - with intercept | 3707.84 | 2187.16 |
| Linear Regression - without intercept | 106775.36150046 | 19099.44684457 |
| Ridge Regression | 2.452\*103 | 2.851\*103 |
| Gradient descent ridge regression | 2451.53 | 2851.33 |
| Non-linear regression at p=1 | 3845.03473017 | 3895.85646447 |
| Non-linear regression at p=4 | 4443.32789181 | 3895.58266828 |

the best metric that can be choosen is ridge regression because ridge regression has least MSE values for test value and Train data.