CSE 574 Assignment 2

Classification and Regression

<u>Group 6:</u>

Akash Mandole

Chintan Thakker

Pranshu Pancholi

Problem 1: Experiment with Gaussian Discriminators

LDA Accuracy: 97%

QDA Accuracy: 96%

The boundary plots for linear and quadratic discriminant are shown below. As we can see from the plot below for LDA the discriminating boundaries for different classes are represented by straight lines because in LDA a single covariance matrix is calculated for all the classes. While for QDA the discriminating boundaries appear curved because in QDA the covariance matrix is calculated or each class.

Fig 1. LDA

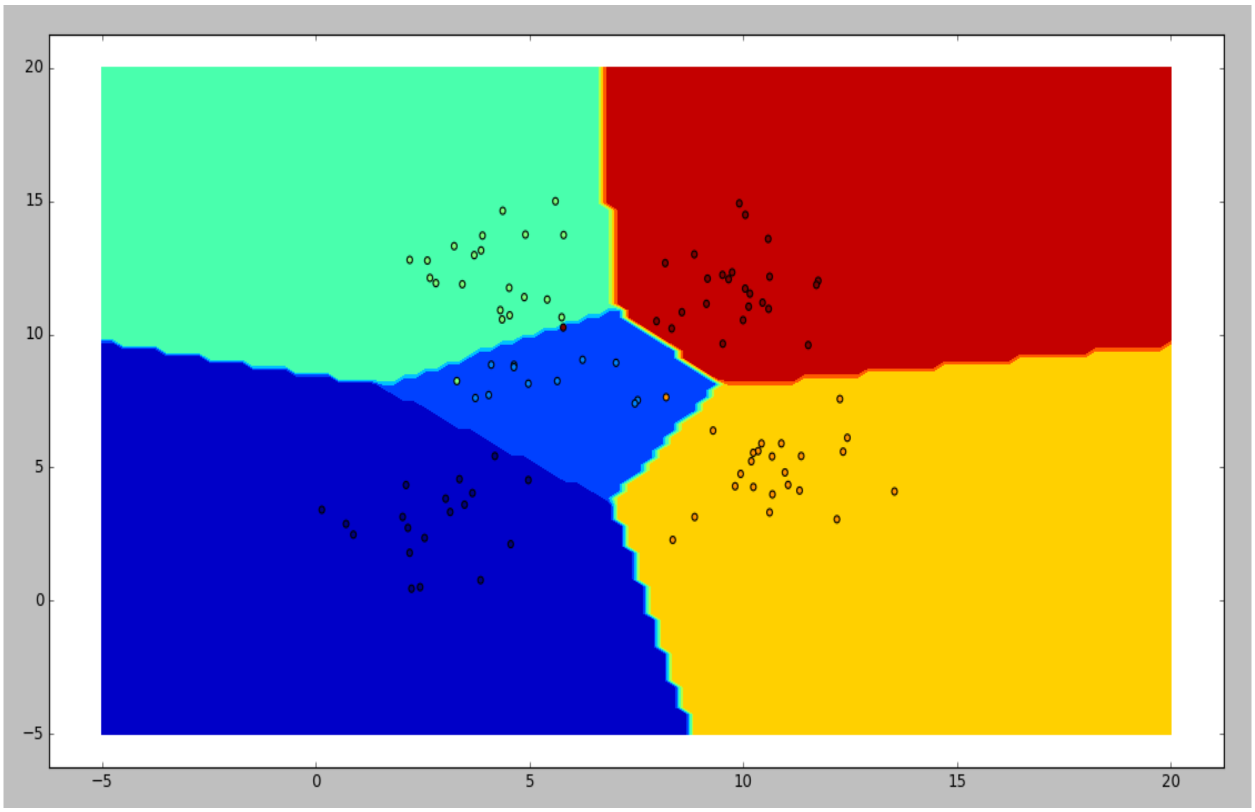
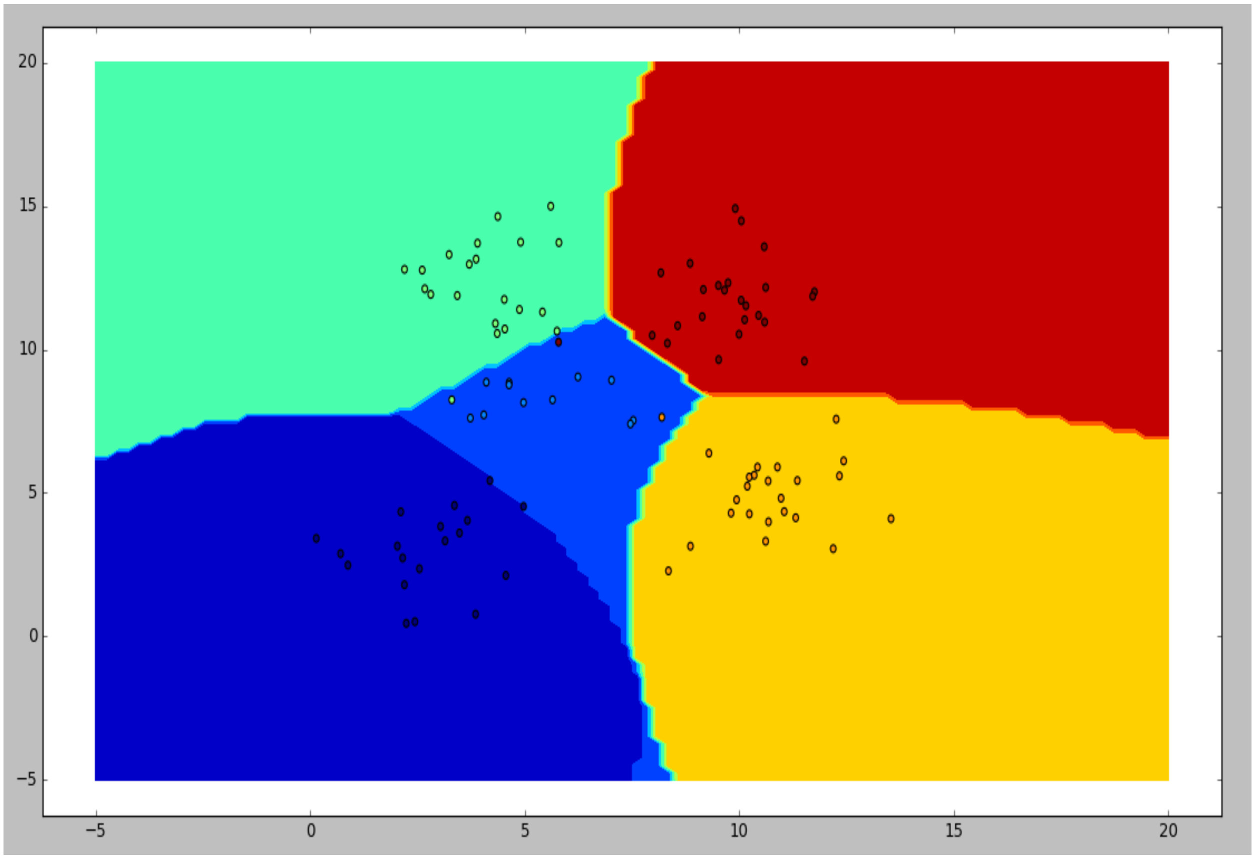


Fig 2. QDA



Problem 2: Experiment with Linear Regression

	Without Intercept	**With Intercept**
RMSE for Training Data	138.20074835	46.76708559
RMSE for Test Data	326.76499438	60.89203709

Table 2.1 RMSE values for Training and Test Data with and without intercept

From the table 2.1, it can be concluded that the root mean square error value is much less after adding an intercept to the input matrix for both training as well as test data. Hence, using an intercept is a better choice.

Problem 3: Experiment with Ridge Regression

Lambda	Test Data Error	Train Data Error
0.0	60.89203709	46.76708559
0.01	54.61177638	48.02949321
0.02	53.86068684	48.51877311
0.03	53.58116823	48.85468415
0.04	53.46026945	49.11332857
0.05	53.41035232	49.32721801
0.06	53.3978484	49.51291236
0.07	53.40739644	49.67974992
0.08	53.43107466	49.83337884
0.09	53.46442201	49.97739748
0.1	53.50474691	50.11419242
0.11	53.55033062	50.24539818
0.12	53.60002129	50.37216394
0.13	53.65301361	50.49531549
0.14	53.70872289	50.6154574
0.15	53.76671005	50.73303941
0.16	53.82663522	50.84840091
0.17	53.88822809	50.9618013
0.18	53.95126849	51.07344124
0.19	54.01557335	51.18347767
0.2	54.08098762	51.2920347

Table 3.1 Training and Test RMSE for different Lambda values.

The optimal value of lambda from Table 3.1 is **0.06** because the Testing Data RMSE value is minimum for lambda equal to 0.06.

RMSE with intercept using OLE Approach

Test Data – 60.89203709

Training Data – 46.76708559

RMSE with intercept using Ridge Regression Approach for optimal lambda (=0.06)

Test Data – 53.3978484

Training Data – 49.51291236

Above RMSE values depicts that testing data error is significantly reduced by using ridge regression approach. Thus, ridge regression is a better approach as compared to OLE approach for the given problem.

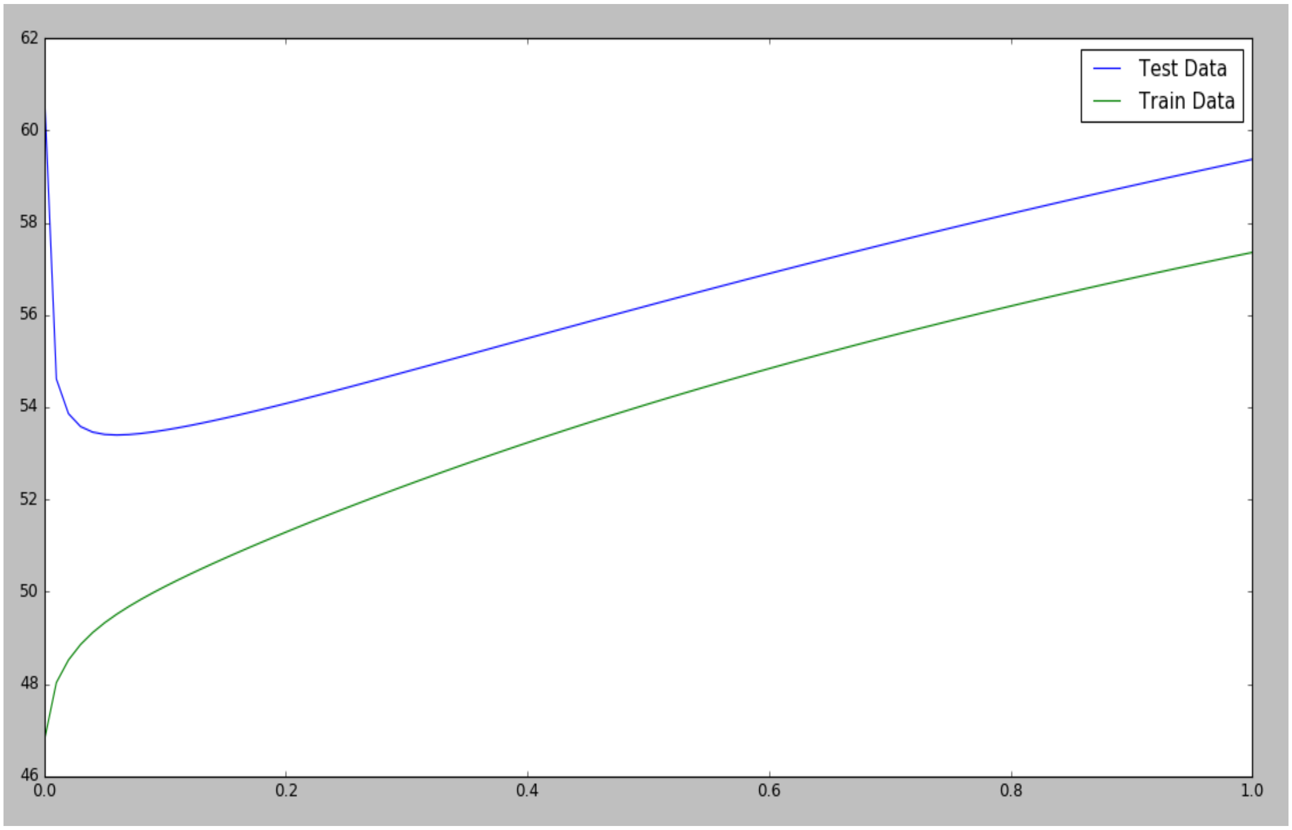


Fig 3.1 Lambda vs Error in Training and Test Data

The mean weight of OLE is always greater than mean weight of ridge regression. Thus, ridge regression is a better choice in terms of relative weights of OLE and ridge regression.

Problem 4: Using Gradient Descent for Ridge Regression Learning

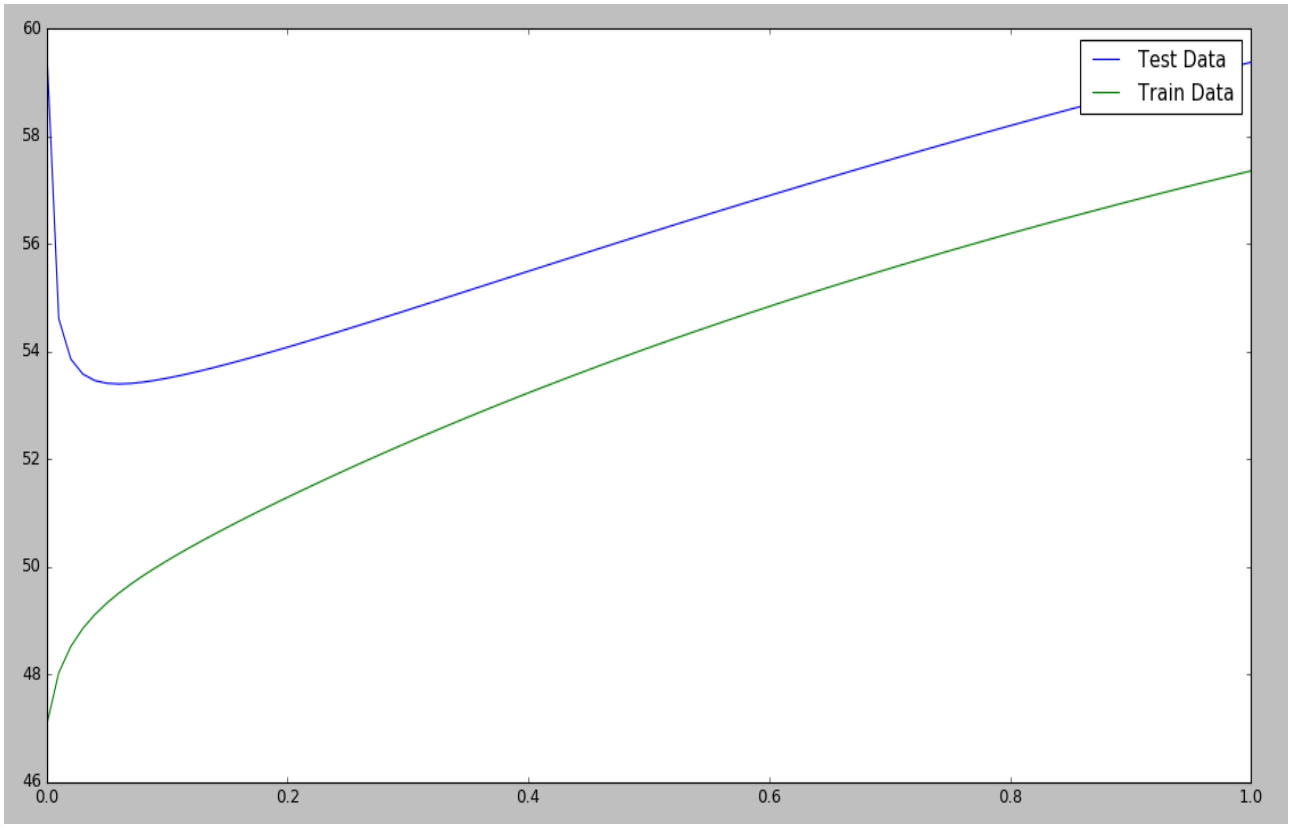


Fig 4.1 Lambda vs Training and Test Data RMSE after Gradient Descent learning

Comparing Fig 3.1 and Fig 4.1, we can depict that both the graphs are almost similar. Therefore, ridge regression approach is much faster in our problem scenario, as for gradient descent approach there is an overhead of computation using minimize function which takes time to converge to a point.

Problem 5: Non-Linear Regression

From the Fig 5.1 we can see that the RMSE value of Testing Data is significantly reduced for higher order polynomials of input data because higher order polynomial provide better fitting of the data points. When no regularization is performed the testing error decreases for p =1,2,3. After p = 3, the testing error increases because of overfitting problem i.e. the boundaries become highly curved and the learned curve performs poorly when a new testing data arises. The overfitting problem can be reduced by performing regularization as can be depicted from the Fig 5.1.

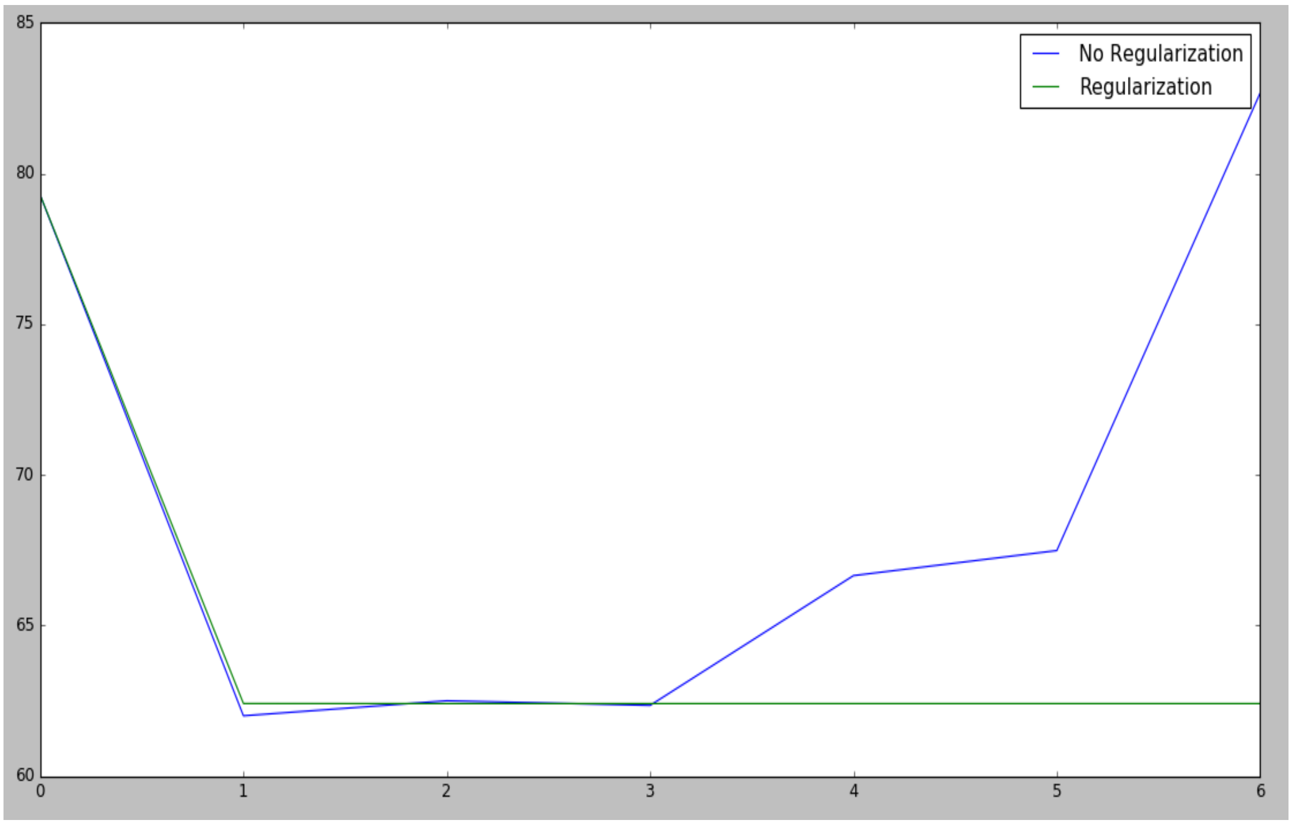


Fig 5.1 Degree of polynomial (p) value vs Testing RMSE for lambda = 0 and lambda = 0.06 (optimal)

p	**Test RMSE with Lambda = 0**	**Test RMSE with Optimal Lambda**
0	79.2868513165	79.2898604296
1	62.0083440367	62.416796333
2	62.5070243981	62.4146141215
3	62.3536329193	62.4146033867
4	66.6582919959	62.4146030051

5	67.4894834581	62.4146030085
6	82.6647394523	62.4146030086

Table 5.1 Testing RMSE values for lambda = 0 and optimal lambda.

From the Table 5.1 we can conclude that optimal value of **p = 1** when lambda = 0 i.e. for non-regularization. While optimal value of **p = 4** when lambda is 0.06(optimal) i.e. for regularization because it gives the least RMSE of testing data.

Problem 6: Interpreting Results

Approach	**Testing RMSE**	**Training RMSE**
OLE Regression without intercept	326.76499438	138.20074835
OLE Regression with intercept	60.89203709	46.76708559
Ridge Regression with optimal Lambda	53.3978484	49.51291236
Ridge Regression using Gradient Descent	53.39784867	47.05868526
Non Linear Regression (No Regularization)	62.0083440367	62.1842701127
Non Linear Regression (Regularization)	62.4146030051	62.8544535824

Table 6.1 Minimum Testing and Training RMSE observed using different approaches.

Conclusion:

From Table 6.1 it can be concluded that **Ridge Regression is best for testing data set**. Although gradient descent for ridge regression provides minimum RMSE on testing data but as discussed in Problem 4, it takes time to converge and provides a miniscule reduction in testing error. Thus, ridge regression using optimal lambda is the best setting if testing error is the metric.

On the other hand, **OLE Regression with intercept is best for training data set** as from the Table 6.1 it provides minimum RMSE on training data.