PA-2 Classification and Regression

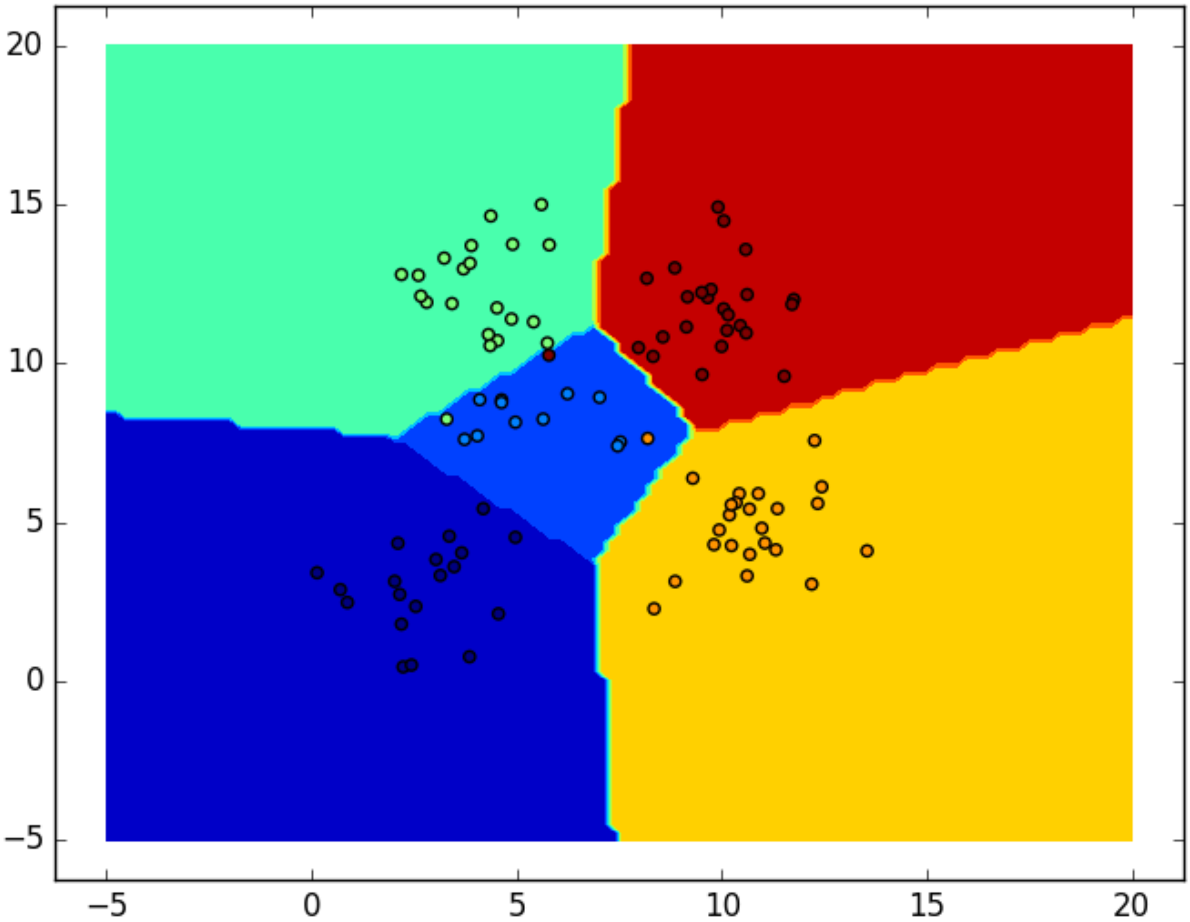
CSE 574- Introduction to Machine Learning

Course Number: CSE 574
Group No: 19
SURBHI BHATNAGAR
VARUN JOSHI
VINAY GOYAL

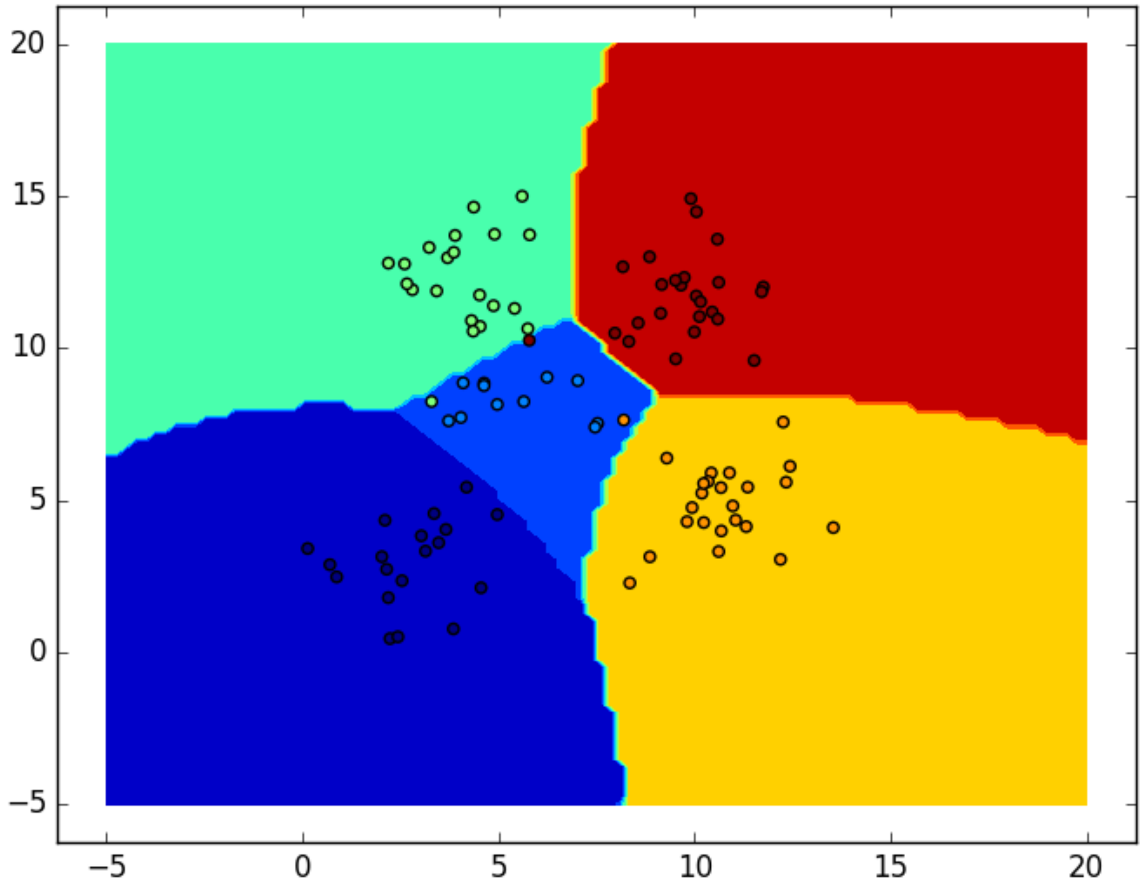
Problem 1: Experiment with Gaussian discriminators

Accuracy for LDA: 97%
Accuracy for QDA: 97%

The discriminating boundaries for linear and quadratic discriminators are plotted below:



Linear Discriminant Analysis



Quadratic Discriminant Analysis

There is a difference between the boundaries as plotted by the Linear Discriminant Analysis and Quadratic Discriminant Analysis because the former assumes the covariance for each class as the same whereas the latter calculates covariance for each class separately.
Also we can observe from the boundaries plotted that the QDA has parabolic boundaries whereas the LDA has linear boundaries.

Problem 2: Experiment with Linear Regression

RMSE Values

	Without Intercept	With Intercept
Training Data	138.20074835	46.7670855937
Test Data	326.764994391	60.892037097

Table 2: RMSE Values for Training and Test Data with and without intercept

The OLE with using intercept is better when compared with without using intercept since the smaller the RMSE, the better it is. It can be seen that the RMSE value with using intercept is lesser for both training and test data.

The RMSE values obtained from the simulations are shown in table 2. We can observe two positive effects of using an intercept for prediction:

- a significant error decrease when considering a single data set in isolation (either training or test), and
- an even more reduction of the error is noticed on the test set compared to the training set.

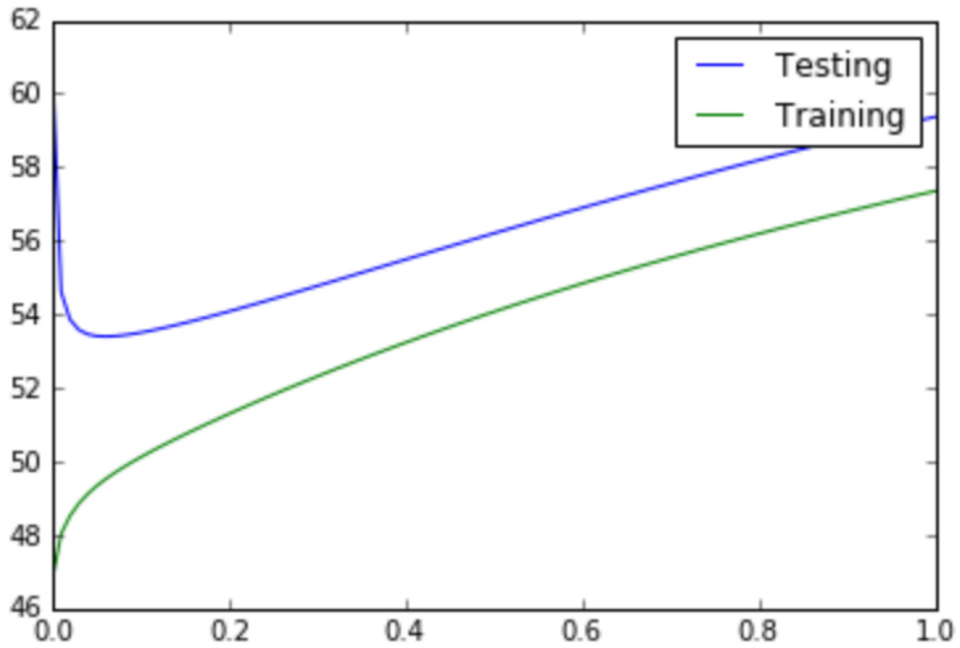
Moreover, when comparing the RMSE for training and test data, we can see that not using an intercept causes the test error to be 2.36 times the error on the training set. However, when using an intercept, the test error is 1.3 times the RMSE on the training set.

Problem 3: Ridge Regression

The RMSE for Ridge Regression using intercept

Training Data: 46.76708559
Testing Data: 53.3978484

Plot of RMSE for Testing vs Training Data for Ridge Regression



Comparison of Mean weights for Ridge Regression and Linear Regression

MEAN_OLE: 882.807625044
MEAN_RIDGE: 17.3219272625

On comparing the mean weights, it can be seen that Ridge Regression is better than OLE as it has lesser mean weight learnt.

Comparison of Error for Ridge Regression and Linear Regression

Model	Testing Data	Training Data
OLE	60.892037097	46.7670855937
Ridge Regression	53.3978484	46.76708559

On comparing the RMSE for the two approaches, while the error is similar for training data for both the models, the Ridge regression is certainly better than OLE regression in term of the error in testing data.

Calculation of Lambda

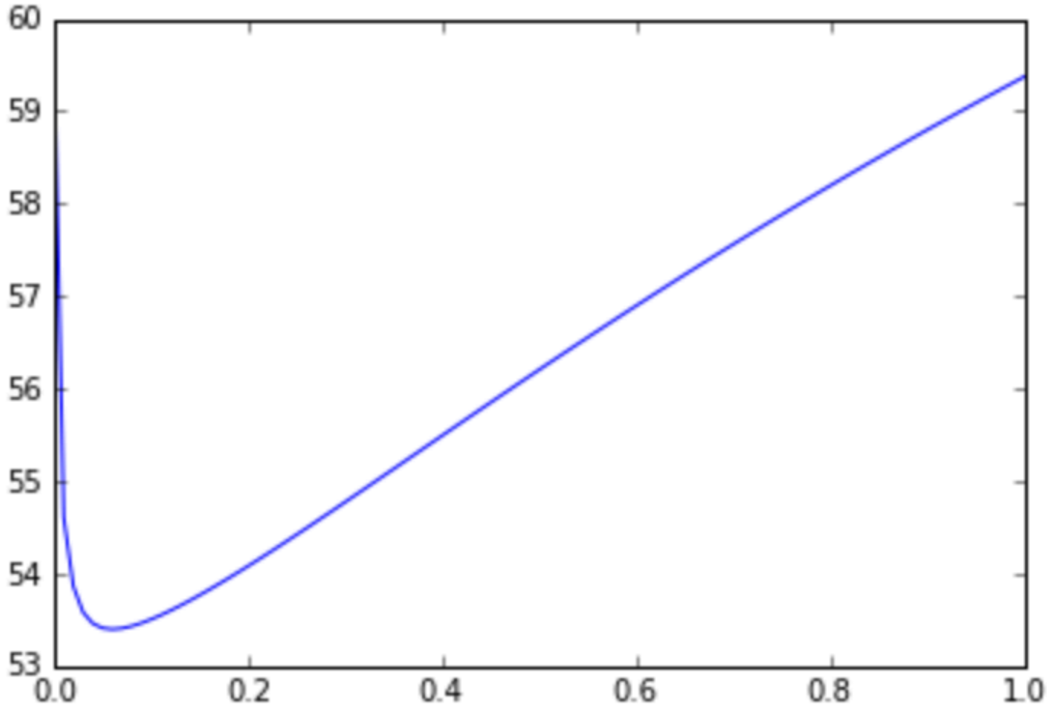
The value of lambda ranges from 0 to 1 increasing in steps of 0.01. In the table below, we have displayed the first 15 values of lambda.

Lambda	Testing Data	Training Data
0	60.8920371	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

On observing the values of RMSE for testing and training data, the lowest testing error is observed at the value of **λ = 0.06.** At this value, the errors converge and start increasing again on increasing values on lambda.

Problem 4: Ridge Regression using Gradient Descent

The plot for RMSE vs lambda for testing data for Ridge Regression using Gradient Descent is as below:



The lowest value for RMSE using gradient descent is 53.39784867. On comparing this value with the min value observed using Ridge regression in problem 3 (53.3978484), the value for RMSE is almost exactly similar.
Also, the optimal value of lambda observed in both the problems is same, i.e. **λ = 0.06.**

Problem 5: Non- Linear Regression

The plot for RMSE vs P-value is shown below:

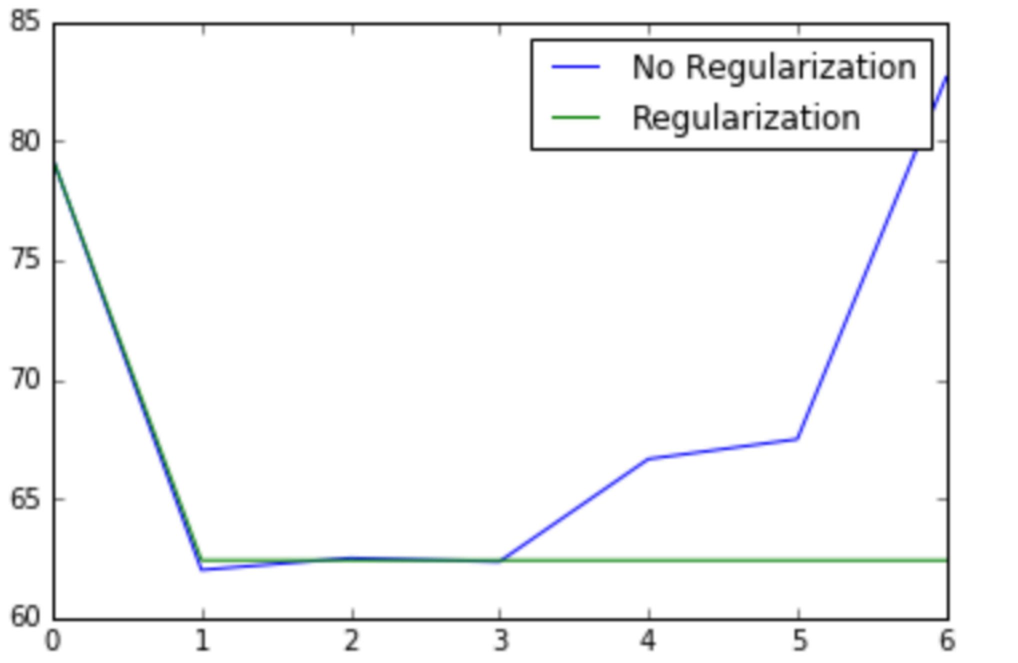


Figure 5: Comparison of p values against testing error with lambda value 0 and optimal value 0.06

P Value	Test Error with zero value of lambda	Test Error with optimal value of lambda
0	79.28685132	79.28986043
1	62.00834404	62.41679633
2	62.5070244	62.41461412
3	62.35363292	62.41460339
4	66.658292	62.41460301
5	67.48948346	62.41460301
6	82.66473945	62.41460301

It can be observed from the graph and the table that the test data optimal p is 1 for zero value of lambda and p is 4 for optimal value (0.06) of lambda.

When we do not use regularization, the error reaches its minimum at p=1. Then it slowly increases until p=3, after which it rapidly grows

and becomes even greater than p=0, case in which we use a horizontal line. The reason is that with higher order polynomials we have over fitting: the training error decreases, as figure 5 shows, but the learned curve is highly bound to the training data, and with a different data set the error steeply increases.
However, when we are using regularization, the error in this case decreases and reaches its minimum at p=5. But, it does not decrease much, as compared to the case with no regularization. This is because the regularization term penalizes high weights: therefore, even with high order polynomials, the correspondent weights with be very low and the resulting curve very smooth and ’linear’. Simpler curves work better in the general case, with data sets other than the train

Problem 6: Interpreting Results

The results obtained using the various regression techniques as summarized as below:

OLE Regression

RMSE for test data without intercept: 326.764994391
RMSE for test data with intercept: 60.892037097
RMSE for training data without intercept: 138.20074835
RMSE for training data with intercept: 46.7670855937

Ridge Regression

Optimal λ **= 0.06.**

Training Error: 46.76708559
Testing Error: 53.3978484

Ridge Regression with Gradient Descent

Optimal λ **= 0.06.**

Testing Error: 53.39784867

Non Linear Regression

P-Value	Test Error
1	62.00834404
4	62.41460301

Conclusion:

On comparing the errors obtained using the various methods, we can say that the Ridge regression or Ridge Regression with gradient descent is the best approach while performing linear regression as they give the lowest testing error (RMSE) on the trained data.