CS5691: Assignment 2

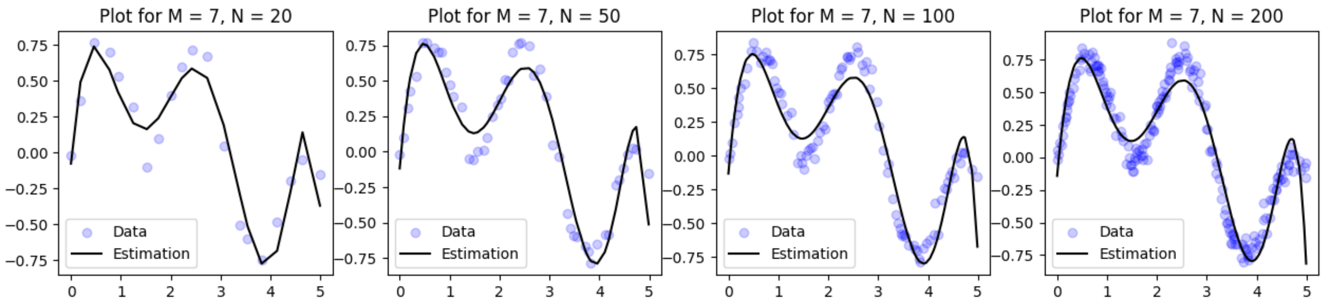
Akshat Meena (CS19B052)
Rohit Bhagat (CS19B038)

A. Regression

1. **1D Dataset**

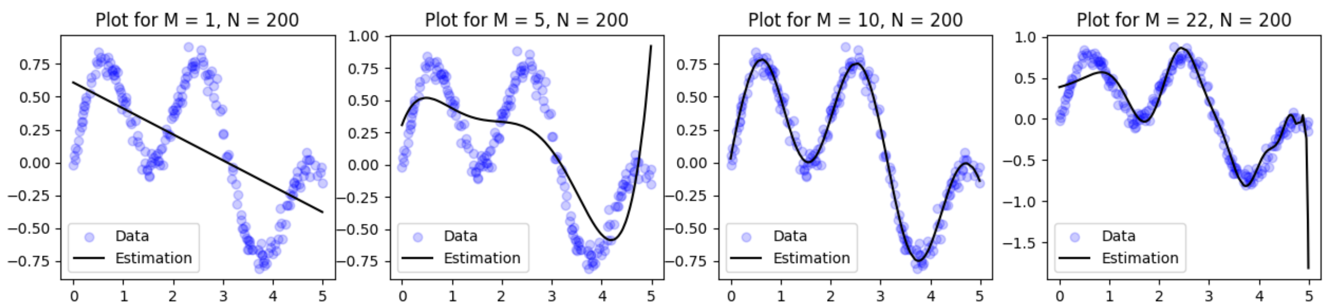
- **Least square regression**
  We experiment with different sample size keeping the complexity constant. We plot the least square regression for N = 20, 50, 100, 200 and M = 7. We observe that for lower value of N, the curve isn't smooth and may over-fit too depending on M. As we start increasing N, the curve starts getting smoother.

Figure 1: Plots for N = 20, 50, 100, 200 and M = 7.



Then we experiment with different model complexity keeping the size same.We plot the least square regression for M = 1, 5, 10 ,22 and N = 1000. We observe the for low model complexity the curve is not able to fit properly but as the M increase it starts fitting the data, if we increase the model complexity more then the curve start over-fitting.

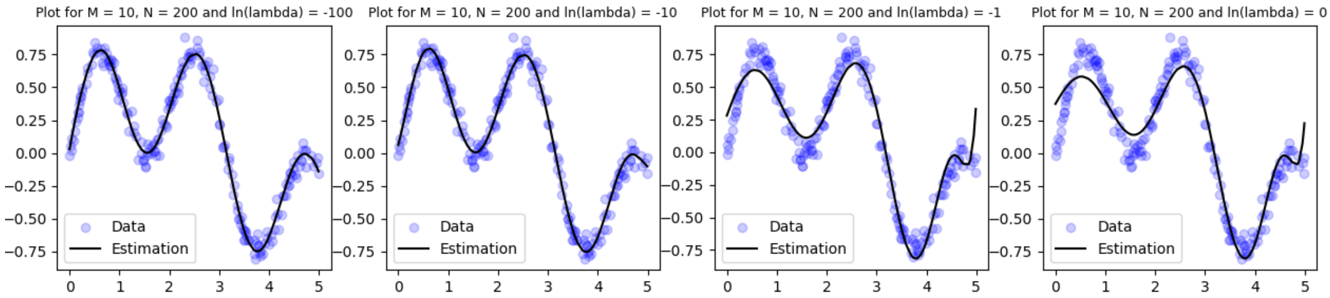
Figure 2: Plots for M = 1, 5, 10, 22 and N = 200.



- **Ridge regression**
  We experiment with different regularization parameter(λ) keeping the M and N constant. We plot the Ridge regression for $\ln(\lambda)$ = -100, -10, -1, 0. The regularization parameter helps dealing with over-fitting.

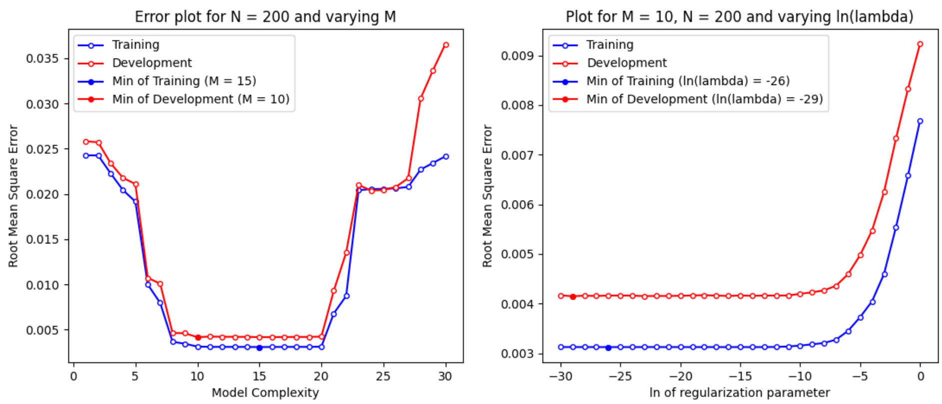
- **Error plots**
  We also plot the error of our model for training and development sets vs model complexity and

Figure 3: Plots for ln(λ) = -100, -10, -1, 0 and M = 10 and N = 200.



the regularization parameter. In the plot with varying M we see that initially the error decreases but later due to over-fitting it starts increasing. Similarly in the plot with varying regularization parameter however the plot shifts from over-fit to under-fit.

Figure 4: Error plots



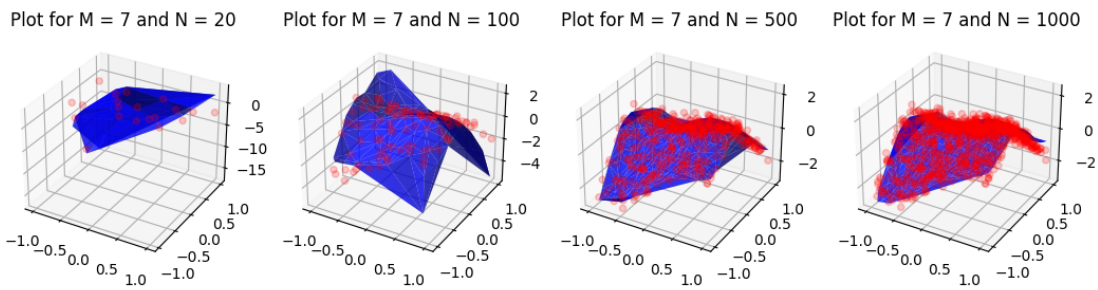
We can see that we can get best trained model for Model complexity(M) = 10 and regularization parameter(λ) = e^{-29}

2. 2D Dataset

- **Least square regression**

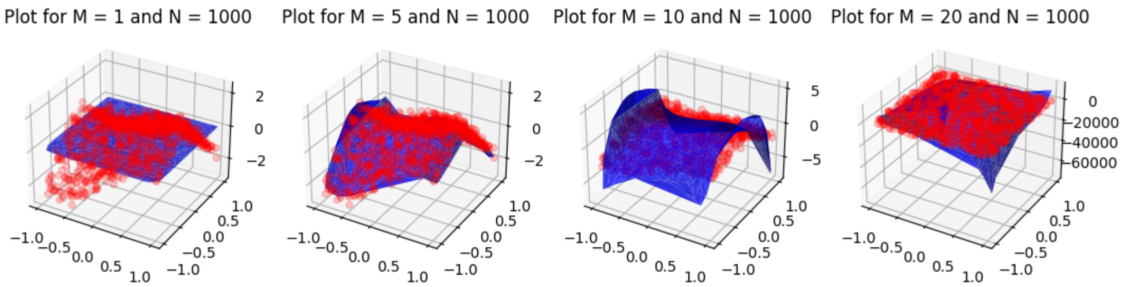
  We experiment with different sample size keeping the complexity constant. We plot the least square regression for N = 20, 100, 500, 1000 and M = 7. We observe that for lower value of N, the curve isn't smooth and may over-fit too depending on M. As we start increasing N, the curve starts getting smoother.

Figure 5: Plots for N = 20, 100, 500, 1000 and M = 7.



Then we experiment with different model complexity keeping the size same.We plot the least square regression for M = 1, 5, 10 ,20 and N = 1000. We observe the for low model complexity the curve is not able to fit properly but as the M increase it starts fitting the data, if we increase the model

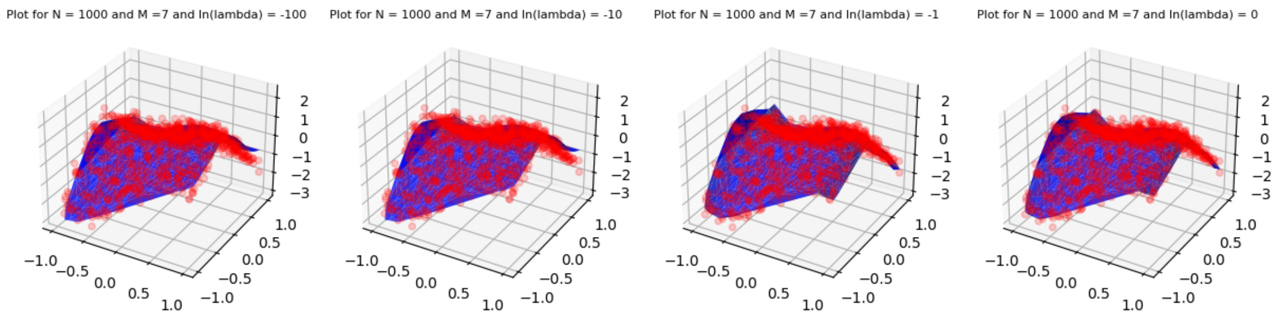
Figure 6: Plots for M = 1, 5, 10, 20 and N = 200.



complexity more then the curve start over-fitting.

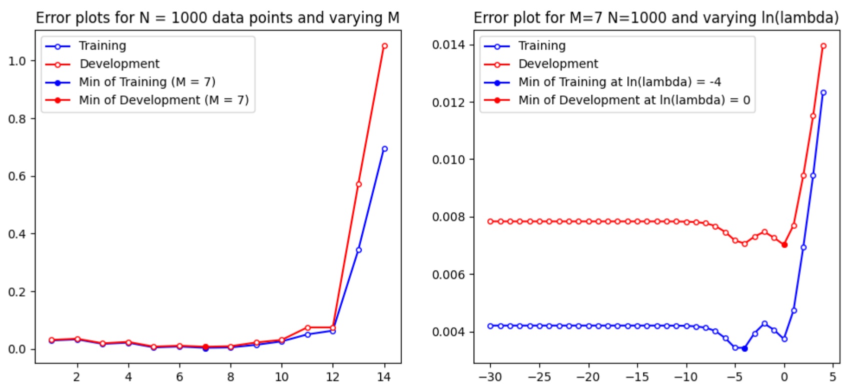
- **Ridge regression**
  We experiment with different regularization parameter(λ) keeping the M and N constant. We plot the Ridge regression for ln(λ) = -100, -10, -1, 0. The regularization parameter helps dealing with over-fitting.

Figure 7: Plots for ln(λ) = -100, -10, -1, 0 and M = 7 and N = 1000.



- **Error plots**
  We also plot the error of our model for training and development sets vs model complexity and the regularization parameter. In the plot with varying M we see that initially the error decreases but later due to over-fitting it starts increasing. Similarly in the plot with varying regularization parameter however the plot shifts from over-fit to under-fit.

Figure 8: Error plots



We can see that we can get best trained model for Model complexity(M) = 7 and regularization parameter(λ) = e^0

CS5691: Assignment 2

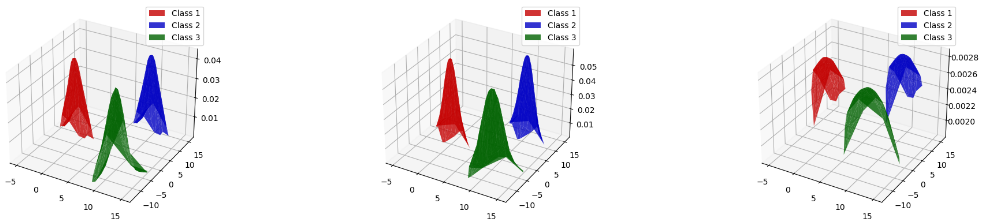
Akshat Meena (CS19B052)
Rohit Bhagat (CS19B038)

B. Bayesian Classifier

1. **Linearly Separable Data**

- We observe that classification of linearly separable data can be done very easily. We get 100% accuracy.

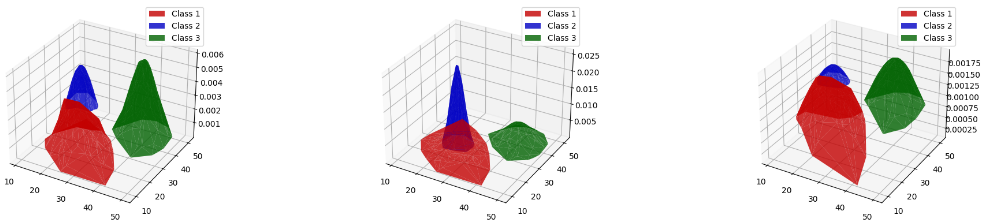
Figure 1: Plot of PDF for case 1, 2, 3 respectively



2. **Non-Linearly Separable Data**

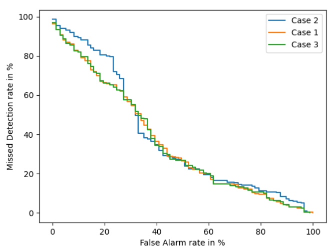
- Below is the Gaussian PDF plot for Non-Linearly separable data for different cases

Figure 2: Plot of PDF for case 1, 2, 3 respectively



- DET curve for Non-Linearly seperable data

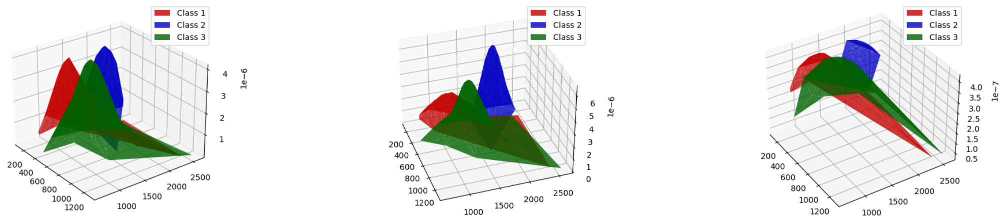
Figure 3: DET Plot



3. **Real Data**

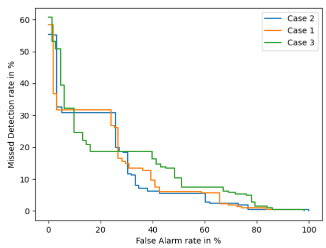
- Below is the Gaussian PDF plot for Real data for different cases

Figure 4: Plot of PDF for case 1, 2, 3 respectively



- DET curve for Real data

Figure 5: DET Plot



- Confusion Matrix for Real data

Figure 6: Confusion Matrix