Homework 2

Neural Network Signal Processing

Rishab Goel

February, 2, 2016

IRIS DataSet plots:

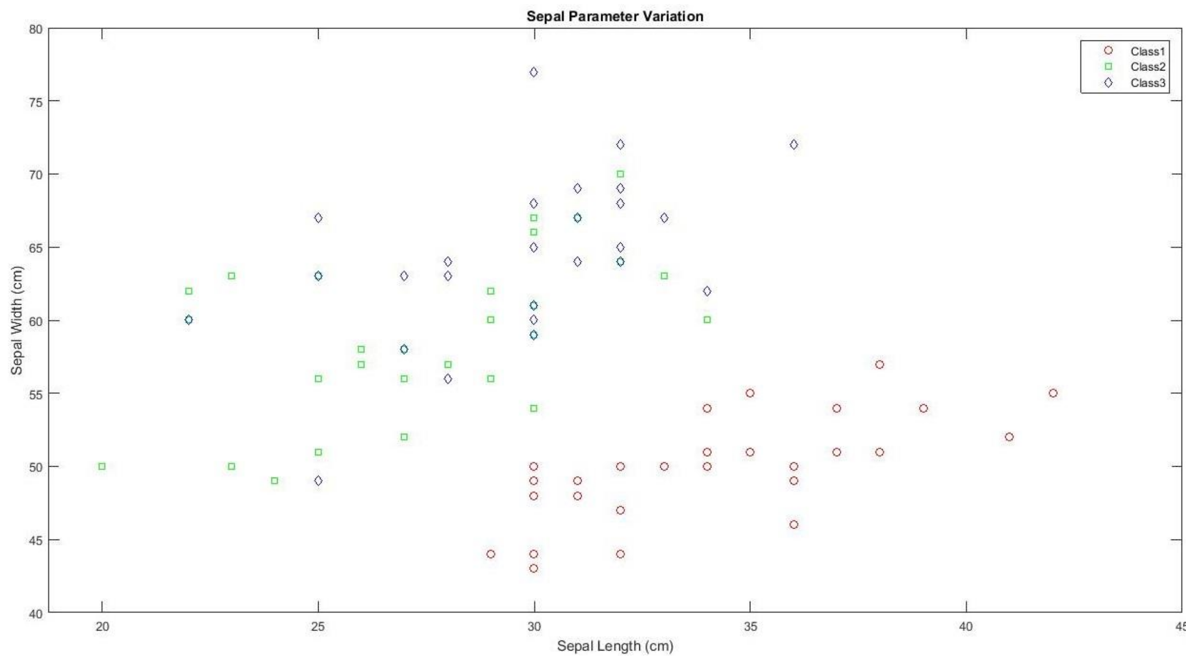

FIG 1

Training Data Plot Description: In Sepal Parameter Plot the class 2 and class 3 data is fused together to be distinguished. Class 1 is completely separable form from class 2 and class 3 on the basis of sepal parameters.

In Petal Parameter plot there appears to have distinct clusters of three classes: Class 1, Class 2 and Class3. Though there is a slight overlap in Class 2 and Class 3, but it seems from the plot there is avery little overlap in Class 2 and Class 3.

It seems clear from the plots that Petal Parameters could better classify the IRIS dataset than Sepal Parameters.

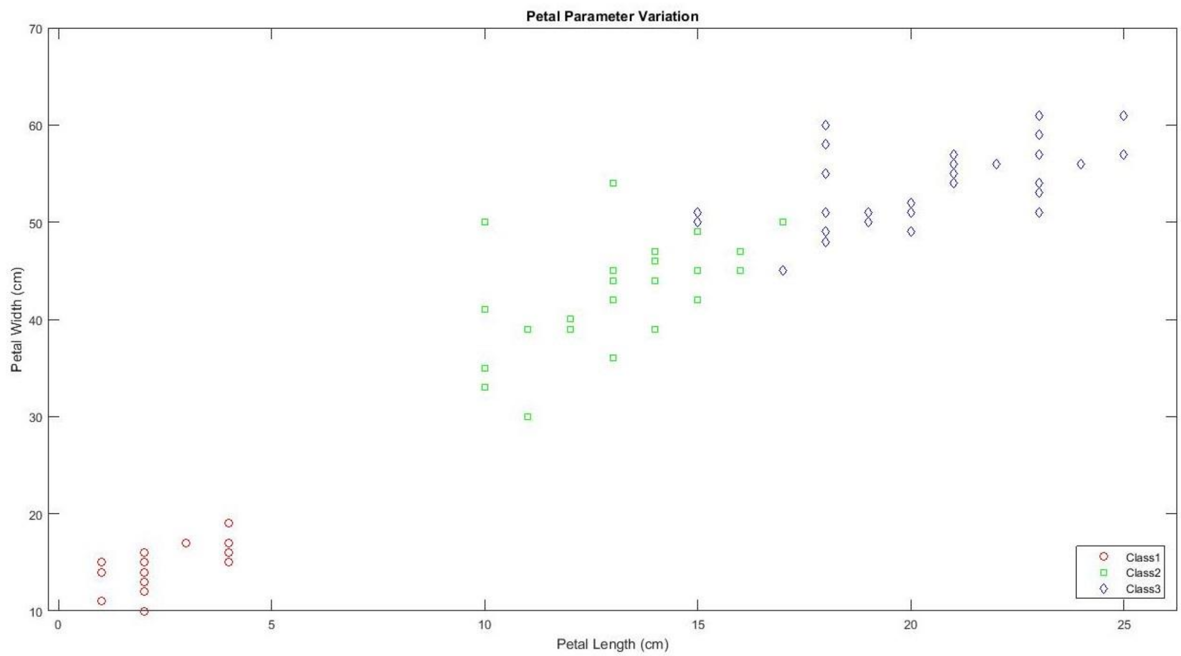

FIG 2

Bayes Classifier:

Assumption Made: It was evident from the FIG 1 and FIG 2 that petal parameter could be better to classify the data set so to minimize the complexity of Bayesian discriminant and visual realization of the decision boundaries we classify only using Petal parameters. The Given Sample was of 50 samples for each class, we have divided it in 60:40 ratio for train dataset and test dataset respectively.

bayes_discriminant: Function is used to calculate the bayes discriminant by calculating the mean and convariance and its covariance determinant are used to design the equation of the quadratic discriminant for the individual classes. The multivariate Gaussian distribution equation is used as the quadratic discriminant. Equating two discriminant equations together we get the decision boundary for each pair of classes. The plot of decision boundaries is shown below (FIG 3):

FIG 3: The plot description shows three decision boundaries: Class1|Class2 (Red) boundary separates Class1 Data from the Class 2 as the data for class 1 lies is separate cluster in bottom right corner. Class3|Class1 (Blue) boundary separates Class1 Data from the Class 3 since Class1|Class2 (Red) boundary already separates Class1 data so this boundary is not necessary for this data. Class2|Class3 (Green) boundary separates the slightly fused Class2 Data from the Class 3 and here we see some false predictions as per the plot.

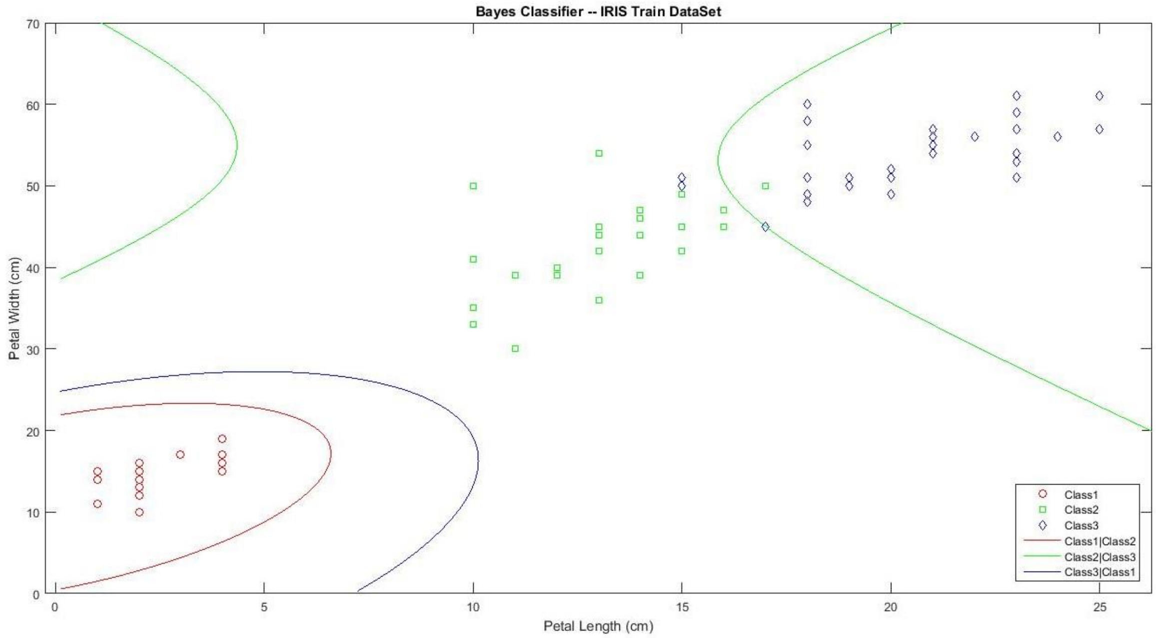

FIG 3

Testing the data (FIG4): We use the decision boundaries to classify the test data of total 60 samples by substituting the input values in the boundary equation the value for whichever class decision boundary is positive, it belongs to that particular class. Test Data plot in FIG4.

- Decision Boundary: Class1|Class2 > 0; Class1,
- Decision Boundary: Class2|Class3 > 0; Class2,
- Decision Boundary: Class3|Class1 > 0; Class3,

We generate that the Confusion Matrix to display the deviation in Actual Class (col) vs Predicted Class (row) and also output the percentage of false predictions for the test sample. The False predictions is quite less and only occurs for Class 2 and Class3 for some samples which crosses the decision boundary.

Output:

Parameters of Petals seems more efficient for classification than parameters of Sepal from the plot.

Confusion Matrix and Percent of False Predictions from IRIS Test Data of 60 samples:

Con_Mat =

20 0 0

0 18 3

0 2 17

False_Predictions = 8.3333 %

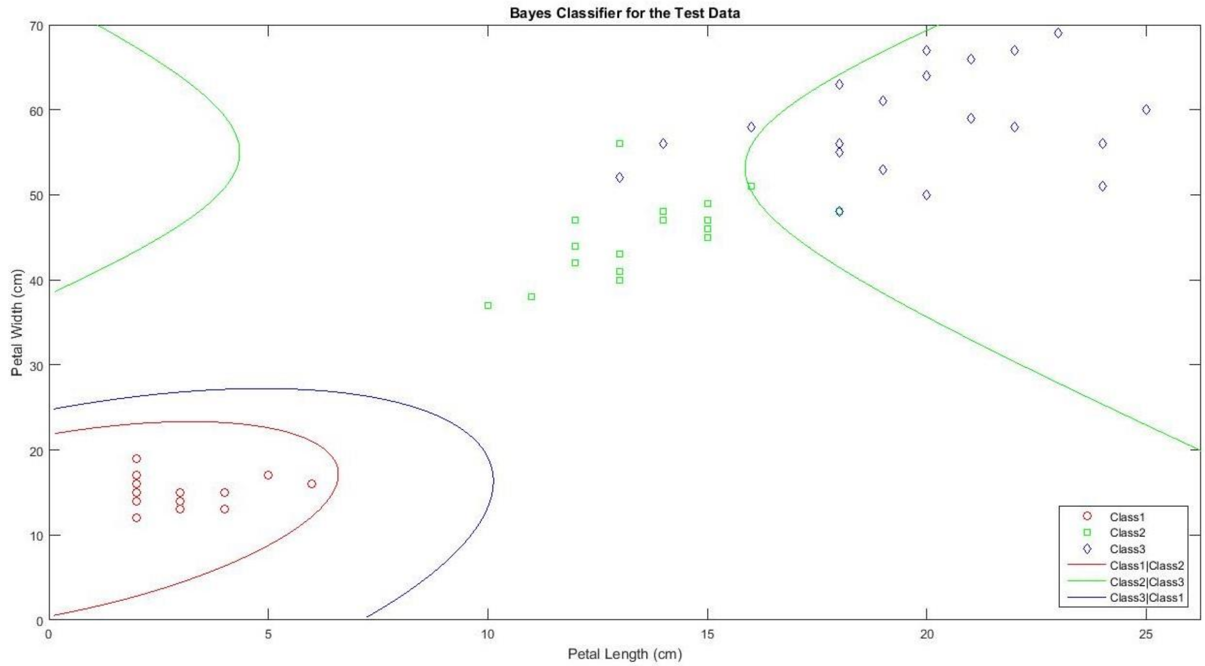

FIG 4

Linear Discriminant Analysis:

The concept of the Linear Discriminant Analysis states that there exists an eigenspace to which if the data transformed could be classified more accurately and efficiently. We use the both the petal and sepal parameter dataset for LDA classification as the transformation would be more accurate with all set of parameters for the data set.

We calculate the mean and covariance of each of the class using get_mean() and within_class_scatter() function followed by overall mean. We calculate the eigenvalues and eigenvector of Inverse(Sw)*Sb where the Sw is weighted sum of the covariance of each class while Sb the between class scatter matrix calculated using weighted sum of square the difference of individual means of each class and overall mean. We assume if two eigenvalues if occupies more than 90% of the eigenvalue vector sum than the eigenspace of their respective eigen vectors is well suited to classify the data. Linear discriminant analysis have an inherent assumption that the classes have the same covariance with a stable Gaussian estimate. We solve the system of equations of linear discriminant functions using mean, covariance and priori probabilities of each class to come out with a linear discriminant function for each class. The decision boundaries are computed by

- Decision Boundary: eq1- eq2,
- Decision Boundary: eq2-eq3,
- Decision Boundary: eq3-eq1

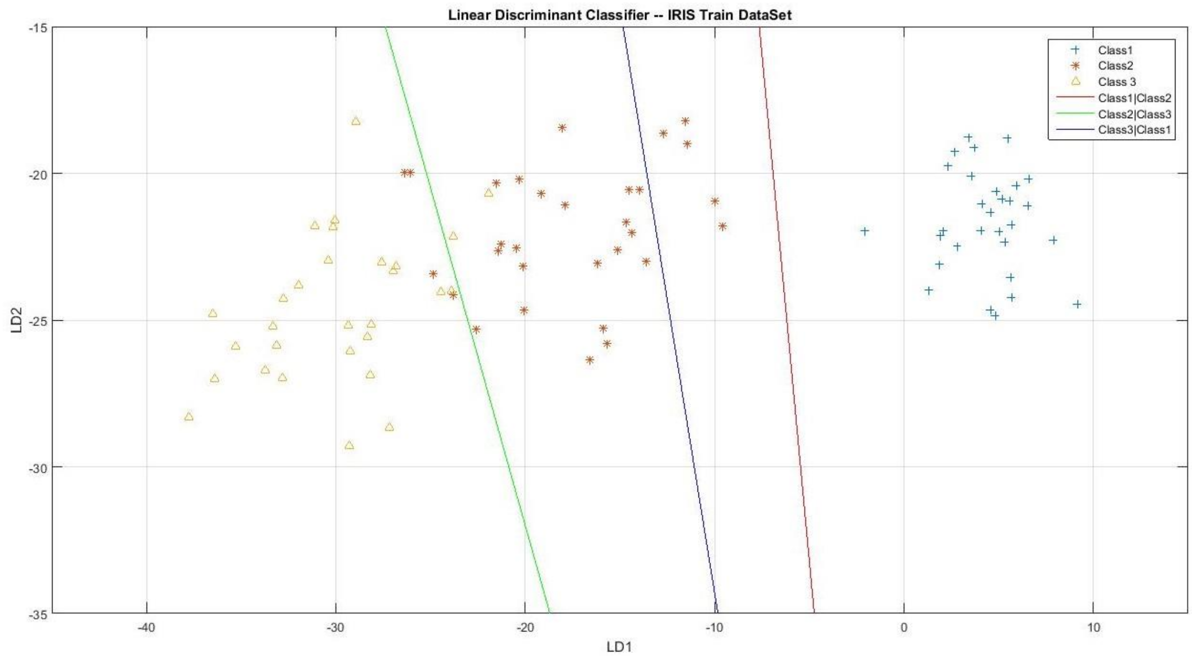

FIG 5

The FIG 5 shows the data in transformed eigenspace with clear and distinct linear decision boundaries for each of the classes.Class1|Class3(Blue) is not at all useful in terms of visualization for classification, the boundaries created by Class1|Class2 and Class2|Class3 decision boundaries are sufficient to classify the regions for each class.

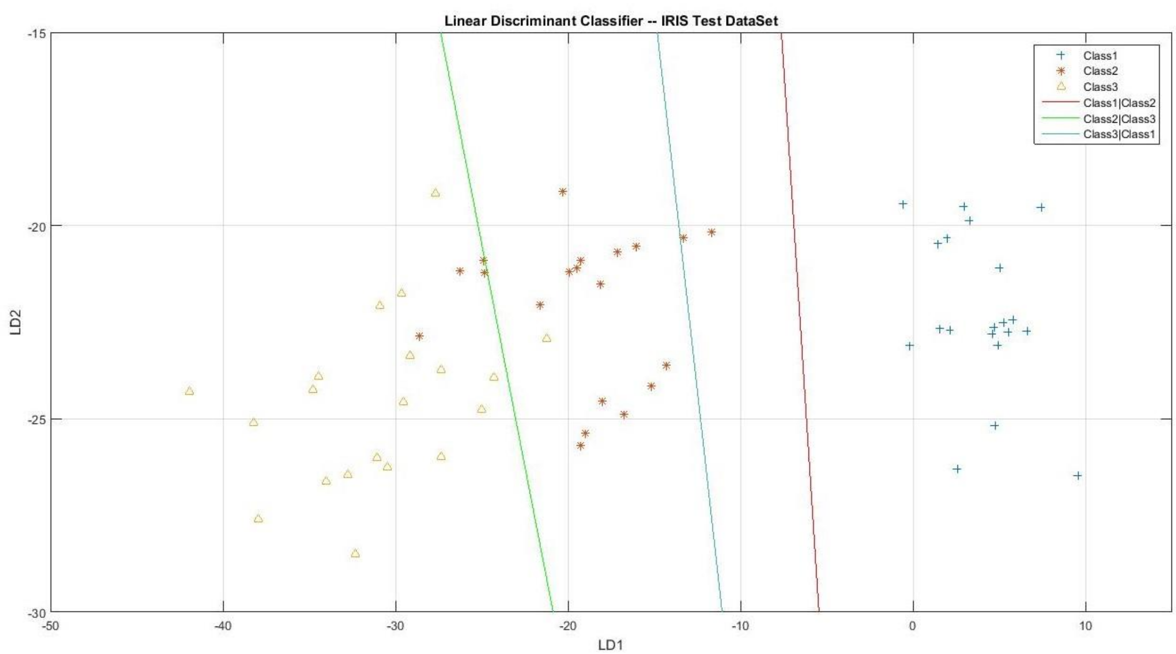

FIG 6

Testing the data :We use the weights biases of each discriminant functions (eq1, eq2, eq2) and for the test data is discriminant value is computed for each set of test data which val(eq1) , val(eq2) , val(eq3) is greater, the current dataset have a higher probability to belong to that class. So the higher probability class is the predicted class. We generate that the Confusion Matrix to display the deviation in Actual Class (col) vs Predicted Class (row) and also output the percentage of false predictions for the test sample. FIG 6 the plot of the test data with the decision boundaries. Class1|Class3(Blue) is not at all useful in terms of visualization for classification, the boundaries created by Class1|Class2 and Class2|Class3 decision boundaries are sufficient to classify the regions for each class. Note that false predictions for the samples are also very close to the decision boundaries while in case of Bayes Classifier there were slightly far away.

Output:

Confusion Matrix and Percent of False Predictions from IRIS Test Data of 60 samples:

Con_Mat =

20	*0*	*0*
0	*16*	*1*
0	*4*	*19*

False_Predictions = 8.3333

Conclusion:

Both LDA and Baye's Classifier have same percentage of false predictions with a slightly different confusion matrix. Optimal Bayes' Classifier Discriminant is considered theoretically more accurate than any linear discriminant analysis. But as evident from this dataset, linear discriminant could also provide competitive enough prediction with slightly reduced complexity of visualization and computation. It is to be noted that Bayes' Classifier is realized here with the reduction of the input dataset a compromise made for better visualization and reduced complexity compared to robustness and accuracy. It could be concluded that though Bayes' Classifier discriminant may be optimal but the complexity in visualization and computations presents the linear discriminant as a good alternative with a slight or no compromise false predictions. Though the dataset with classes having nearly same covariance, stable Gaussian estimates and linear decision boundaries are rare, but Linear discriminant analysis could be used if the dataset is closer to meeting the above criteria. We could not say which one is better for this IRIS dataset as false predictions are same, but we could LDA is slightly easy to visualize.