# Part 3

The final section of the lab involves classifying two non-gaussian distributions using a sequential discriminant approach of the MED classifier in order to create a non-linear, composite MED decision boundary. This involves choosing a prototype from each class to create a possible MED decision boundary and reiterate this process until the two classes are fully classified.

We ran the sequential algorithm three times to understand how the discriminants are developed. For each sequence, a series of discriminants is created. The slight variations in the discriminants are a result of new prototypes being randomly chosen, creating slight variations in the parameters of the discriminant. Figures \_ through \_ show the resulting discriminants.

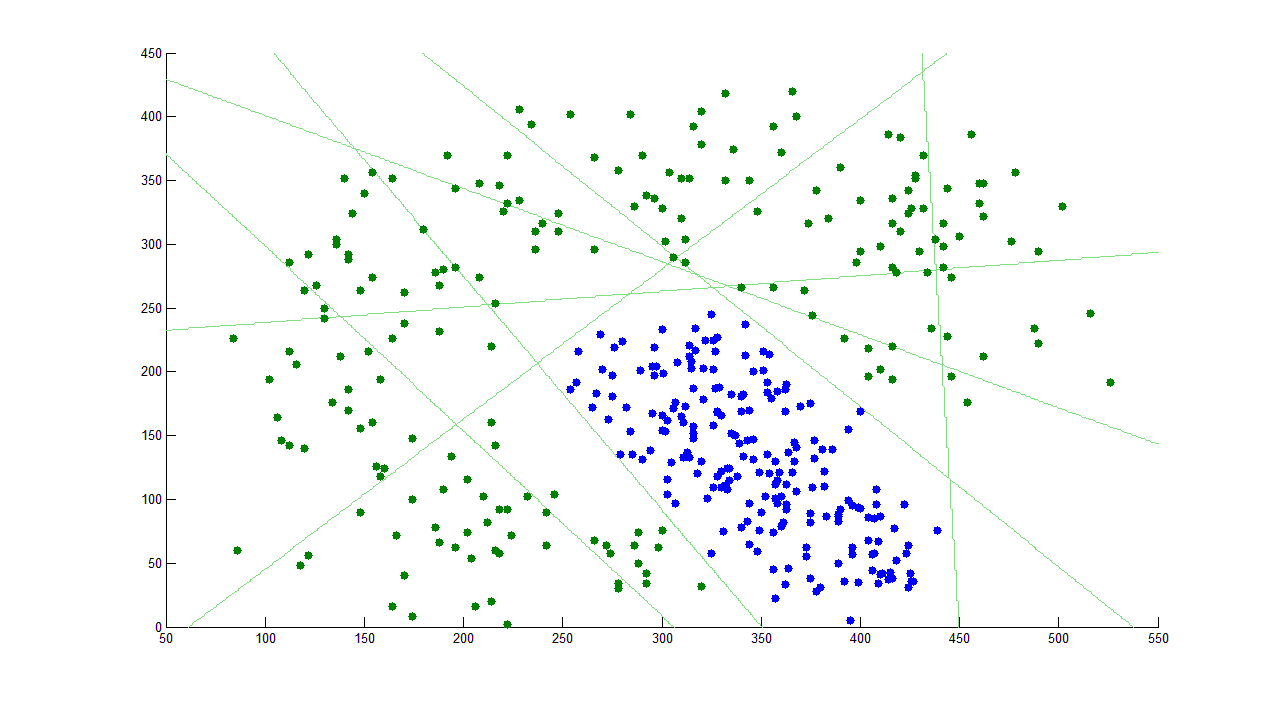


Figure : first attempt at sequential classifier

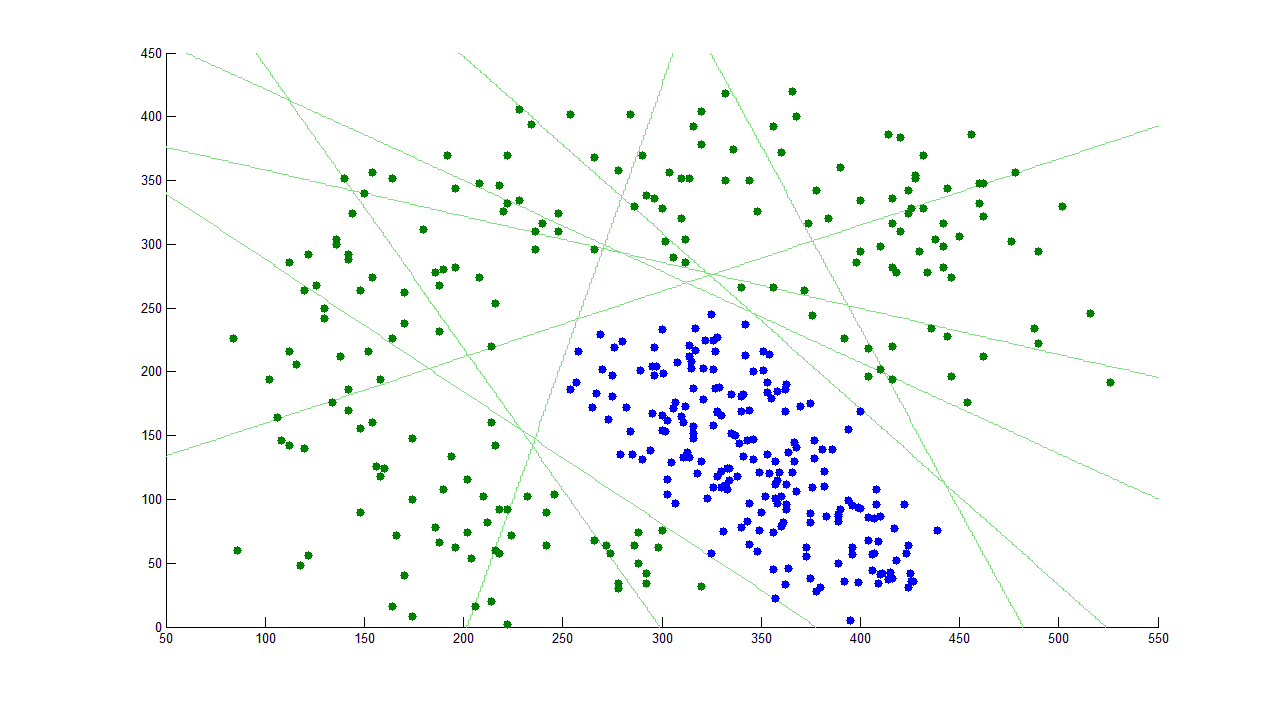


Figure : Second attempt at sequential classifier

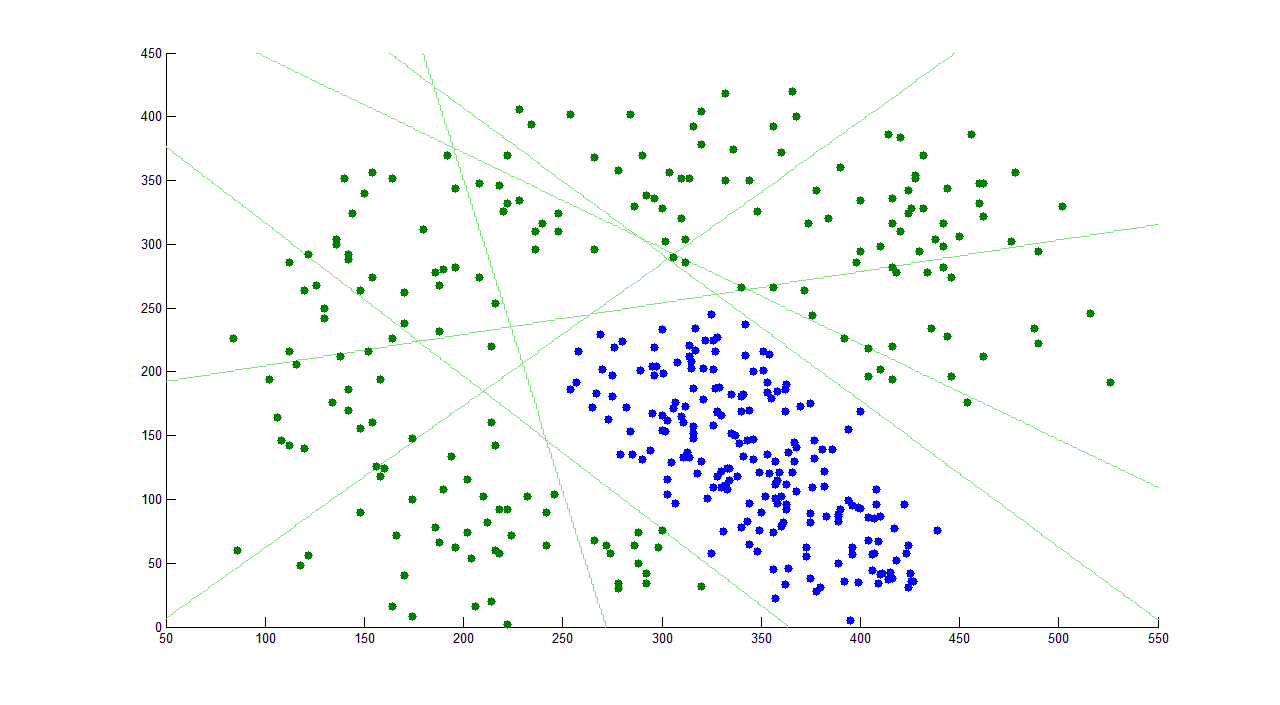


Figure : Third attempt at sequential classifier

Testing the classifier using the training data should always give zero errors. This is a natural consequence of the discriminants being created explicitly to fit the training data. When creating the filter, new discriminants are created, eliminating the successfully classified points, until there are no more points to classify.

To understand the the impact of the number of classifiers on the error rates, we run the sequential classifier with a limited number of classifiers J=1,2,...,5. Repeating the process 20 times. This results in the data in table 1.

Table : Errors for each increment of the number of classifiers

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| J | 1 | 2 | 3 | 4 | 5 |
| Average error rate | 0.2884 | 0.3528 | 0.4259 | 0.4645 | 0.4881 |
| Minimum error rate | 0.2175 | 0.2475 | 0.3125 | 0.4200 | 0.4650 |
| Maximum error rate | 0.3375 | 0.4275 | 0.4850 | 0.4950 | 0.4975 |
| Standard deviation of error rates | 0.0430 | 0.0471 | 0.0447 | 0.0169 | 0.0096 |

The relation between the error rates as a function of J classifiers is shown in Figure 4 below. Although the average error rate increases as J increases, the standard deviation decreases. This is also reflected in how the minimum and maximum error rates begin to converge. This indicates that the MED classifier has a constant error rate as J approaches infinity.

Figure : Error as a function of J

By limiting the number of point pairs to be used as prototyping, the amount of error will increase because the MED classifier will be unable to classify all test points perfectly as J approaches zero.

## Conclusion

The sequential classifier was learned three times to show the variant decision boundaries that are created as a composite of all the MED decision boundaries. The final decision boundary created a non-linear boundary that successfully classified all the data points of each non-Gaussian distribution.

When limiting the number of prototype pairs used to test MED classifiers, the average error rate increased, however the standard deviation decreased with it, showing less variation in the error.