Challenge 5 - Support Vector Machine

Kanchan Satpute
Texas A&M University

I. INTRODUCTION

A Support Vector Machine (SVM) can be used as a discriminative classifier formally defined by a separating hyperplane. In other words, given labeled training data, the algorithm outputs an optimal hyperplane which categorizes the data. If the data is separable, we will have a linear hyperplane. But if the data is non-separable, we have to use kernel trick which does some extremely complex data transformations and converts the non-separable problem to a separable problem.

II. MODEL

We have built a binary classifier using SVM on a dataset consisting of 200 observations and 2 class labels. We have explored various kernel methods to see how they perform on this dataset. We have divided the entire data into training and validation dataset and used grid search method to optimize the parameters of the model. We found that the RBF kernel gave best performance on our model with minimum misclassification on the validation dataset.
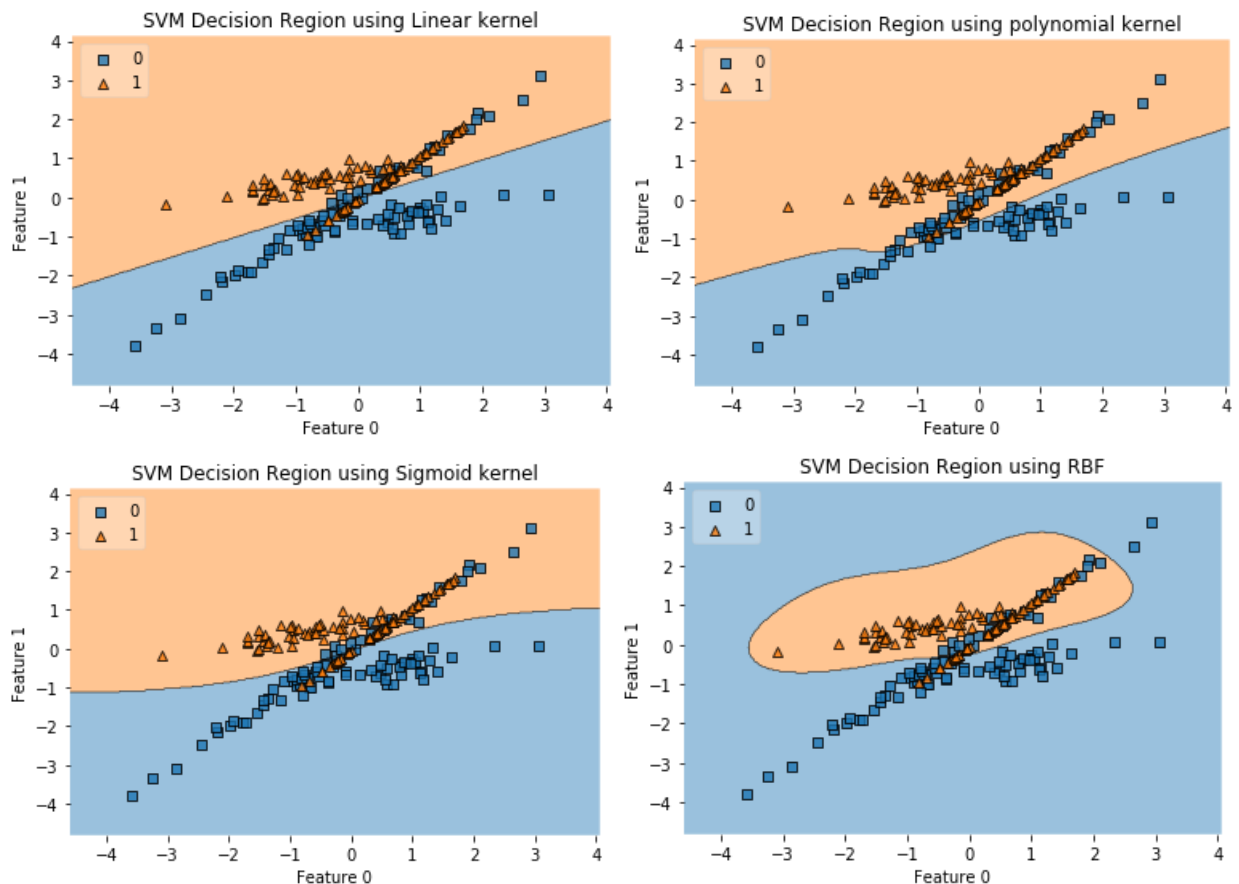
From the figure above, we can visualize the performance of each kernel function on the data. The data certainly isn't linearly separable and hence the linear kernel and the sigmoid kernel were bound to perform bad compared to the other kernels. As the order of polynomial kernel increases, the size of the function class increases. The radial basis function kernel (gaussian kernel) uses a squared exponential term in its function and hence it defines a function space that is a lot larger than that of a linear or polynomial kernel. So, in some sense, the RBF kernel is as powerful as an infinite order polynomial kernel.

III. CONCLUSION

RBF kernel is a general-purpose kernel which is used when there is no prior knowledge about the data. It performed better than the other kernels and hence we used it for building our model and for prediction on the test data.

REFERENCES

[1] https://medium.com/machine-learning-101/chapter-2-svm-support-vector-machine-theory-f0812effc72
[2] https://www.quora.com/Why-does-RBF-kernel-generally-outperforms-linear-or-polynomial-kernels