

ECEN 689 - Applied Information Science

Challenge 5 - Binary SVM Classifier

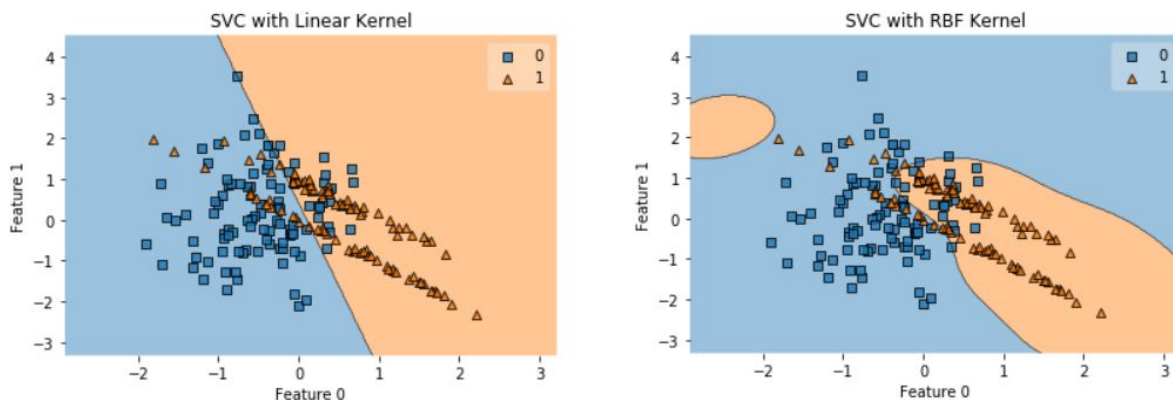
Mahalakshmi Sridharan

Abstract

The objective of this challenge was to build a binary classifier based on Support Vector Machine by experimenting with different kernels for optimal boundary decision. The training dataset consisted of 200 data points with 2 features each.

Performance of different Kernels

The four most common kernels for SVC viz. linear, polynomial, sigmoid and RBF were employed with their default parameters initially to decide on which kernel provides the best classification boundary. The accuracy scores for the kernels were found to be 0.78 for linear kernel, 0.725 for polynomial, 0.6 for sigmoid and 0.79 for RBF. It is observed that the accuracy scores for linear and RBF are the highest with RBF outperforming linear by a very small margin. A simple observation of the plot of the dataset shows that the linear kernel could not be optimized further with considerable improvement in the accuracy whereas there is scope for optimizing the performance of the RBF kernel.



Optimizing the RBF Kernel

The c and γ parameters of the RBF kernel were tuned using cross validation and grid search. The parameters after optimization were 1.0 for c and 10.0 for γ with a cross validated accuracy score of 0.835. Applying this optimized parameters for the RBF kernel gave an model accuracy of 0.895. This optimized RBF kernel SVC was used to classify the test data set.

