

ECEN 689

Challenge 5: Support Vector Machines

Devarsh Jhaveri

Abstract- The report describes how the Support Vector Machines are used in the most basic sense. Different parameters are used to optimize the performance resulting in the final model.

I. INTRODUCTION

The problem is a binary classification task, where the training data consists of 200 data points with 2 features which are classified into two classes i.e. 1 and 0 respectively. The problem focus on the optimization of the SVM using the kernel trick.

II. METHODS

The Scikit-Learn Support Vector Machine package was used in the analysis and to implement the SVM task [1]. A 10-Fold Cross Validation technique was used in order to approximate the test accuracy using different kernel's. Linear, polynomial with degree 3, radial basis function (RBF) and sigmoid kernel were used for comparison.

III. UNITS

The radial basis function (RBF) kernel was the best performer. A parameter grid search [1] with 10-fold cross validation was conducted over the parameters C and gamma between the range of 0.1 and 20. The best cross-validation accuracy achieved was 0.79. The cross-validation accuracy never increased above the 0.79 threshold for the training data irrespective of the different kernels used. Hence, the parameters were fixed and the SVM with RBF kernel was applied on the testing data in order to predict the classes.

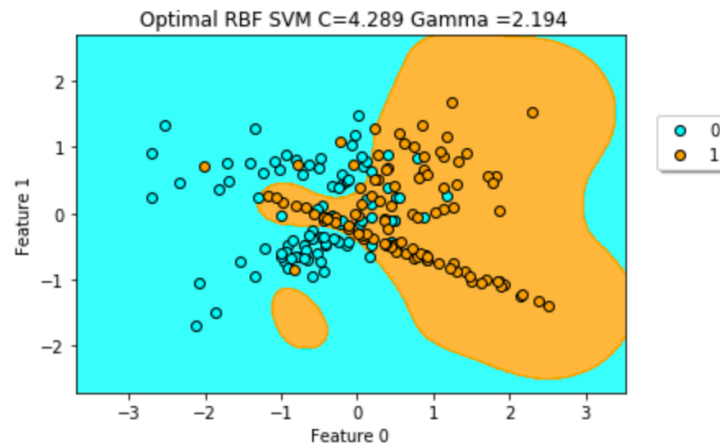


Fig 1: SVM decision boundary juxtaposed with the training data

REFERENCES

- [1] [Scikit-learn: Machine Learning in Python](#), Pedregosa *et al.*, JMLR 12, pp. 2825-2830, 2011.