Project 2
Classification

- CAP6610 - Machine Learning - Spring 2018 -

Project Report

Group 3

Connor, McCurley
Dylan, Stewart
Charles, Thelijjagoda Subodha
Jianfeng, Chen

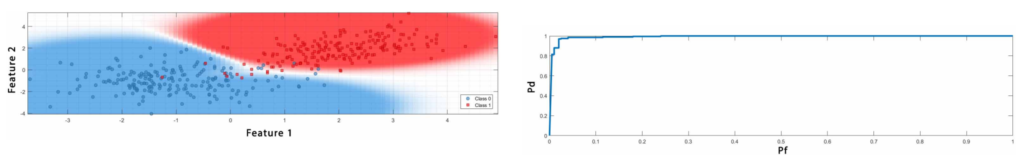
University of Florida

1 EXPERIMENTS

For each algorithm we performed 5 fold cross validation and trained and validated on 80% of the data and saved 20% for testing. All algorithms classified the features we were given well for each of the two class problems as evident from results in Section 2.

1.1 Relevance Vector Machines (RVM)

Verification of RVM functionality was performed using a synthetic data set provided in the Matlab Pattern Recognition Toolbox. Figure 1 demonstrates a non-linear RVM classifier on two classes, red and blue. The classifier exhibited reveals a correct-classification accuracy of 98.0%. Following performance verification, the RVM classifier was trained with the 60-dimensional data set provided.



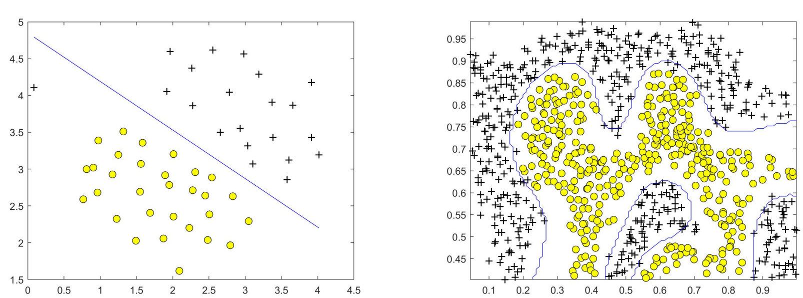
(a) Relevance Vector Machine (prtDataGenUnimodal)

(b) ROC

Fig. 1. Non-linear RVM classifier on 2D data. The white space between the two classes (red and blue) demonstrates the decision boundary.

1.2 Support Vector Machines (SVM)

To visualize classification using SVM, we tested the algorithm with two synthetic test cases - a linear classifier was trained using SVM on one dataset and a non-linear classifier was developed using a Gaussian Kernal on an SVM [1–3]. The performance of SVM on the two classification problems are shown in Figure 2.



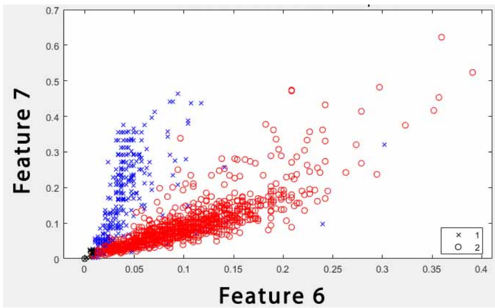
(a) Linear classifier performance. Accuracy : 98.03%.

(b) Non-linear classifier performance. Accuracy : 98.96%.

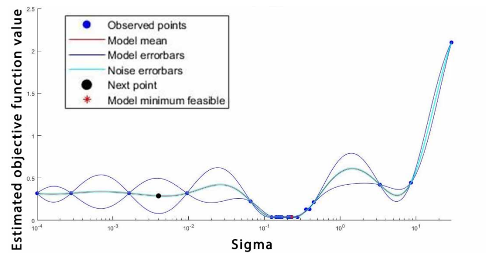
Fig. 2. SVM Results on 2D data. Blue line denotes the classification boundary of the trained model.

1.3 Gaussian Process Regression (GPR)

Figure 3a demonstrates a GPR classifier on two classes - red and blue which are randomly selected from the data. Features 6 and 7 are selected to visualize the result. Black points are mislabeled data points. The classifier exhibited a correct-classification accuracy of 98.6%. Figure 3b shows how the matlab function *fitrgp* optimizes the hyper-parameters.



(a) GPR classifier on two classes



(b) Results of Matlab function fitrgp which optimizes hyper-parameters

Fig. 3. GPR Results on 2D data

2 OBSERVATIONS

As 5-fold cross-validation is required, and we used all-pairs method, it gives $\binom{5}{2} = 10$ classifiers for each algorithm. The accuracy (Table 1) and confusion matrices (Table 2) for each classifier with each algorithm are shown below.

Table 1. Accuracy of the 10 classifiers for each algorithm for the given dataset

Algorithm	1	2	3	4	5	6	7	8	9	10
RVM	98.7	98.5	98.8	98.2	98.6	95.8	99.3	99.0	98.4	98.1
SVM	99.1	98.3	99.1	99.3	99.2	92.9	99.6	98.4	97.1	97.3
GPR	99.3	99.5	99.6	99.5	99.7	97.6	99.8	99.4	99.3	99.1

Table 2. Confusion matrices of the 10 classifiers for each algorithm for the given dataset

Algorithm	1	2	3	4	5
RVM	[484,6;7,503]	[495,6;9,490]	[489,7;5,499]	[485,11;7,497]	[499,7;7,487]
SVM	[491,4;5,500]	[479,16;1,504]	[490,6;3,501]	[485,6;1,508]	[491,3;5,501]
GPR	[502,4;3,491]	[477,2;3,518]	[494,2;2,502]	[490,4;1,505]	[504,0;3,493]

Algorithm	6	7	8	9	10
RVM	[488,13;29,470]	[484,0;7,509]	[492,4;6,498]	[503,4;12,481]	[477,6;13,504]
SVM	[458,38;33,471]	[516,1;3,480]	[484,14;2,500]	[493,19;10,478]	[493,8;19,480]
GPR	[479,10;14,497]	[505,0;2,493]	[501,4;2,493]	[504,2;5,489]	[490,1;8,501]

The results show that all three algorithms perform well and give comparable results. We noticed that RVM gives the final classification decision faster than SVM. This is consistent with previous studies [4]. Generally, RVM and SVM both have comparable accuracy. Yet, RVM is much sparser and when the training set size grows, the number of relevance vectors increases slower than that of SVM.

GPR can learn the kernel parameters automatically from data, no matter how flexible we wish to make the kernel. It can incorporate interpretable noise models and priors over functions. Moreover, it can sample from the prior to get intuitions about the model assumptions. Compared to SVMs, GPR offers several advantages: learning the kernel and regularization parameters, fully probabilistic predictions, and interpretability.

REFERENCES

[1] Ashis Pradhan. Support vector machine-a survey. *International Journal of Emerging Technology and Advanced Engineering*, 2(8):82–85, 2012.

[2] Mathias M Adankon and Mohamed Cheriet. Support vector machine. In *Encyclopedia of biometrics*, pages 1303–1308. Springer, 2009.

[3] Shan Suthaharan. Support vector machine. In *Machine learning models and algorithms for big data classification*, pages 207–235. Springer, 2016.

[4] Xu Xiang-min, Mao Yun-feng, Xiong Jia-ni, and Zhou Feng-le. Classification performance comparison between rvm and svm. In *Anti-counterfeiting, Security, Identification, 2007 IEEE International Workshop on*, pages 208–211. IEEE, 2007.