

CSSE463 Image Recognition

Lab 5 SVM

Zesun Yang

yangz@rose-hulman.edu

December 24, 2017

1. Introduction

In this lab, I studied and analyzed the demo function in the SVM toolbox, and learned how to train SVMs for classifying points. By testing data with different kernel functions, I obtained different testing results. In the lab report, I included the best results I can get with specific kernel settings, even though my classifier could not achieve 100% accuracy.

2. Simple Test Set Result

a. Parameter settings:

Kernel function: rbf with kernel scale of 4 with boxConstraint of 80

b. Table:

	True Yes	No
Detected Yes (1)	39 (true pos)	1 (false pos)
No (-1)	4 (false neg)	86 (true neg)

$$\text{TPR} = 39 / (39 + 4) = 90.7\%$$

$$\text{FPR} = 1 / (1 + 86) = 1.14\%$$

$$\text{Accuracy} = (39 + 86) / (39 + 1 + 4 + 86) = 96.15\%$$

c. Figures

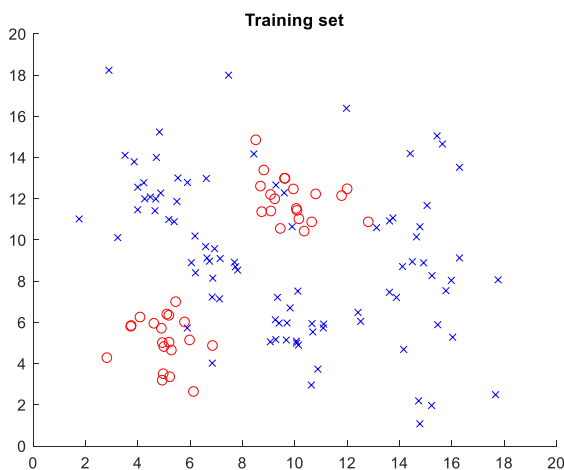


Figure 1. Simple Training Set

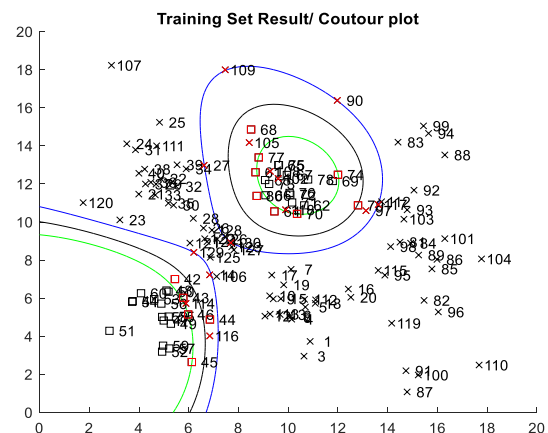


Figure 2. Simple Decision boundary

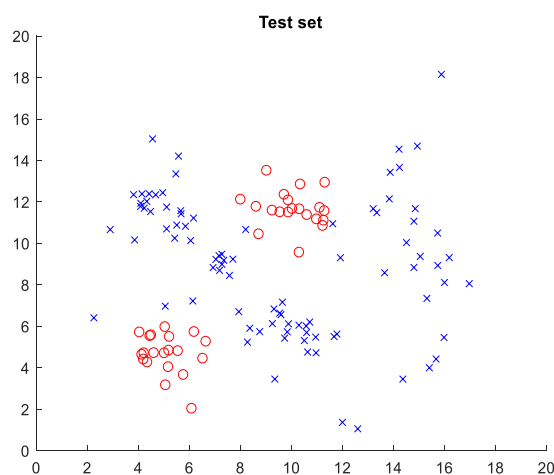


Figure 3. Simple Test Set

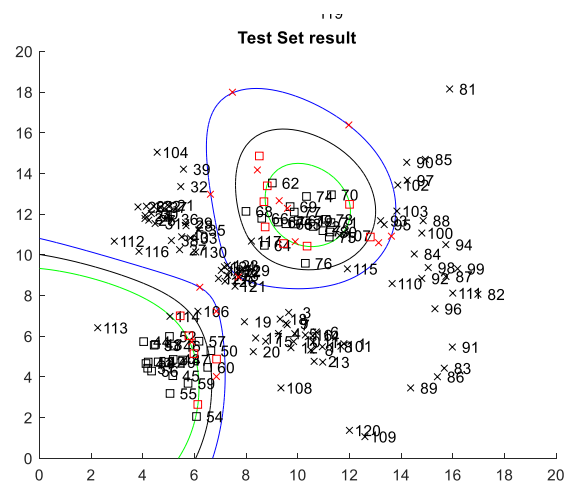


Figure 4. Test set result

3. Harder Test Set Result

a. Parameter Setting

Kernel function: rbf with kernel scale of 3 with boxConstraint of 40

b. Table

Detected \ True	Yes	No
	Yes (1)	No (-1)
Yes (1)	80 (true positive)	20 (false positive)
No (-1)	4 (false negative)	96 (true negative)

TPR: $80/84=95.2\%$

FPR: $20/116=17.2\%$

Accuracy= $(80+96)/(4+20+80+96)=88\%$

c. Figures

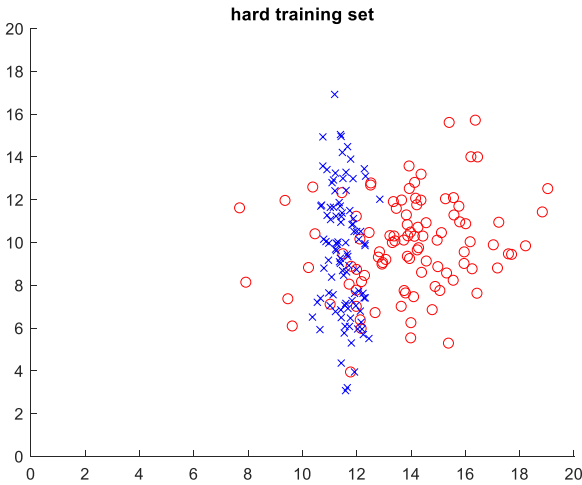


Figure 5. Hard Training Set

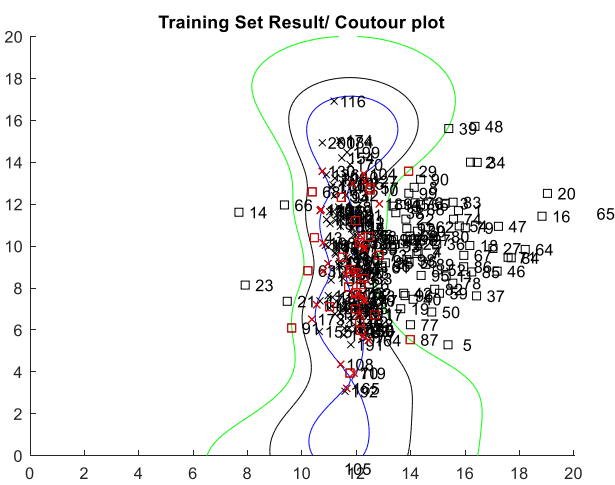
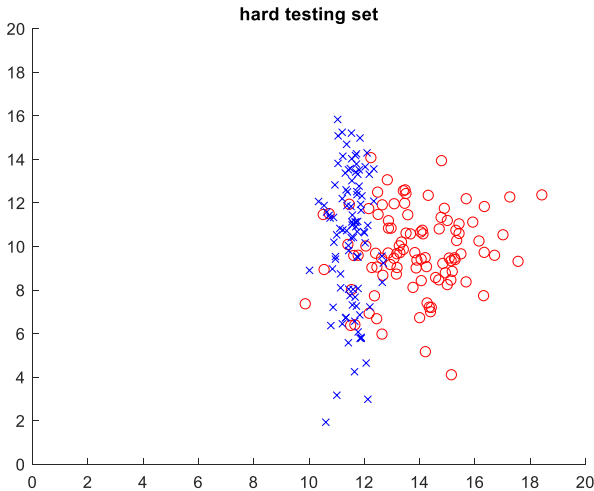


Figure 6. Training Result/



Contour plot

Figure 7. Hard Testing Set

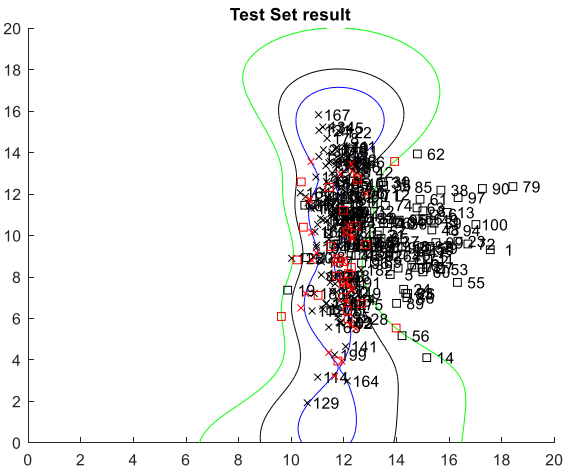


Figure 8. Hard sample result

4. Discussion

Overall my classifier was successfully. The classifier can achieve a high accuracy on simple data set (96.15%) with an adjusted kernel function, and it achieved a lower accuracy (88%) on the hard data set because points are in cluster.

I have also tried the polynomial kernel function (as listed on MatLab website), but it did not provided better accuracy compared to the rbf kernel.

In both data sets, some data did not have clear boundary and thus, made the SVM not achieving 100% accuracy. Smaller box constrains can be helpful when data points are clustered (as applied in the hard data set). I also verified that my classifier was not “over-fitting” any of these data sets. Another reason of the classifier not achieving 100% accuracy could be the ROC curve threshold setting. The default threshold is 0, and there should exist a better threshold to obtain higher accuracy.