Nicholas Kamper and Eric Henderson

CSSE463 – Image Recognition

Lab 5 – Support Vector Machines

# Abstract/Procedure

For this lab, we utilized the sample code provided in toyProblem.m as a starting point. We ran the SVM on the clustered and Gaussian data sets, changing the parameters to the kernel function as necessary to optimize the training for the type of data set.

For both data sets, we found that the ‘rbf’ (Radial Basis Function) kernel worked reasonably well. As noted in toyProblem.m, a linear kernel does not work and we found the results provided by the ‘poly’ (polynomial) kernel were not able match the accuracy of the Radial Basis Function.

# Results

We performed SVM training and testing on both the clustered and Gaussian data sets, with better results for the clustered data set.

## Clustered

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Total Points | True Negative | True Positive | False Negative | False Positive | Recall (TPR) | FPR | Accuracy |
| RBF/p=10 | 130 | 86 | 40 | 0 | 4 | 100% | 4.44% | **96.92%** |
| Poly/n=2 | 130 | 74 | 37 | 3 | 16 | 92.5% | 17.78% | 85.38% |
| Poly/n=3 | 130 | 82 | 38 | 2 | 8 | 95% | 8.89% | 92.31% |

## 

Figure /2 - RBF (left) and Polynomial w/ n=2 (right)

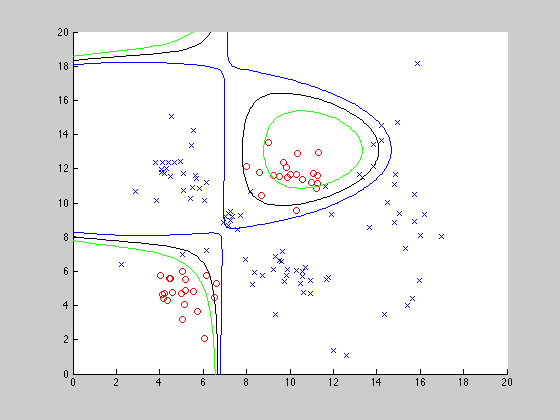


Figure 3 – Polynomial w/ n=3

## Gaussian

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Total Points | True Negative | True Positive | False Negative | False Positive | Recall (TPR) | FPR | Accuracy |
| RBF/p=10 | 200 | 97 | 77 | 23 | 3 | 77% | 3.0% | **87.0%** |
| Poly/n=2 | 200 | 100 | 0 | 100 | 0 | 73% | 1.0% | 86.0% |
| Poly/n=3 | 200 | 100 | 1 | 99 | 0 | 1% | 0% | 50.5% |

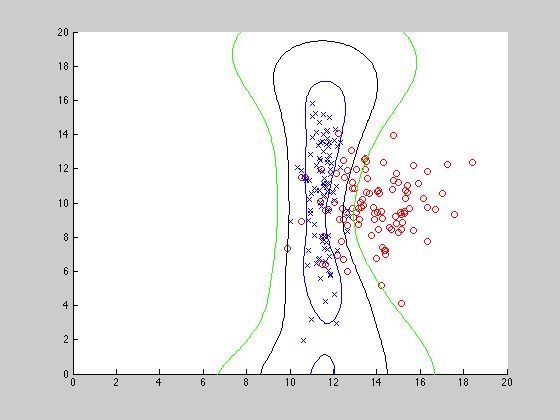
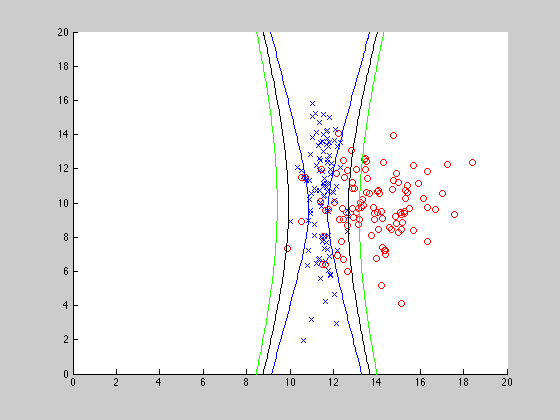


Figure 4/5 – RBF (left) and Polynomial w/ n=2 (right)

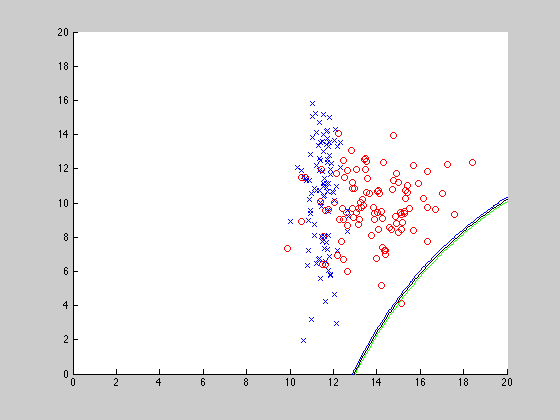


Figure 6 – Polynomial w/ n=3

## Conclusion

As expected, the RBF kernel provided the best accuracy out of all of the kernels used, yielding 96.92% and 86% accuracy on the clustered and Gaussian datasets, respectively.

Surprisingly, the polynomial kernel with n=3 worked significantly better than we initially expected on the clustered dataset, coming in at 92.31% accuracy, while performing poorly on the Gaussian dataset, coming in at 50.5% accuracy. Inversely, the polynomial kernel with n=2 worked well on the Gaussian dataset, coming in at 85% accuracy, while only achieving 85.38% accuracy on the clustered dataset.

The poor performance of the polynomial kernel with n=3 on the Gaussian dataset is a bit troubling, considering that it only correctly classified one of the points in the positive dataset. We feel this might be the result of a defect in the SVM library.