

## SVM: Guiding Questions

Comment on the testing and training accuracy of your SVM classifier on the Spam dataset. Discuss how kernels (and their hyperparameters) affect the classifiers accuracy.

The training accuracy is higher than the testing accuracy for the Spam dataset no matter which kernel is used. For the linear kernel the train accuracy is 97.5 % and the test accuracy 88.75 %. For the RBF kernel the train accuracy is 95.0 % and the test accuracy 91.25 %. For the RBF kernel the train accuracy is 99.69 % and the test accuracy 86.25 %. You can see that as you pick a kernel with higher complexity the training error decreases. However, by the time you are using a polynomial kernel, we see the testing accuracy decrease, meaning that we are likely overfitting to our training data. I also did a few tests to investigate the effect of changing the hyperparameters. When you increase to  $d = 3$  in the polynomial kernel, increasing the complexity, we see both training and testing accuracy increase.

## Decision Trees: Guiding Questions

I found that in non-pruned trees, the training error is always lower than the testing error. After adding pruning, however, the testing error decreases, and the training error only increases slightly. This makes sense because pruning combats over-fitting, which causes a decision tree to perform better on the training data but worse in general. The pruned testing error is smaller in all 6 cases, so pruning was effective! The gini index reduced error most effectively, although entropy also does very well.

As expected, as the depth increases, the loss decreases. That is to say, we do better when the depth is bigger, which makes sense because the tree is able to use more features. However, in general the decrease in loss for each added depth begins to taper for bigger depths. Again, this makes sense because the decision tree is built using a greedy algorithm so the most helpful features are chosen first so adding depth becomes less helpful as the depth gets bigger.

