When classifying the virginica dataset vs the versicolor dataset.   
Example of why data is not linearly seperable.  
  
  
Legend- the red circles are the training datasets, the stars are the testing datasets  
for class virginica, whist the blue pluses and stars are the training and teting datasets for the versicolor dataset.  
  
Feature 2  
When classifying the virginica dataset vs the versicolor dataset.   
Example of why data is not linearly seperable.  


Feature 3: Example of features that are partially seperable but in the 1 feature space  


Feature 4  
  


~~~~~  
When combining two features we get the following curves  
Features 1 and 2  
  
  
  
Features 1 and 3  


Features 1 and 4  


Features 2 and 3  


Features 2 and 4  


Features 3 and 4  
  
  
As can be seen some features are more linearly seperable with others the goal of using the kernels of the SVM is to increase this seperation.   
  
  
The best ROC curve base on the graph. (Feature 2 without using the kernel)  


The worst ROC curve based on the graph. (Feature 3 and 4 after using a polynomial kernel)  
