# Math 5364 Homework 12

1. (a) Create a version of the `iris` data set, where the class labels "versicolor" and "virginica" are replaced by "nonsetosa". This problem involves building an SVM to classify iris flowers as setosa or nonsetosa.
   (b) Fit a linear support vector machine to the data with `cost=1000`, and plot the SVM.
   (c) Is this data set linearly separable?
   (d) What is the classification accuracy for this model?
   (e) How many support vectors are there?
   (f) Find the parameters $w$ and $b$ that define the decision boundary $w \cdot x + b = 0$.

2. Split `wdbc.data` into 70% training and 30% test data.
   (a) Fit an SVM to the training data.
   (b) What type of kernel was used?
   (c) Find the classification accuracy of this SVM on the training and test data.
   (d) Use the `tune.svm` command to tune the values of `cost` and `gamma`. It may take some experimentation to find suitable ranges for these parameters.
   (e) Refit the SVM using the tuned `cost` and `gamma` values.
   (f) Find the classification accuracy of the tuned SVM on the training and test data.

3. (a) Create a data set similar to the one below, where there are four normally distributed clusters, each containing 50 points, centered at $(0,0)$, $(6,0)$, $(0,6)$, and $(6,6)$. For all four clusters, $\sigma_x = \sigma_y = 1.5$.
   (b) Create an SVM for distinguishing between the black circles and red triangles, plot the SVM, and calculate its classification accuracy.