STT 811

In-Class Assignment 17

This problem will use the Heart dataset, with the numerical y as the target.

(1) Do a 70/30 train/test split

```
heart <- read.csv("data/Heart.csv")
heart$AHD <- as.numeric(as.factor(heart$AHD))
split_pct <- 0.7
n <- length(heart$AHD)*split_pct # train size
row_samp <- sample(1:length(heart$AHD), n, replace = FALSE)
train <- heart[row_samp,]
test <- heart[-row_samp,]
```

(2) Build a SVM model using RestBP and MaxHR. Try different kernels, with default hyperparameter values. Which is best according to accuracy?

```
svm_mod <- svm(AHD ~ RestBP + MaxHR, data = train, type = 'C-classification', kernel =
'linear', cost = 1, gamma = 0.5)
```

(3) For the best, try varying the hyperparameters. Is there a better set?

```
svm_mod <- svm(AHD ~ RestBP + MaxHR, data = train, type = 'C-classification', kernel =
'linear', cost = 4, gamma = 1)
```

(4) Add other inputs, seeing if you can improve the accuracy. What's the best model you can find?