**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,]

1. 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)

1. 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)

1. Add other inputs, seeing if you can improve the accuracy. What’s the best model you can find?