STT 811

In-Class Assignment 11

This problem will use the OJ dataset. Note that you will use Purchase as the target (no need to convert to 0/1)

1. Split the data into training and test datasets (with a 75/25 split).

```
oj <- ISLR2::OJ
split_pct <- 0.75
n <- length(oj$Purchase)*split_pct # train size
row_samp <- sample(1:length(oj$Purchase), n, replace = FALSE)
train <- oj[row_samp,]
test <- oj[-row_samp,]
```

2. Build a KNN model for your target based on PriceDiff and LoyalCH, varying K. How does the accuracy compare on the test dataset? Which K works best here?

```
train.Y = train$Purchase
test.Y = test$Purchase
train_scale = scale(train[,c(2,3)])
test_scale = scale(test[,c(2,3)])
knn_mod<-knn(train = train_scale, test = test_scale, cl = train.Y, k=5)

cm<-confusionMatrix(knn_mod, reference = as.factor(test.Y))
cm$table
```