```
> library(caret)
> data(iris)
> dim(iris)
[1] 150   5
> head(iris)
  Sepal.Length Sepal.Width Petal.Length Petal.Width Species
1          5.1         3.5          1.4         0.2  setosa
2          4.9         3.0          1.4         0.2  setosa
3          4.7         3.2          1.3         0.2  setosa
4          4.6         3.1          1.5         0.2  setosa
5          5.0         3.6          1.4         0.2  setosa
6          5.4         3.9          1.7         0.4  setosa
> unique(iris$Species)
[1] setosa     versicolor virginica 
Levels: setosa versicolor virginica
> iris <- iris[-which(iris$Species=='setosa'),]
> dim(iris)
[1] 100   5
> unique(iris$Species)
[1] versicolor virginica 
Levels: setosa versicolor virginica
> iris$Species <- iris$Species %>% factor()
> unique(iris$Species)
[1] versicolor virginica 
Levels: versicolor virginica
> y <- iris$Species
> y
  [1] versicolor versicolor versicolor versicolor versicolor versicolor
  [7] versicolor versicolor versicolor versicolor versicolor versicolor
 [13] versicolor versicolor versicolor versicolor versicolor versicolor
 [19] versicolor versicolor versicolor versicolor versicolor versicolor
 [25] versicolor versicolor versicolor versicolor versicolor versicolor
 [31] versicolor versicolor versicolor versicolor versicolor versicolor
 [37] versicolor versicolor versicolor versicolor versicolor versicolor
 [43] versicolor versicolor versicolor versicolor versicolor versicolor
 [49] versicolor versicolor virginica  virginica  virginica  virginica 
 [55] virginica  virginica  virginica  virginica  virginica  virginica 
 [61] virginica  virginica  virginica  virginica  virginica  virginica 
 [67] virginica  virginica  virginica  virginica  virginica  virginica 
 [73] virginica  virginica  virginica  virginica  virginica  virginica 
 [79] virginica  virginica  virginica  virginica  virginica  virginica 
 [85] virginica  virginica  virginica  virginica  virginica  virginica 
 [91] virginica  virginica  virginica  virginica  virginica  virginica 
 [97] virginica  virginica  virginica  virginica 
Levels: versicolor virginica
> 
> 
> set.seed(2, sample.kind="Rounding")
Warning message:
In set.seed(2, sample.kind = "Rounding") :
  non-uniform 'Rounding' sampler used
> test_index <- createDataPartition(y,times=1,p=0.5,list=FALSE)
> test <- iris[test_index,]
> train <- iris[-test_index,]
> dim(train)
[1] 50  5
> dim(test)
[1] 50  5
> 
> 
> cutoff <- seq(range(train$Sepal.Length)[1], range(train$Sepal.Length)[2], 0.1)
> cutoff
 [1] 5.0 5.1 5.2 5.3 5.4 5.5 5.6 5.7 5.8 5.9 6.0 6.1 6.2 6.3 6.4 6.5 6.6 6.7 6.8
[20] 6.9 7.0 7.1 7.2 7.3 7.4 7.5 7.6 7.7 7.8 7.9
> accuracy <- map_dbl(cutoff, function(x) {
+ y_hat <- ifelse(train$Sepal.Length > x, "virginica", "versicolor")
+ mean(y_hat == train$Species)
+ })
> max(accuracy)
[1] 0.7
```

```
>
>
> # Efficient solution for Question 8 rather than repeating above code block for every feature
> foo <- function(x) {
+ cutoff <- seq(range(x)[1], range(x)[2], by=0.1)
+ sapply(cutoff, function(c) {
+ y_hat <- ifelse(x > c, 'virginica','versicolor')
+ mean(y_hat==train$Species)
+ })
> apply(train[,-5], 2, foo)    # apply foo function on first 4 columns (represented by 2) of train(features)
$Sepal.Length
 [1] 0.52 0.54 0.54 0.54 0.54 0.62 0.64 0.66 0.66 0.68 0.66 0.68 0.70 0.68 0.68
[16] 0.62 0.66 0.66 0.64 0.58 0.60 0.58 0.58 0.58 0.56 0.56 0.56 0.52 0.52 0.50

$Sepal.Width
 [1] 0.52 0.52 0.54 0.58 0.62 0.62 0.60 0.60 0.54 0.62 0.54 0.54 0.54 0.52 0.52
[16] 0.52 0.52 0.52 0.50

$Petal.Length
 [1] 0.52 0.52 0.52 0.52 0.52 0.54 0.54 0.56 0.58 0.60 0.68 0.70 0.72 0.76 0.82
[16] 0.88 0.90 0.96 0.96 0.94 0.90 0.82 0.80 0.76 0.74 0.70 0.70 0.66 0.62 0.58
[31] 0.58 0.56 0.56 0.56 0.54 0.54 0.54 0.52 0.52 0.50

$Petal.Width
 [1] 0.58 0.62 0.66 0.80 0.90 0.94 0.94 0.94 0.86 0.78 0.72 0.66 0.64 0.54 0.52
[16] 0.50

> accuracies <- apply(train[,-5], 2, foo)
> sapply(accuracies, max)
Sepal.Length  Sepal.Width Petal.Length  Petal.Width 
        0.70         0.62         0.96         0.94 
>
>
> cutoff <- seq(range(train[,3])[1], range(train[,3])[2], by=0.1)
> cutoff
 [1] 3.0 3.1 3.2 3.3 3.4 3.5 3.6 3.7 3.8 3.9 4.0 4.1 4.2 4.3 4.4 4.5 4.6 4.7 4.8
[20] 4.9 5.0 5.1 5.2 5.3 5.4 5.5 5.6 5.7 5.8 5.9 6.0 6.1 6.2 6.3 6.4 6.5 6.6 6.7
[39] 6.8 6.9
> smart_cutoff <- cutoff[which.max(foo(train[,3]))]
> smart_cutoff
[1] 4.7
> y_hat <- ifelse(test[,3] > smart_cutoff,'virginica','versicolor')
> mean(y_hat == test$Species)
[1] 0.9
>
>
>
```