

K-Nearest Neighbor in R

Ryan Honea

11/20/2017

K-Nearest Neighbor

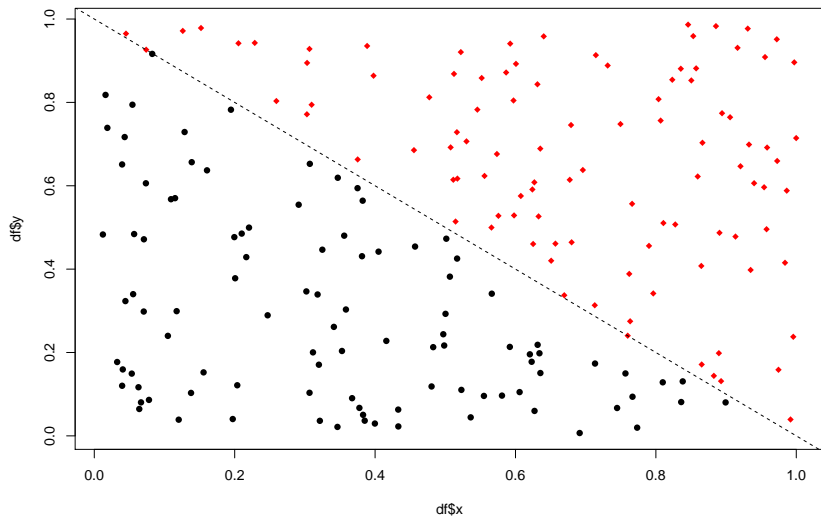
- ▶ Uses `knn` function from `class` library.
- ▶ Function takes in four non-default values:
 - ▶ training set
 - ▶ testing set
 - ▶ labels
 - ▶ k number of neighbors
- ▶ Looks like `knn(train, test, labels, k)`

Creating a Dataset

```
df <- data.frame(  
  x = x <- runif(200,0,1),  
  y = y <- runif(200,0,1)  
)  
df$labels <- factor(apply(df, 1, function(x)  
  if (sum(x) > 1) 1 else 0))
```

Plot of Data

```
plot(df$x, df$y, pch = c(16,18)[as.numeric(df$labels)], col =  
abline(1, -1, lty = 2)
```



KNN Training

```
require(class)
```

```
## Loading required package: class
```

```
model <- knn(train = df[1:150,1:2], test = df[151:200,1:2],  
             cl = df$labels[1:150], k = 5)  
correct <- factor(model == df$labels[151:200])
```

Plotting Results

```
plot(df$x[151:200], df$y[151:200], col = c("red", "green")[a
      pch = c(16, 18)[as.numeric(model)]]
abline(1, -1, lty = 2)
legend("bottomright", c("Correct", "Incorrect"), col = c("r
```

