CS513 HW3: knn

Aidan Fischer - 10447681

2023-10-07

I pledge my honor that I have abided by the Stevens Honor System.

Creation

```
rm(list=ls())
library(caTools)
library(class)
data = read.csv("breast-cancer-wisconsin.csv")
data$F6 <- suppressWarnings(as.numeric(data$F6))
data = data[complete.cases(data), ]
#Convert categories to the factor data type

for(i in 1:9){
  col = paste("F",i,sep='')
  data[col] <- factor(data[[col]], levels = 1:10)
}

data$Class <- factor(data$Class, levels=c(2,4))
set.seed(255)
split = sample.split(data$Class, SplitRatio=0.7)
train = subset(data, split == TRUE)
test = subset(data, split == FALSE)

test_3 = knn(train = train, test = test, cl = train$Class, k=3)
test_5 = knn(train = train, test = test, cl = train$Class, k=5)
test_10 = knn(train = train, test = test, cl = train$Class, k=10)
```

Evaluation

```
actual <- test$Class
cm_3 <- table(actual,test_3)
cm_5 <- table(actual,test_5)
cm_10 <- table(actual,test_10)
```

K = 3

```
print(cm_3)
```

```
##       test_3
## actual   2   4
##      2 100  33
##      4  52  20
```

```
accuracy <- sum(diag(cm_3))/length(actual)
sprintf("Accuracy: %.2f%%", accuracy*100)
```

```
## [1] "Accuracy: 58.54%"
```

K = 5

```
print(cm_5)
```

```
##       test_5
## actual   2   4
##      2 108  25
##      4  56  16
```

```
accuracy <- sum(diag(cm_5))/length(actual)
sprintf("Accuracy: %.2f%%", accuracy*100)
```

```
## [1] "Accuracy: 60.49%"
```

K = 10

```
print(cm_10)
```

```
##       test_10
## actual   2   4
##      2 111  22
##      4  58  14
```

```
accuracy <- sum(diag(cm_10))/length(actual)
sprintf("Accuracy: %.2f%%", accuracy*100)
```

```
## [1] "Accuracy: 60.98%"
```