181058_DSLab11

Classification

Importing the file

```
# message warning False for any output while importing library

library(caret)
heart = read.csv('heart.csv')
set.seed(12) # random value for seed
```

Encoding output

Now we would be encoding output via factoring

```
heart$output = as.factor(heart$output)
```

Splitting data to training and test

```
library(caTools)
set.seed(123)
split = sample.split(heart$output, SplitRatio = 0.8)
trainingset <- subset(heart, split == TRUE)
testset <- subset(heart, split == FALSE)
paste("Training Set", dim(trainingset))
```

```
## [1] "Training Set 242" "Training Set 14"
```

```
paste("Test set", dim(testset))
```

```
## [1] "Test set 61" "Test set 14"
```

Setting cross-validation

```
trControl <- trainControl(method = "repeatedcv", number = 10, repeats = 10)
```

Training model

```
model <- train(output ~ ., method='knn', data = trainingset, metric='Accuracy', trControl = trControl)
```

Evaluating model

```
pred <- predict(model, testset)

factor_pred <- as.factor(pred)
factor_truth <- as.factor(testset$output)
```

Now, calculating precision

```
precision <- posPredValue(factor_truth, factor_pred)
#recall
recall <- sensitivity(factor_truth, factor_pred)
```

Applying confusion Matrix

```
cm <- confusionMatrix(pred, testset$output)
Accuracy <- cm$overall[1]
confusion_matrix <- cm$table
confusion_matrix
```

```
##           Reference
## Prediction  0  1
##          0 13 11
##          1 15 22
```