

# Part One

## Decision Trees

### Assessing Model Prediction ¶

```
treePredictions = decisionTree.predict(x_test)
print(confusion_matrix(y_test, treePredictions))
```

```
[[ 7 11]
 [ 8 29]]
```

```
print(classification_report(y_test, treePredictions))
```

|              | precision | recall | f1-score | support |
|--------------|-----------|--------|----------|---------|
| 0            | 0.47      | 0.39   | 0.42     | 18      |
| 1            | 0.72      | 0.78   | 0.75     | 37      |
| accuracy     |           |        | 0.65     | 55      |
| macro avg    | 0.60      | 0.59   | 0.59     | 55      |
| weighted avg | 0.64      | 0.65   | 0.65     | 55      |

#### Confusion matrix

**Top left:** The model correctly predicted 7 persons as DID NOT survive.

**Top right:** The model wrongly predicted 11 people as DID NOT survive, when in reality they survived.

**Bottom left:** The model wrongly predicted 8 people as SURVIVED, when in reality they did not survive.

**Bottom right:** The model correctly predicted 29 persons as SURVIVED.

#### Accuracy

The model's prediction for **non-survival** was 47% accurate

While the prediction for **survival** was 72% accurate

# Part Two

## Random Forests

### Assessing model predictions

```
forestPredictions = forest.predict(x_test)
print(confusion_matrix(y_test, forestPredictions))
```

```
[[ 8 10]
 [ 6 31]]
```

```
print(classification_report(y_test, forestPredictions))
```

|              | precision | recall | f1-score | support |
|--------------|-----------|--------|----------|---------|
| 0            | 0.57      | 0.44   | 0.50     | 18      |
| 1            | 0.76      | 0.84   | 0.79     | 37      |
| accuracy     |           |        | 0.71     | 55      |
| macro avg    | 0.66      | 0.64   | 0.65     | 55      |
| weighted avg | 0.70      | 0.71   | 0.70     | 55      |

### Confusion matrix

**Top left:** The model correctly predicted 8 persons as DID NOT survive.

**Top right:** The model wrongly predicted 10 people as DID NOT survive, when in reality they survived.

**Bottom left:** The model wrongly predicted 6 people as SURVIVED, when in reality they did not survive.

**Bottom right:** The model correctly predicted 31 persons as SURVIVED.

### Accuracy

The model's prediction for **non-survival** was 57% accurate

While the prediction for **survival** was 76% accurate

### Conclusion

The precision rates for both models are close. However, the random forest model has better accuracy due to a high number of estimators.