

## 1- Decision Tree Classification

Accuracy: 0.5303030303030303

Precision: 1.0

Recall: 0.4918032786885246

## 2- Naive Bayes Classifier

Accuracy: 0.5303030303030303

Precision: 1.0

Recall: 0.4918032786885246

## 3 - Comparing Scikit-learn Classifiers and Random Forest Classifier

### Decision Tree Classifier

Accuracy: 0.5303030303030303

Precision: 1.0

Recall: 0.4918032786885246

### Naive Bayes Classifier

Accuracy: 0.5303030303030303

Precision: 1.0

Recall: 0.4918032786885246

### Random Forest Classifier

Best parameters: {'max\_depth': 5, 'min\_samples\_leaf': 4, 'min\_samples\_split': 2, 'n\_estimators': 30}

Accuracy: 0.5303030303030303

Precision: 1.0

Recall: 0.4918032786885246

## Inference

All the models have identical performance metrics.

The confusion matrix for all classifiers is the same:

	Actual Positive	Actual Negative
Predicted Positive	30	0
Predicted Negative	31	5

The False Negative classifications are for **adult, males** in the test dataset.

In the training dataset there **1324** adult males who did not survive and **322** adult males who survived Titanic disaster. Since there is a large proportion of adult males who did not survive in the training dataset, the models classify the adult males in the test dataset as not survived. But nearly half of the samples in the test dataset are adult males who survived. This drastic difference between training data and test data is the reason for a large number of False Negatives.