Data Mining & Machine Learning
Computer Exercise 10 - Boosting

Steinarr Hrafn Höskuldsson

October 2022

Section 1.2

The missing ages of the passengers were replaced with the mean age of all known passenger ages, which turned out to be 29.9

Section 2.2

Setting `n_estimators` as 50 gave the results:

$$accuracy = 81.34\%, precision = 78.49\%, recall = 70.87\%$$

Section 2.4

Setting `n_estimators` as 150 and `learning_rate` to 0.2 gave the results:

$$accuracy = 81.71\%, precision = 78.72\%, recall = 71.84\%$$

Section 2.5

The best parameters found by `RandomSearchCV` were: `n_estimators = 20`, `max_depth = 7` and `learning_rate=0.098`
   Resulting in :

$$accuracy = 82.46\%, precision = 79.78\%, recall = 72.82\%$$

Section 3.2

The predictions were submitted into the Titanic competition on Kaggle resulting in

```
7982     Steinarr Hrafn Höskuldsson      0.77511      11s
```