

Machine Learning Assignment 2 Report

111062117, Hsiang-Sheng Huang

October 12, 2024

1 Attribute Settings of the Random Forest Model

1.1 Number of Trees

I used 5 trees in the random forest model.

1.2 Number of Features

I used $\lfloor \sqrt{n_{\text{features}}} \rfloor = \lfloor \sqrt{30} \rfloor = 5$ features for each decision tree.

1.3 Number of Instances

The model was trained on $\lfloor n_{\text{samples}} \times 0.8 \rfloor = 5440$ instances, using 80% of the available data.

2 Difficulty Encountered

One challenge I encountered was that the F1-score submitted to Kaggle did not reach a satisfactory level (the baseline score is 0.7), which stuck with me for a long time.

3 Solutions and Reflections

After experimenting with a variety of hyperparameters, the best result was achieved with 5 trees and a maximum depth of 3 for each tree.