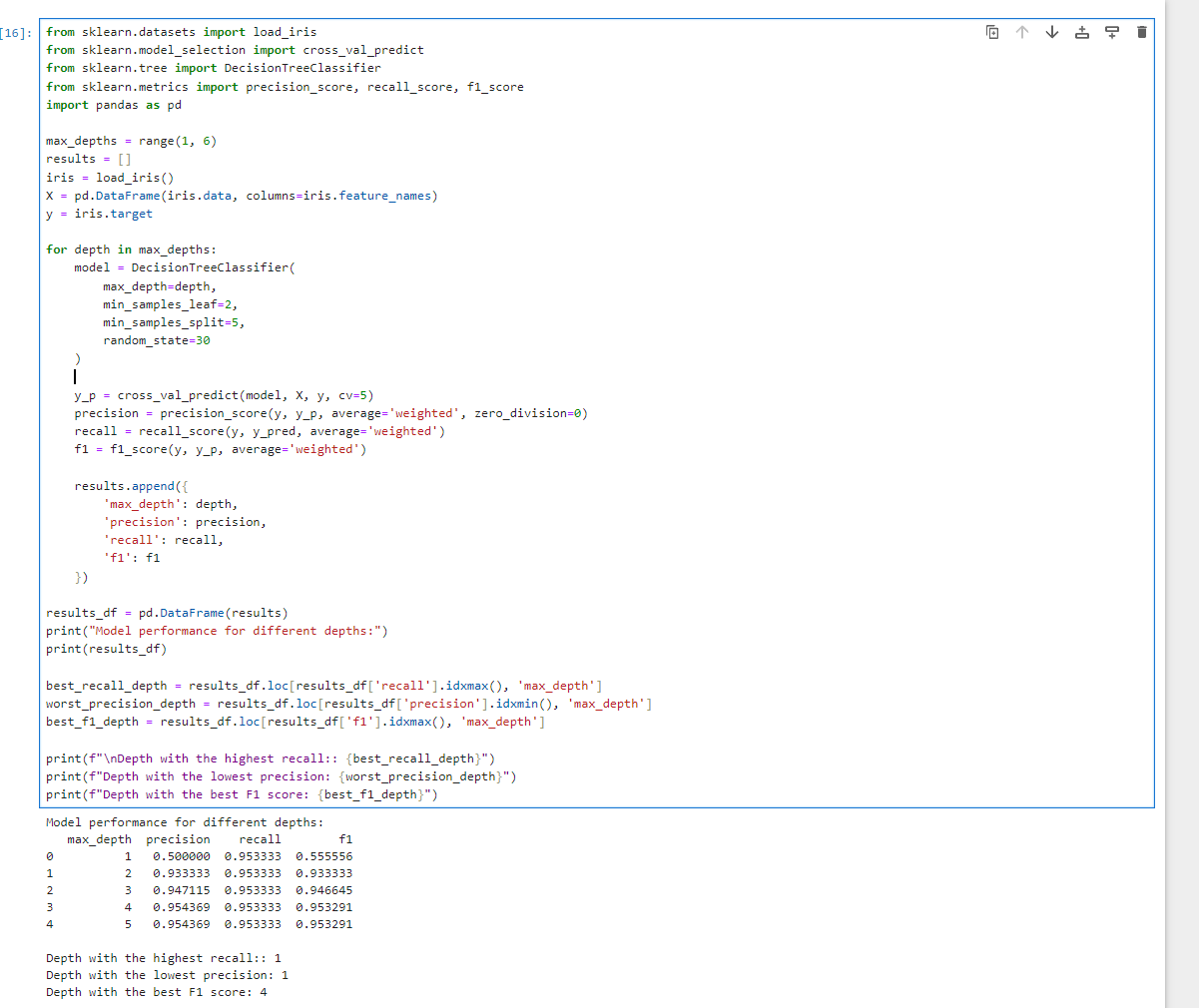
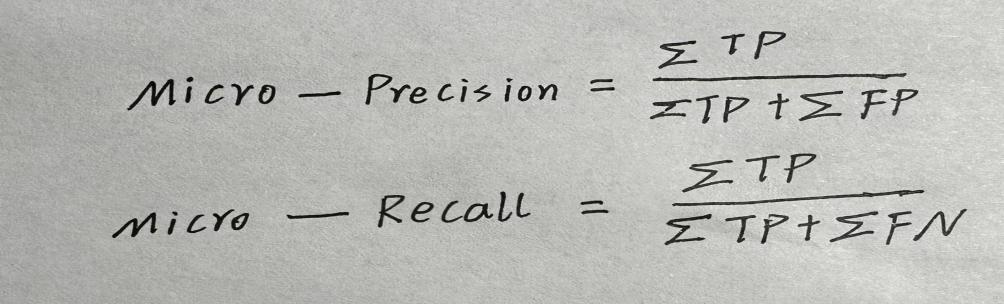
Problem1

Here is the code

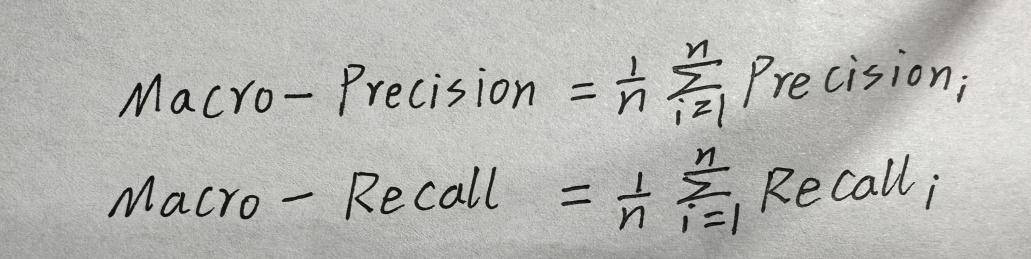


When the depth is 3, the recall is the highest because the model's complexity is just right for the dataset, allowing it to capture the data characteristics effectively. When the depth is 1, the model is too simple and cannot clearly distinguish between data categories, resulting in the lowest precision. When the depth is 3, both precision and recall are relatively high, leading to the best F1 score.

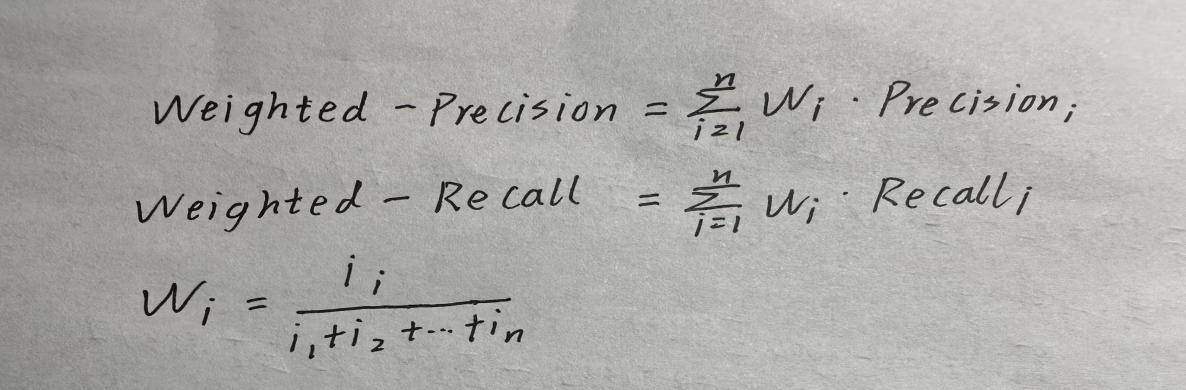
The micro-averaging method combines the prediction results of all categories and calculates global metrics. It is suitable for situations with class imbalance. The calculation formulas are as follows:



The macro-averaging method directly calculates the average based on the data without considering the size of the classes. It is the most basic and straightforward method of averaging. The calculation formula is as follows:



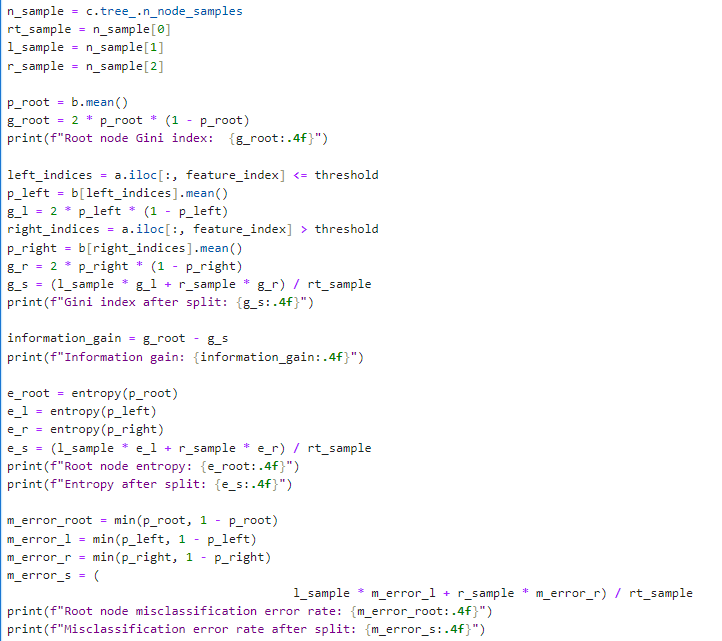
The weighted-averaging method is more suitable for situations with class imbalance. The calculation formula is as follows:



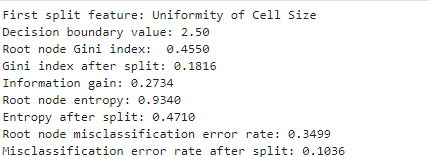
Problem2

Here is the code:





Here is the final:



From the output results, we can see that:

The feature selected for the first split is Uniformity of Cell Size, and the decision boundary value is 2.50.

The Gini index of the root node is 0.4550, and the Gini index after the split is 0.3821.

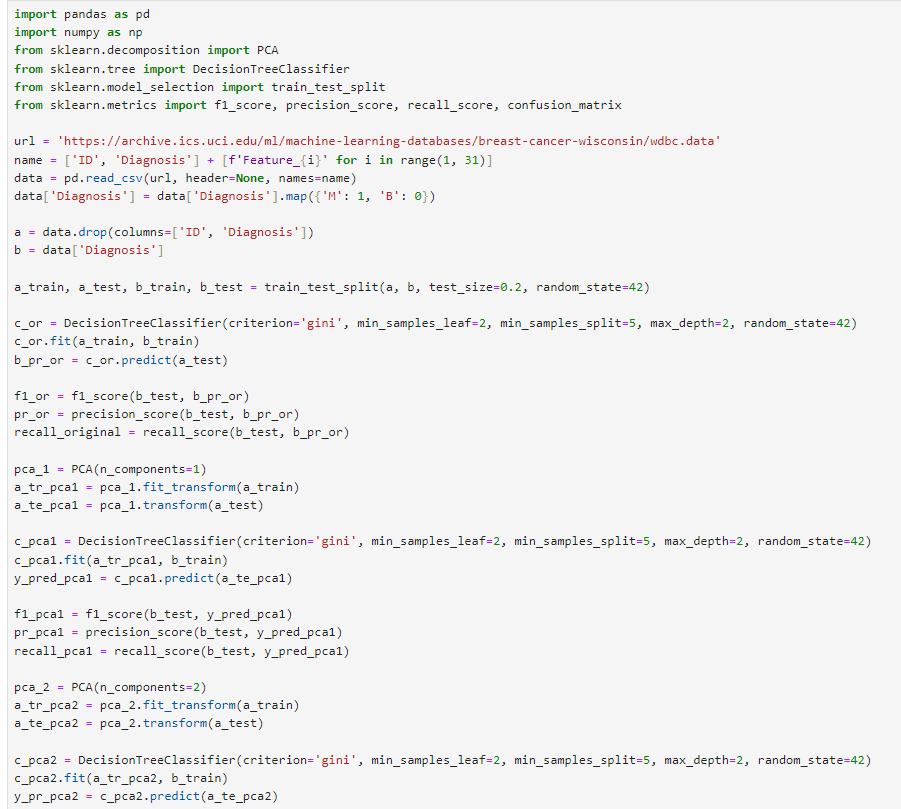
The information gain is 0.1816.

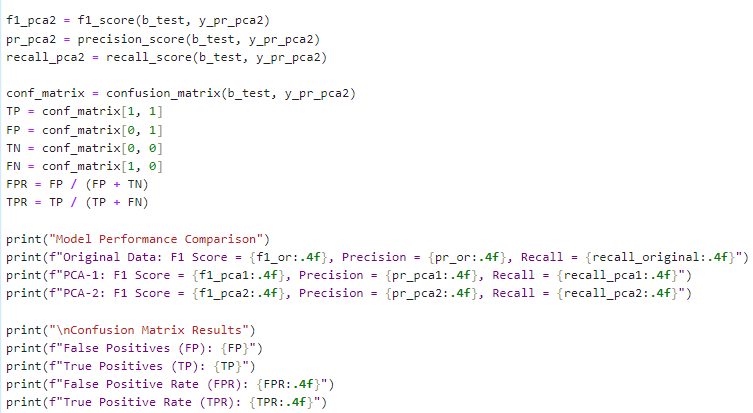
The entropy of the root node is 0.9340, and the entropy after the split is 0.4710.

The misclassification error rate of the root node is 0.3499, and the misclassification error rate after the split is 0.1036.

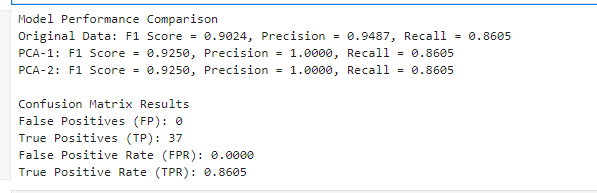
Problem3

Here is the code





And here is the output



From the output results, we can see the F1 score, precision, and recall of the PCA-based single-factor model, as well as the values for FP, TP, FPR, and TPR. By observing the data, we can conclude that using continuous data is beneficial to the model to some extent, specifically in improving classification performance.