1.

Decision Tree

Table 1 Summary Table for Decision Tree

Classifier	Accuracy	Precision	Recall	F1	Training Time
{Entropy, depth=5}	0.5158	0.5034	0.5138	0.4819	0.0106s
{Entropy, depth=10}	0.5311	0.5220	0.5311	0.5221	0.0186s
{Entropy, depth=15}	0.5545	0.5498	0.5545	0.5520	0.0233s
{Entropy, depth=20}	0.5841	0.5797	0.5841	0.5812	0.0240s
{Gini, depth = 5}	0.5280	0.5169	0.5280	0.4932	0.0079s
{Gini, depth = 10}	0.5566	0.5487	0.5566	0.5469	0.0146s
{Gini, depth = 15}	0.5864	0.5799	0.5861	0.5824	0.0188s
{Gini, depth = 20}	0.6075	0.6057	0.6075	0.6063	0.0199s

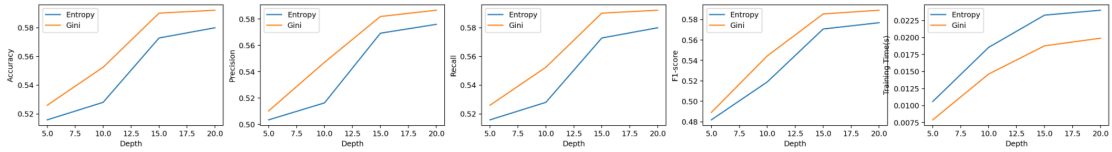

Figure 1 Visualization of Summary Table

In general, all of the accuracy, precision, recall and F1-score for all the classifiers are not satisfied (in range of (0.5,0.6)).

From the perspective of splitting criterion, no matter the maximum depth, the performance of Gini decision tree in all perspective is greater than entropy decision Tree.

From the perspective of depth, with the increase of depth. The accuracy, precision, recall, F1-score, and Training Time increase monotonically.

KNN & Random Forest

Table 2 Summary Table for KNN & Random Forest

Classifier	Accuracy	Precision	Recall	F1	Training Time
5-NN	0.4873	0.4683	0.4873	0.4714	0s
21-NN	0.4577	0.4425	0.4577	0.4133	0s
5-NN with weighted distance	0.5902	0.5882	0.5902	0.5799	0s
Random Forest	0.6544	0.6701	0.6595	0.6443	0.4801s

Note that KNN is lazy trainer, its training time should be zero.

Compared to 5-NN classifier, 21-NN classifier losses about 3% accuracy, 2.5% precision, 3% recall and 6% F1-Score. In our case, the more neighbors we use, the worse the performance.

When I remain the number of neighbors used and use the weighted distance, each criterion increases about 10%. It demonstrates that the closer neighbors are more perusable.

Compared to three K-NN classifiers, the random forest has much better result. When the baseline is 5-NN with weighted distance, the accuracy of random forest increases about 7%, precision increases about 9%, recall increases about 7%, and F1-score increases about 7%. It is the best classifiers in Question 1.

2.

$C^*(x) = \text{sign}\,[0.4236 * C_1(x) + 0.2428 * C_2(x) + 0.2098 * C_3(x) + 0.5734 * C_4(x) + 0.5064 * C_5(x) + 0.2772 * C_6(x) + 0.3687 * C_7(x) + 0.2603 * C_8(x)]$

$$C_1(x) = \begin{cases} +1 & x \leq 1.0 \\ -1 & x > 1.0 \end{cases} \qquad C_2(x) = \begin{cases} +1 & x \leq 6.0 \\ -1 & x > 6.0 \end{cases} \qquad C_3(x) = \begin{cases} -1 & x \leq -0.5 \\ +1 & x > -0.5 \end{cases}$$

$$C_4(x) = \begin{cases} -1 & x \leq 4.0 \\ +1 & x > 4.0 \end{cases} \qquad C_5(x) = \begin{cases} +1 & x \leq 1.0 \\ -1 & x > 1.0 \end{cases} \qquad C_6(x) = \begin{cases} -1 & x \leq -0.5 \\ +1 & x > -0.5 \end{cases}$$

$$C_7(x) = \begin{cases} -1 & x \leq 8.0 \\ +1 & x > 8.0 \end{cases} \qquad C_8(x) = \begin{cases} -1 & x \leq 4.0 \\ +1 & x > 4.0 \end{cases}$$

The final accuracy on the training data is 1.0.