Impurity Metric	Max Depth	Precision	Recall	F1 Score
Gini Index	2	0.758	0.965	0.849
	3	0.780	0.950	0.857
	4	0.802	0.965	0.876
	5	0.821	0.980	0.893
Entropy	2	0.758	0.965	0.849
	3	0.772	0.960	0.856
	4	0.794	0.960	0.869
	5	0.810	0.975	0.885

Table 1 – Model performance metrics for "no-recurrence-events" class

Impurity Metric	Max Depth	Precision	Recall	F1 Score
Gini Index	2	0.767	0.271	0.400
	3	0.756	0.365	0.492
	4	0.841	0.435	0.574
	5	0.913	0.494	0.641
Entropy	2	0.767	0.271	0.400
	3	0.778	0.329	0.463
	4	0.814	0.412	0.547
	5	0.886	0.459	0.605

Table 2 – Model performance metrics for "recurrence-events" class

Question 1 (1 point): Based upon the model performance metrics, which class value is DecisionTreeClassifier **"better"** at predicting? Be sure to specifically mention the metric(s) you're using to quantify your findings.

>>Based on the provided metrics, the DecisionTreeClassifier consistently demonstrates better predictive performance for the class value 'no-recurrence-events'. This conclusion is drawn from higher precision, recall, and F1 score values observed across various criteria and maximum depths compared to the 'recurrence-events' class. Thus, the model is more effective at predicting instances of 'no-recurrence-events' based on the given evaluation metrics.

Question 2 (1 point): Which impurity metric provides higher quality predictions? Be sure to specifically mention the metric(s) you're using to quantify your findings.
>> The entropy impurity metric generally yields slightly higher precision, recall, and F1 scores compared to the Gini impurity metric across various maximum depths. This observation holds true for both classes ('no-recurrence-events' and 'recurrence-events'). Therefore, based on the evaluation metrics, the entropy impurity metric provides slightly higher quality predictions compared to the Gini impurity metric.

Question 3 (1 point): Regardless of the impurity metric, what happens as max depth increases? Be sure to specifically mention the metric(s) you're using to quantify your findings.
>>As the maximum depth increases:
1. Precision tends to rise for both classes, indicating fewer false positives.
2. Recall initially increases but may stabilize or decrease slightly at higher depths.

3. F1 score generally mirrors the trend of recall, increasing initially and then stabilizing or decreasing.
4. The model becomes better at capturing complex relationships within the data.
5. However, increasing depth may lead to overfitting, compromising generalization to unseen data.
6. Monitoring the F1 score is crucial to balance model complexity and performance.

Question 4 (1 point): Try incrementally increasing the max depth value up to 20 for each of the impurity metrics. What happens to the performance metrics when you do this? Why is this trend actually problematic for the model? Be sure to specifically name the phenomenon that is occurring as the max depth is increased.

For Gini :

```
MINGW64:/c/Users/vpark/Vee/DSCI/DSCI-633/assignments/DecisionTree

vpark@Vee_17 MINGW64 ~/Vee/DSCI/DSCI-633/assignments/DecisionTree (main)
$ python A6.py
Criterion: gini, Max Depth: 2
{'no-recurrence-events': {'prec': 0.7578125, 'recall': 0.9651741293532339, 'f1':
 0.8490153172866521}, 'recurrence-events': {'prec': 0.7666666666666667, 'recall'
: 0.27058823529411763, 'f1': 0.39999999999999997}}

Criterion: gini, Max Depth: 3
{'no-recurrence-events': {'prec': 0.7795918367346939, 'recall': 0.9502487562189055, 'f1': 0.8565022421524663}, 'recurrence-events': {'prec': 0.7560975609756098,
 'recall': 0.36470588235294116, 'f1': 0.49206349206349204}}

Criterion: gini, Max Depth: 4
{'no-recurrence-events': {'prec': 0.8016528925619835, 'recall': 0.9651741293532339, 'f1': 0.8758465011286681}, 'recurrence-events': {'prec': 0.8409090909090909,
 'recall': 0.43529411764705883, 'f1': 0.5736434108527132}}

Criterion: gini, Max Depth: 5
{'no-recurrence-events': {'prec': 0.8106995884773662, 'recall': 0.9800995024875622, 'f1': 0.8873873873873874}, 'recurrence-events': {'prec': 0.9069767441860465,
 'recall': 0.4588235294117647, 'f1': 0.609375}}

Criterion: gini, Max Depth: 6
{'no-recurrence-events': {'prec': 0.8298755186721992, 'recall': 0.9950248756218906, 'f1': 0.904977375565611}, 'recurrence-events': {'prec': 0.9777777777777777,
'recall': 0.5176470588235295, 'f1': 0.676923076923077}}

Criterion: gini, Max Depth: 7
{'no-recurrence-events': {'prec': 0.8614718614718615, 'recall': 0.9900497512437811, 'f1': 0.9212962962962963}, 'recurrence-events': {'prec': 0.9636363636363636,
 'recall': 0.6235294117647059, 'f1': 0.7571428571428571}}

Criterion: gini, Max Depth: 8
{'no-recurrence-events': {'prec': 0.8701298701298701, 'recall': 1.0, 'f1': 0.9305555555555556}, 'recurrence-events': {'prec': 1.0, 'recall': 0.6470588235294118,
 'f1': 0.7857142857142858}}

Criterion: gini, Max Depth: 9
{'no-recurrence-events': {'prec': 0.881578947368421, 'recall': 1.0, 'f1': 0.9370629370629371}, 'recurrence-events': {'prec': 1.0, 'recall': 0.6823529411764706,
'f1': 0.8111888111888113}}

Criterion: gini, Max Depth: 10
{'no-recurrence-events': {'prec': 0.881578947368421, 'recall': 1.0, 'f1': 0.9370629370629371}, 'recurrence-events': {'prec': 1.0, 'recall': 0.6823529411764706,
'f1': 0.8111888111888113}}

Criterion: gini, Max Depth: 11
{'no-recurrence-events': {'prec': 0.9162790697674419, 'recall': 0.9800995024875622, 'f1': 0.9471153846153845}, 'recurrence-events': {'prec': 0.9436619718309859,
 'recall': 0.788235294117647, 'f1': 0.8589743589743589}}

Criterion: gini, Max Depth: 12
{'no-recurrence-events': {'prec': 0.9136363636363637, 'recall': 1.0, 'f1': 0.9548693586698338}, 'recurrence-events': {'prec': 1.0, 'recall': 0.7764705882352941,
 'f1': 0.8741721854304636}}
```

```
Criterion: gini, Max Depth: 13
{'no-recurrence-events': {'prec': 0.9259259259259259, 'recall': 0.99502487562189
06, 'f1': 0.9592326139088729}, 'recurrence-events': {'prec': 0.9857142857142858,
 'recall': 0.8117647058823529, 'f1': 0.8903225806451613}}

Criterion: gini, Max Depth: 14
{'no-recurrence-events': {'prec': 0.9305555555555556, 'recall': 1.0, 'f1': 0.964
0287769784173}, 'recurrence-events': {'prec': 1.0, 'recall': 0.8235294117647058,
 'f1': 0.9032258064516129}}

Criterion: gini, Max Depth: 15
{'no-recurrence-events': {'prec': 0.9436619718309859, 'recall': 1.0, 'f1': 0.971
0144927536231}, 'recurrence-events': {'prec': 1.0, 'recall': 0.8588235294117647,
 'f1': 0.9240506329113923}}

Criterion: gini, Max Depth: 16
{'no-recurrence-events': {'prec': 0.9436619718309859, 'recall': 1.0, 'f1': 0.971
0144927536231}, 'recurrence-events': {'prec': 1.0, 'recall': 0.8588235294117647,
 'f1': 0.9240506329113923}}

Criterion: gini, Max Depth: 17
{'no-recurrence-events': {'prec': 0.9615384615384616, 'recall': 0.99502487562189
06, 'f1': 0.9779951100244498}, 'recurrence-events': {'prec': 0.9871794871794872,
 'recall': 0.9058823529411765, 'f1': 0.9447852760736197}}

Criterion: gini, Max Depth: 18
{'no-recurrence-events': {'prec': 0.970873786407767, 'recall': 0.995024875621890
6, 'f1': 0.9828009828009829}, 'recurrence-events': {'prec': 0.9875, 'recall': 0.
9294117647058824, 'f1': 0.9575757575757576}}

Criterion: gini, Max Depth: 19
{'no-recurrence-events': {'prec': 0.970873786407767, 'recall': 0.995024875621890
6, 'f1': 0.9828009828009829}, 'recurrence-events': {'prec': 0.9875, 'recall': 0.
9294117647058824, 'f1': 0.9575757575757576}}

Criterion: gini, Max Depth: 20
{'no-recurrence-events': {'prec': 0.975609756097561, 'recall': 0.995024875621890
6, 'f1': 0.9852216748768473}, 'recurrence-events': {'prec': 0.9876543209876543,
'recall': 0.9411764705882353, 'f1': 0.963855421686747}}
```

For Entropy:

MINGW64:/c/Users/vpark/Vee/DSCI/DSCI-633/assignments/DecisionTree

```
Criterion: entropy, Max Depth: 2
{'no-recurrence-events': {'prec': 0.7578125, 'recall': 0.9651741293532339, 'f1':
 0.8490153172866521}, 'recurrence-events': {'prec': 0.7666666666666667, 'recall'
: 0.27058823529411763, 'f1': 0.39999999999999997}}

Criterion: entropy, Max Depth: 3
{'no-recurrence-events': {'prec': 0.772, 'recall': 0.9601990049751243, 'f1': 0.8
558758314855875}, 'recurrence-events': {'prec': 0.7777777777777778, 'recall': 0.
32941176470588235, 'f1': 0.46280991735537186}}

Criterion: entropy, Max Depth: 4
{'no-recurrence-events': {'prec': 0.7942386831275721, 'recall': 0.96019900497512
43, 'f1': 0.8693693693693694}, 'recurrence-events': {'prec': 0.813953488372093,
'recall': 0.4117647058823529, 'f1': 0.546875}}

Criterion: entropy, Max Depth: 5
{'no-recurrence-events': {'prec': 0.8, 'recall': 0.9751243781094527, 'f1': 0.878
9237668161435}, 'recurrence-events': {'prec': 0.8780487804878049, 'recall': 0.42
35294117647059, 'f1': 0.5714285714285715}}

Criterion: entropy, Max Depth: 6
{'no-recurrence-events': {'prec': 0.8163265306122449, 'recall': 0.99502487562189
06, 'f1': 0.8968609865470852}, 'recurrence-events': {'prec': 0.975609756097561,
'recall': 0.47058823529411764, 'f1': 0.6349206349206349}}

Criterion: entropy, Max Depth: 7
{'no-recurrence-events': {'prec': 0.8410041841004184, 'recall': 1.0, 'f1': 0.913
6363636363637}, 'recurrence-events': {'prec': 1.0, 'recall': 0.5529411764705883,
 'f1': 0.7121212121212122}}

Criterion: entropy, Max Depth: 8
{'no-recurrence-events': {'prec': 0.8410041841004184, 'recall': 1.0, 'f1': 0.913
6363636363637}, 'recurrence-events': {'prec': 1.0, 'recall': 0.5529411764705883,
 'f1': 0.7121212121212122}}

Criterion: entropy, Max Depth: 9
{'no-recurrence-events': {'prec': 0.8701298701298701, 'recall': 1.0, 'f1': 0.930
5555555555556}, 'recurrence-events': {'prec': 1.0, 'recall': 0.6470588235294118,
 'f1': 0.7857142857142858}}

Criterion: entropy, Max Depth: 10
{'no-recurrence-events': {'prec': 0.8777292576419214, 'recall': 1.0, 'f1': 0.934
8837209302325}, 'recurrence-events': {'prec': 1.0, 'recall': 0.6705882352941176,
 'f1': 0.8028169014084507}}

Criterion: entropy, Max Depth: 11
{'no-recurrence-events': {'prec': 0.9013452914798207, 'recall': 1.0, 'f1': 0.948
1132075471698}, 'recurrence-events': {'prec': 1.0, 'recall': 0.7411764705882353,
 'f1': 0.8513513513513514}}

Criterion: entropy, Max Depth: 12
{'no-recurrence-events': {'prec': 0.9095022624434389, 'recall': 1.0, 'f1': 0.952
6066350710901}, 'recurrence-events': {'prec': 1.0, 'recall': 0.7647058823529411,
 'f1': 0.8666666666666666}}
```

```
Criterion: entropy, Max Depth: 13
{'no-recurrence-events': {'prec': 0.9259259259259259, 'recall': 0.99502487562189
06, 'f1': 0.9592326139088729}, 'recurrence-events': {'prec': 0.9857142857142858,
 'recall': 0.8117647058823529, 'f1': 0.8903225806451613}}

Criterion: entropy, Max Depth: 14
{'no-recurrence-events': {'prec': 0.9259259259259259, 'recall': 0.99502487562189
06, 'f1': 0.9592326139088729}, 'recurrence-events': {'prec': 0.9857142857142858,
 'recall': 0.8117647058823529, 'f1': 0.8903225806451613}}

Criterion: entropy, Max Depth: 15
{'no-recurrence-events': {'prec': 0.9521531100478469, 'recall': 0.99004975124378
11, 'f1': 0.9707317073170731}, 'recurrence-events': {'prec': 0.974025974025974,
'recall': 0.8823529411764706, 'f1': 0.9259259259259258}}

Criterion: entropy, Max Depth: 16
{'no-recurrence-events': {'prec': 0.9476190476190476, 'recall': 0.99004975124378
11, 'f1': 0.9683698296836983}, 'recurrence-events': {'prec': 0.9736842105263158,
 'recall': 0.8705882352941177, 'f1': 0.9192546583850931}}

Criterion: entropy, Max Depth: 17
{'no-recurrence-events': {'prec': 0.9436619718309859, 'recall': 1.0, 'f1': 0.971
0144927536231}, 'recurrence-events': {'prec': 1.0, 'recall': 0.8588235294117647,
 'f1': 0.9240506329113923}}

Criterion: entropy, Max Depth: 18
{'no-recurrence-events': {'prec': 0.9571428571428572, 'recall': 1.0, 'f1': 0.978
1021897810218}, 'recurrence-events': {'prec': 1.0, 'recall': 0.8941176470588236,
 'f1': 0.9440993788819877}}

Criterion: entropy, Max Depth: 19
{'no-recurrence-events': {'prec': 0.970873786407767, 'recall': 0.995024875621890
6, 'f1': 0.9828009828009829}, 'recurrence-events': {'prec': 0.9875, 'recall': 0.
9294117647058824, 'f1': 0.9575757575757576}}

Criterion: entropy, Max Depth: 20
{'no-recurrence-events': {'prec': 0.9617224880382775, 'recall': 1.0, 'f1': 0.980
4878048780488}, 'recurrence-events': {'prec': 1.0, 'recall': 0.9058823529411765,
 'f1': 0.9506172839506173}}
```

When incrementally increasing the max depth value up to 20 for each of the impurity metrics (Gini and entropy), several trends are visible:

1. Precision: Generally increases or remains stable with increasing max depth.
2.Recall: Initially increases but then stabilizes or decreases as max depth increases.
3.F1 Score: Initially improves but then stabilizes or decreases with higher max depth.

The phenomenon occurring here is known as overfitting. Overfitting happens when a model learns to capture noise in the training data rather than the underlying pattern. In the context of decision trees (and tree-based models in general), increasing the max depth allows the model to create more complex decision boundaries that can perfectly separate the training data. However, this may not generalize well to unseen data.