|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Impurity Metric** | **Max Depth** | **Precision** | **Recall** | **F1 Score** |
| Gini Index | 2 | 0.802 | 0.983 | 0.884 |
| 3 | 0.779 | 0.96 | 0.863 |
| 4 | 0.740 | 0.919 | 0.820 |
| 5 | 0.76 | 0.919 | 0.826 |
| Entropy | 2 | 0.7625 | 0.983 | 0.859 |
| 3 | 0.769 | 0.967 | 0.857 |
| 4 | 0.814 | 0.919 | 0.863 |
| 5 | 0.783 | 0.935 | 0.852 |

**Table 1 – Model performance metrics for “no-recurrence-events” class**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Impurity Metric** | **Max Depth** | **Precision** | **Recall** | **F1 Score** |
| Gini Index | 2 | 0.9 | 0.375 | 0.529 |
| 3 | 0.777 | 0.291 | 0.424 |
| 4 | 0.5 | 0.2083 | 0.294 |
| 5 | 0.545 | 0.25 | 0.342 |
| Entropy | 2 | 0.833 | 0.208 | 0.333 |
| 3 | 0.75 | 0.25 | 0.375 |
| 4 | 0.687 | 0.458 | 0.549 |
| 5 | 0.666 | 0.333 | 0.444 |

**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.

<your answer goes here>

For 'no-recurrence-events' prediction, the DecisionTreeClassifier performs better. The F1 score, which measures the harmonic mean of recall and precision and denotes a balance between the two, is used to quantify this. The F1 scores for "no-recurrence-events" are consistently higher than those for "recurrence-events" at all depths and impurity measurements.

**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.

<your answer goes here>

The impurity metric providing higher quality predictions varies; however, for “no-recurrence-events”, gini at max depth 2 shows the highest F1 score, indicating a slightly better performance. For 'recurrence-events', entropy at max depth 4 yields the best F1 score.

**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.

<your answer goes here>

The model's initial capacity to predict "recurrence-events" improves with increasing max depth; this is demonstrated by an increase in F1 scores up to a certain point before decreasing or declining, indicating overfitting. This pattern is concerning because it may result in a model that performs badly when applied to fresh data.

**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.

<your answer goes here>

When the maximum depth is increased to 20, performance measurements may show an early trend of improvement, which could be followed by a decline or plateauing. This is indicative of overfitting, where the model learns the training data too well, including noise, which harms its performance on unseen data.