Te-Jou Hou(Carol)
AU ID: 5545955

Part 1: Iris

1. The samples where the classifiers agree

```
The sample input: [6.4 2.8 5.6 2.1]
KNN Prediction: virginica
DT Prediction: virginica

The sample input: [4.8 3.4 1.6 0.2]
KNN Prediction: setosa
DT Prediction: setosa

The sample input: [7.7 2.6 6.9 2.3]
KNN Prediction: virginica
DT Prediction: virginica

The sample input: [5.7 4.4 1.5 0.4]
KNN Prediction: setosa
DT Prediction: setosa

The sample input: [6.3 2.7 4.9 1.8]
KNN Prediction: virginica
DT Prediction: virginica

The sample input: [7.7 3.  6.1 2.3]
KNN Prediction: virginica
DT Prediction: virginica

The sample input: [5.2 4.1 1.5 0.1]
KNN Prediction: setosa
DT Prediction: setosa

The sample input: [4.9 3.  1.4 0.2]
KNN Prediction: setosa
DT Prediction: setosa

The sample input: [6.5 3.  5.5 1.8]
KNN Prediction: virginica
DT Prediction: virginica

The sample input: [4.8 3.  1.4 0.3]
KNN Prediction: setosa
DT Prediction: setosa

The sample input: [5.  3.5 1.3 0.3]
KNN Prediction: setosa
DT Prediction: setosa

The sample input: [6.4 3.2 5.3 2.3]
KNN Prediction: virginica
DT Prediction: virginica

The sample input: [5.  3.4 1.6 0.4]
KNN Prediction: setosa
DT Prediction: setosa

The sample input: [5.2 3.4 1.4 0.2]
KNN Prediction: setosa
DT Prediction: setosa

The sample input: [6.7 3.  5.2 2.3]
KNN Prediction: virginica
DT Prediction: virginica

The sample input: [6.2 2.9 4.3 1.3]
KNN Prediction: versicolor
DT Prediction: versicolor

The sample input: [6.  2.2 4.  1. ]
KNN Prediction: versicolor
DT Prediction: versicolor

The sample input: [7.7 2.8 6.7 2. ]
KNN Prediction: virginica
DT Prediction: virginica

The sample input: [6.8 3.  5.5 2.1]
KNN Prediction: virginica
DT Prediction: virginica

The sample input: [7.2 3.2 6.  1.8]
KNN Prediction: virginica
DT Prediction: virginica

The sample input: [5.4 3.4 1.5 0.4]
KNN Prediction: setosa
DT Prediction: setosa

The sample input: [6.3 2.5 5.  1.9]
KNN Prediction: virginica
DT Prediction: virginica
```

2. The samples where the classifiers disagree

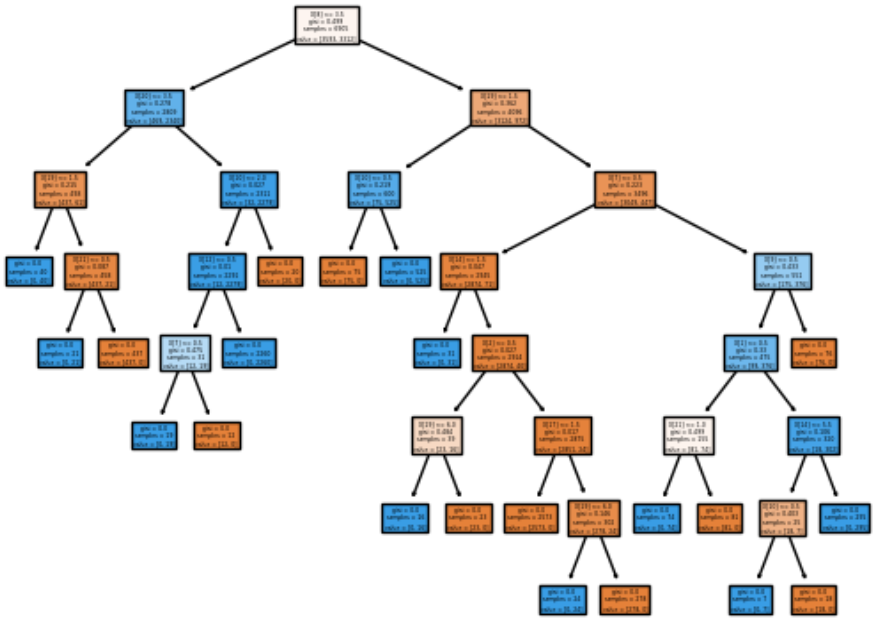
```
The sample input: [6.7 3.  5.  1.7]
KNN Prediction: versicolor
DT Prediction: virginica
```

KNN (K-Nearest Neighbors) makes predictions based on the proximity of data points, whereas decision trees rely on rules derived from data splits. Among the five nearest data points to [6.7, 3.0, 5.0, 1.7], the majority may belong to the "versicolor" class. However, according to the splitting rules learned by the decision tree, the group that [6.7, 3.0, 5.0, 1.7] falls into has a higher proportion of "virginica" samples, which leads to different predictions by the two methods.

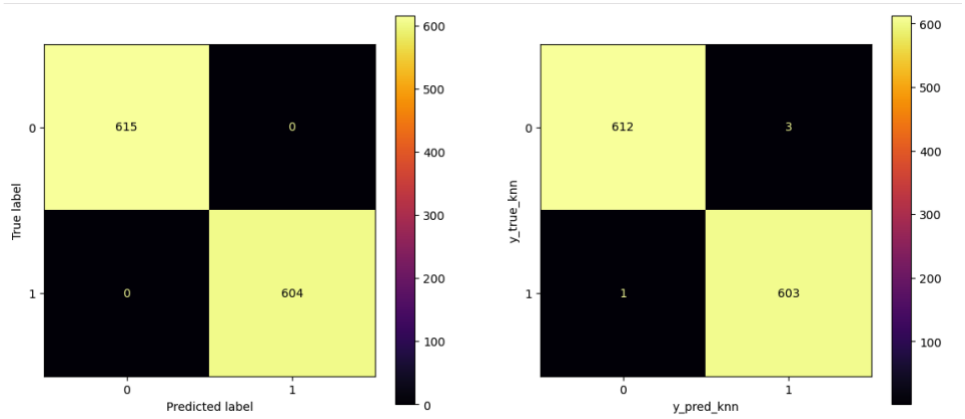
Te-Jou Hou(Carol)
AU ID: 5545955

Part 2: Mushrooms

1. Decision Tree



2. Confusion matrix for both the decision tree and knn



From the confusion matrix, we can see that the decision tree classifier performs very well with no misclassifications. However, the K-NN classifier has one false negative and three false positives. This is likely due to the choice of a large K value, which causes the K-NN model to rely too much on overall trends, leading to overly simplistic and underfitting results.

Therefore, by observing the numbers of true positives, false positives, true negatives, and false negatives in the confusion matrix, we can assess the model's accuracy and misclassification behavior, providing guidance for selecting the right model and parameters.

Te-Jou Hou(Carol)
AU ID: 5545955

Part 3: Heart Failure

1. The classification report for both the decision tree and knn

```
Full set of features(Decision Tree):
              precision    recall  f1-score   support

           0       0.71      0.86      0.78        29
           1       0.60      0.38      0.46        16

    accuracy                           0.69        45
   macro avg       0.66      0.62      0.62        45
weighted avg       0.67      0.69      0.67        45

Reduce set of features(Decision Tree):
              precision    recall  f1-score   support

           0       0.77      0.83      0.80        29
           1       0.64      0.56      0.60        16

    accuracy                           0.73        45
   macro avg       0.71      0.70      0.70        45
weighted avg       0.73      0.73      0.73        45

Full set of features(KNN):
              precision    recall  f1-score   support

           0       0.67      0.83      0.74        29
           1       0.44      0.25      0.32        16

    accuracy                           0.62        45
   macro avg       0.56      0.54      0.53        45
weighted avg       0.59      0.62      0.59        45

Reduce set of features(KNN):
              precision    recall  f1-score   support

           0       0.77      0.83      0.80        29
           1       0.64      0.56      0.60        16

    accuracy                           0.73        45
   macro avg       0.71      0.70      0.70        45
weighted avg       0.73      0.73      0.73        45
```

Using a simplified feature set (serum creatinine and ejection fraction) improved prediction performance in both classifiers. For the Decision Tree, accuracy increased from 0.69 to 0.73. Similarly, for KNN, accuracy improved from 0.62 to 0.73. These results support the paper's hypothesis.