Quiz 9 - Model Evaluation in KNIME and Spark

1. KNIME: In the confusion matrix as viewed in the Scorer node, low_humidity_day is:

- **the target class label**
- the predicted class label
- the only input variable that is categorical

2. KNIME: In the confusion matrix, what is the difference between low_humidity_day and Prediction(low_humidity_day)?

- **low_humidity_day is the target class label, and Prediction(low_humidity_day) is the predicted class label**
- low_humidity_day is the predicted class label, and Prediction(low_humidity_day) is the target class label
- There is no difference. The two are the same

3. KNIME: In the Table View of the Interactive Table, each row is color-coded. Blue specifies:

- **that the target class label for the sample is humidity_not_low**
- that the target class label for the sample is humidity_low
- that the predicted class label for the sample is humidity_not_low
- that the predicted class label for the sample is humidity_low

4. KNIME: To change the colors used to color-code each sample in the Table View of the Interactive Table node:

- **change the color settings in the Color Manager node**
- change the color settings in the Interactive Table dialog
- It is not possible to change these colors

5. KNIME: In the Table View of the Interactive Table, the values in RowID are not consecutive because:

- **the RowID values are from the original dataset, and only the test samples are displayed here**
- the samples are randomly ordered in the table
- only a few samples from the test set are randomly selected and displayed here

6. Spark: To get the error rate for the decision tree model, use the following code:

```
print ("Error = %g " % (1.0 - accuracy)) [X]
evaluator = MuticlassClassificationEvaluator(
    labelCol="label",
    predictionCol="prediction",
    metricName="error")
error = evaluator.evaluate(1 - predictions)
```

7. Spark: To print out the accuracy as a percentage, use the following code:

```
print ("Accuracy = %.2g" % (accuracy * 100)) [X]
print ("Accuracy = %100g" % (accuracy))
print ("Accuracy = %100.2g" % (accuracy))
```

8. Spark: In the last line of code in Step 4, the confusion matrix is printed out. If the “transpose()” is removed, the confusion matrix will be displayed as:

```
array([[87., 14.],  [X]
       [26., 83.]])
array([[83., 26.],
       [14., 87.]])
array([[83., 87.],
       [14., 26.]])
```