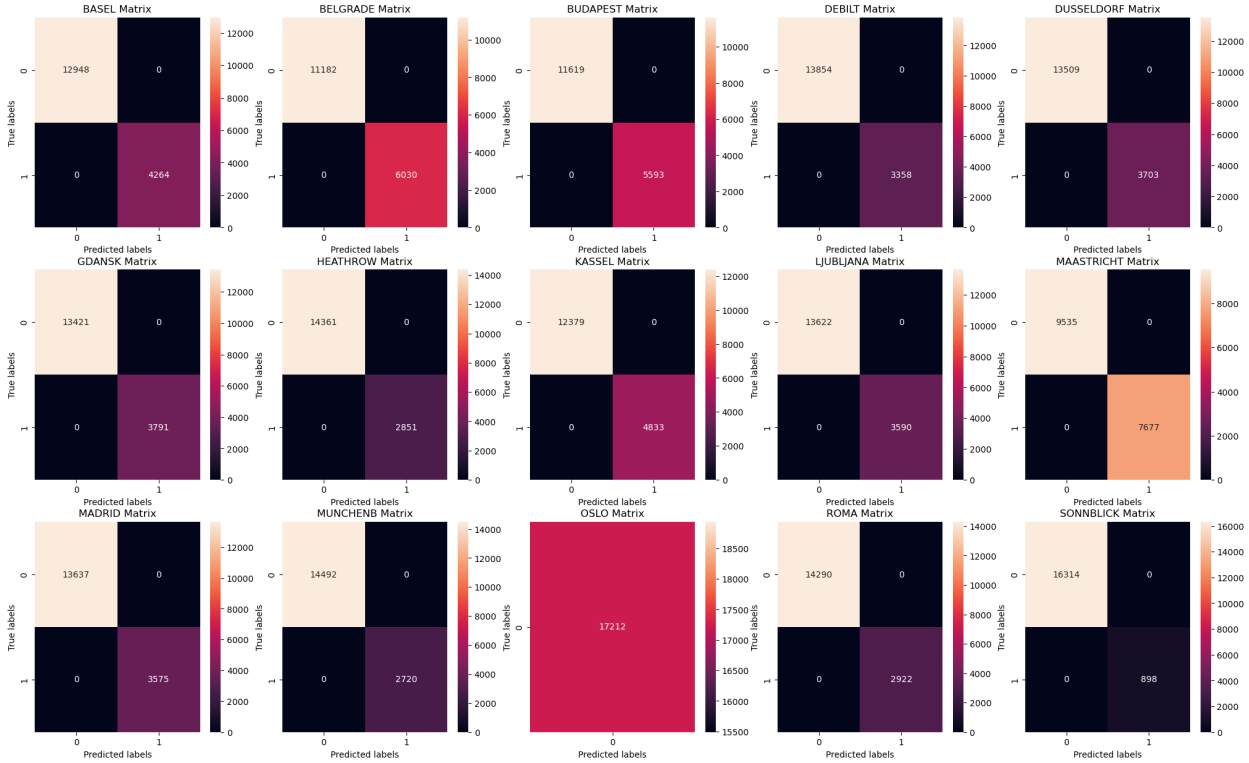
Shaquille Obomeghie
Exercise 1.5

1. Decision tree

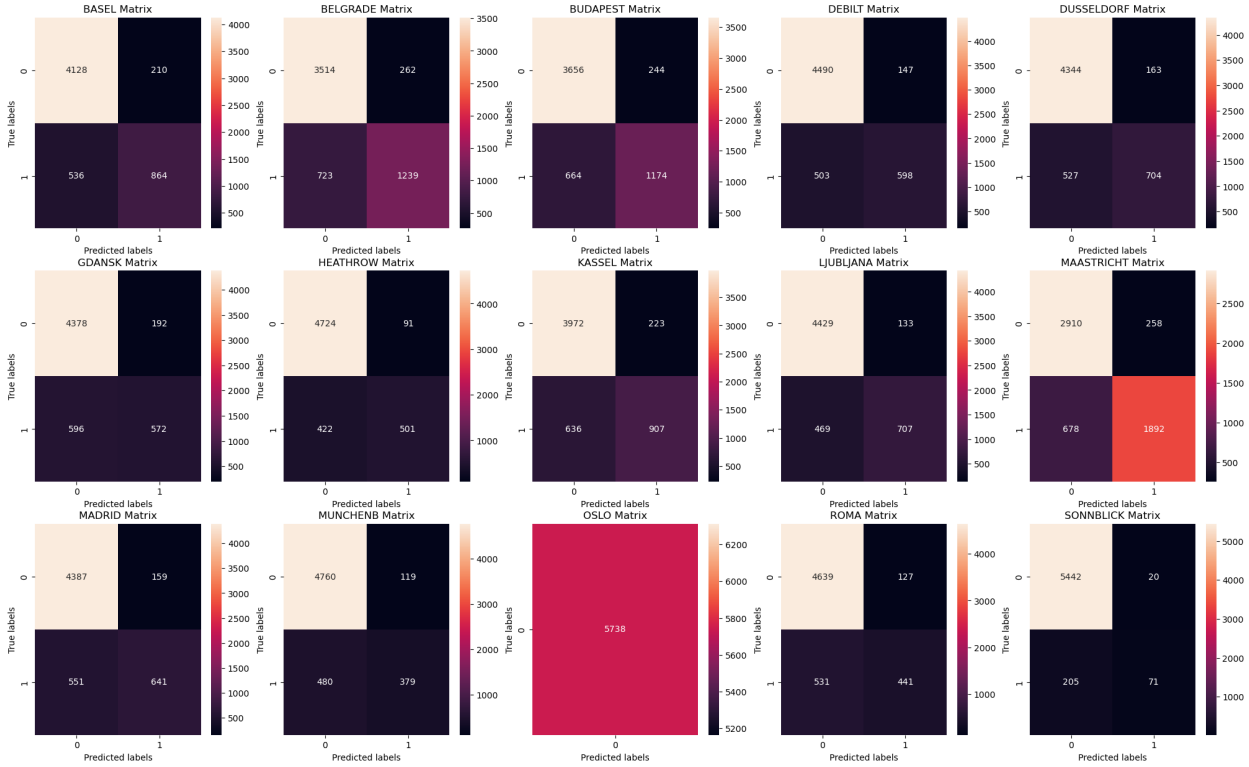
Training accuracy: 0.60 or 60%
Testing accuracy: 0.63 or 63%

Does the decision tree need to be pruned?
Yes, because the decision tree for this data is quite complex. It has many nodes which can be difficult to understand and interpret. Pruning will simplify and reduce the size of the tree making it easier to understand.

Training



Testing

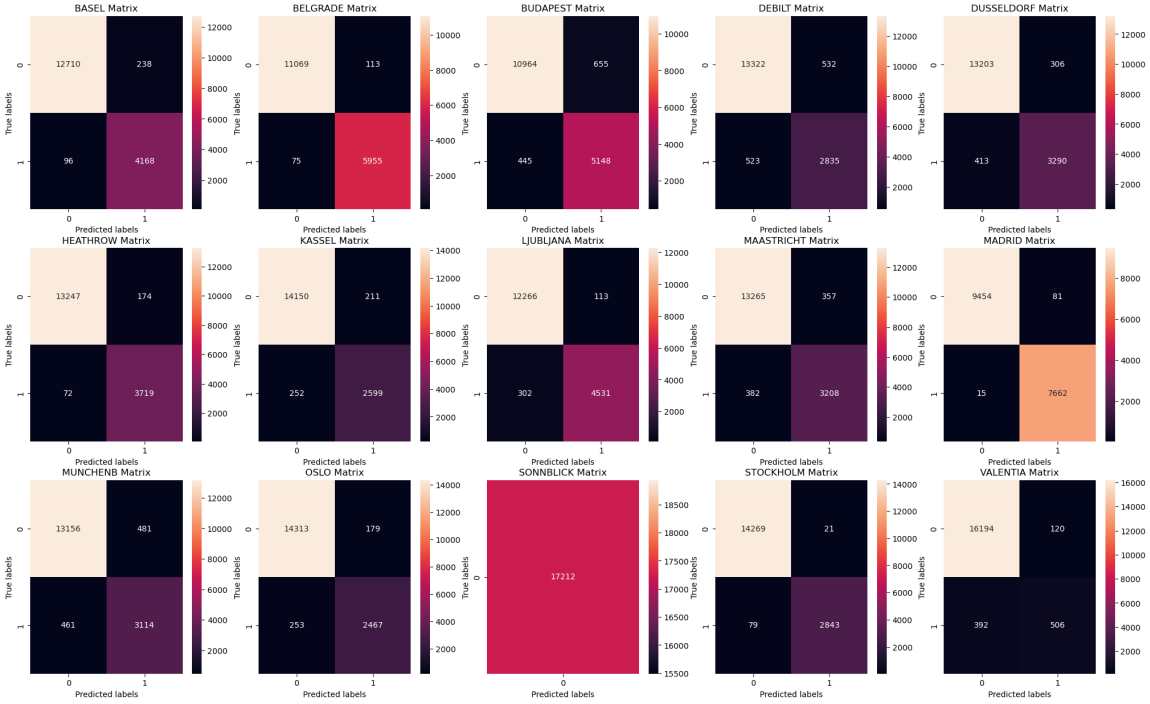


2. Artificial neural network model

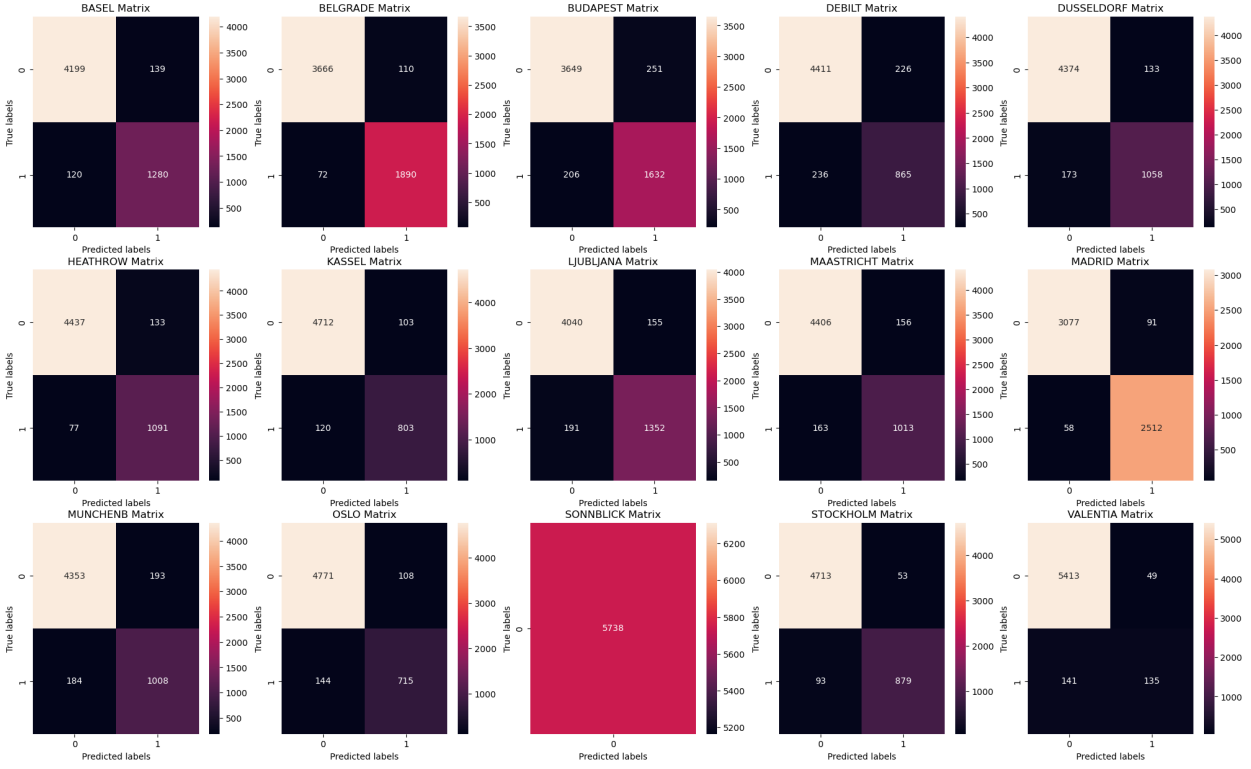
Scenario 1

- MLPClassifier(hidden_layer_sizes=(50, 10), max_iter=500, tol=0.0001)
  Training accuracy: 71%
  Testing accuracy: 62%

Training data



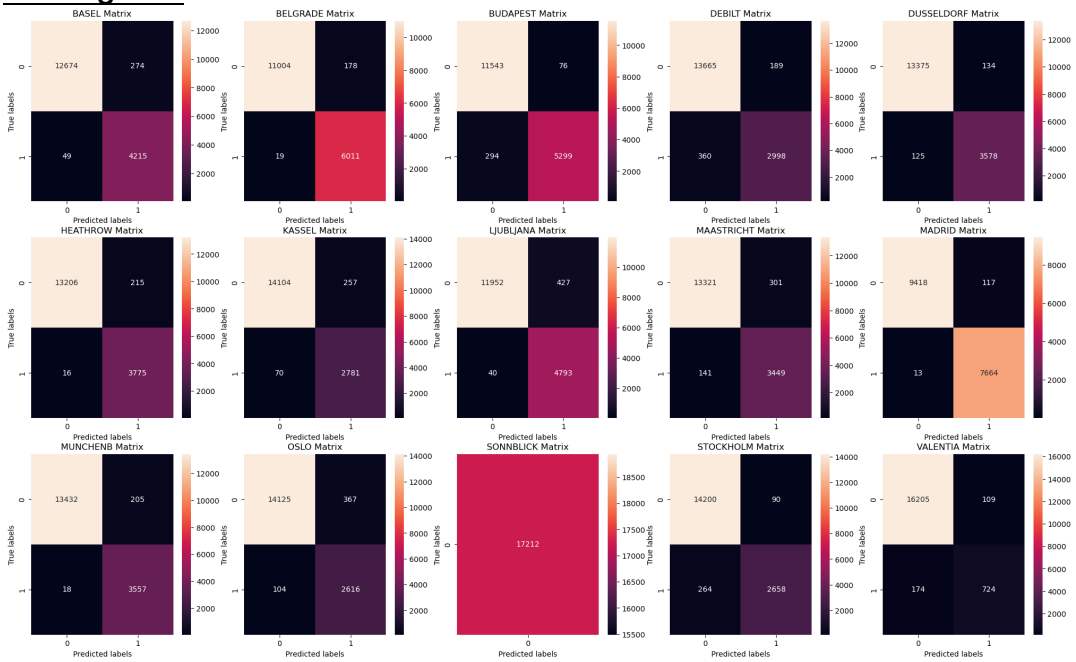
Test data



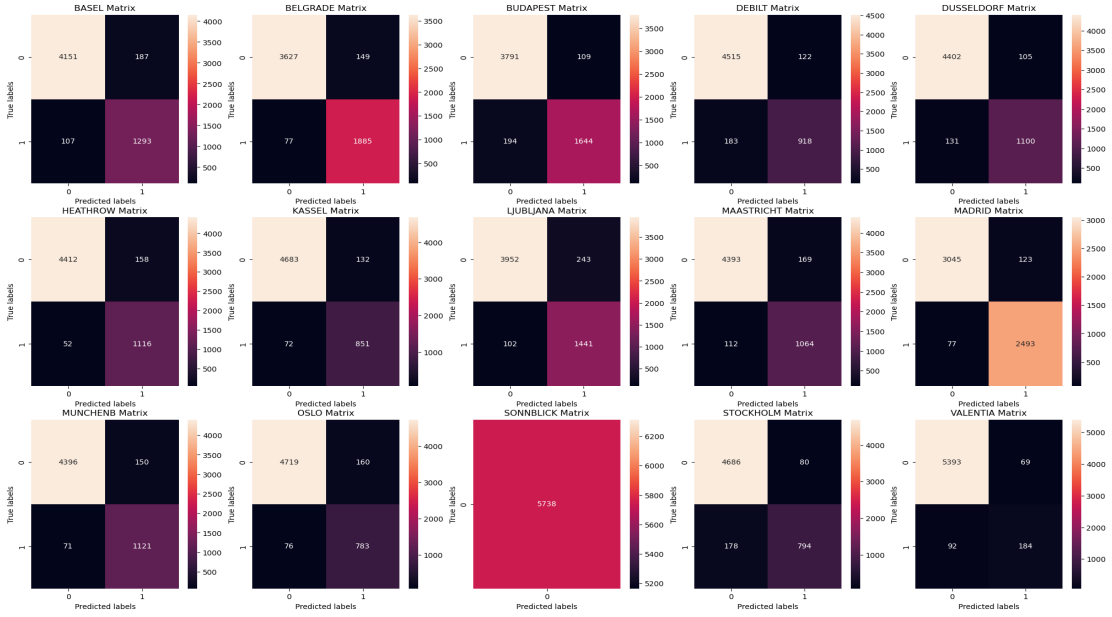
- **Scenario 2**

- MLPClassifier(hidden_layer_sizes=(60, 30, 20), max_iter=1200, tol=0.0001)
  Training data: 79%
  Testing data: 62%

Training data



Testing data



There is an overfitting for Sonnblick whereby there is a 100% accuracy.

3. Based on the accuracy of each model, the **Artificial neural network model** has the highest accuracy model. It is also better with complex data sets due to the ability to manually manipulate.