Task 1.5
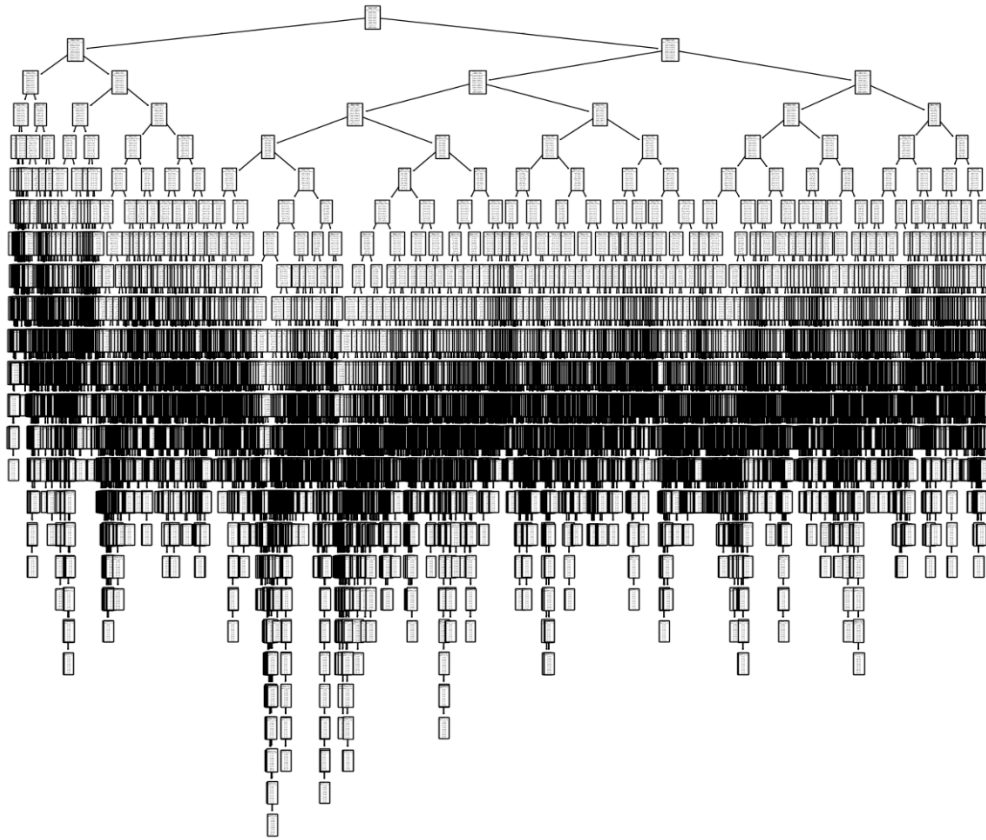
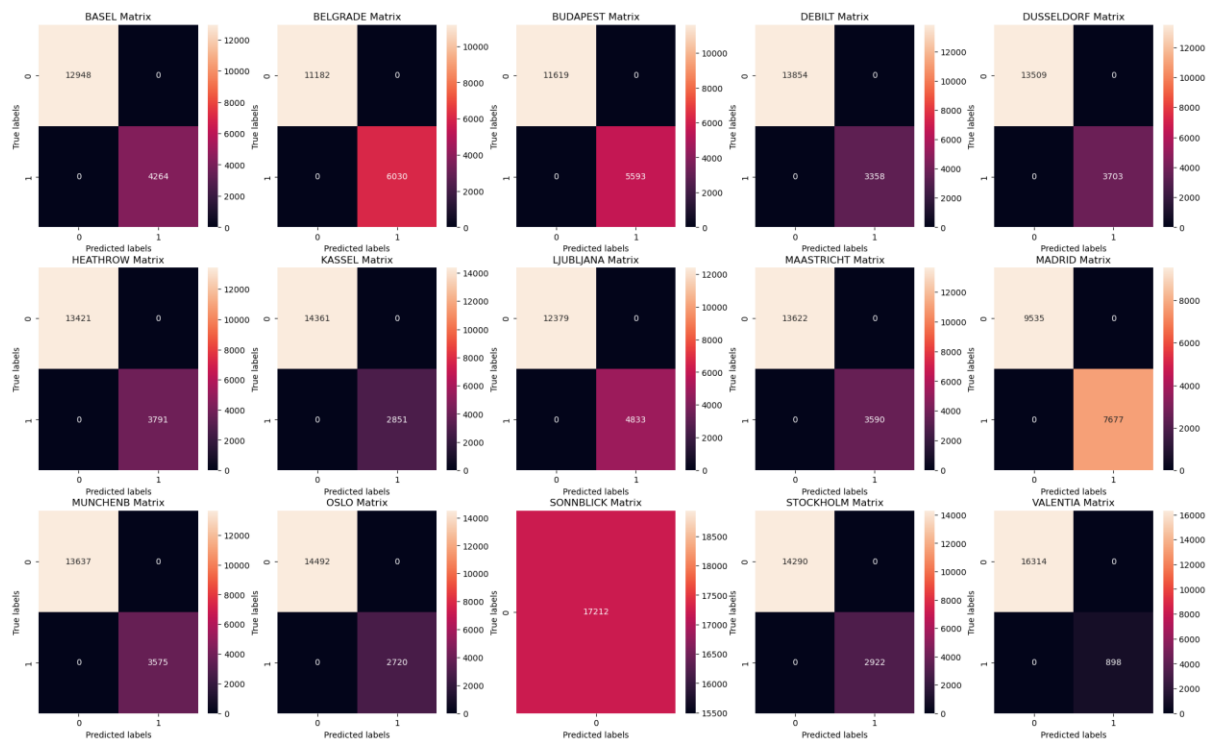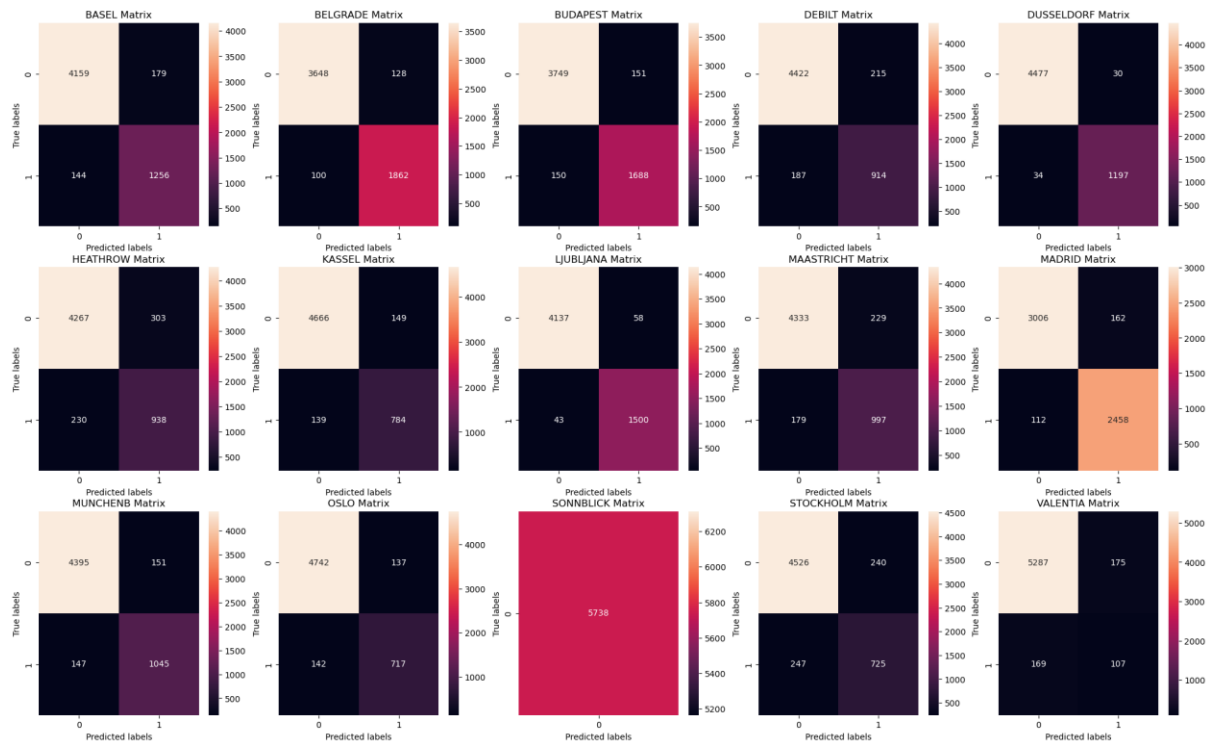**Decision Tree Model – Scaled Data**
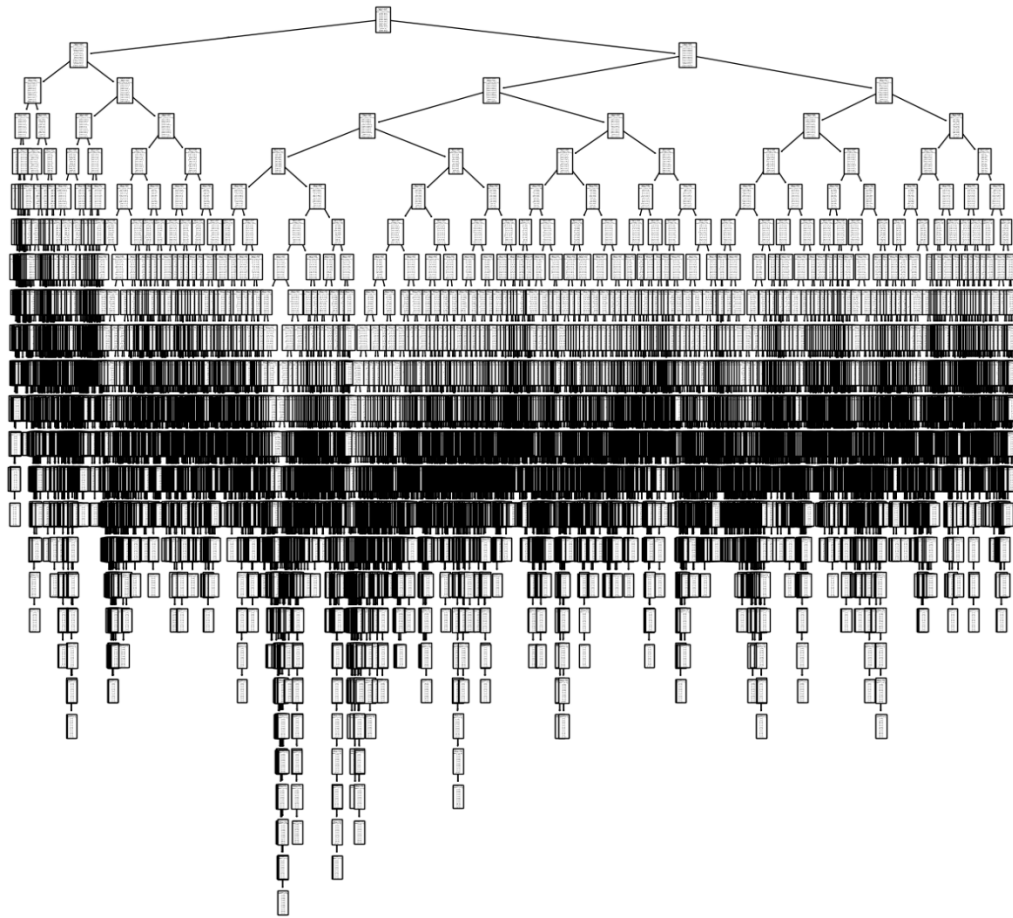

Train accuracy: 60.7%

Test accuracy: 56.1%

BASEL Matrix

	Predicted 0	Predicted 1
True 0	4159	179
True 1	144	1256

BELGRADE Matrix

	Predicted 0	Predicted 1
True 0	3648	128
True 1	100	1862

BUDAPEST Matrix

	Predicted 0	Predicted 1
True 0	3749	151
True 1	150	1688

DEBILT Matrix

	Predicted 0	Predicted 1
True 0	4422	215
True 1	187	914

DUSSELDORF Matrix

	Predicted 0	Predicted 1
True 0	4477	30
True 1	34	1197

HEATHROW Matrix

	Predicted 0	Predicted 1
True 0	4267	303
True 1	230	938

KASSEL Matrix

	Predicted 0	Predicted 1
True 0	4666	149
True 1	139	784

LJUBLJANA Matrix

	Predicted 0	Predicted 1
True 0	4137	58
True 1	43	1500

MAASTRICHT Matrix

	Predicted 0	Predicted 1
True 0	4333	229
True 1	179	997

MADRID Matrix

	Predicted 0	Predicted 1
True 0	3006	162
True 1	112	2458

MUNCHENB Matrix

	Predicted 0	Predicted 1
True 0	4395	151
True 1	147	1045

OSLO Matrix

	Predicted 0	Predicted 1
True 0	4742	137
True 1	142	717

SONNBLICK Matrix

	Predicted 0
True 0	5738

STOCKHOLM Matrix

	Predicted 0	Predicted 1
True 0	4526	240
True 1	247	725

VALENTIA Matrix

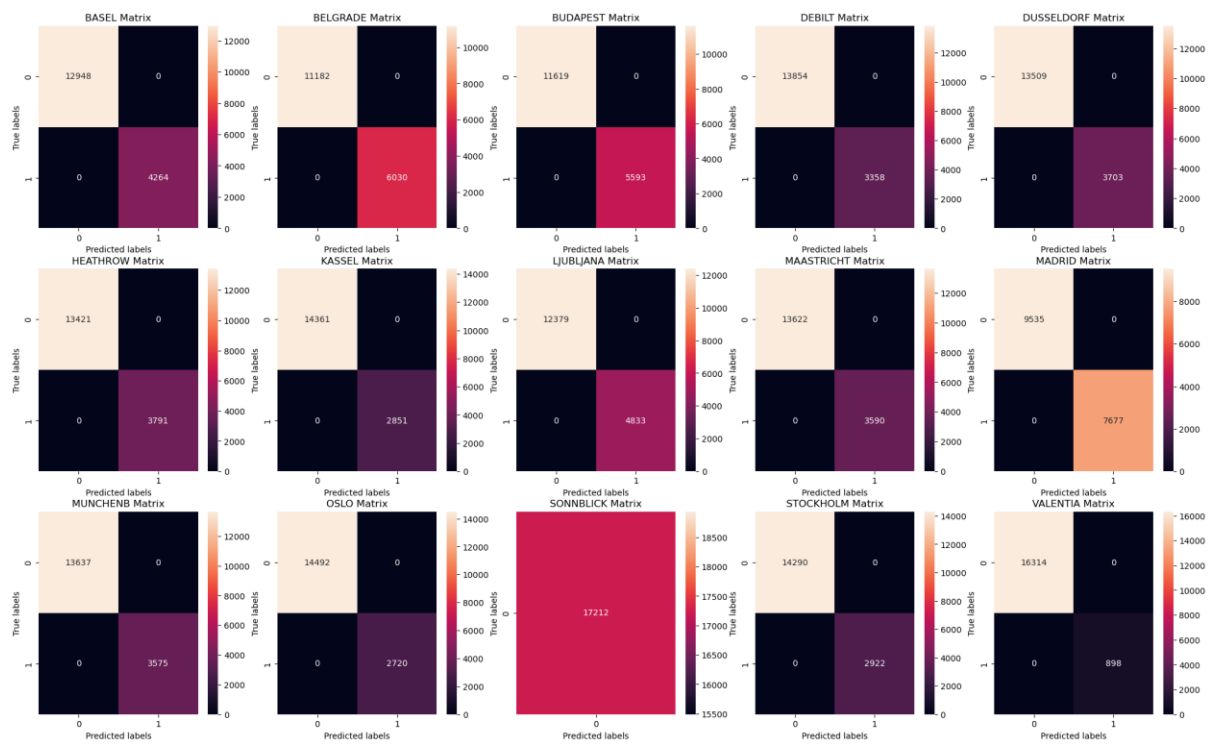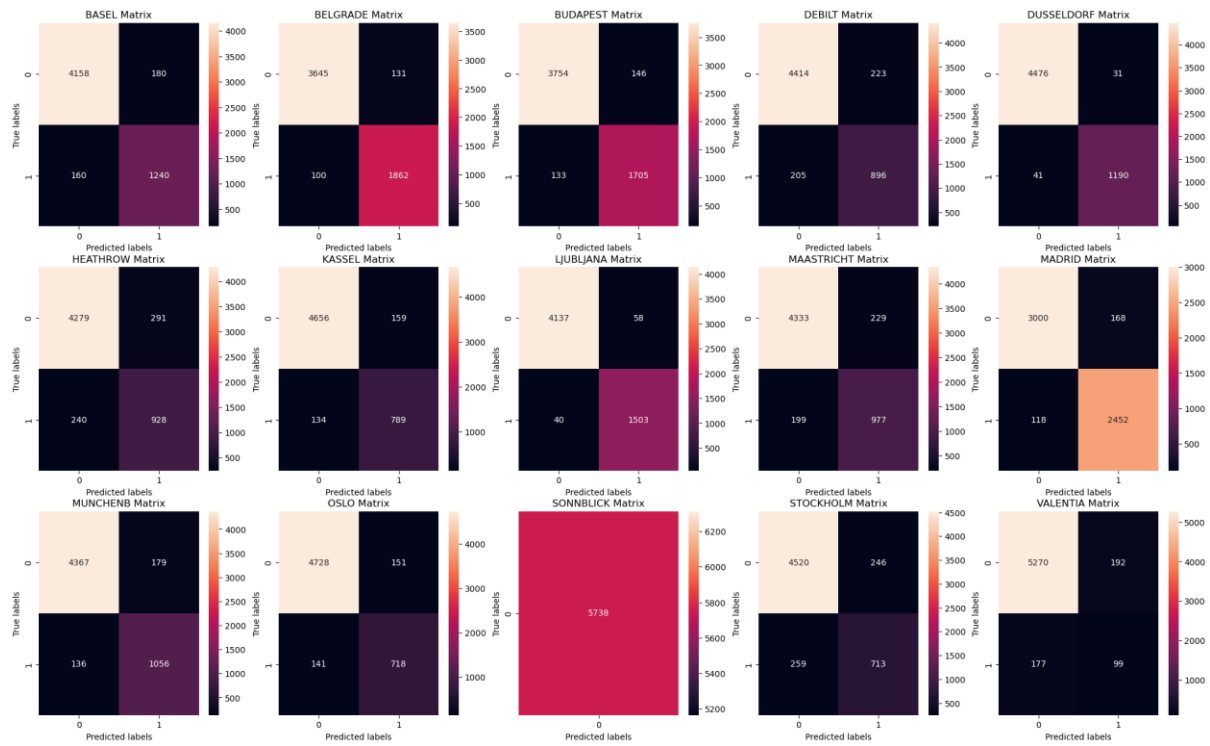|	Predicted 0	Predicted 1
True 0	5287	175
True 1	169	107

Decision Tree – Unscaled Data

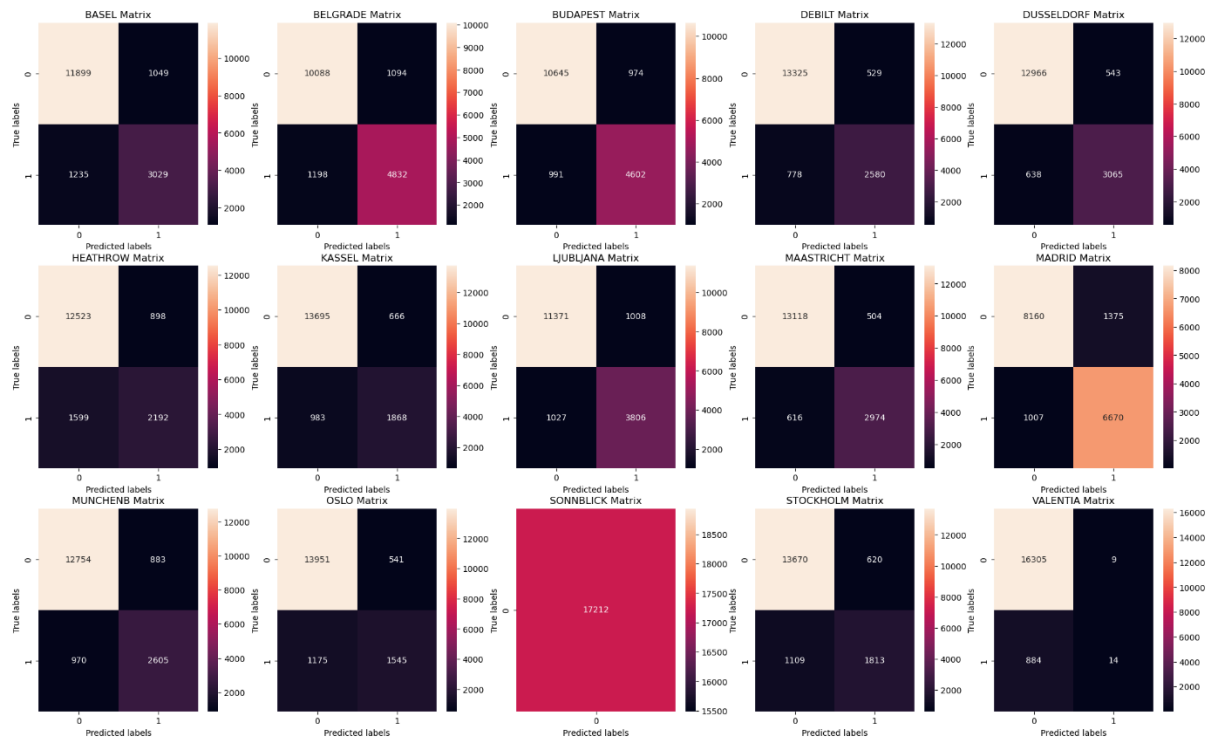Train accuracy: 61%

Test accuracy: 56%

ANN Model – Unscaled Data
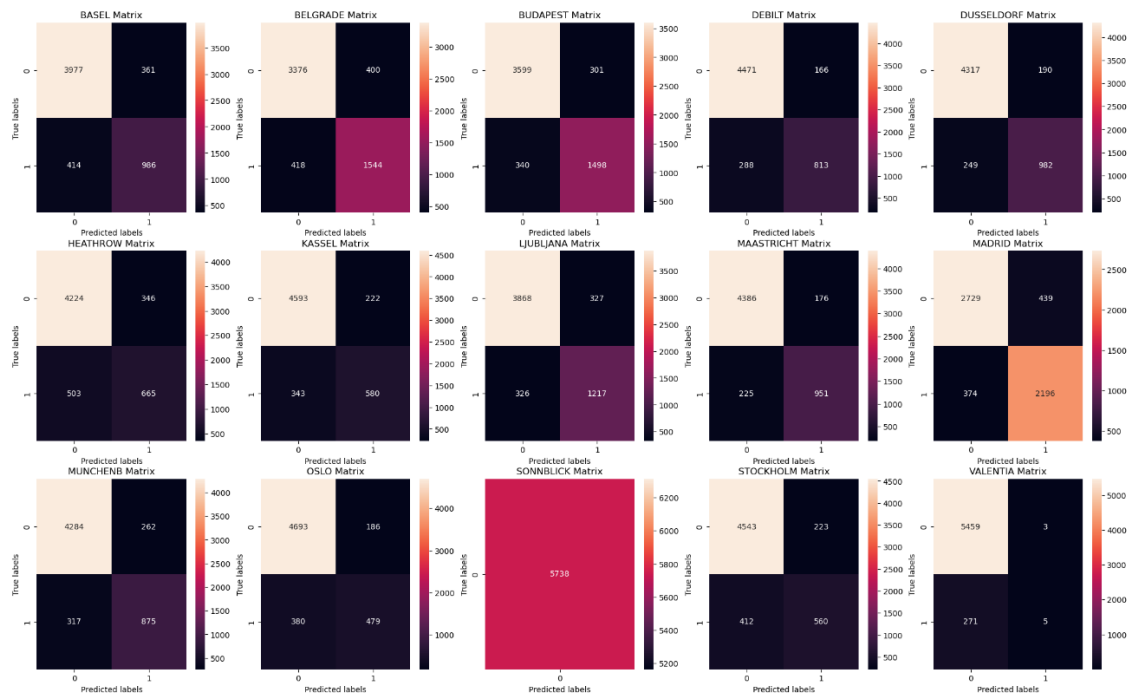
Hidden layer sizes: 5, 5, 5

Max. iter: 500

Tol: 0.0001
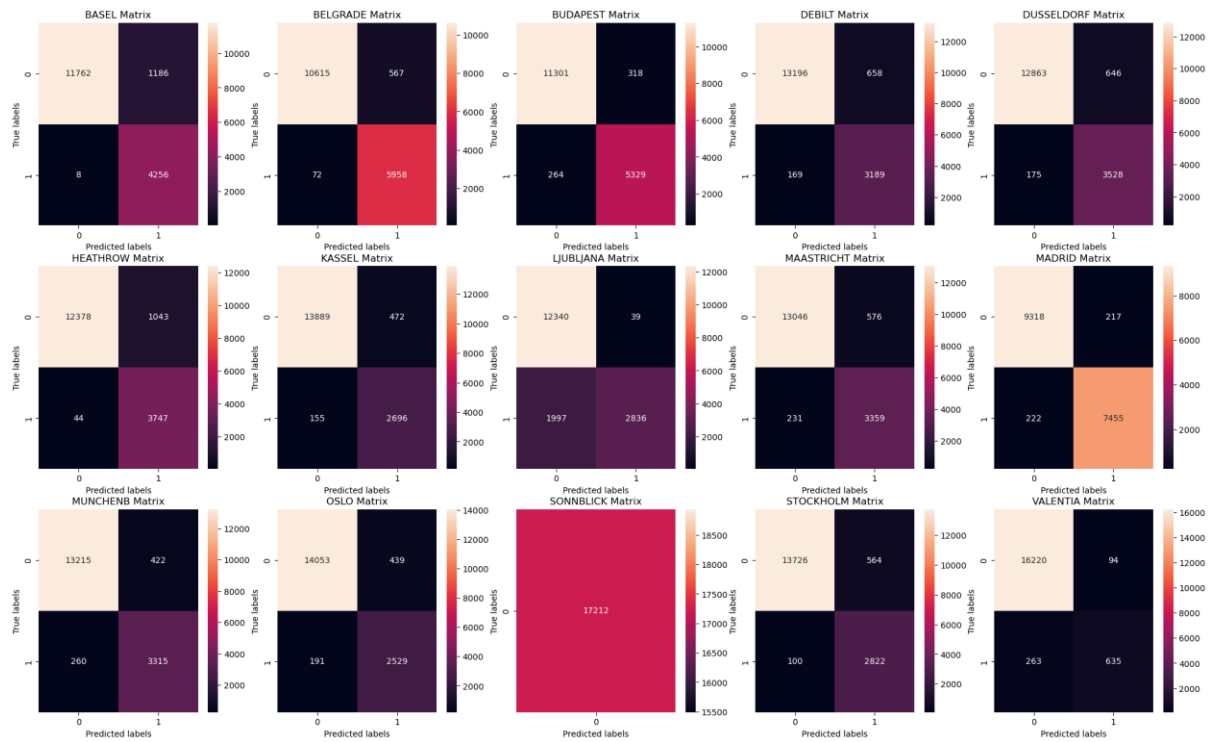
Train accuracy: 45.3%

Test accuracy: 45.2%

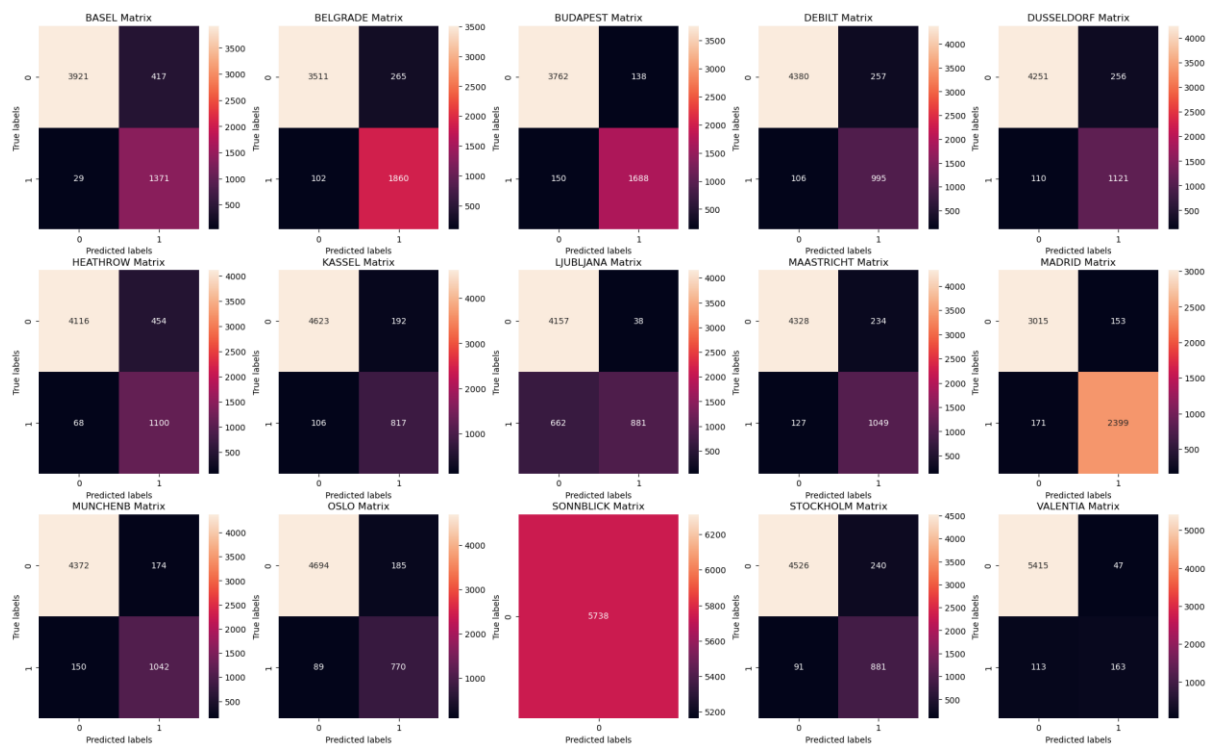Hidden layer sizes: 100, 50, 50

Max iter: 1000

Tol: 0.0001

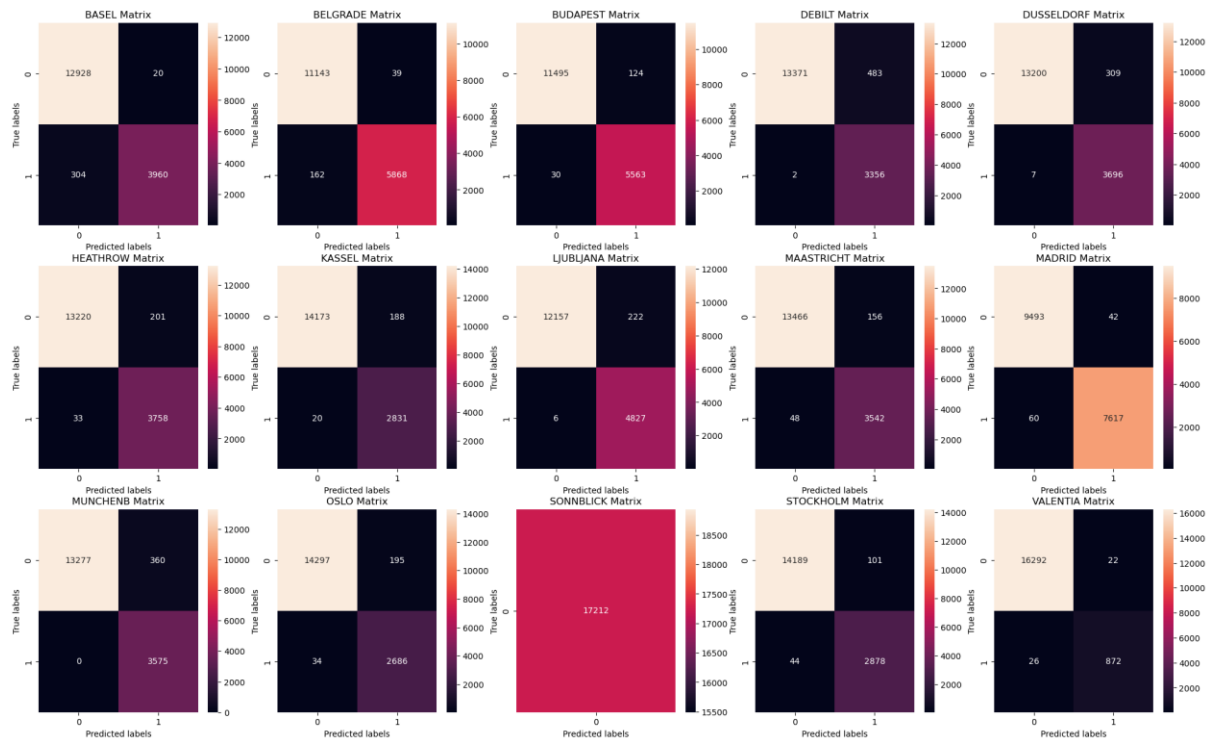Train accuracy: 62.3%
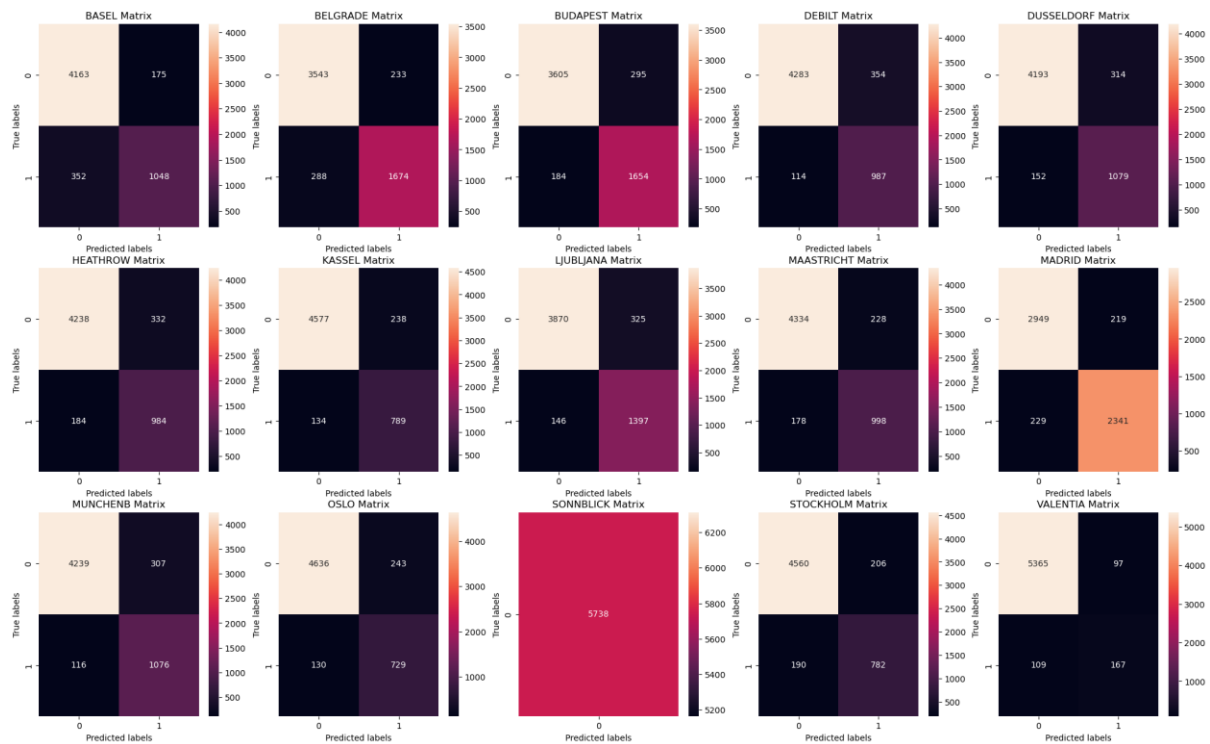

Test accuracy: 53.8%

Hidden layer sizes: 500, 250, 250

Max iter: 1000

Tol: 0.0001
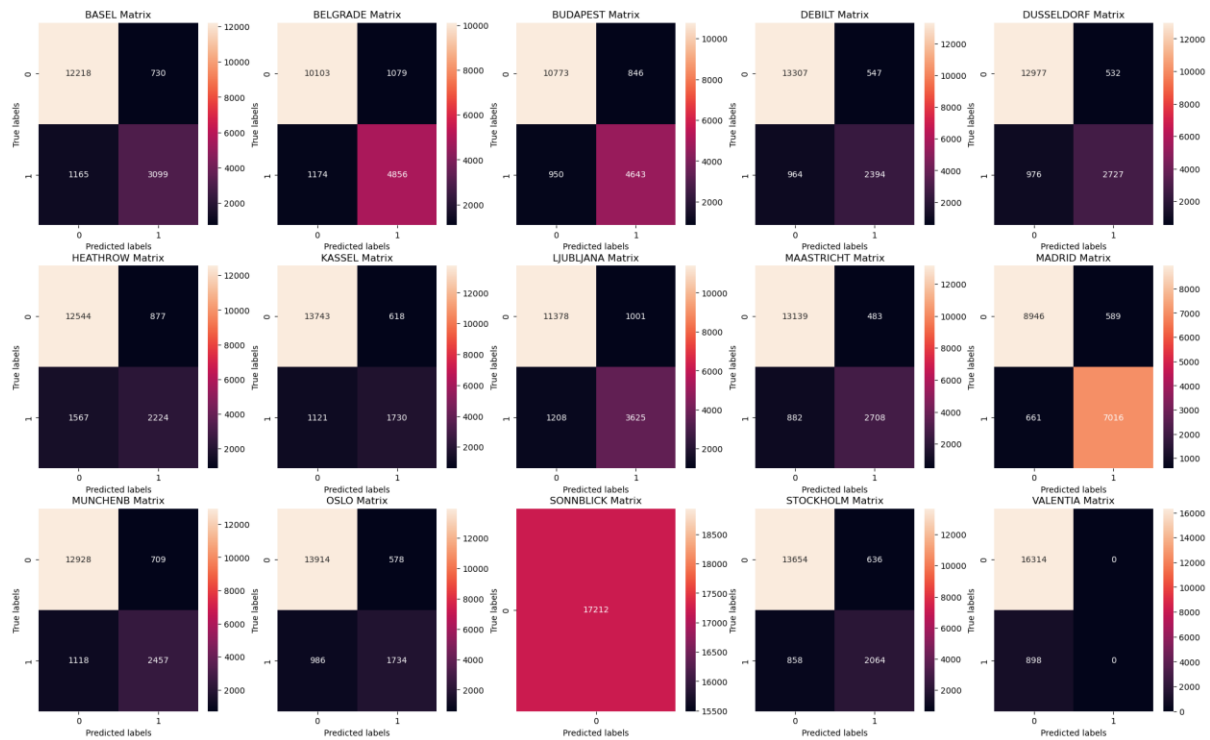
Train accuracy: 85.1%

Test accuracy: 51.2%

ANN Model – Scaled Data

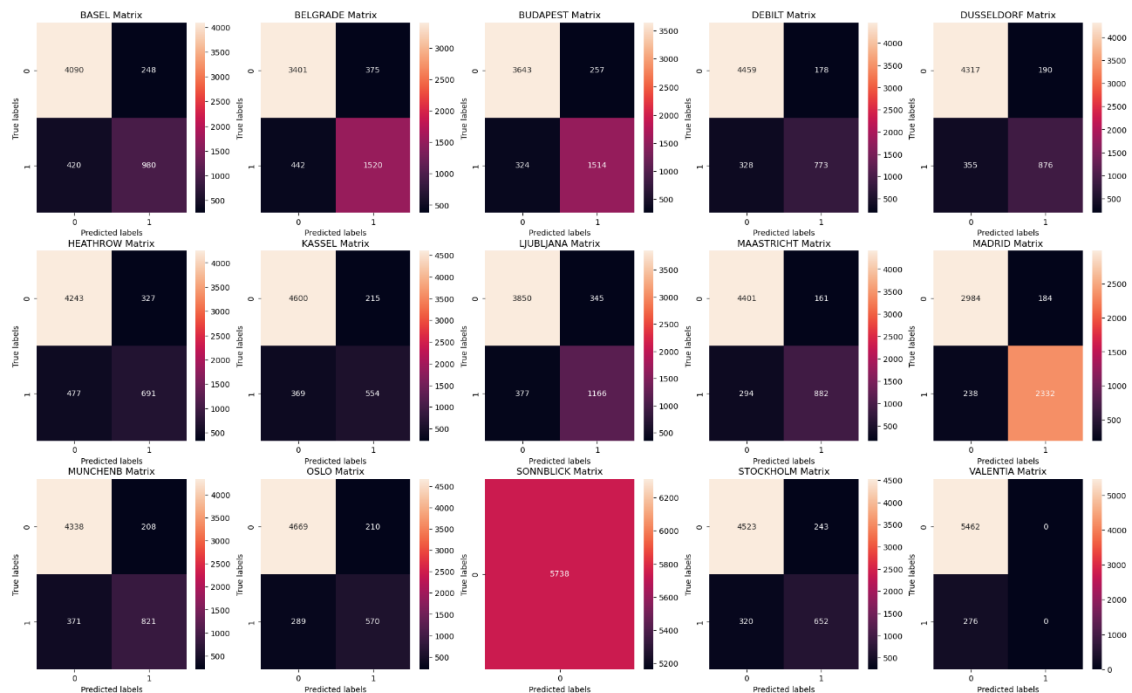Hidden layer sizes: 5, 5, 5

Max iter: 500

Tol: 0.0001

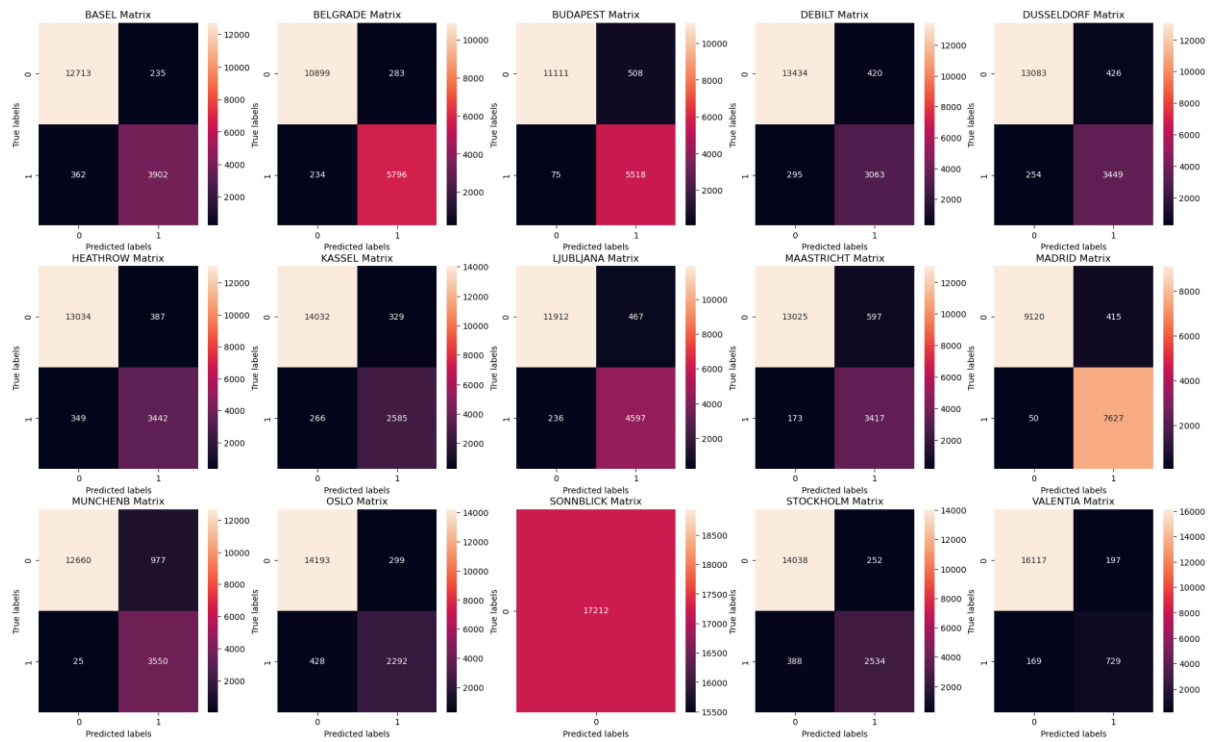Train accuracy: 46.8%

Test accuracy: 47.7%

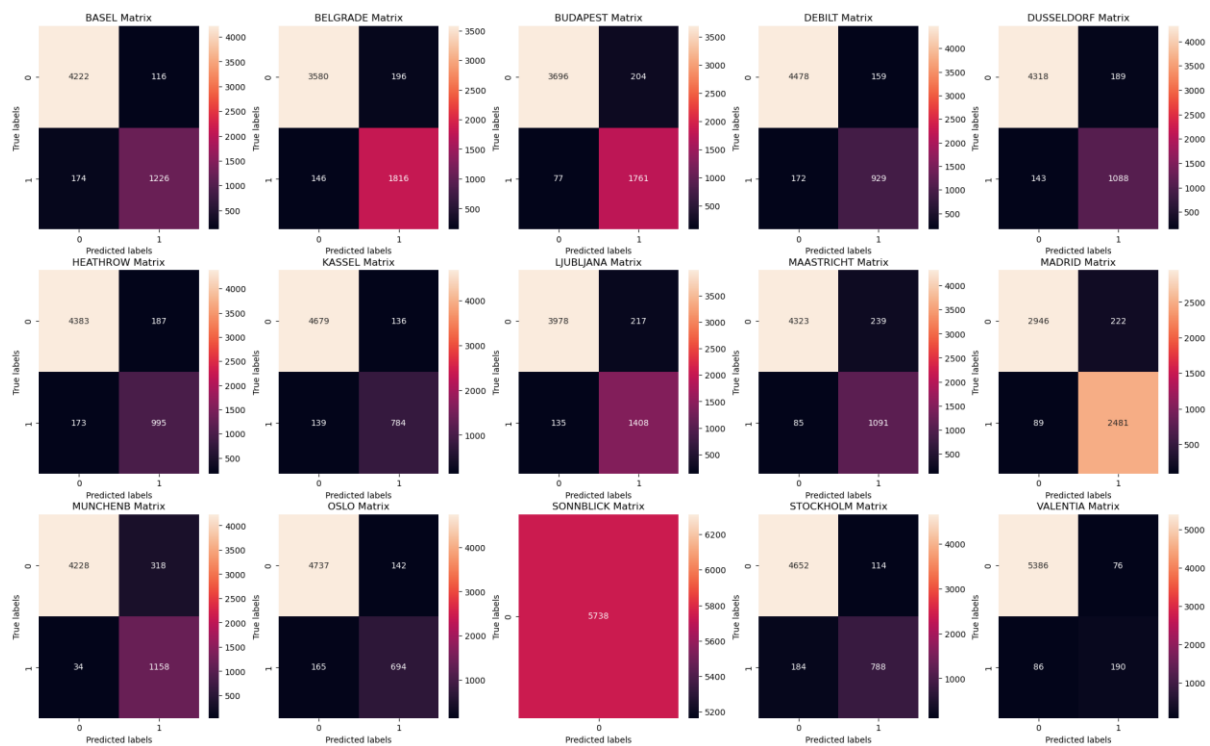Hidden layer sizes: 100, 50, 50

Max iter: 1000

Tol: 0.0001

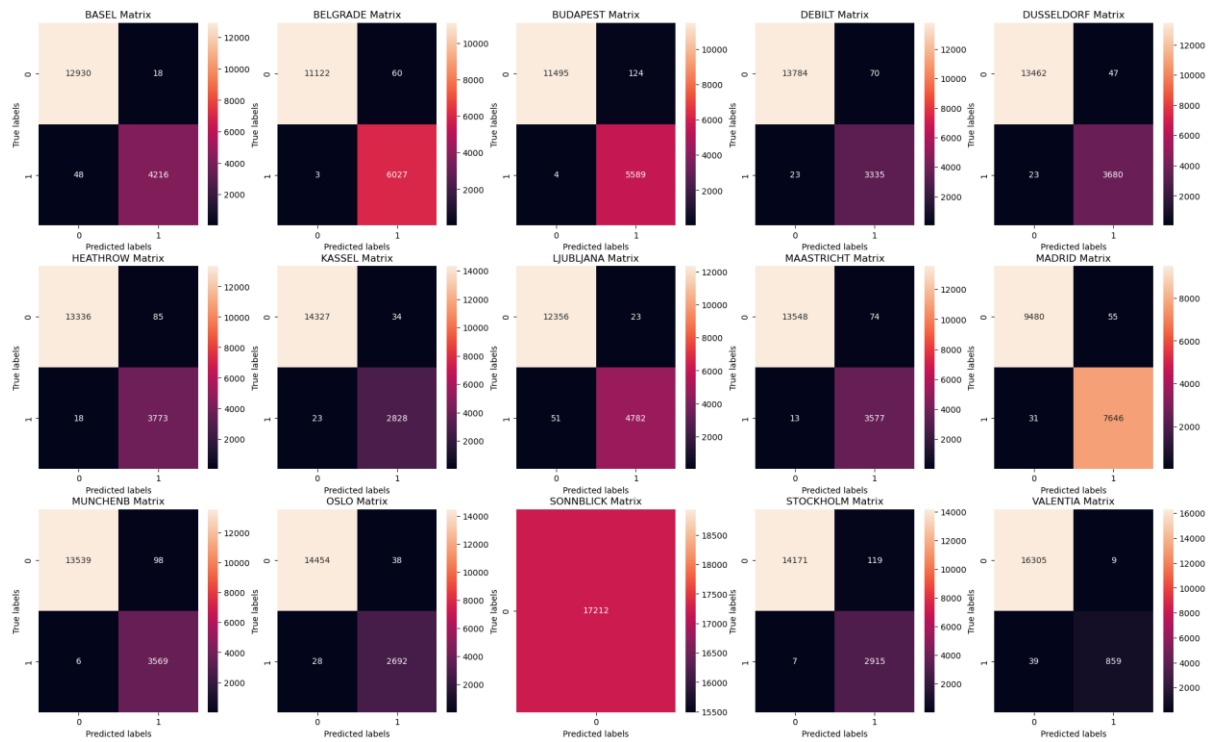Train accuracy: 64.2%

Test accuracy: 56.4%

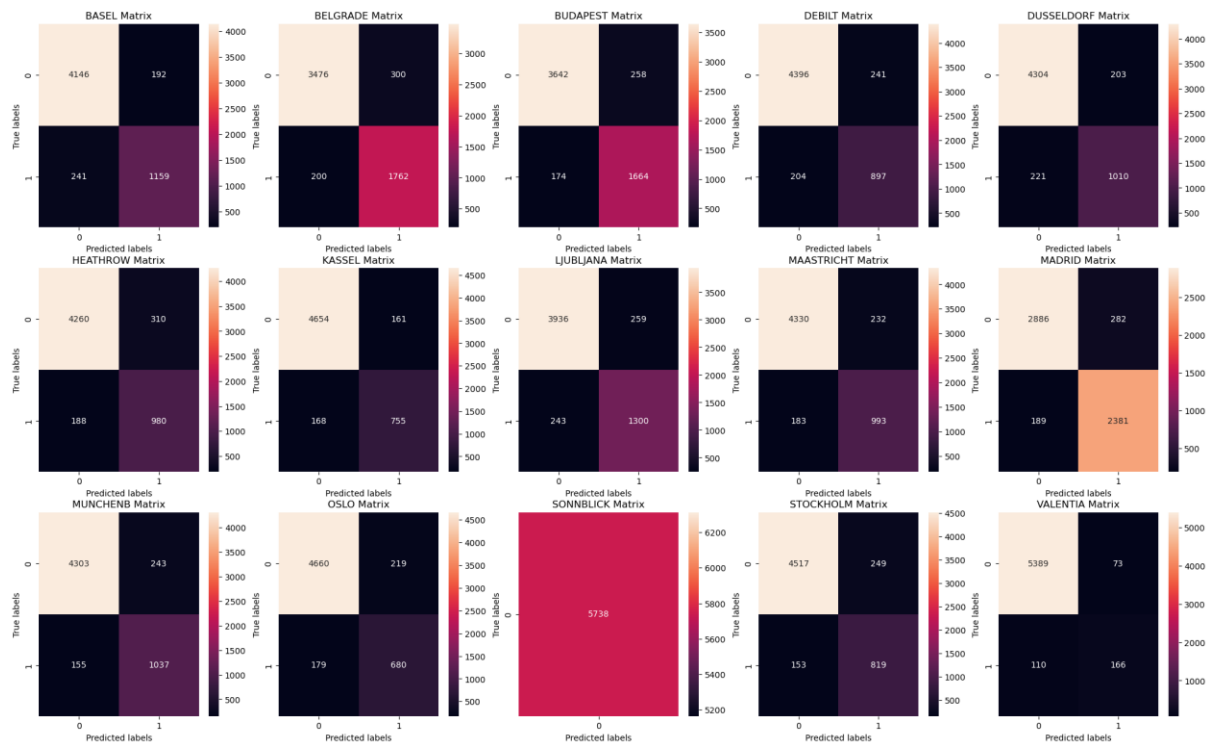Hidden layer sizes: 500, 250, 250

Max iter: 1000

Tol: 0.0001

Train accuracy: 94.1%

Test accuracy: 50.9%

Observations

The most accurate models using test data were the decision tree and an ANN with 3 layers of 100, 50 and 50 nodes. The ANN with the largest number of nodes per layer showed almost perfect accuracy with the training data, yet much lower accuracy with the test data. This means that these layer sizes resulted in overfitting the training data.

Scaling the data does not make a significant difference in the accuracy of either type of model. In both cases, scaling the data led to slightly higher accuracy with training data.

A decision tree (perhaps with fewer branches) or an ANN model with the appropriate number of nodes per layer would be useful for predicting pleasant weather.