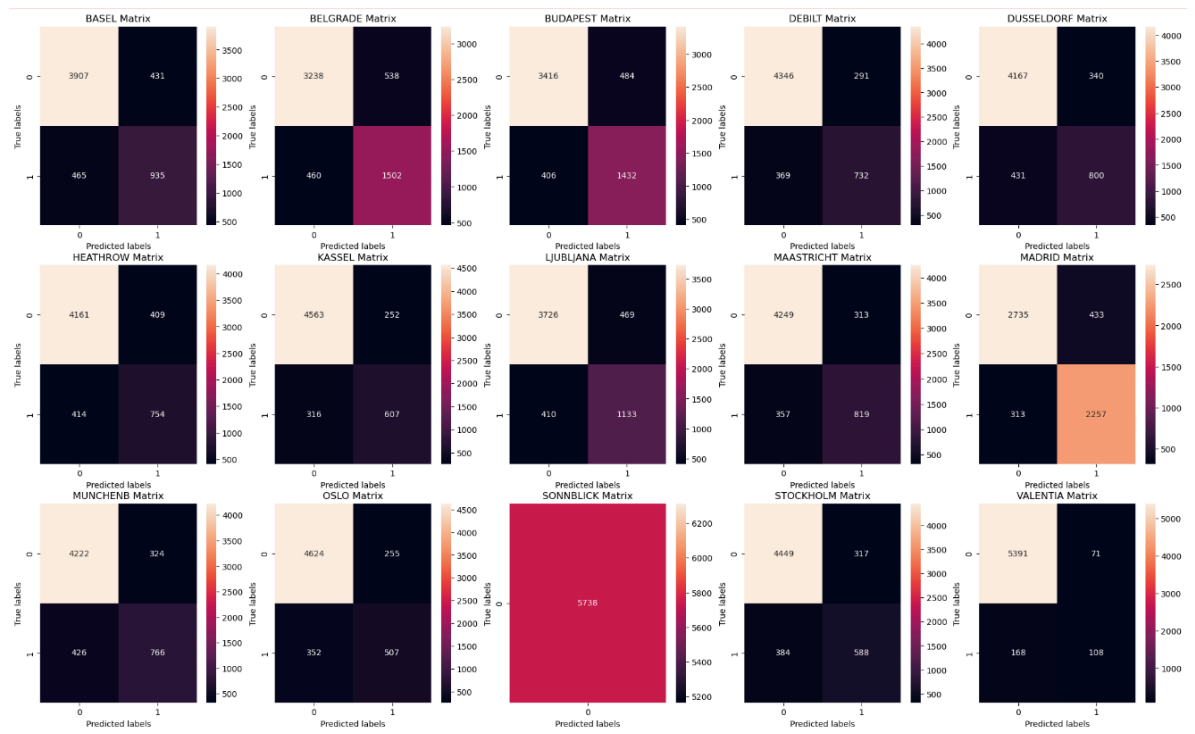


## Exercise 1.4: Supervised Learning Algorithms part 1



Weather Station	Accurate Predictions		False Positive	False Negative	Accuracy Rate
Basel	3907	935	465	431	84%
Belgrade	3238	1502	460	538	83%
Budapest	3416	1432	406	484	84%
Deblit	4346	732	369	291	88%
Dusseldorf	4167	800	431	340	87%
Heathrow	4161	754	414	409	86%
Kassel	4563	607	316	252	90%
Ljubljana	3726	1133	410	469	85%
Maastricht	4249	819	357	313	88%
Madrid	2735	2257	313	433	87%
Munchenb	4222	766	426	324	87%
Oslo	4624	507	352	255	89%
Sonnblick	5738		0	0	100%

Stockholm	4449	588	384	317	88%
Valentia	5391	108	168	71	96%

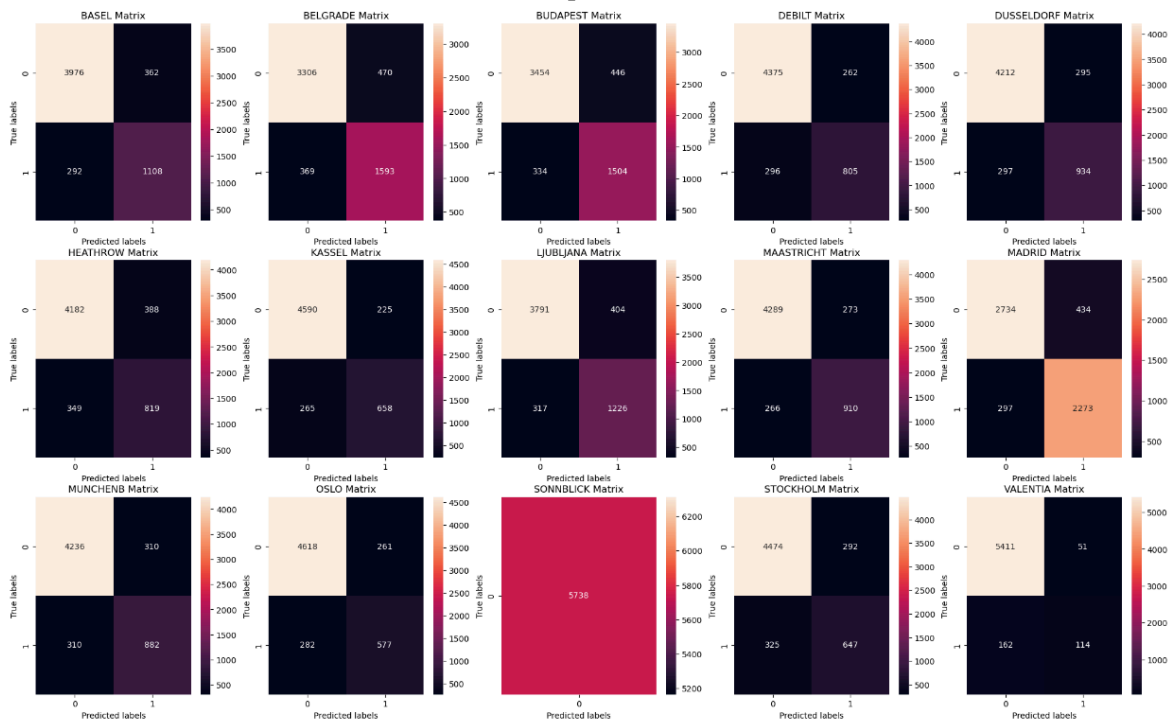
The average accuracy of this algorithm is 88%, which can be considered a high accuracy rate. Sonnblick shows a 100% accuracy, and while at first it might seem positive, when investigating further we can see that all values in the 'SONNBlick\_pleasant\_weather' column are '0', therefore the algorithm could not be properly trained because a lack of variety in the data.

## Exercise 1.5: Supervised Learning Algorithms part 2

### Decision Tree

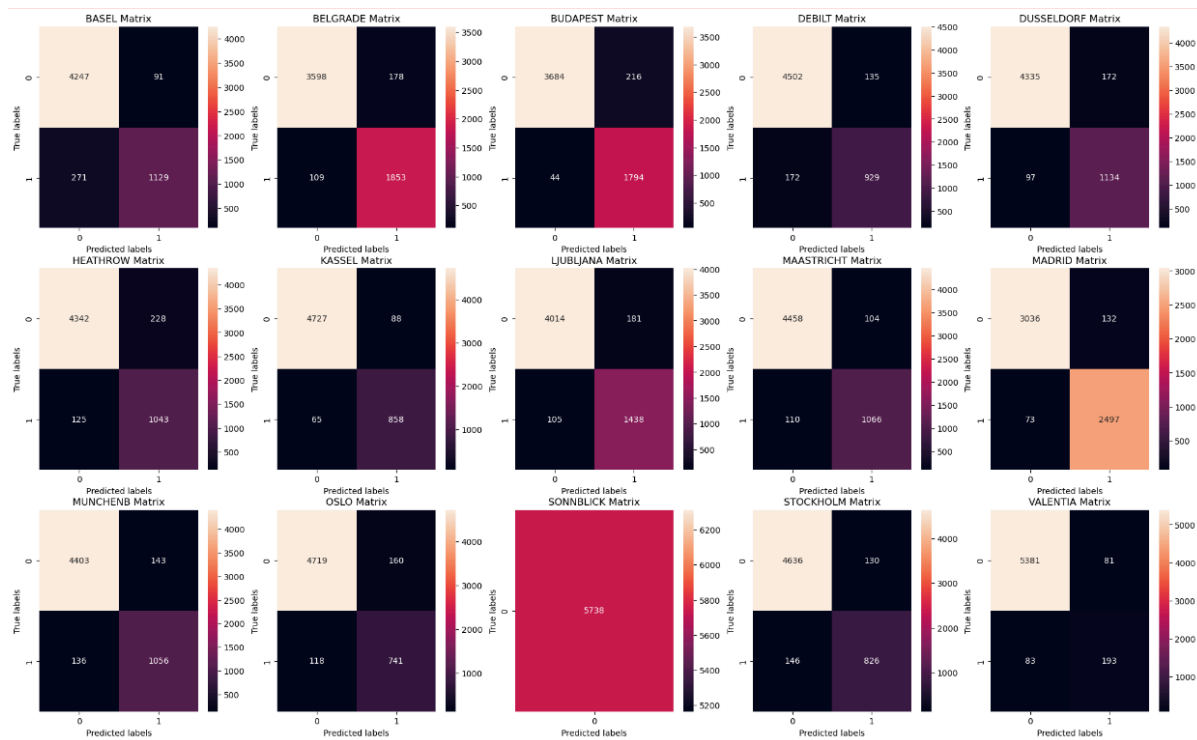
The accuracy of this Decision tree was of 0.602.

I believe pruning this tree would be highly beneficial. A quick look at the plot shows how confusing and inconclusive it is when we use this number of splits.



## Artificial Neural Network

	Hidden Layers	Max itineration	Tolerance	Accuracy
TEST 1	10, 5	500	0.0001	0.5216 0.5203
TEST 2	5, 5	500	0.0001	0.4917 0.4956
TEST 3	30, 50, 60	500	0.0001	0.7214 0.6341
TEST 4	30, 50, 60	500	0.0003	0.6975 0.6094



The algorithm that best represents the current data is the KNN model.

There is one particular station, Sonnblick, that is experiencing overfitting. There isn't any case of full accuracy for the rest of the Stations.

I would recommend ClimateWins to use the ANN model because it works best with complex data such as this, and further adjustments in the parameters would probably lead to a more accurate outcome.