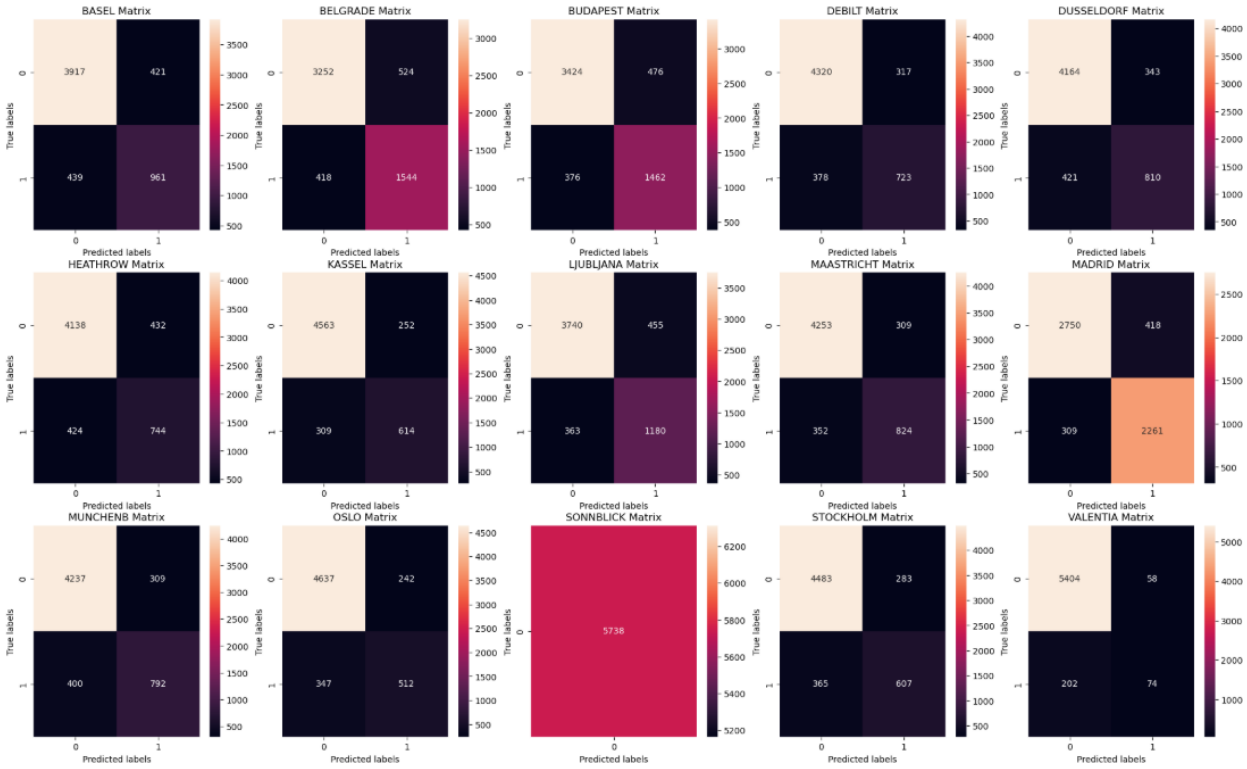
ML Exercise 1.4
Esteban Torres

Supervised Learning Algorithms Pt 1



Weather Station	Positive 0	Positive 1	False 0	False 1	Accuracy Rate
BASEL	3917	961	421	439	85.1%
BELGRADE	3252	1544	524	418	83.7%
BUDAPEST	3424	1462	476	376	85.2%
DEBILT	4320	723	317	378	87.9%
DUSSELDORF	4164	810	343	421	86.7%
HEATHROW	4138	744	432	424	85.1%
KASSEL	4563	614	252	309	90.2%
LJUBLJANA	3740	1180	455	363	85.7%
MAASTRICHT	4253	824	309	352	88.5%
MADRID	2750	2261	418	309	87.3%
MUNCHENB	4237	792	309	400	87.7%
OSLO	4637	512	242	347	89.7%
SONNBLICK	5738	0	0	0	100%
STOCKHOLM	4483	607	283	365	88.7%
VALENTIA	5404	74	58	202	95.5%
				Average	88.5%

- How well does this algorithm predict the current data?

The weather prediction model shows various levels of accuracy, but it seems to work relatively well with an average accuracy prediction rate of 88.5 percent. In this case 5 of the stations had above average rates of accuracy.

- Are any weather stations fully accurate? Is there any overfitting happening?

Yes, there is one station that is fully accurate, Sonnblick. The model in the case of this station may be at risk of overfitting because it has limited exposure to different weather, which may reduce the model's ability to perform.

- Are there certain features of the data set (such as particular weather status) that might contribute to overall accuracy or inaccuracy?

With stable temperatures like Sonnblick the models can match patterns in the training data well, but when there are large variations or outliers the accuracy can reduce. Seasonality can help the KNN distinguish patterns, while missing or inconsistent patterns can increase the misclassification. In short, accuracy is higher when the features cleanly separate classes and the data is consistent. It drops when there's high variability, noise, or rare events.