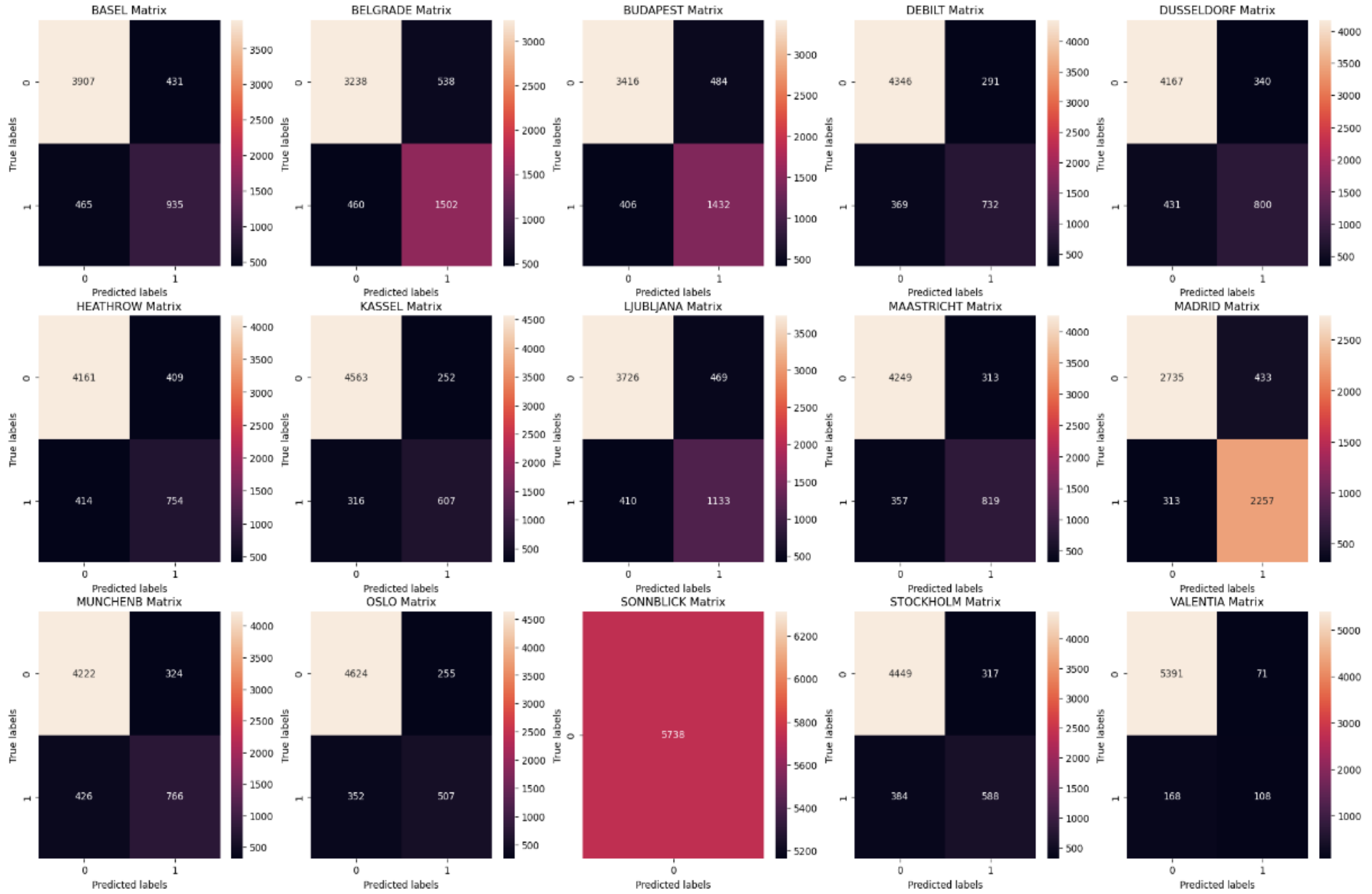
Exercise 1.4


BASEL Matrix
3907
431
465
935
BELGRADE Matrix
3238
538
460
1502
BUDAPEST Matrix
3416
484
406
1432
DEBILT Matrix
4346
291
369
732
DUSSELDORF Matrix
4167
340
431
800
HEATHROW Matrix
4161
409
414
754
KASSEL Matrix
4563
252
316
607
LJUBLJANA Matrix
3726
469
410
1133
MAASTRICHT Matrix
4249
313
357
819
MADRID Matrix
2735
433
313
2257
MUNCHENB Matrix
4222
324
426
766
OSLO Matrix
4624
255
352
507
SONNBLICK Matrix
5738
STOCKHOLM Matrix
4449
317
384
588
VALENTIA Matrix
5391
71
168
108
True labels
Predicted labels

Weather Station	Accurate Positive	Accurate Negative	False Positive	False Negative	Accuracy
Basel	935	3907	431	465	84%
Belgrade	1502	3238	538	460	83%
Budapest	1432	3416	484	406	84%
Debilt	732	4346	291	369	88%
Dusseldorf	800	4167	340	431	87%
Heathrow	754	4161	409	414	86%
Kassel	607	4563	252	316	90%
Ljubjana	1133	3726	469	410	85%
Maastricht	819	4249	313	357	88%
Madrid	2257	2735	433	313	87%
Munchen	766	4222	324	426	87%
Oslo	507	4624	255	352	89%
Sonnblick	X	5738	X	X	100%
Stockholm	588	4449	317	384	88%
Valentia	108	5391	71	168	96%

The model has an overall accuracy of 88%. Sonnblick was 100% accurate, but every point was classified as negative. This could be the model being overfit to the training set. An overfit model could give inaccurate predictions to the real world. Sourcing more information for Sonnblick could be helpful. Or increasing the number of K-neighbors. The latter may cause the model to be underfit, so a balance is necessary.