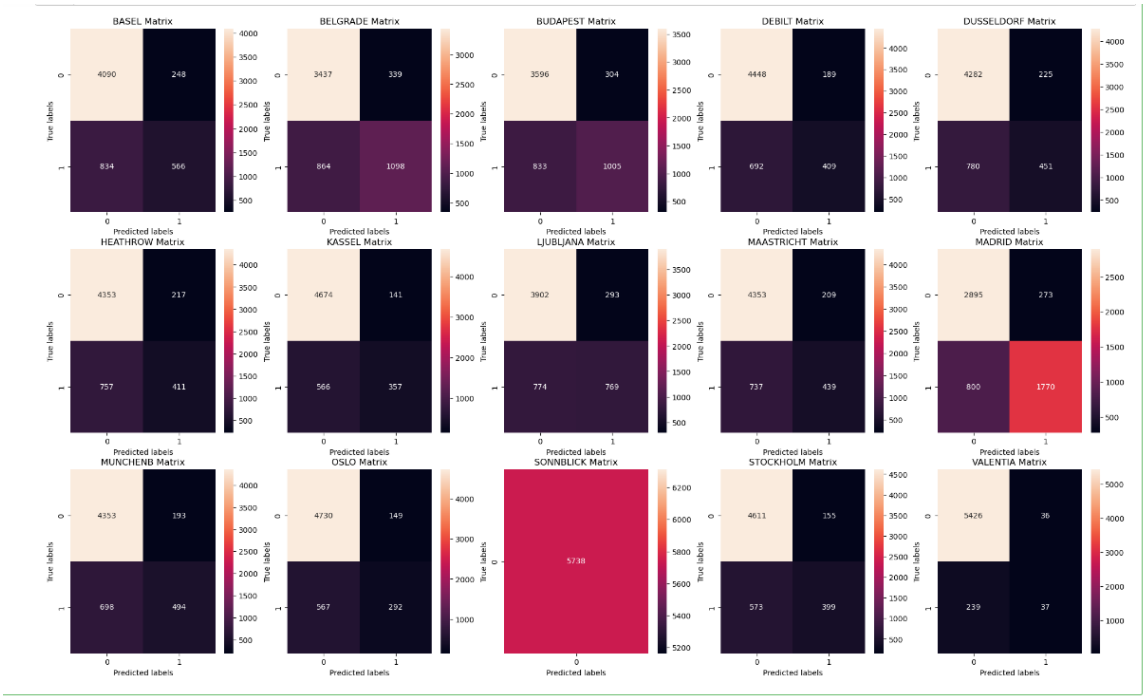
Keanu Gomes
Exercise 1.4 March 12, 2024

Basics of Machine Learning for Analysts
1.4: Supervised Learning Algorithms Part 1

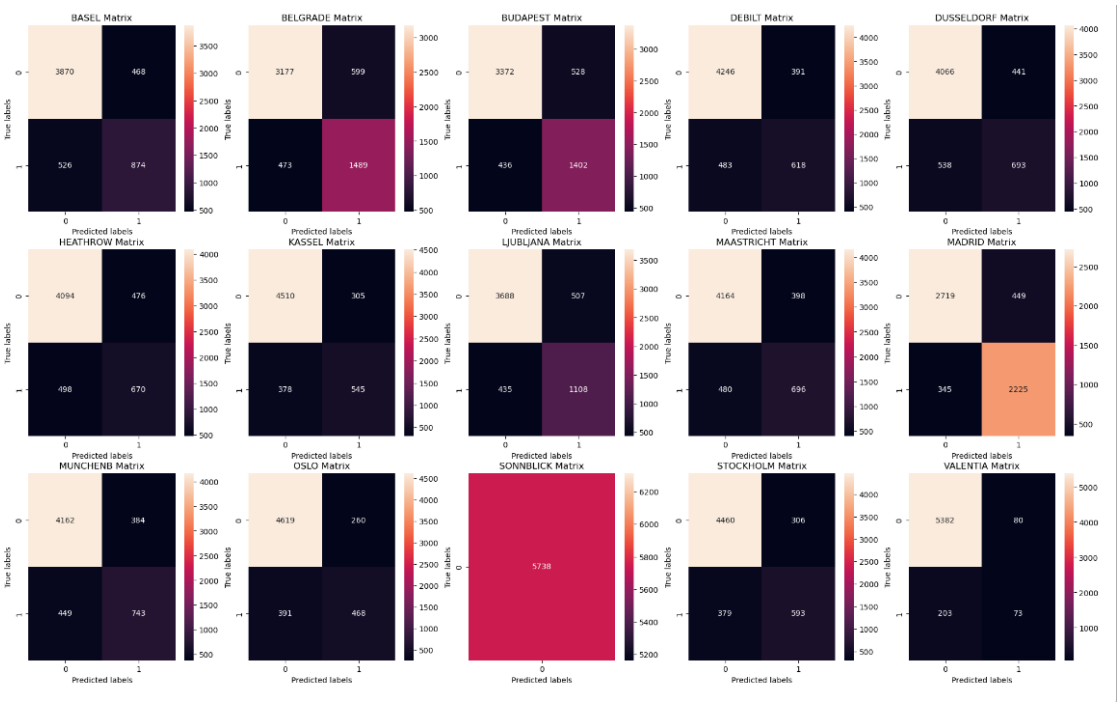
Started parameters here:

- **K_range = (1,3)**

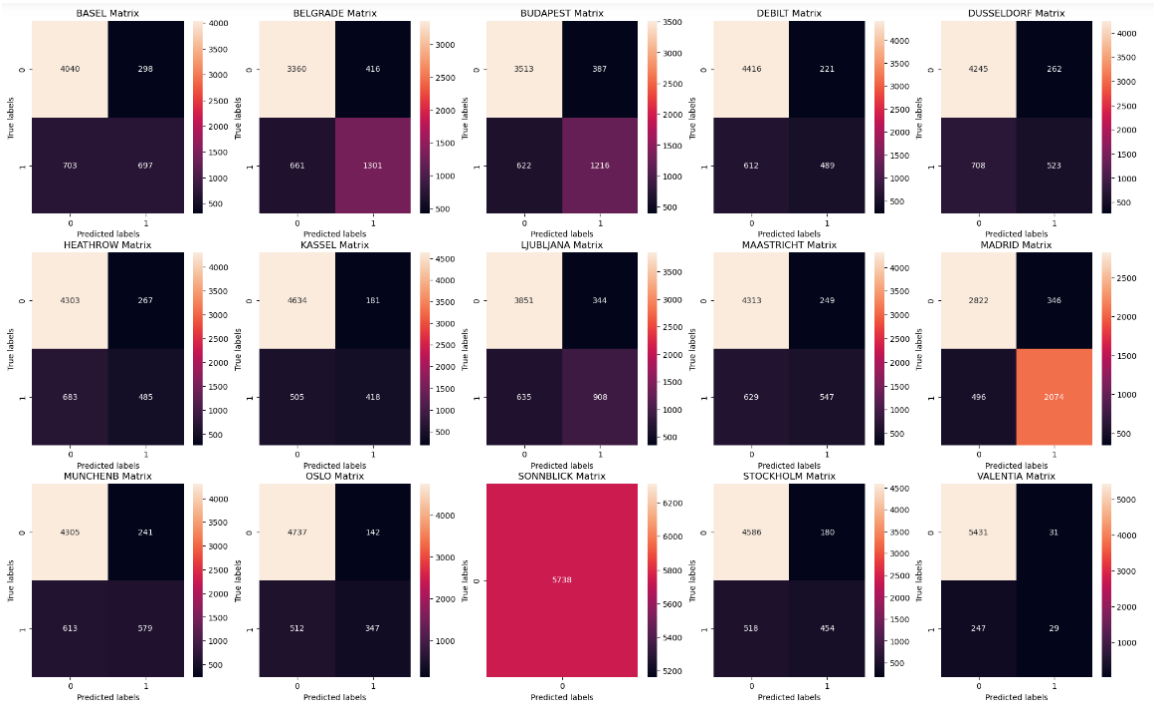


- - **note:** error occurring with SONNBLICK_pleasant_weather_answer labeling due to having no ('pleasant weather attributes' = 1)
- (0 = unpleasant, 1 = pleasant)

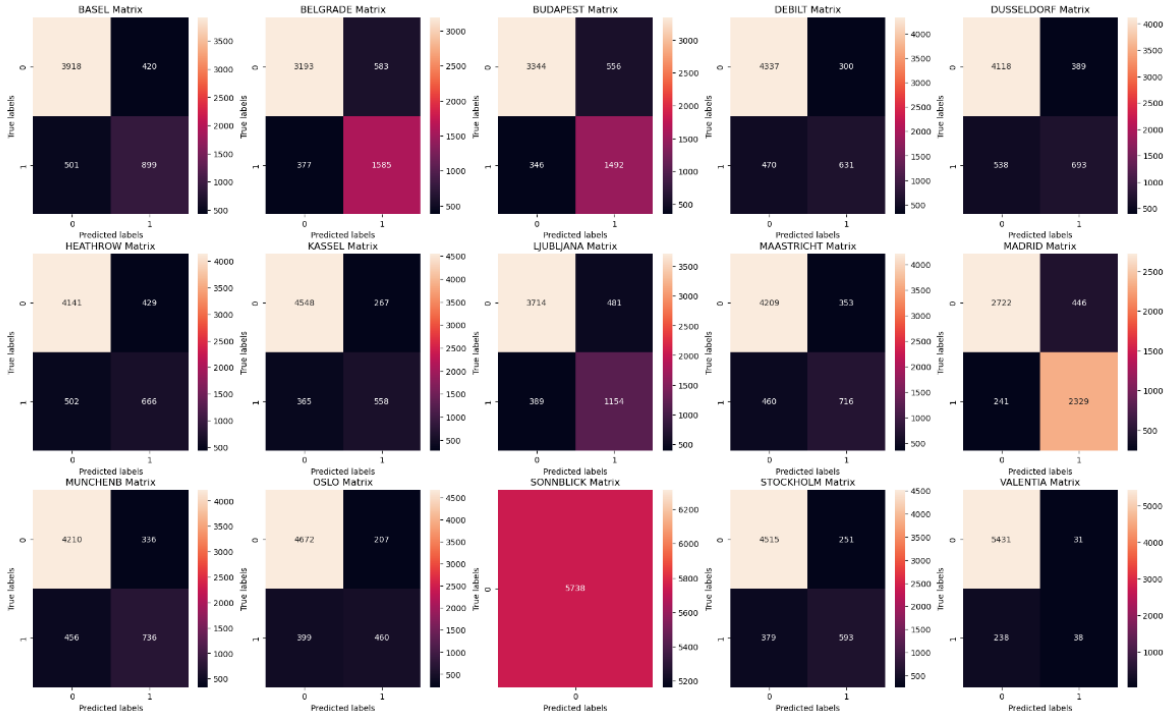
- **K_range = (1,4)**


BASEL Matrix
3870
468
526
874
BELGRADE Matrix
3177
599
473
1489
BUDAPEST Matrix
3372
528
436
1402
DEBILT Matrix
4246
391
483
618
DUSSELDORF Matrix
4066
441
538
693
HEATHROW Matrix
4094
476
498
670
KASSEL Matrix
4510
305
378
545
LJUBLJANA Matrix
3688
507
435
1108
MAASTRICHT Matrix
4164
398
480
696
MADRID Matrix
2719
449
345
2225
MUNCHENB Matrix
4162
384
449
743
OSLO Matrix
4619
260
391
468
SONNBLICK Matrix
5738
STOCKHOLM Matrix
4460
306
379
593
VALENTIA Matrix
5382
80
203
73
True labels
Predicted labels

- **K_range = (1,5)**

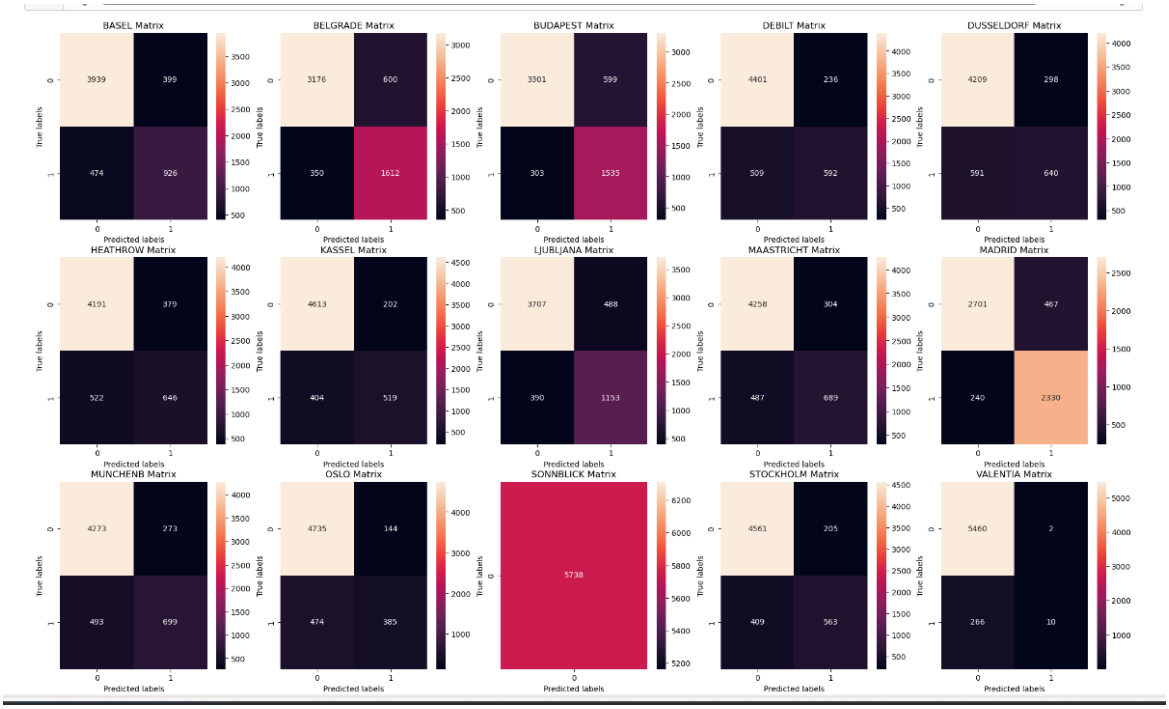

BASEL Matrix
4040
298
703
697
BELGRADE Matrix
3360
416
661
1301
BUDAPEST Matrix
3513
387
622
1216
DEBILT Matrix
4416
221
612
489
DUSSELDORF Matrix
4245
262
708
523
HEATHROW Matrix
4303
267
683
485
KASSEL Matrix
4634
181
505
418
LJUBLJANA Matrix
3851
344
635
908
MAASTRICHT Matrix
4313
249
629
547
MADRID Matrix
2822
346
496
2074
MUNCHENB Matrix
4305
241
613
579
OSLO Matrix
4737
142
512
347
SONNBLICK Matrix
5738
STOCKHOLM Matrix
4586
180
518
454
VALENTIA Matrix
5431
31
247
29
True labels
Predicted labels

- **K_range = (1,10)**


BASEL Matrix
3918
420
501
899
BELGRADE Matrix
3193
583
377
1585
BUDAPEST Matrix
3344
556
346
1492
DEBILT Matrix
4337
300
470
631
DUSSELDORF Matrix
4118
389
538
693
HEATHROW Matrix
4141
429
502
666
KASSEL Matrix
4548
267
365
558
LJUBLJANA Matrix
3714
401
389
1154
MAASTRICHT Matrix
4209
353
460
716
MADRID Matrix
2722
446
241
2329
MUNCHENB Matrix
4210
336
456
736
OSLO Matrix
4672
207
399
460
SONNBLICK Matrix
5738
STOCKHOLM Matrix
4515
251
379
593
VALENTIA Matrix
5431
31
238
38
True labels
Predicted labels

Parameters ended here:

- **K_range = (1,60)**



- **VALENTIA** seems to have the least false positives and negatives, and the highest number of true positives out of every k_range instance, this indicates that it may be the most accurate.

1. **True Positives (TP):** The top left cell (5460) represents the number of instances where the actual class was "Unpleasant", and the model correctly predicted it as "Unpleasant."

2. **False Positives (FP):** The top right cell (2) represents the number of instances where the actual class was "Unpleasant", but the model incorrectly predicted it as "Pleasant."

3. **False Negatives (FN):** The bottom left cell (266) represents the number of instances where the actual class was "Pleasant", but the model incorrectly predicted it as "Unpleasant."

4. **True Negatives (TN):** The bottom right cell (10) represents the number of instances where the actual class was "Pleasant", and the model correctly predicted it as "Pleasant."

VALENTIA PREDICTION METRICS for 60 neighbors

1. **Accuracy:** 95.34%

2. **Precision:** 99.96%

3. **Recall (Sensitivity):** 95.37%

4. **F1 Score:** 97.61%