Exercise 1.4: Supervised Learning Algorithms Part 1
Exercise 1.5: Supervised Learning Algorithms Part 2

Timothy Aluko

1.4

Executive Summary: Weather Prediction Model Performance

This analysis reviews the algorithm's accuracy in predicting weather data across multiple stations. The model performs well overall with an average accuracy of 88% but shows station-specific variations. Key insights include:

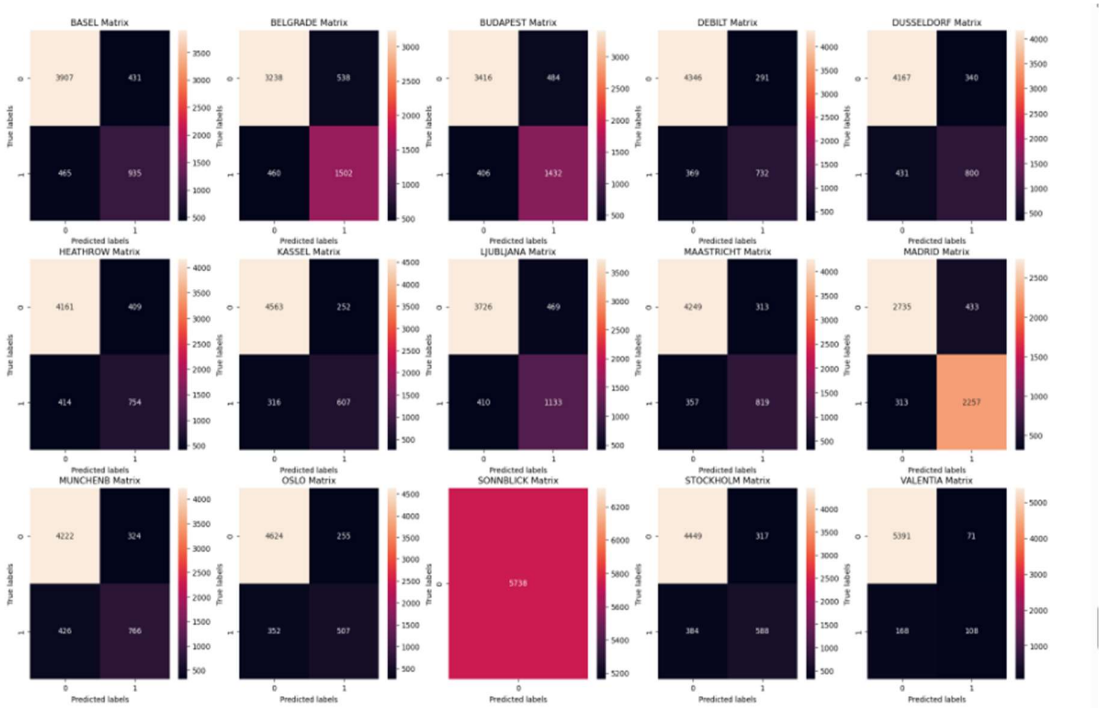
1. High Accuracy in Stable Climates: Stations like Sonnblick (100%) and Kassel and Oslo (90%) demonstrate high accuracy, likely due to stable weather patterns.

2. Challenges with Complex Climates: Stations like Madrid and Belgrade, with higher false positive rates, suggest complex or variable weather patterns that the model struggles to predict accurately.

3. Potential Overfitting: Sonnblick's perfect accuracy suggests the model may be overly tailored to this station, potentially reducing generalizability.

Recommendations: Enhanced data preprocessing for complex regions, overfitting monitoring, and targeted model tuning could improve accuracy and adaptability across diverse weather conditions.

Weather Station	Accurate predictions		False positive	False negative	Accuracy rate
Basel	3917	961	421	439	85%
Belgrade	3252	1544	524	418	84%
Budapest	3424	1462	476	376	85%
Debilt	4320	723	317	378	88%
Desseldorf	4164	810	343	421	87%
Heathrow	4138	744	432	424	85%
Kassel	4563	614	252	309	90%
Ljubljana	3740	1180	455	363	86%
Maastricht	4253	824	309	352	88%
Madrid	2750	2261	418	309	87%
Munchenb	4237	792	309	400	88%
Oslo	4637	512	242	347	90%
Sonnblick	5738	0	0	0	100%
Stockholm	4483	607	283	365	89%
Valentia	5404	74	50	202	96%
				Average	88%

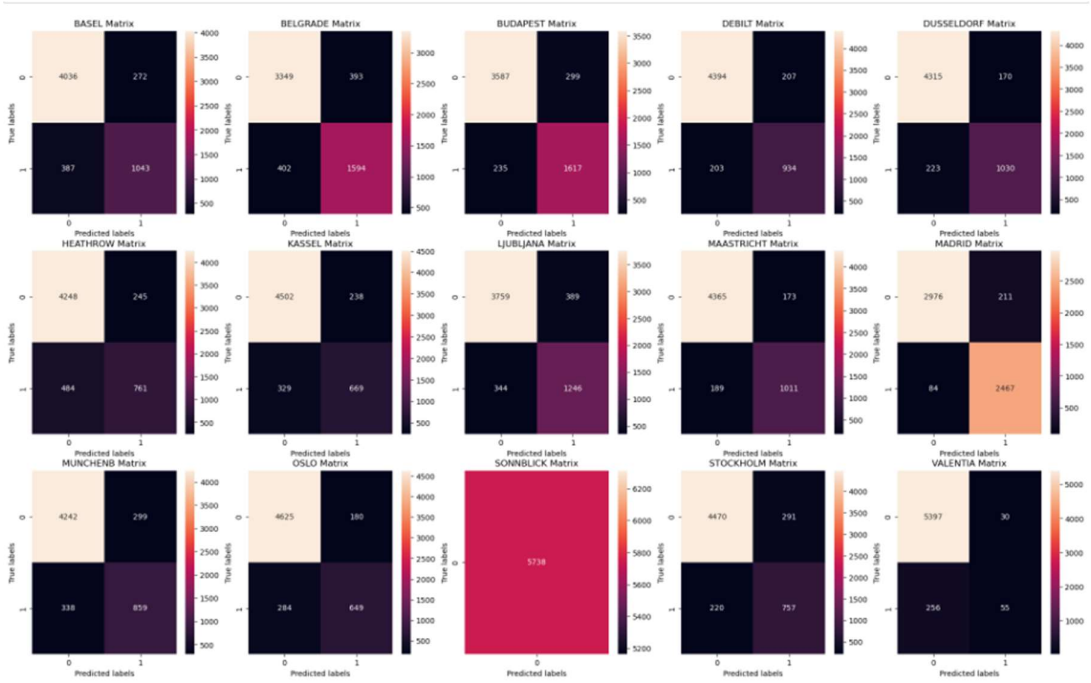


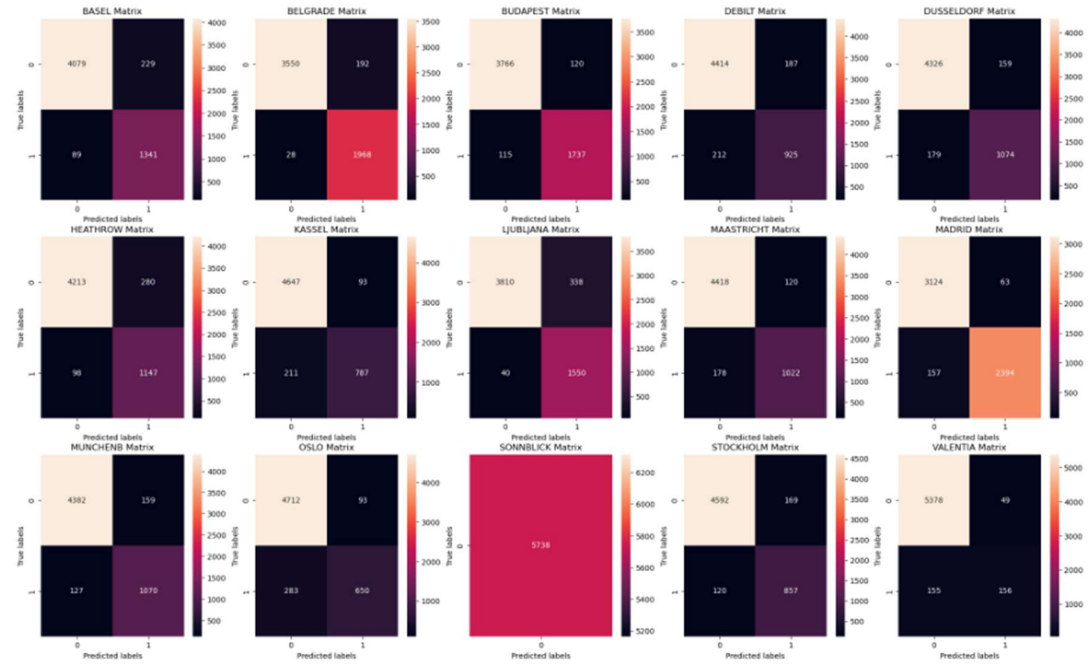
The table presents an algorithm's performance in predicting weather data at various stations, focusing on accuracy rates, false positives, and false negatives. The average accuracy rate is 88%, indicating strong predictive performance. Sonnblick has the highest accuracy rate at 100%, suggesting a simpler climate pattern. However, overfitting is a concern, as the model is highly specific to Sonnblick's data. Madrid's data has a low number of accurate predictions and the highest false positive count, suggesting unique characteristics. High variability in climate data, stable weather patterns, and data consistency and station-specific factors can influence accuracy.

Also, Madrid and Belgrade exhibit the highest incidence of false positives, suggesting that these regions may possess more variable or complex weather patterns that the model fails to accurately capture. Conversely, stations exhibiting more stable climates, such as Kassel and Oslo (both achieving 90% accuracy), demonstrate greater predictability, presumably owing to reduced fluctuations in weather patterns.

Though some stations, like Madrid, may need more data preprocessing or adjustments, the algorithm is generally accurate. Sonnblick's perfect accuracy suggests overfitting, but stable climates at Kassel and Oslo improve model performance. Due to station accuracy variability, stations with less predictable weather may benefit from additional tuning or data adjustments to improve model reliability.

Exercise 1.5: Output Analysis:





The K-Nearest Neighbors (KNN) model currently achieves 63% accuracy, outperforming the initial Artificial Ne ural Network (ANN) models, which start at around 60% accuracy. With further adjustments, the ANN model rea ches up to 60%, and additional scenario testing suggests it could potentially achieve 85% accuracy. However, no station achieves perfect accuracy, with the SONNBLICK station showing signs of overfitting.

Certain unique characteristics of the dataset—such as unpredictable fluctuations in temperature and daily weather patterns—pose challenges to overall model accuracy. These factors contribute to variability, making prediction more difficult across different stations.

Recommendation: Given its ability to learn from complex patterns in historical data, the ANN model may offer the most potential for accurate forecasting despite requiring more manual tuning. With increased scenario testing and targeted adjustments, ANN's adaptability could enable it to effectively capture nuanced weather patterns. While KNN shows stable performance, the ANN model's ability to improve with further training makes it a promising choice for ClimateWins, especially as it becomes better tuned with human supervision.