CLIMATEWINS:

CLIMATE CHANGES PREDICTION

By Rose Alappat Joy

Objective

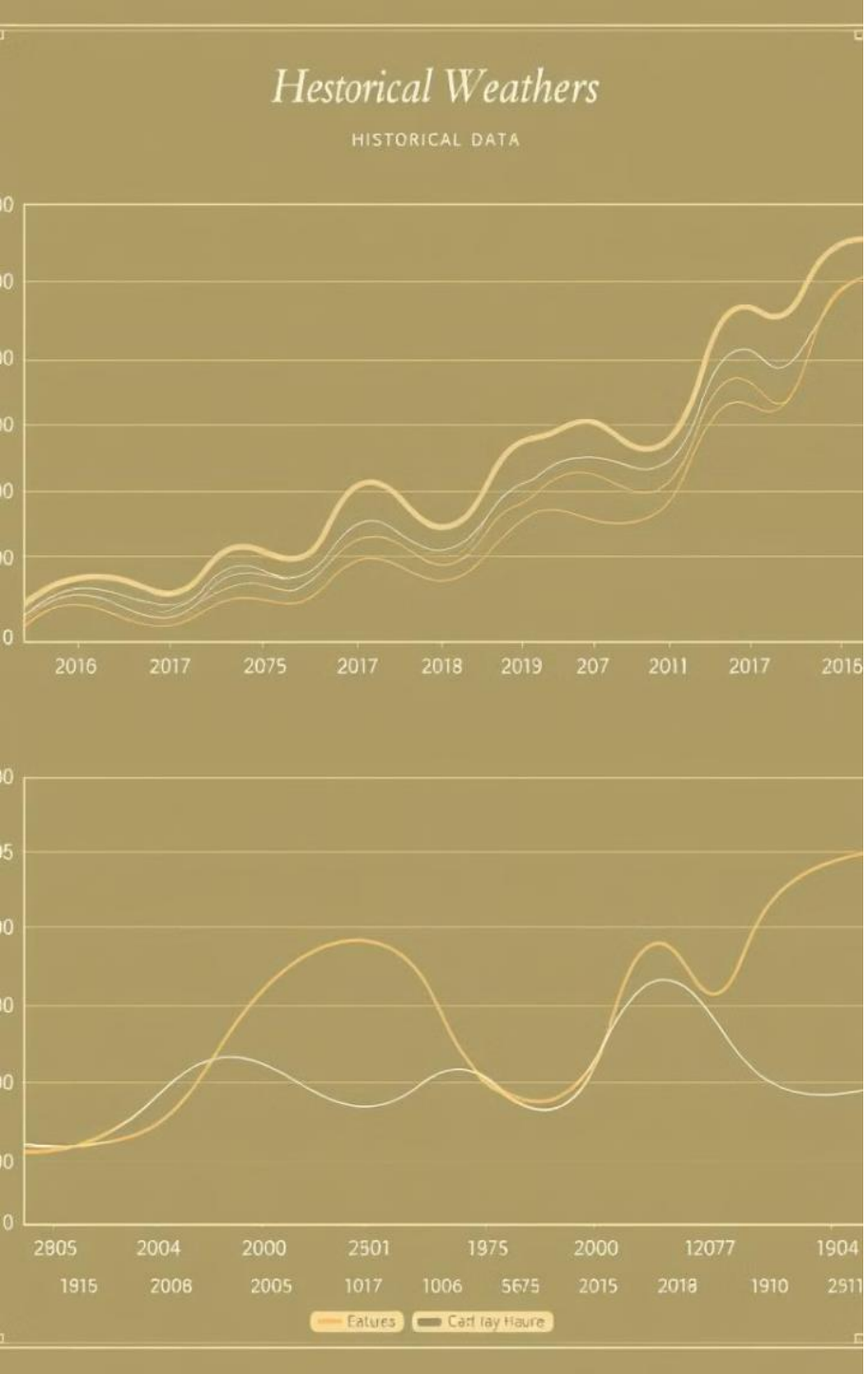
ClimateWins aims to use machine learning to predict the impacts of climate change across Europe and potentially worldwide

Hypothesis

Machine learning can predict the effects of climate change using weather data.

Machine learning models could have ethical issues, like bias or lack of transparency

Machine learning can predict whether the weather will be good or bad on a specific day.



Data Source & Bias

Data Origin

The European Climate Assessment & Data Set project provided the primary dataset, which includes comprehensive weather observations from 18 weather stations across Europe, meticulously recorded from 1960 to 2022..

Data Variety

Includes temperature, wind speed, snow, and global radiation

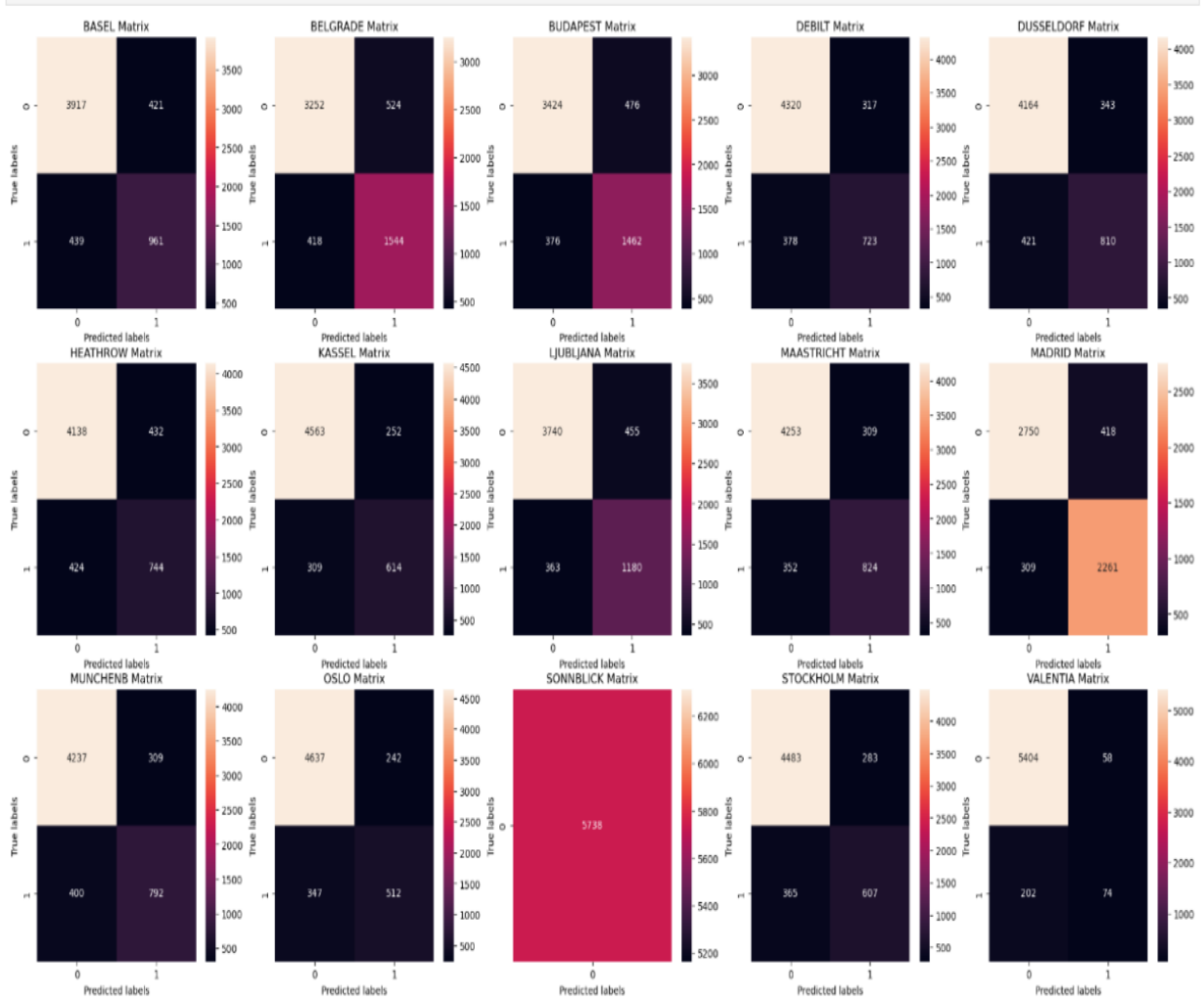
Data Bias

Collection Bias: Variations in instrumentation, measurement techniques, or the relocation of stations over time can lead to biased data.

Sampling Bias: Gathering data from only 18 specific stations out of 26,321 across Europe may result in an unbalanced representation of climate patterns.

Supervised Learning: K-Nearest Neighbor

The K-Nearest Neighbours (KNN) algorithm, a supervised machine learning method is used here, which classifies data points based on their distance to nearby points in other categories.

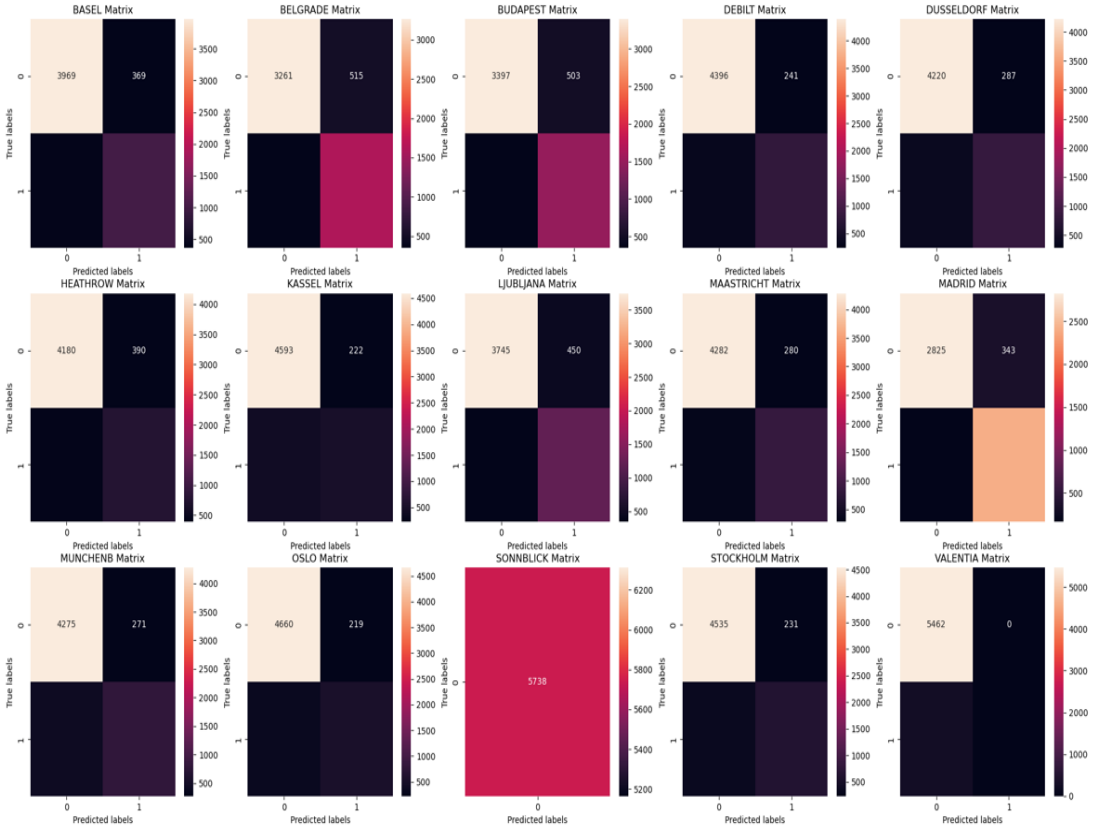


Weather Station	Accuracy Rate
Sonnblick	100%
Oslo	90%
Basel	85%

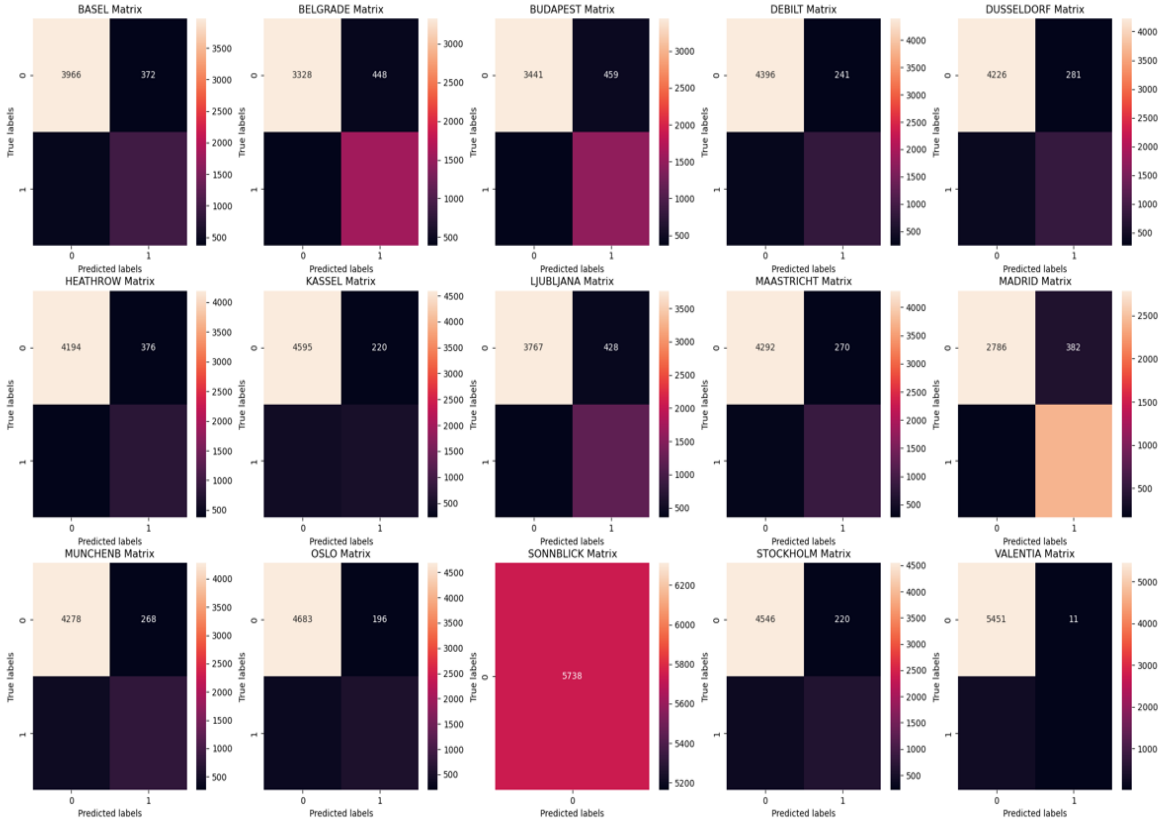
Artificial Neural Network (Confusion Matrix)

We used an Artificial Neural Network (ANN) to identify complex patterns in large datasets, achieving 50% accuracy, which reflects the challenges of predicting the unpredictable nature of weather

Test 1 (45.5% Accuracy)

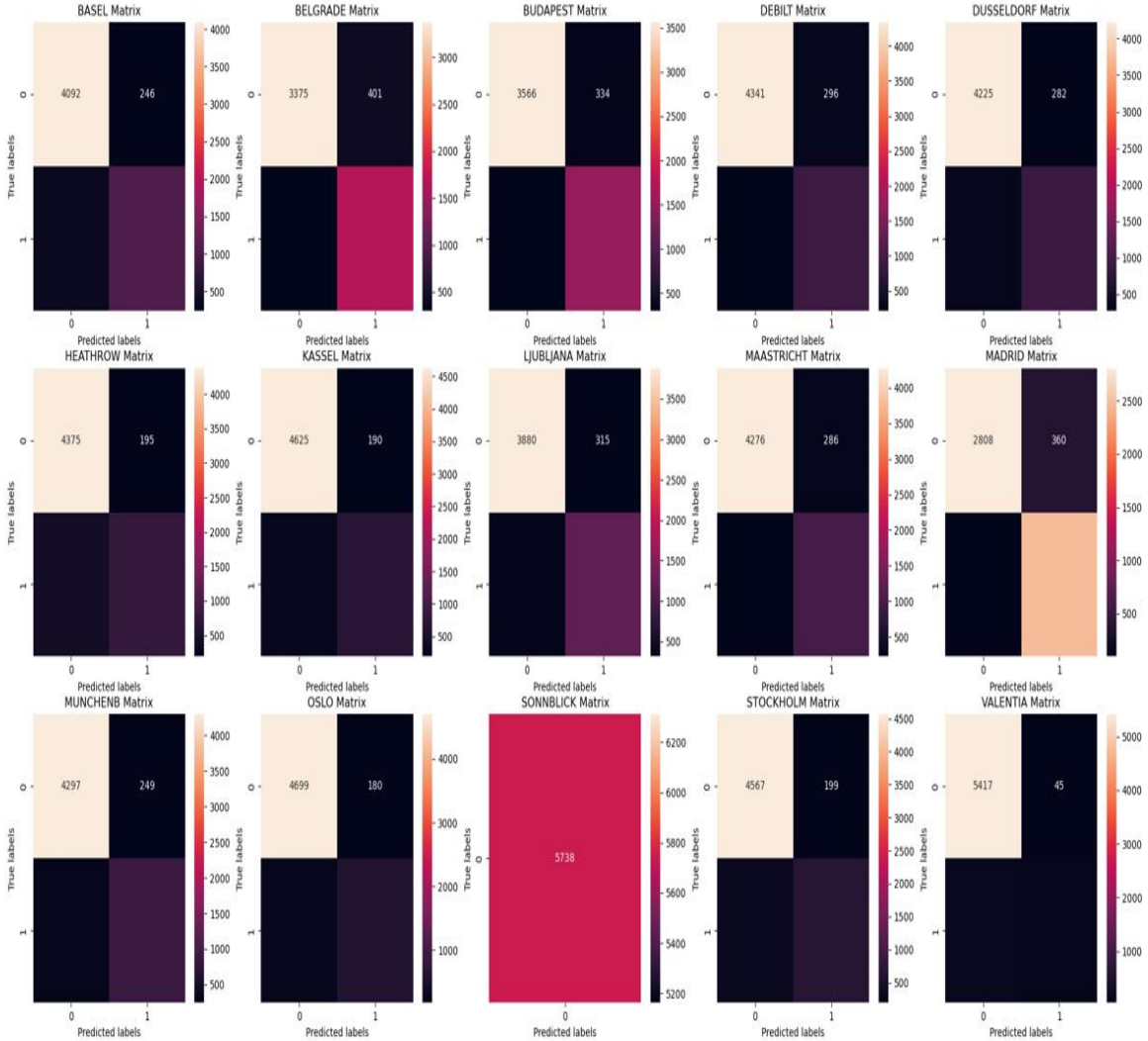


Test 2 (47% Accuracy)



Artificial Neural Network (Confusion Matrix)

Test 3 (50% Accuracy)



Optimal architecture and training

Minimize FP,FN

Maximizing TP,TN

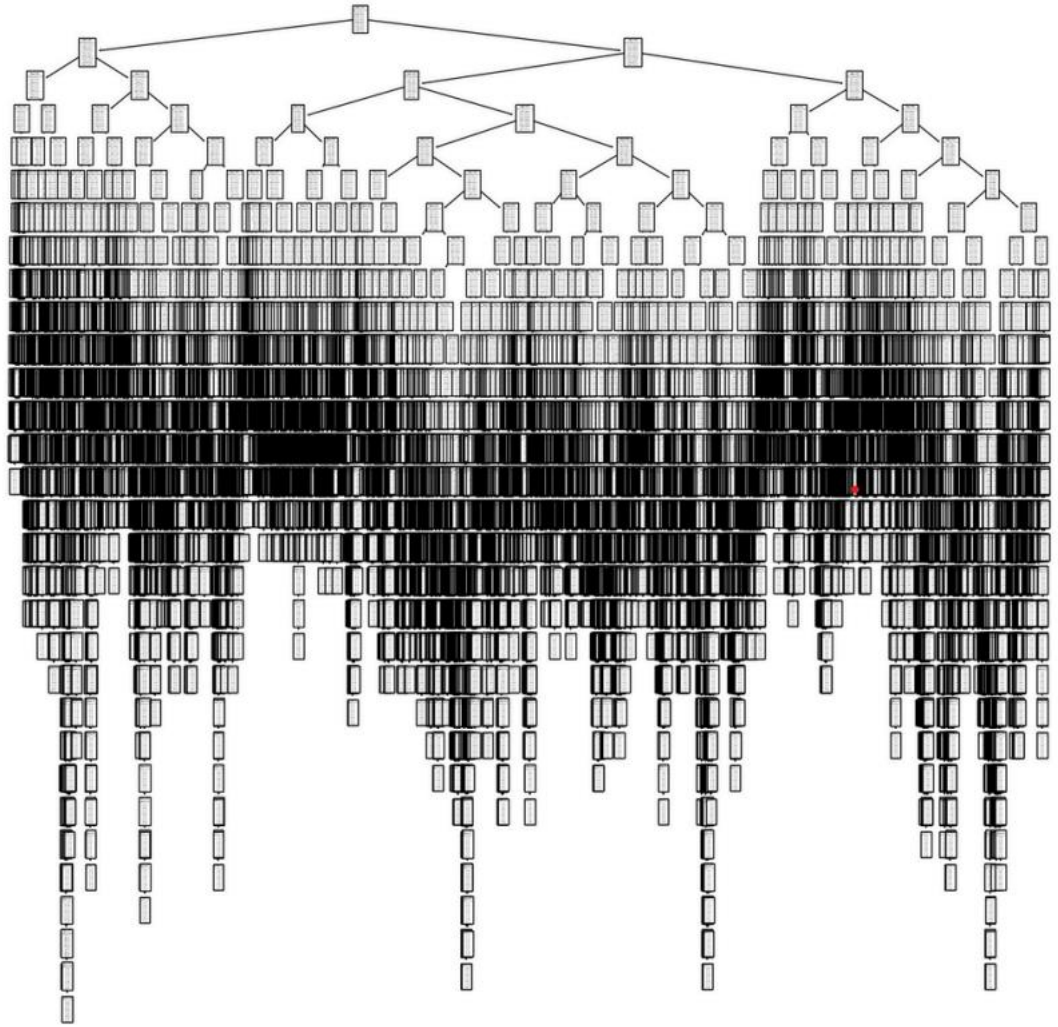
FP- False Positive

FN-False Negative

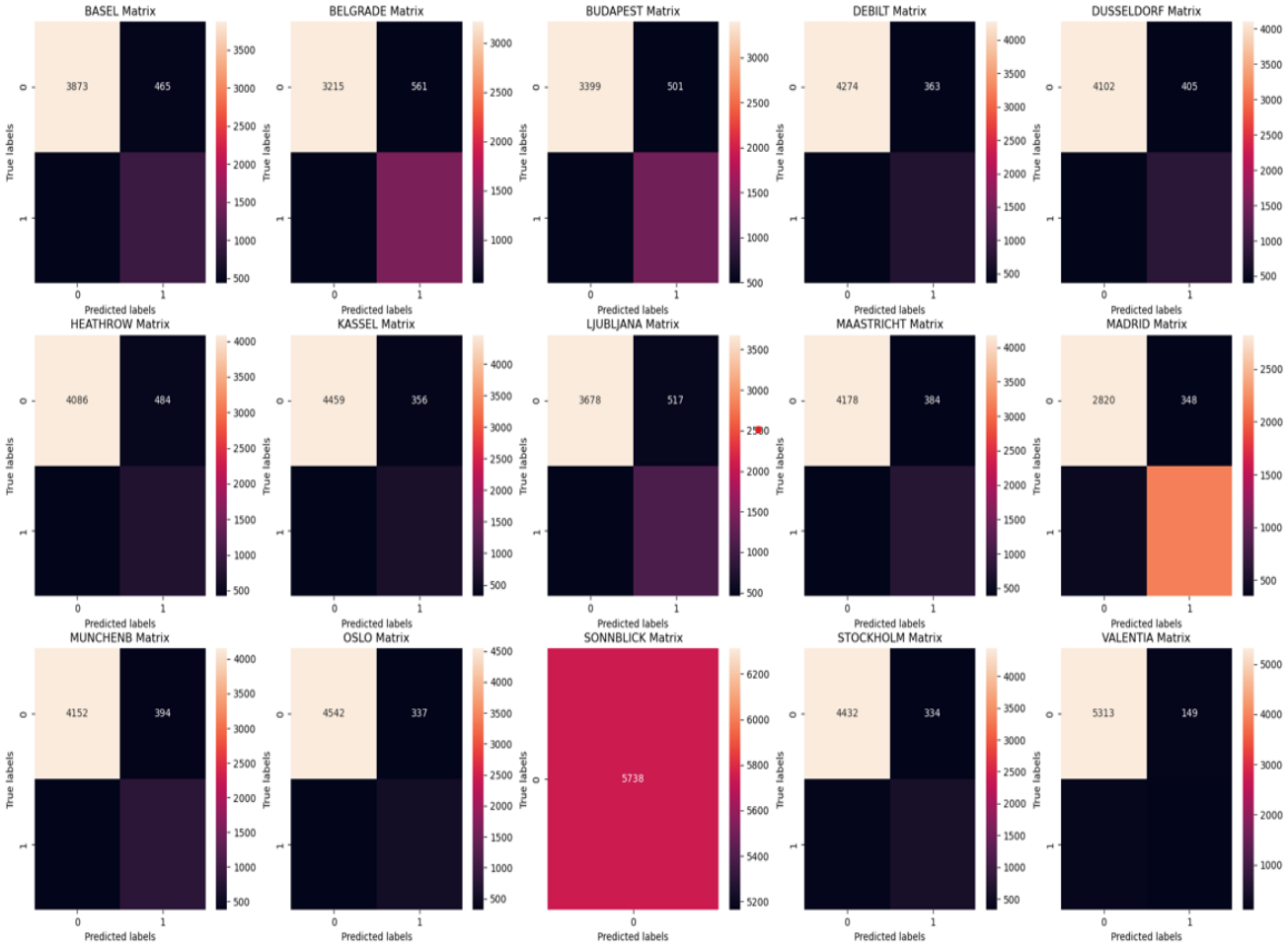
TP- True positive

TN-True Negative

Decision Tree



Test set- 46% Accuracy



The decision tree needs to be pruned as it is too deep and complex

Conclusion

- The KNN algorithm achieved the highest performance with an average accuracy of 88% on the test set.
- The decision tree requires pruning to improve its accuracy.
- The ANN demonstrated a maximum accuracy of only 50%
- Therefore, KNN is the best model for weather prediction in this dataset

Future Directions

1. Model Integration

Future research will focus on combining models like KNN and ANN to create a more robust and accurate climate prediction system

2. Decision Tree Pruning

We will apply pruning techniques to optimize the decision tree, simplifying its structure to prevent overfitting and improve accuracy on new data.

.

3. Future Exploration

We will use uncategorized data and unsupervised learning to find patterns and improve weather predictions.

Thank you