

Europe Climate Changes and Prediction Model Report

Jinwoo Chung (Data Analyst)
ClimateWins
2024. 3. 18

Table of Contents

1. Background
2. Objective & Hypothesis
3. Data & Biases
4. Global Warming
5. Correlations
6. Machine Learning Algorithms
7. ANNs Model Confusion Matrix
8. Summary

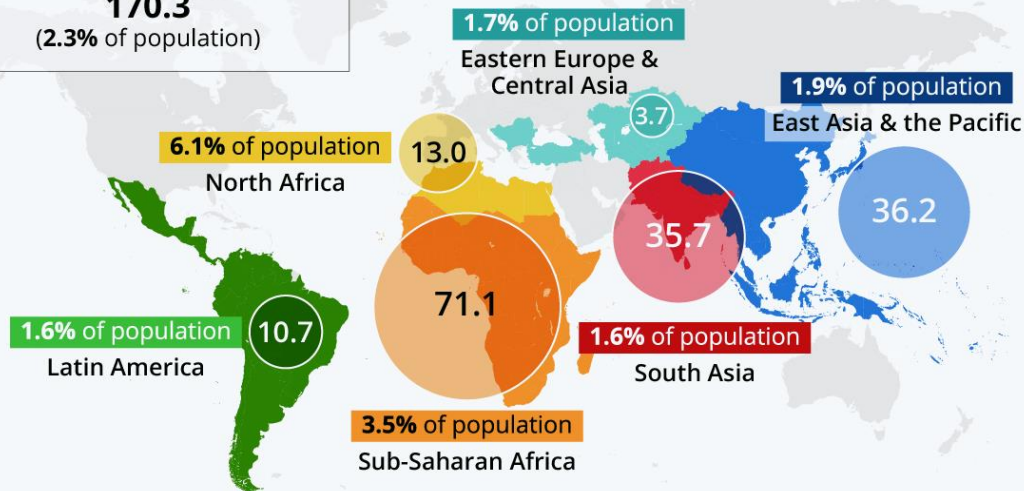
1. Background - Climate Crisis

Climate Change, the Great Displacer

Average number of internal climate migrants by 2050 per region (in millions)*



Total in surveyed regions
170.3
(2.3% of population)

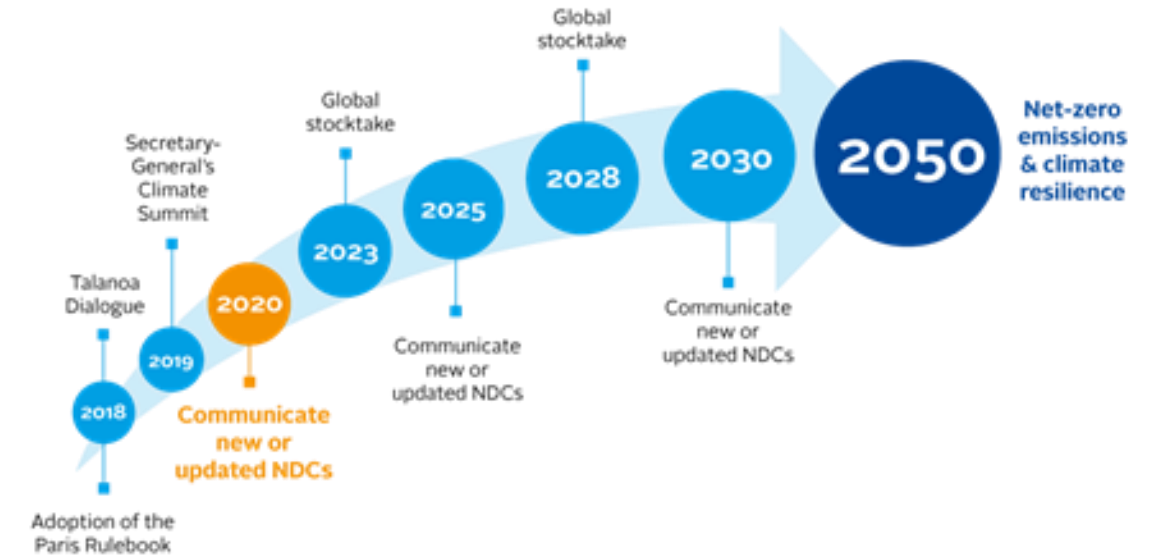


* Modeled on pessimistic reference = High emission & unequal development scenarios concerning water availability, crop productivity and sea-level rise
Source: World Bank



statista

AMBITION MECHANISM IN THE PARIS AGREEMENT



Source: [un.org/publications/ndc-enhancement-by-2020](https://www.un.org/publications/ndc-enhancement-by-2020)

Source: Statista

<https://www.statista.com/chart/26117/average-number-of-internal-climate-migrants-by-2050-per-region/>

2. Objective & Hypothesis

Climate Wins, a European nonprofit organization, aims to help predict the climate change by using machine learning.

Hypothesis

1. The warming of the last 20 years is worse than before (1980-2000 vs 2000-2022)
2. Climate elements such as temperature, humidity, and wind speed are strongly correlated.
3. Based on historical climate data, machine learning can predict future climate change.

3. Data & Biases

The data includes daily temperature, humidity, wind speed, and other climate factors from 1960 to 2022 at stations in 18 European locations.

Biases

- Spatial Bias: The distribution of weather stations may not be uniform across Europe, leading to gaps in coverage and potential underrepresentation of certain regions.
- Temporal Bias: Changes in instrumentation, observation methods, or station locations over time could introduce biases in the historical data.
- Selection Bias: The dataset may primarily include data from urban or populated areas, potentially skewing the representation of rural or remote regions.

How to Optimize the data?

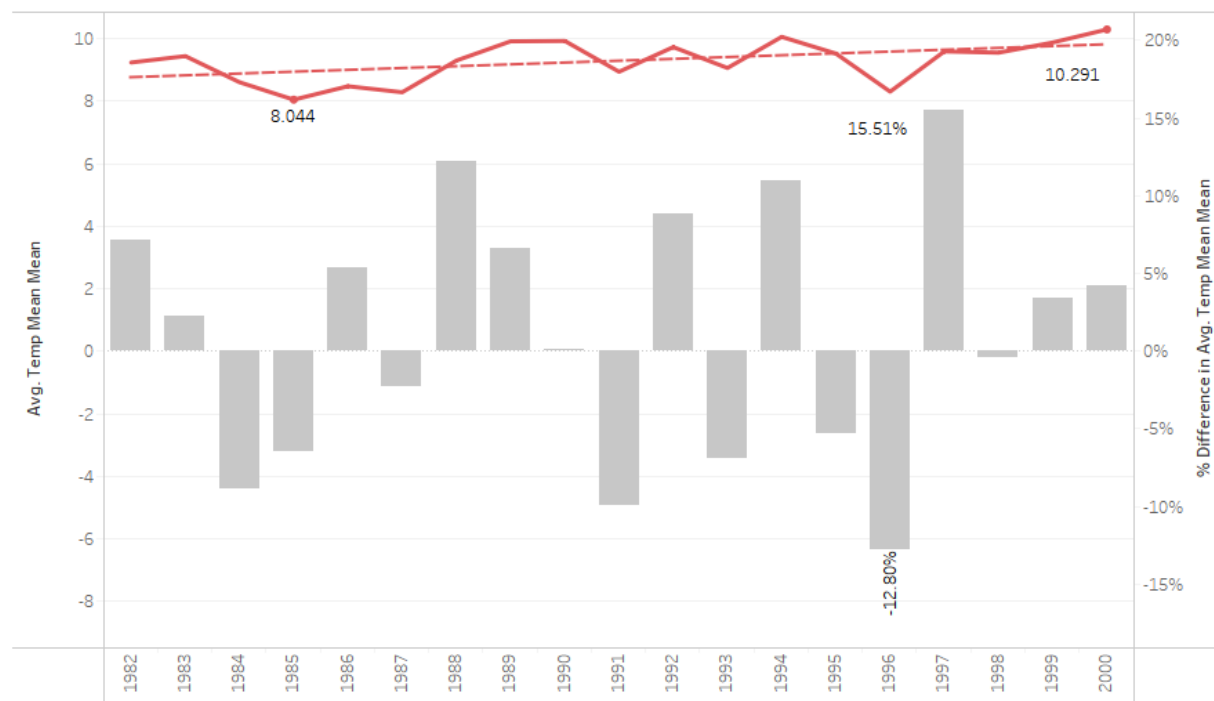
Standardization: transforms data to have a mean of 0 and a standard deviation of 1. It makes all features consistent in scale by subtracting the mean and dividing by the standard deviation. This stabilizes model training and improves performance.

Advantages: Ensures consistent feature scales | Mitigates outlier impact | Improves model generalization | Enhances algorithm performance

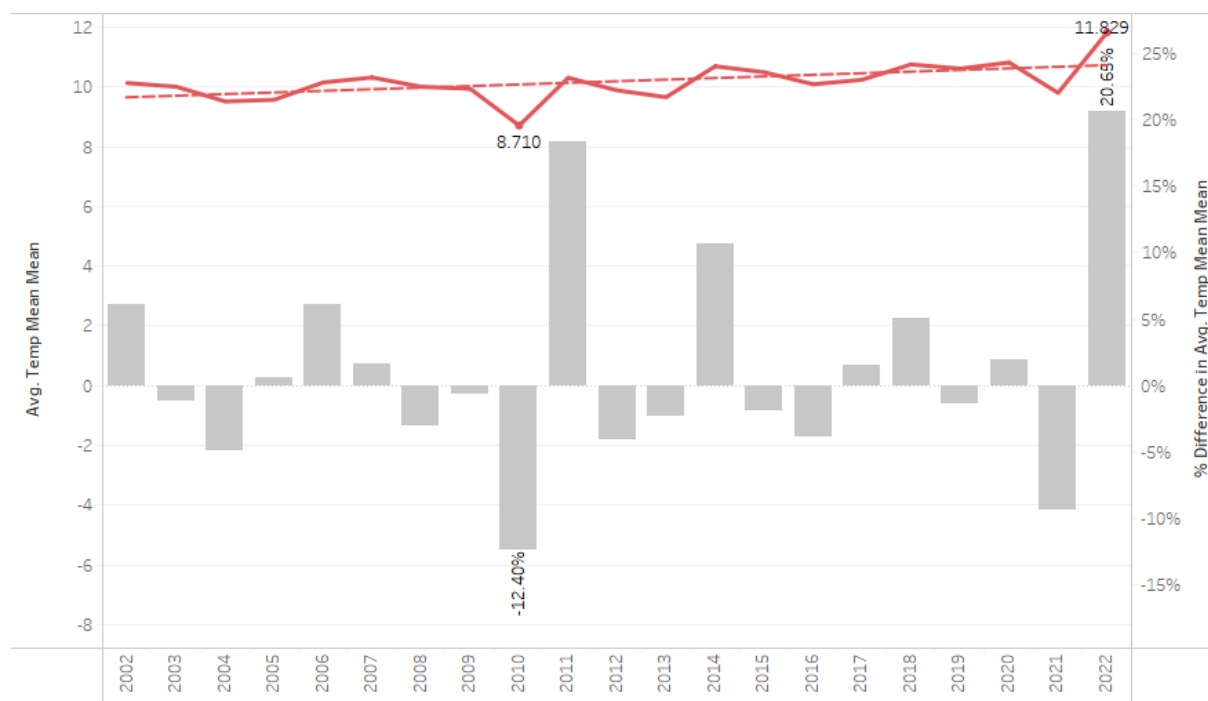
4. Global Warming

Temp-Avg	Min	Max	Slope
2000 – 2022	8.7	11.8	0.053
1980 – 2000	8.0	10.2	0.058

1980 - 2000



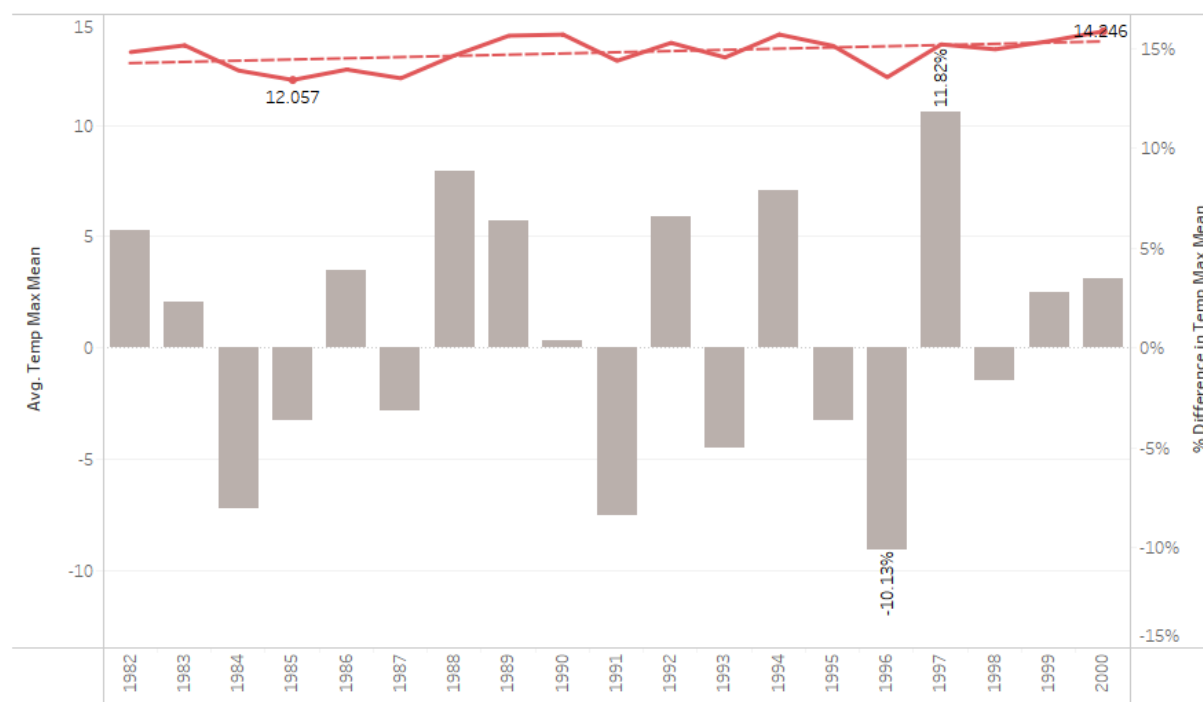
2001 - 2020



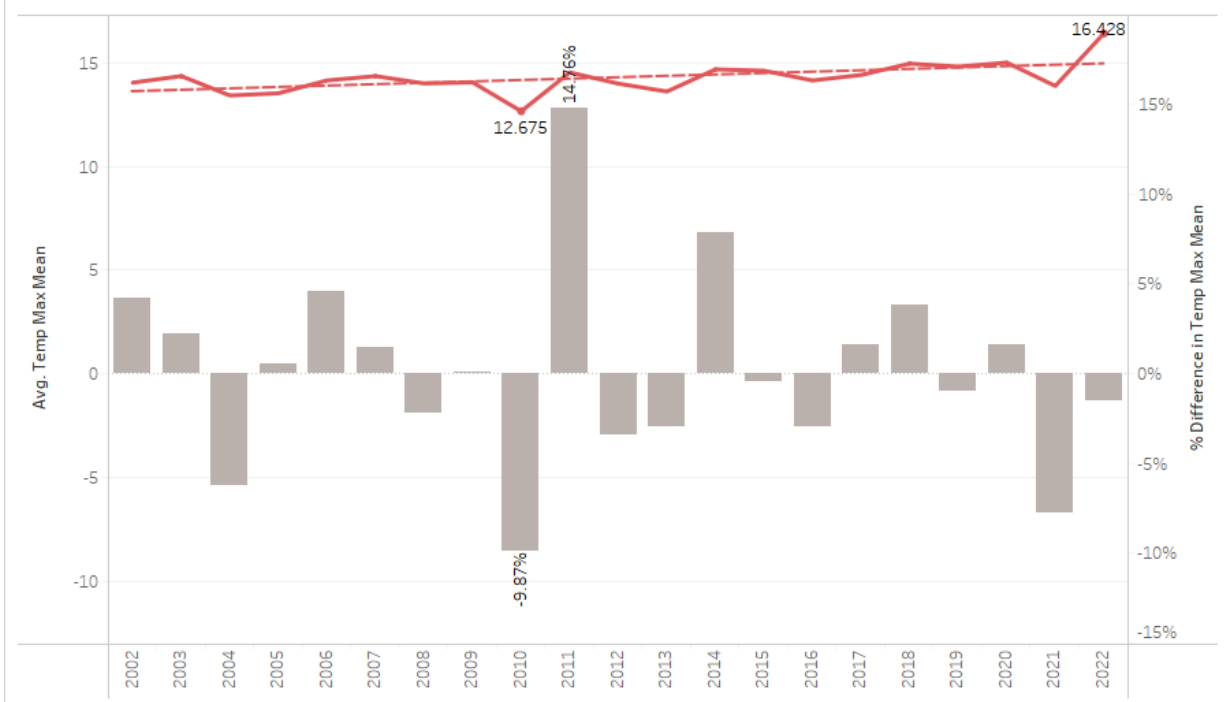
4. Global Warming

Temp-Max	Min	Max	Slope
2000 – 2022	12.6	16.4	0.067
1980 – 2000	12.1	14.2	0.054

1980 - 2000

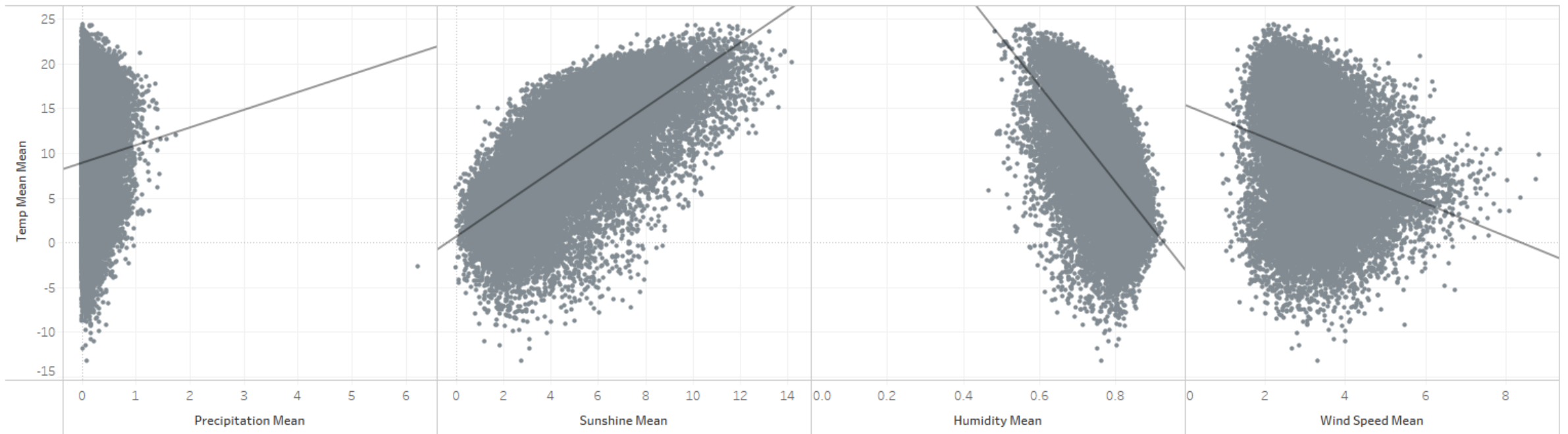


2000 - 2022



5. Correlations

The correlation between average temperature and precipitation, sunshine, humidity, and wind.



6. Machine Learning Algorithms

Gradient Descent

- a way to adjust a model's parameters to minimize errors gradually.
- by repeatedly tweaking parameters in the direction that reduces errors the most.

K-Nearest Neighbors

- predictions based on the majority vote of its nearest neighbors.
- find the K closest data points and predict based on their labels.

Decision Trees

- predictions by following a series of yes/no questions based on features.
- Each question splits the data until a prediction is made.

Artificial Neural Networks models → Highest Accuracy

- ANNs are computer systems inspired by the human brain's structure.
- They consist of layers of interconnected nodes that process data and make predictions.

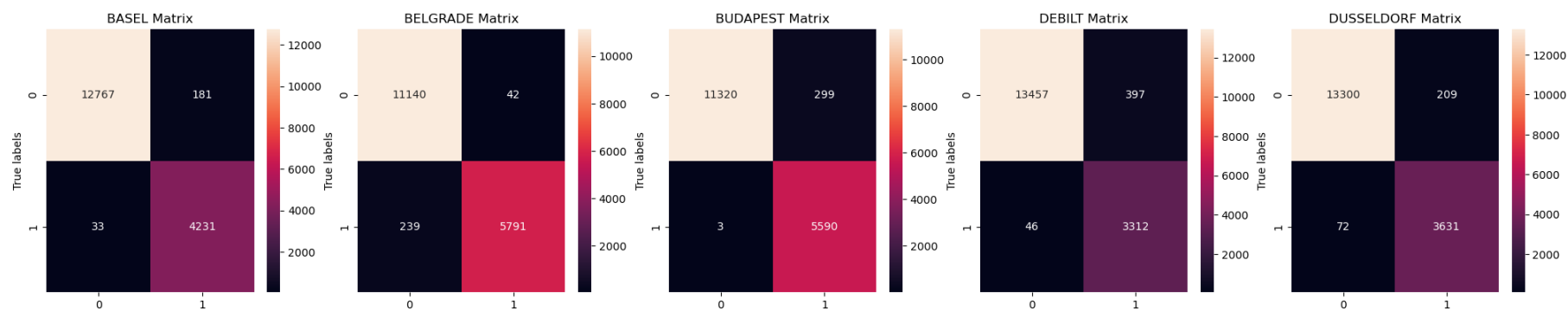
7. ANNs Model Confusion Matrix

Accuracy

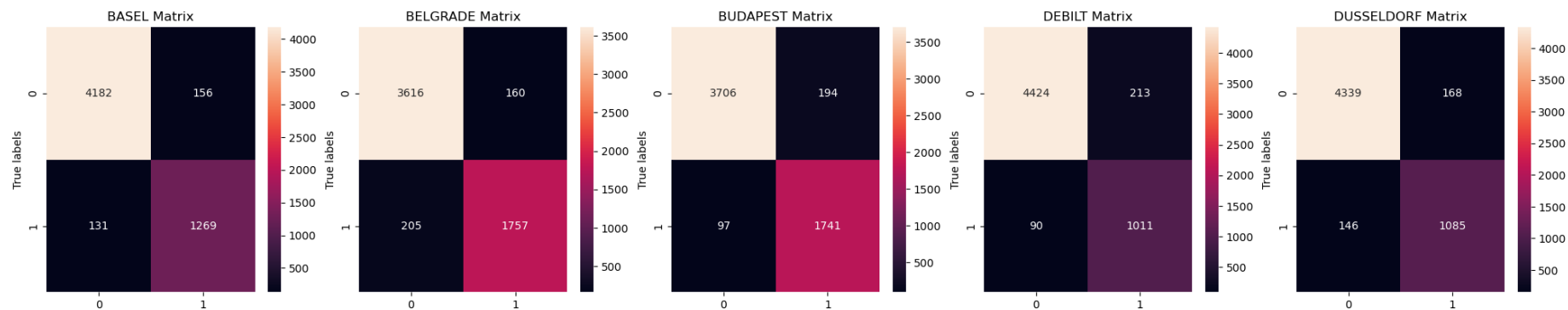
Train Set: 0.81

Test Set: 0.59

ANNS scaled Confusion Matrix - train set



ANNS scaled Confusion Matrix - test set



8. Summary

We identified the urgency of addressing the climate crisis, showcasing international responses, dataset characteristics, and machine learning models' application in predicting climate change consequences.

It emphasized the significance of understanding global warming trends, correlations among climate factors, and the potential of machine learning for climate prediction.

The analysis revealed insights into temperature variations, extreme weather events, and the complexities of climate data correlation. Additionally, machine learning models, particularly ANNs, showed promise in predicting pleasant days based on historical data.

Next Steps & Future Analysis

1. **Refine Machine Learning Models:** Further refine and optimize machine learning models, particularly ANNs, to improve accuracy in predicting climate change consequences.
2. **Expand Dataset:** Expand the dataset to include more locations, diverse climate factors, and longer time periods to enhance model robustness and generalization.
3. **Address Biases:** Address biases in the dataset, such as spatial, temporal, and selection biases, through careful data preprocessing and validation techniques.

Thank You For Your Attention.

Further analysis will be updated and reported soon.
Feel free to reach out with any further questions or inquiries.

[jинwoochung@outlook.com](mailto:jinwoochung@outlook.com)

<http://www.jinwoochung.com>

[Linkedin.com/in/ChungJinwoo](https://www.linkedin.com/in/ChungJinwoo)

Jinwoo Chung (Data Analyst)
ClimateWins
2024. 3. 18