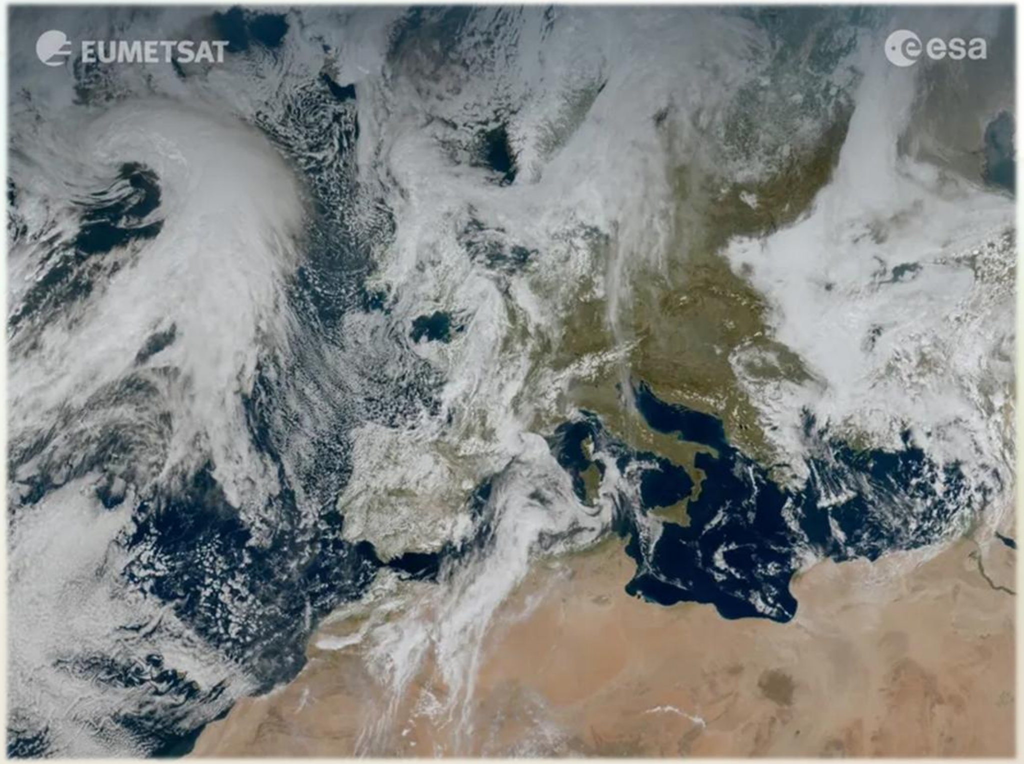
Machine Learning: Weather Conditions and Climate Change For ClimateWins

Esteban Torres-Flores
February 2026

Objective

ClimateWins aims to understand and help predict climate patterns and the impacts of climate change across Europe. To achieve this, a data analyst will examine historical climate data collected by the European Climate Assessment & Data Set project. By applying machine learning to global temperatures and related climate data, ClimateWins seeks to identify patterns, trends, and changes occurring across the world.



Hypothesis

The **average temperatures** across Europe have increased , with higher rate of increase observed in recent years compared to earlier periods.

Extreme weather events have become more frequent over time, indicating increasing climate variability across Europe.

Machine learning models can predict temperature trends more accurately than traditional methods by capturing nonlinear relationships in the data.

Data and Biases

European Climate Assessment & Data Set project

- 18 different weather stations across Europe
- Ranging from 1960’s to 2022
- Values of temperature, wind speed, snow, and global radiation

Biases

- Geographical bias
- Measurement and Instrument bias
- Temporal bias
- Missing data bias

Data Accuracy

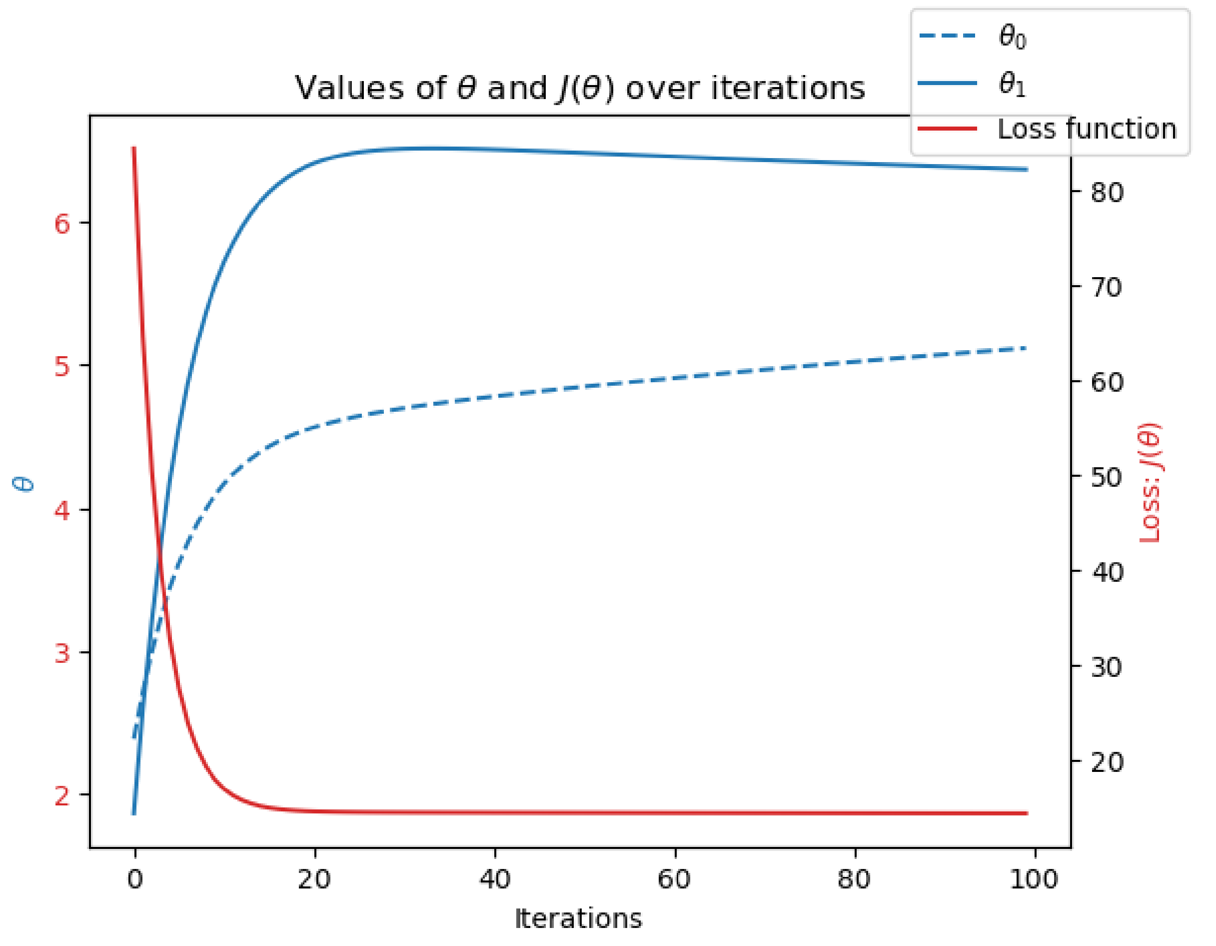
Dataset	Stations
Weather Prediction Processed	18
Weather Prediction Pleasant	15

Differences and Limitations

- The pleasant dataset includes fewer weather stations, reducing coverage.
- The pleasant dataset doesn’t include month variable , limiting temporal analysis.
- The presence of a redundant month column requires dataset alignment during preprocessing.

Data Optimization

- Gradient descent was used to minimize the model’s loss function by reducing prediction error.
- Model parameters were continuously adjusted in the direction that lowers error.
- Each update moved the model closer to a stable minimum, improving accuracy over time.



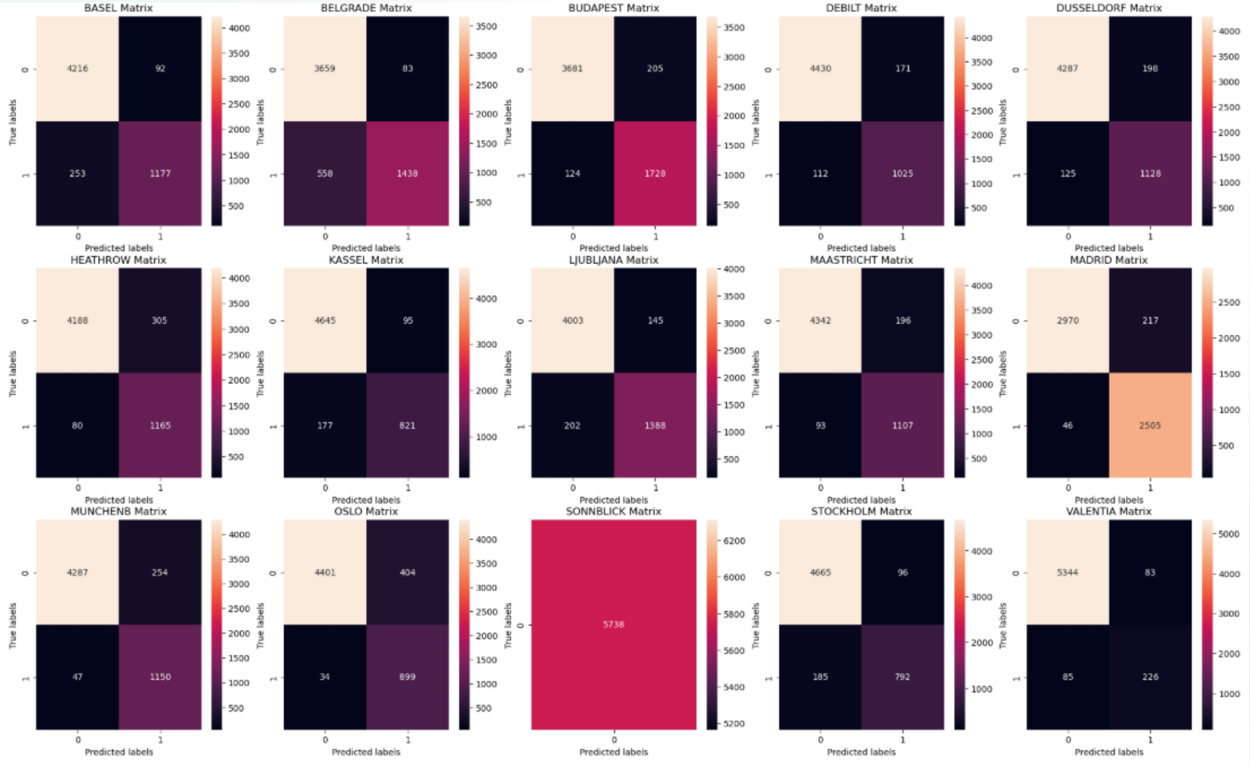
K-Nearest Neighbor

- Weather observations are classified by comparing their conditions to the average weather profile of each class.
- All variable contribute collectively, rather than being evaluated independently.
- The model achieved an average accuracy of 88.5%
- To enhance accuracy and reduce dominance, feature scaling would be applied to prevent any single variable from disproportionately influencing distance calculations.

Weather Station	Positive 0	Positive 1	False 0	False 1	Accuracy Rate
BASEL	3917	961	421	439	85.1%
BELGRADE	3252	1544	524	418	83.7%
BUDAPEST	3424	1462	476	376	85.2%
DEBILT	4320	723	317	378	87.9%
DUSSELDORF	4164	810	343	421	86.7%
HEATHROW	4138	744	432	424	85.1%
KASSEL	4563	614	252	309	90.2%
LJUBLJANA	3740	1180	455	363	85.7%
MAASTRICHT	4253	824	309	352	88.5%
MADRID	2750	2261	418	309	87.3%
MUNCHENB	4237	792	309	400	87.7%
OSLO	4637	512	242	347	89.7%
SONNBLICK	5738	0	0	0	100%
STOCKHOLM	4483	607	283	365	88.7%
VALENTIA	5404	74	58	202	95.5%
				Average	88.5%

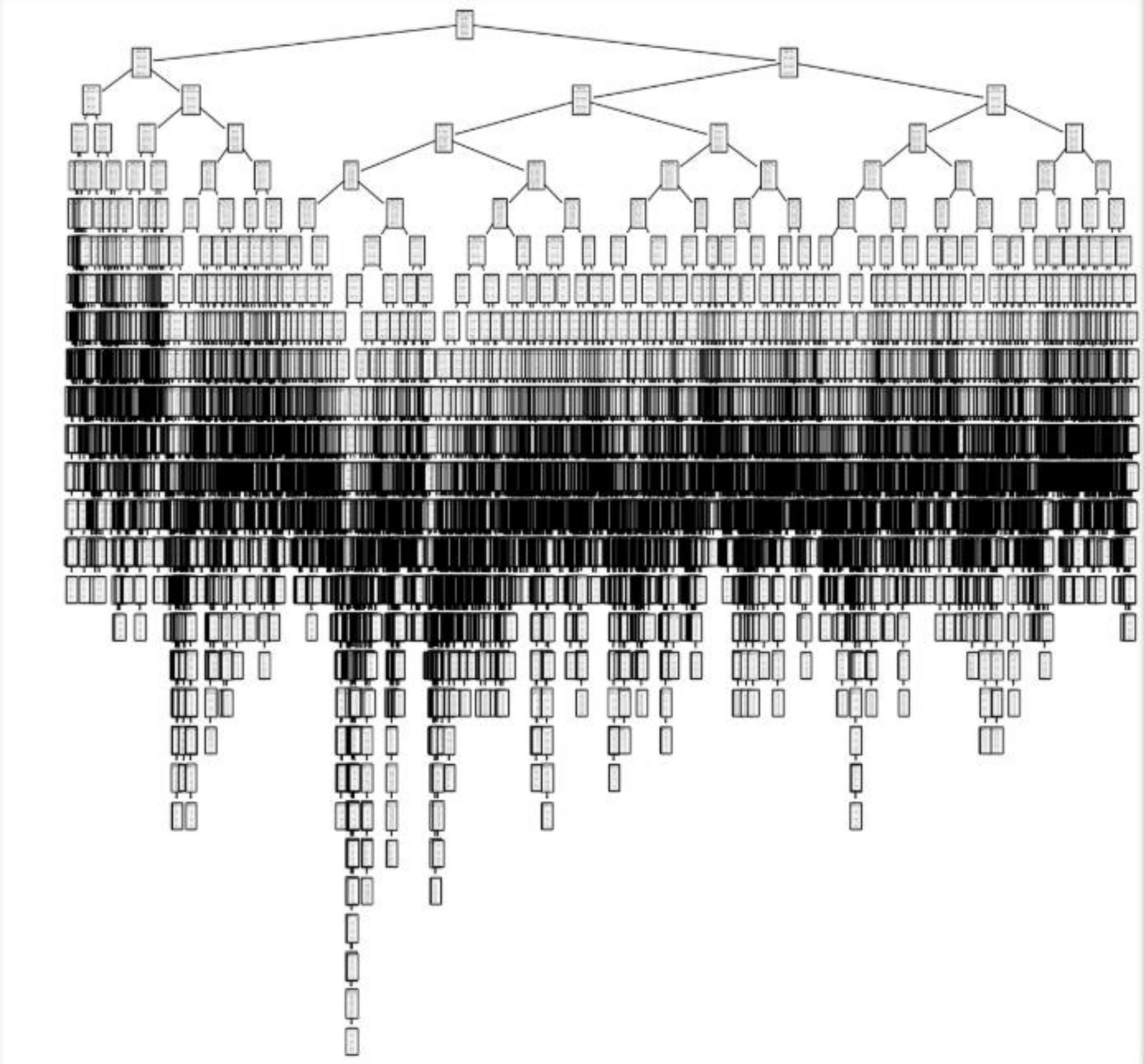
Artificial Neural Network

- Weather observations are processed through multiple interconnected layers, allowing the model to learn complex, nonlinear relationships.
- Model weights are iteratively updated using gradient descent to minimize the loss function and reduce prediction error.
- The model achieved an average accuracy of 62.5%
- ANN achieved the lowest performance among the three models tested.



Decision Tree

- Weather observations are classified by sequentially splitting data based on feature thresholds.
- Each split is chosen to maximize information gain and reduce classification error.
- The model achieved an average accuracy of 63.9%
- The model likely overfits, as many splits capture noise rather than meaningful patterns; pruning is needed to improve generalization.



Model Selection

Supervised learning models including KNN, Decision Tree, and ANN were evaluated to analyze climate trends across Europe. Based on accuracy, KNN was selected as the preferred model due to its ability to capture nonlinear patterns by comparing weather observations based on overall similarity. This makes KNN well suited for testing temperature trends and changes in extreme weather frequency, supporting the hypothesis that machine learning models can better model complex climate behavior than traditional linear approaches.

Summary

Average Temperatures
All models captured rising temperature trends, with KNN most accurately reflecting recent increases.

Extreme Weather Events
KNN best identified patterns linked to increasing climate variability and extreme events.

ML Models

KNN outperformed other models, supporting the hypothesis that machine learning captures non-linear climate behavior more effectively.

Next Steps

- Expand the dataset by incorporating additional weather stations and longer time spans to improve model generalization and trend detection.
- Apply feature scaling, feature selection, and hyperparameter tuning to further improve KNN and ANN performance.
- Explore additional supervised models to better capture extreme weather patterns and reduce overfitting.
- Strengthen temporal analysis of climate trends.

Thank You for Your Time

I’m happy to take any questions.

Esteban Torres-Flores
esteban.a.torres22@gmail.com