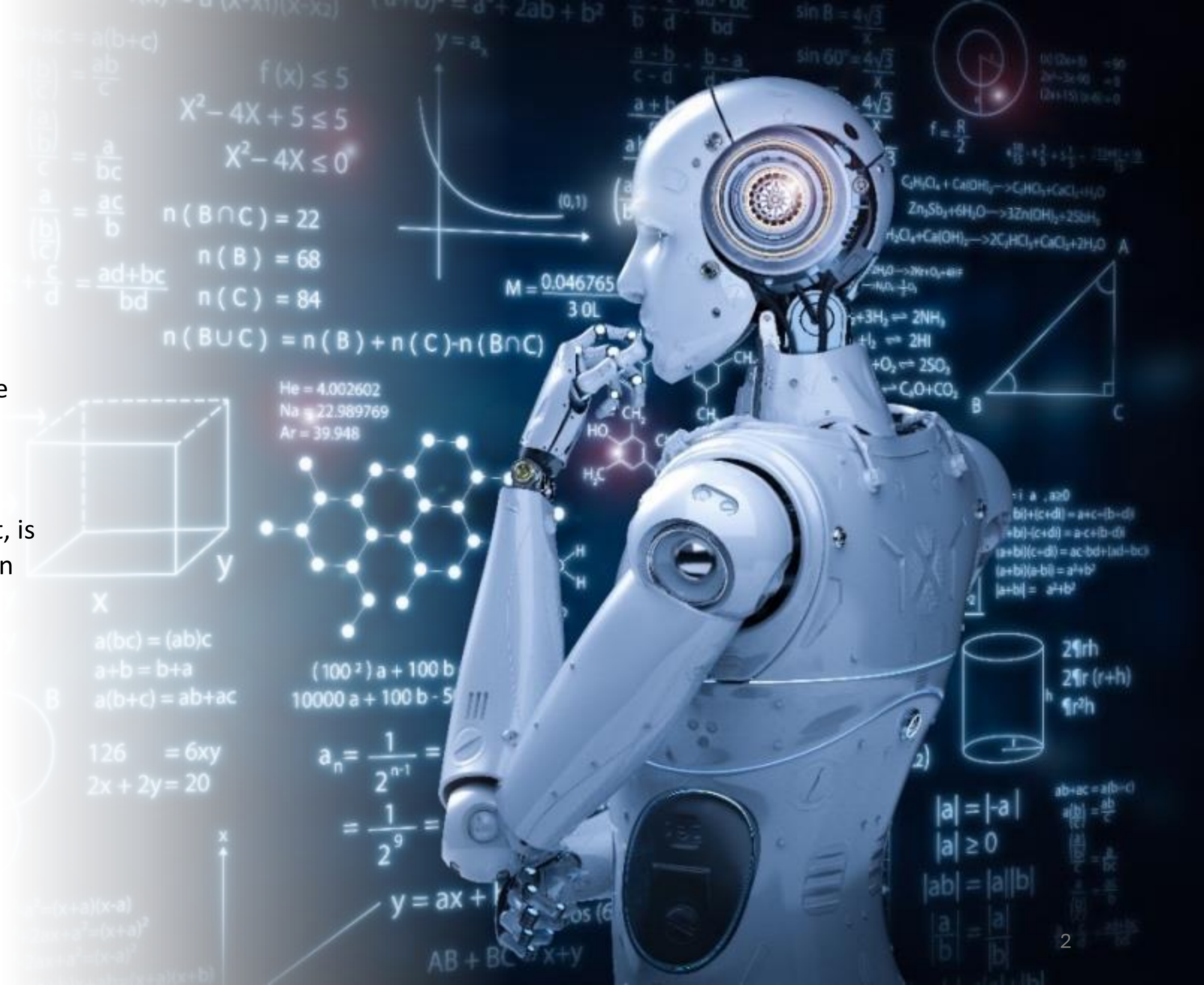
Weather Conditions & Climate Change with ClimateWins

An Interim Report by Dr Yasemin Mustafa

Presented on 23rd July 2025

Why Does It Matter?

- Extreme weather is becoming more frequent and severe across Europe and the world.
- ClimateWins, a European nonprofit, is exploring how machine learning can help predict and understand the impacts of climate change.
- This project uses data science to support smarter, more adaptive responses to our changing environment.



Inside the Figures

- Daily weather records from 18 stations across Europe
- Sourced from the European Climate Assessment & Dataset (ECA&D)
- Covers the period 1960–2022
- Includes temperature, humidity, precipitation, snow, and more
- Complete dataset with no missing values
- Available as a downloadable .csv file

Not All Data is Created Equal

- Temporal Bias: Older records (e.g. 1960s) may reflect outdated tools and methods
- Historical Bias: Long-term data can mask recent climate shifts
- Selection Bias: Stations are concentrated in developed areas—underrepresenting others

What Might We Discover?

1

If historical climate data across Europe is used, then machine learning models can accurately predict daily weather conditions.

2

If geographic and climatic variability differs by region, then prediction accuracy will vary accordingly.

3

If local climate patterns and data structure vary by region, then no single algorithm will consistently outperform others across all locations.

Fine-Tuning the Forecast

- Gradient Descent was used to optimise the weather prediction model
- This method adjusts parameters to minimise the loss function (i.e. the error in prediction)
- Through iterative updates, the model learns to make more accurate predictions
- Temperature data is highly suitable for predictive modelling using differentiable functions

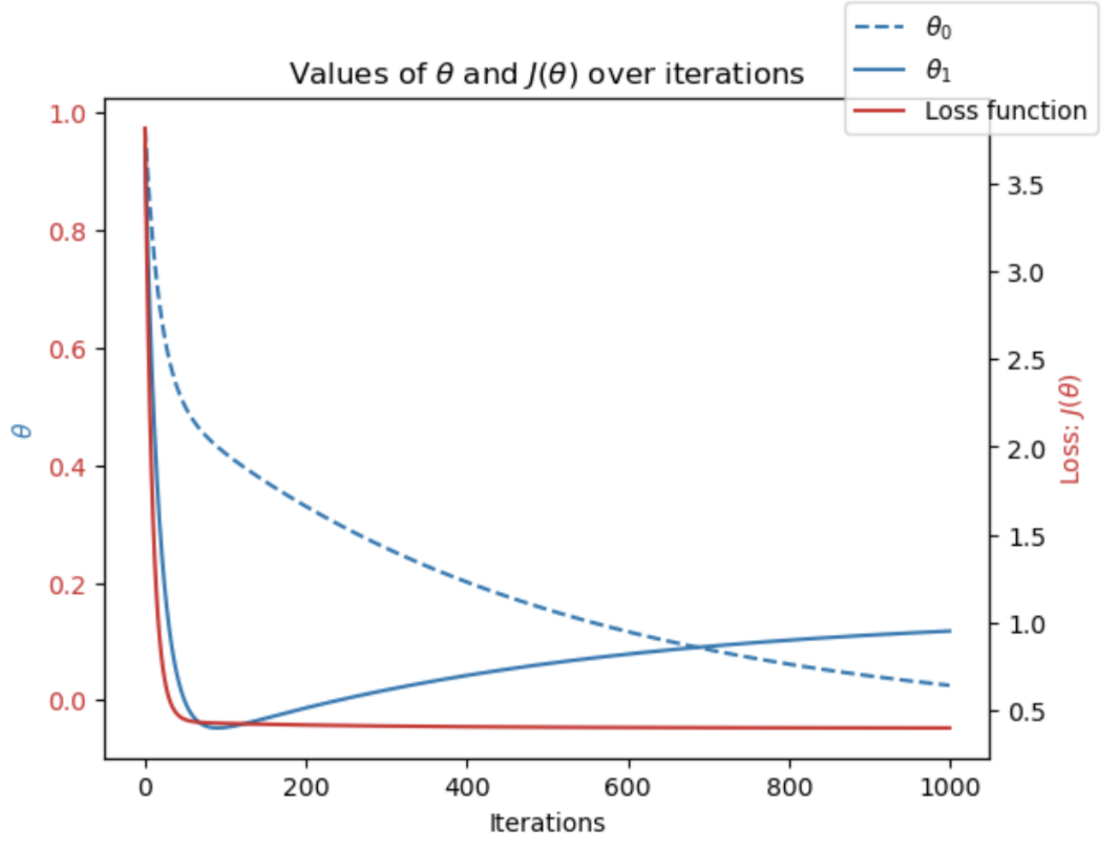


Figure 1. Gradient descent loss function for Dusseldorf's average daily temperatures in 2020.

Method 1: K-Nearest Neighbor (KNN)

- KNN predicts weather by comparing each new data point to its closest neighbors
- The model groups similar data points, using a k-value of 3
- Applied to 15 European weather stations, it predicted pleasant weather days with up to 88% accuracy
- KNN works well but can overfit, especially in stations with extreme or limited weather patterns

Table 1. KNN Predictions for 15 different weather stations

Weather Station	Accurate Predictions		False Positive	False Negative	Accuracy (%)
Basel	3907	935	431	465	85
Belgrade	3238	1502	538	460	81
Budapest	3416	1432	484	406	85
Deblit	4346	732	291	369	85
Dusseldorf	4167	800	340	431	87
Heathrow	4161	754	409	414	86
Kassel	4563	607	252	316	90
Ljubljana	3726	1133	469	410	85
Maastricht	4249	819	313	357	88
Madrid	2735	2257	433	313	87
Munchenb	4222	766	324	426	87
Oslo	4624	507	255	352	89
Sonnblick	5738	0	0	0	100
Stockholm	4449	588	317	384	88
Valentia	5391	108	71	168	96
Average					88

Method 2: Decision Tree

- Make predictions by asking a series of yes/no questions
- The algorithm splits data into branches until it reaches a final classification (leaf)
- Accuracy: ~85% (training and testing)
- The model became too complex, with many branches causing overfitting and long training times
- To improve performance, the tree likely needs to be pruned

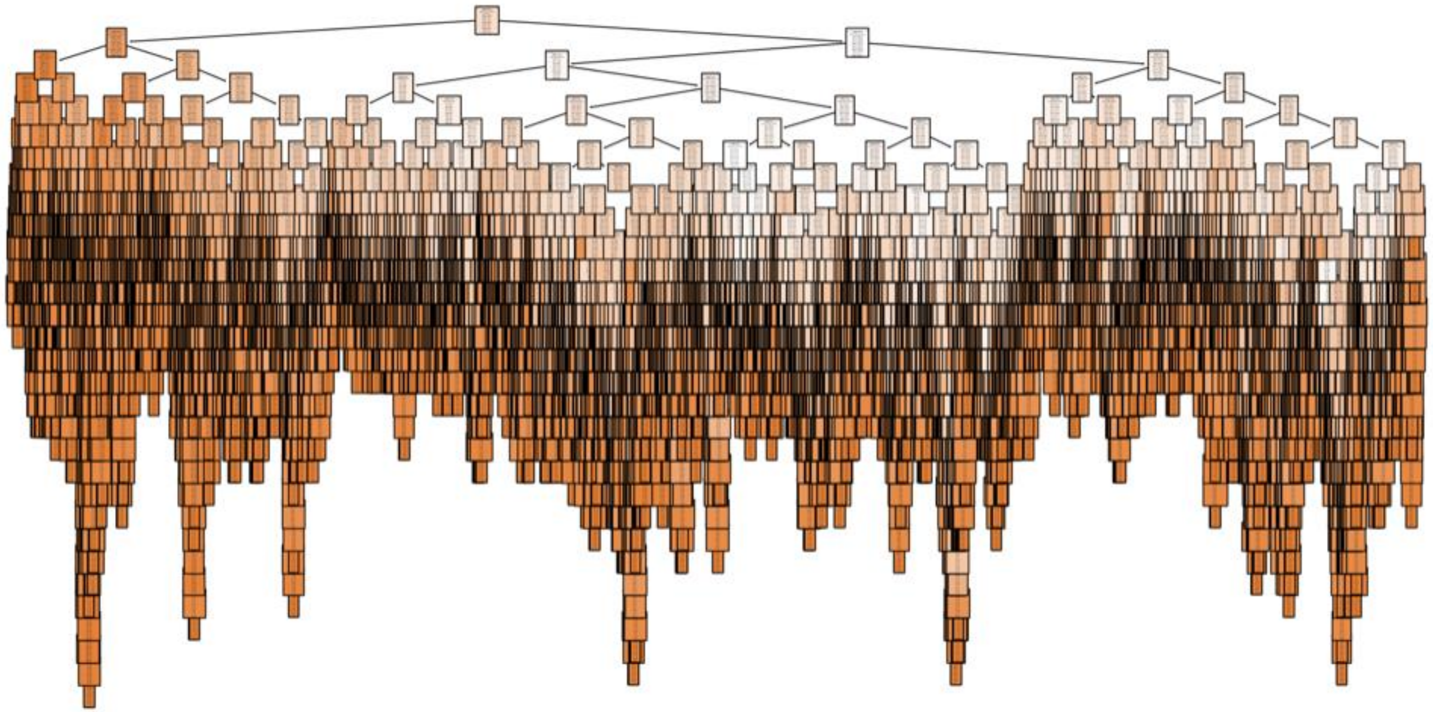


Figure 2. Decision tree used on 15 European weather stations to predict pleasant weather days.

Method 3: Artificial Neural Network (AANs)

- Mimics how the human brain learns
- Uses layers of “neurons” to find patterns in data
- Predicts pleasant weather based on past conditions
- Tested on 3 European stations
- Accuracy: 89% (training), 87% (testing)
- Tweaked layer size and iterations during training
- Overfitting likely—model performed better on training than testing data

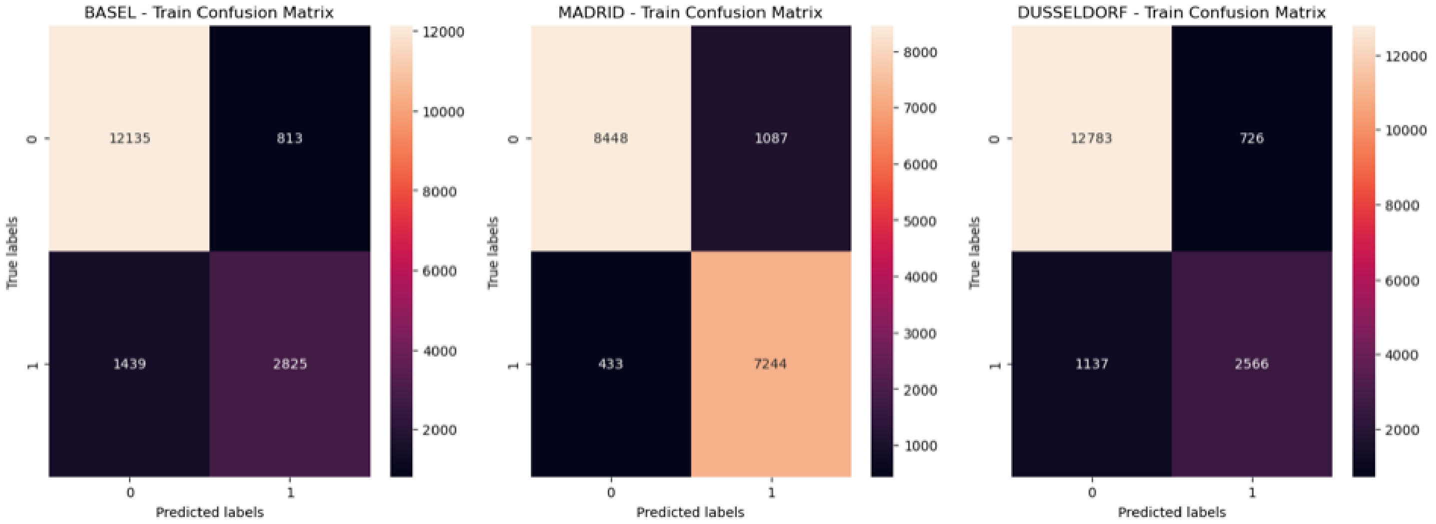


Figure 3. Confusion matrices showing accurate (top-left, bottom-right) and inaccurate (top-right, bottom-left) predictions for pleasant and unpleasant days, respectively for training data.

Which Algorithm Performs Best?

Model	Accuracy	Strengths	Limitations	Notes
KNN	88% (highest)	Most consistent and reliable across stations	May underperform in highly nonlinear contexts	Recommended default model for ClimateWins
Decision Tree	~85%	Performs well; interpretable results	Complex, less generalisable	Suitable where model interpretability is critical
ANN	Varies (some high)	Excels in nonlinear pattern detection	Signs of mild overfitting	Valuable in regions with complex climate dynamics
All Models	—	—	—	Sonnblick's perfect scores suggest possible data imbalance

Summary

Machine learning can predict daily weather conditions in Europe accurately, with some models reaching up to 88% accuracy.

Prediction accuracy varies by location, influenced by regional climate patterns.

Over time, machine learning models can help detect signs of climate change by tracking increases in unpleasant weather events.

Next Steps

- Explore unsupervised machine learning to uncover hidden patterns
- Expand data inputs (e.g. wind speed, snow) for deeper analysis
- Continue tuning model parameters to improve accuracy and reduce overfitting
- Include additional weather stations to boost regional diversity and generalisability
- Begin tracking long-term weather trends for early signs of climate change

Thank You

I appreciate your time and interest in this project.

If you'd like to explore the code, data, and visualisations further, please check out my GitHub:

👉 **https://github.com/YLMustafa/climate_wins**

Feel free to reach out with any questions or feedback!

Youtube Video of Presentation: https://youtu.be/La4G0uLam38