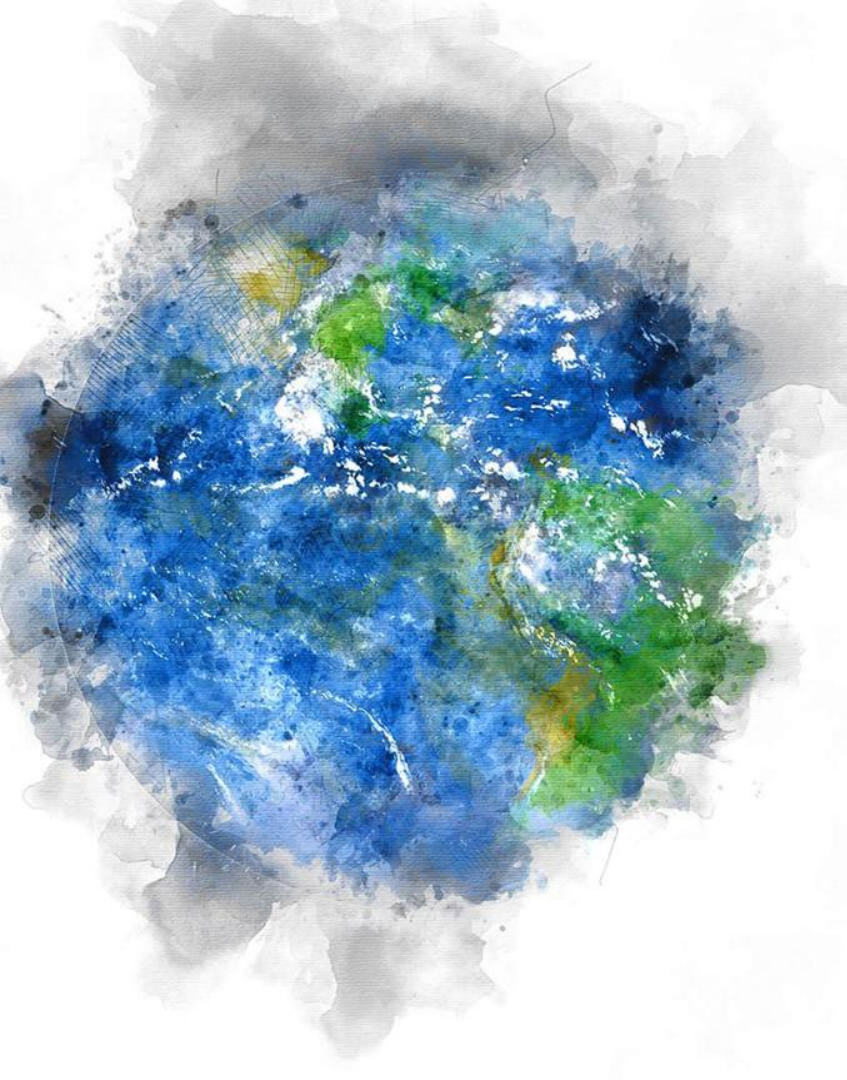
Weather Conditions and Climate Change Analysis with ClimateWins

ANDY CEN



Contents

Project Goals

Goal: *ClimateWins* is interested in using machine learning to **help predict the consequences of climate change** around Europe and, potentially, the world.

Motivation: ClimateWins has been sorting through **hurricane predictions** from The National Oceanic and Atmospheric Administration (NOAA) in the U.S., typhoon data from The Japan Meteorological Agency (JMA) in Japan, world temperatures, and a great deal of other data.

Objective: In this project, I will use **machine learning** to **help predict the consequences of climate change** while working as a **data analyst** at a European nonprofit organization.

Hypotheses

- The accuracy of weather predictions will vary depending on the geographic location, as climate conditions differ from each region.
- Machine Learning Algorithms can accurately predict future weather conditions.
- Extreme weather events, which have increased over the past 10 to 20 years, can be predicted and planned for using advanced tools like machine learning.

Data Set

Date Range

Data is collected from the late 1800s to 2022.

Data Collection

Data is collected by the European Climate Assessment & Data Set Project.

Data is Based on weather observations from 18 different weather stations across Europe

Data Geography

Data consists of wind speed, temperature, global radiation, and more.

Data Content

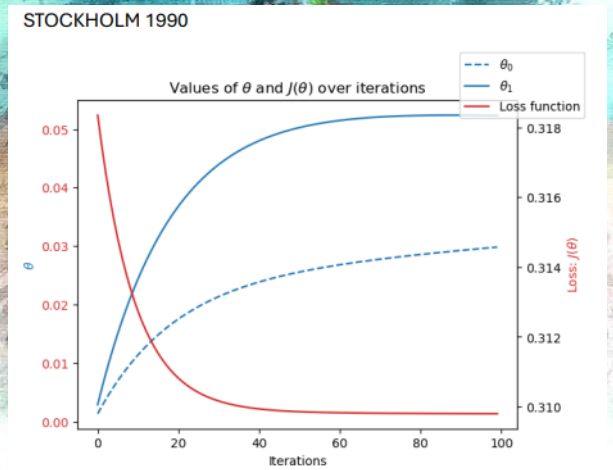
Data Biases

Privacy and Ethics: Using location data in climate change models requires caution to avoid unintentional exposure of personal information.

Regional Bias: Data from one region may not apply universally, so diverse regional datasets are necessary to avoid skewed predictions.

Human Bias: Bias in data collection or reporting could influence machine learning models, leading to inaccurate predictions.

Data Quality: Incomplete or biased data can result in incorrect weather predictions, making supervision and diverse data essential for accurate forecasts.



Data Optimization

This data set was optimized through **Gradient Descent**.

- One of the simplest ways to find a **local minimum** (or valley), as well as one of the most common.
- Adjusts the parameters to **reduce** a loss function to its minimum
- Can be used in both **linear** and **nonlinear** cases.

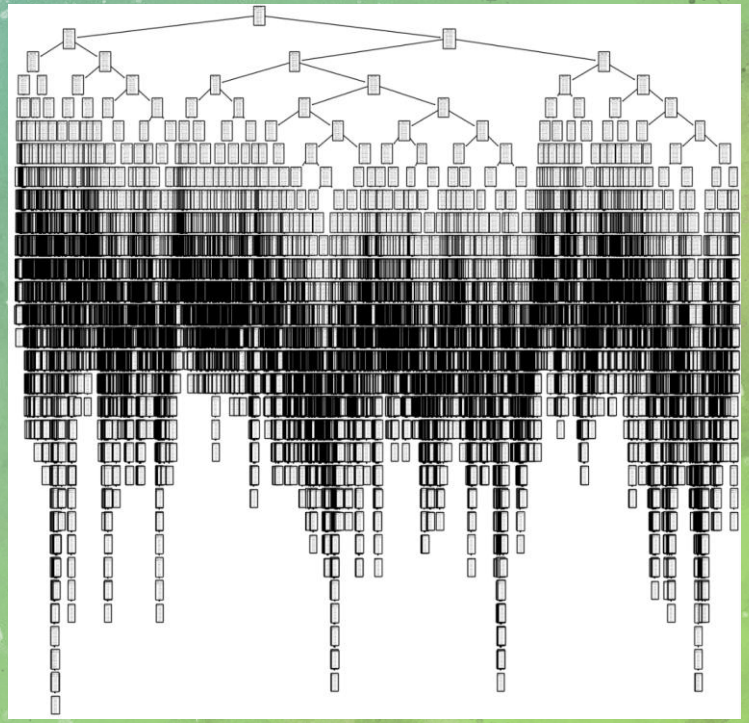
K-Nearest Neighbors

Weather Station	Accurate Predictions		False Positive	False Negative	Accuracy Rate
BASEL	3907	935	431	465	84.8%
BELGRADE	3238	1502	538	460	82.6%
BUDAPEST	3416	1432	484	406	84.5%
DEBILT	4346	732	291	369	88.6%
DUSSELDORF	4167	800	340	431	86.6%
HEATHROW	4161	754	409	414	85.7%
KASSEL	4563	607	252	316	90.1%
LJUBLJANA	3726	1133	469	410	84.7%
MAASTRICHT	4249	819	313	357	88.3%
MADRID	2735	2257	433	313	87.0%
MUNCHENB	4222	766	324	426	86.9%
OSLO	4624	507	255	352	89.4%
SONNBLICK	5738	0	0	0	100%
STOCKHOLM	4449	588	317	384	87.8%
VALENTIA	5391	108	71	168	95.8%
				AVERAGE %	88.2%

- *KNN classifies data based on the majority class of its nearest neighbors.*
- *Used to predict pleasant weather in 15 European stations.*
- *Achieved 88.2% average accuracy.*

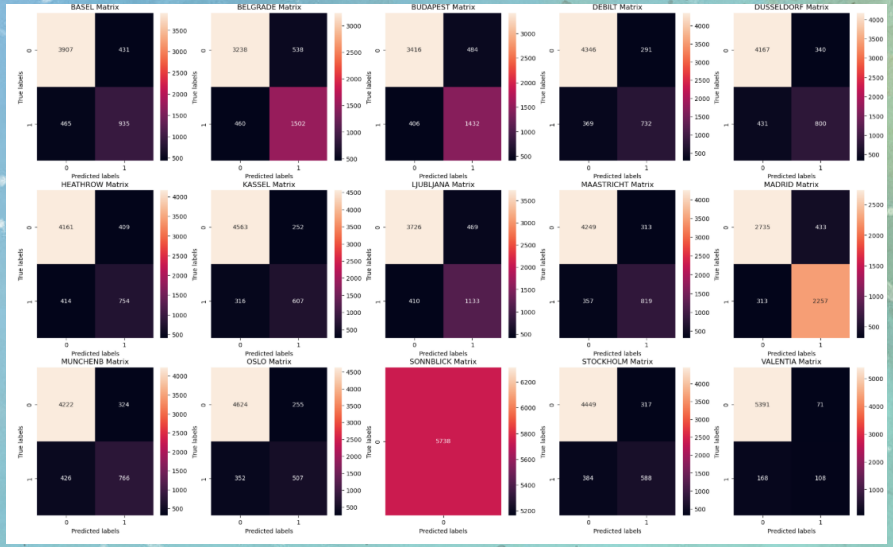
Decision Tree

- The decision tree algorithm classifies data by branching based on features, ending at a classification leaf.
- Used to predict pleasant weather in 15 European stations with 46.1% train accuracy.
- Likely overfitting; pruning may improve accuracy.



Artificial Neural Network

- ANN mimics the brain using layers of interconnected neurons to learn patterns.
- Learns by adjusting neuron connections; optimized via parameters like iterations, layer size, and learning rate.
- Achieved a highest test accuracy of 67%.



Best Algorithm for ClimateWins Analysis

- The decision tree needs further pruning for more accurate predictions.
- The K-Nearest Neighbors Algorithm had the highest accuracy of 88.2% for predicting the climate temperature.
- The Artificial Neural Network had a test accuracy of 67% after 3 tests.
- The ANN model allows for more complexity in its analysis which sets it apart from the KNN model.
- I believe that with further tests and adjustments with more added variables, the ANN would be the better algorithm.

Summary

The accuracy of weather predictions does vary depending on the geographic location, as climate conditions differ from each region.

> Sonnblick had 5,738 accurate predictions whereas Basel only had 3,907 accurate predictions.

Machine Learning Algorithms can accurately predict future weather conditions.

> KNN had an average accuracy of 88.2%

Extreme weather events, which have increased over the past 10 to 20 years, can be predicted and planned for using advanced tools like machine learning.

> With high accuracies from the ANN and KNN models, it is safe to say that extreme weather events can be predicted using machine learning.

Next Steps

Diversify variables other than temperature to find more insights and trends.

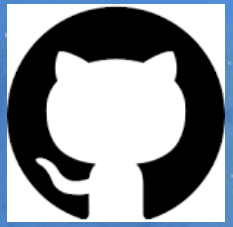
Incorporate more weather stations to expand machine learning capabilities.

Continue to run further tests and make adjustments and improvements to bolster model accuracies.

Thank you!

For any further questions, please contact me at my email

andycen7@gmail.com



Check out my GitHub for the Python scripts and datasets used in this analysis

GitHub Link