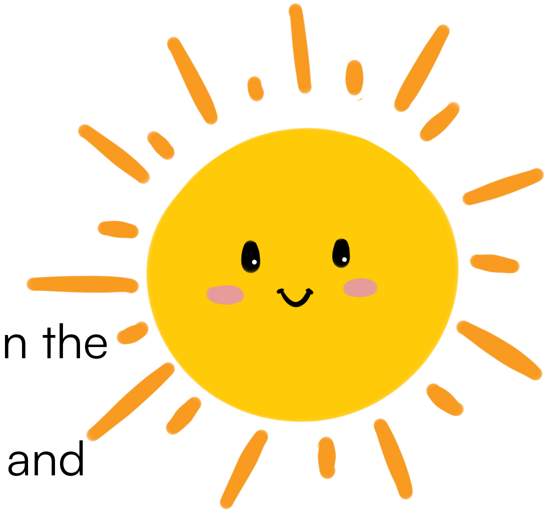
Using Machine Learning for Weather Prediction with ClimateWin

Shaquille Obomeghie

November 2024

Project Objective



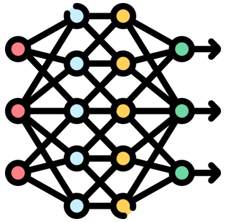
ClimateWin is concerned with increased extreme weather events, especially in the last 10 to 20 years. ClimateWin is interested in the applications of machine learning to help predict the consequences of climate change around Europe and potentially the world.

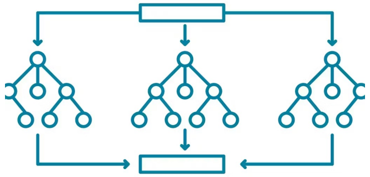
ClimateWin is interested in:

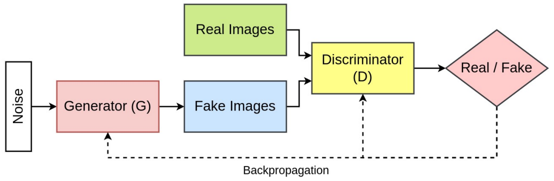
- Finding new patterns in weather changes over the last 60 years.
- Identifying Weather patterns outside the norm in Europe.
- Determining whether unusual weather patterns are increasing.
- Generation possibilities for future weather conditions over the next 25 to 50 years based on current trends.
- Determining the safest places for people to live in Europe within the next 25 to 50 years.

Machine Learning Options

There are multiple Machine Learning Algorithms that could be used for this project, but these were highlighted.







CNN & RNN:
CNN is highly efficient in processing spatial data, while RNN is suited for modelling sequential data.

How it works: CNN handles spatial data, while RNN captures temporal dependencies in sequential data (e.g, historical weather trends)

Random Forest:
This is an ensemble learning method that builds multiple decision trees and combines their outputs to improve prediction accuracy and control overfitting

How it works: This works by combining multiple decision trees to provide accurate predictions for both regression and classification problems

GANs (Generative Adversarial Networks):
This uses two neural networks (generator and discriminator) working against each other.

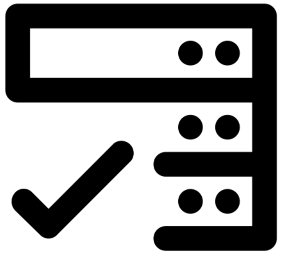
How it works: The generator creates synthetic data and tries to fool the discriminator by generating realistic weather data. While the discriminator differentiates between real weather data and synthetic data

Data Beyond Historical Weather Set

1. **Satellite Imagery:** High-resolution images of cloud cover, land surface and vegetation could be essential especially when working with CNN.

2. **Socioeconomic Data:** Data set that consists of information about energy consumption and population data affecting weather demand. This would be essential when running the RNN model.

3. **Extreme Events Record:** Data on hurricane, flood, droughts and their impact over time



Slide 1: Discovering New Patterns in Historical Weather Changes

Link to ClimateWins Goals:
- Uncover new patterns in weather changes over the last 60 years.

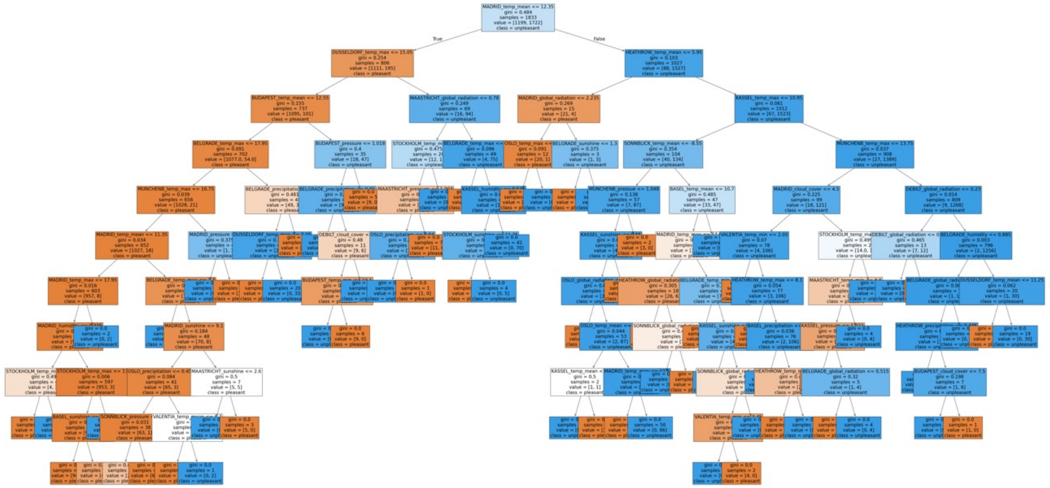
Machine Learning Methods:
- **Random Forest**: Identify key drivers of historical weather changes by analyzing variable importance and uncovering complex feature interactions.

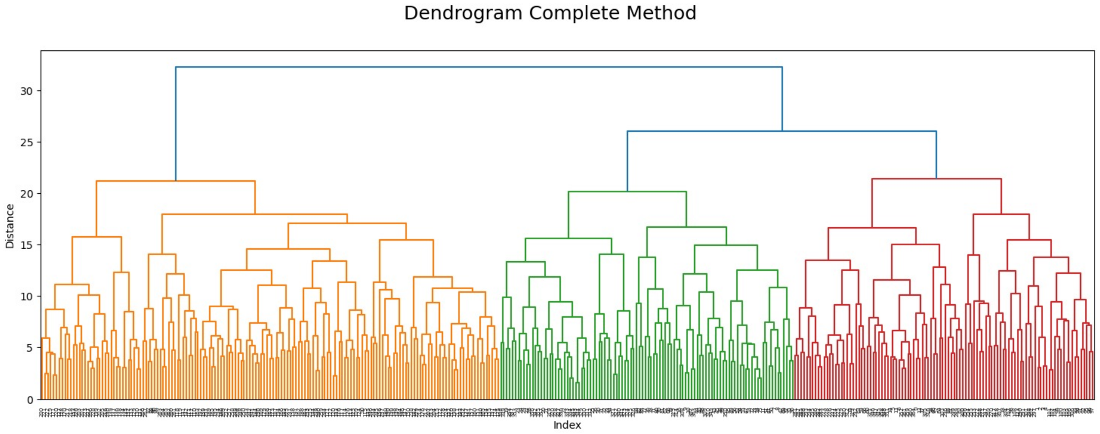
- **Dendrogram:** Perform hierarchical clustering to group similar weather trends across regions or time periods.

Data Needed:
- Historical meteorological data (temperature, humidity, precipitation, wind speeds) from global sources (1960—present).

Outcome:
- Detailed insights into long-term weather trends, enabling ClimateWins to identify critical drivers of climate shifts.

Application





- This is a random forest created for 15 weather stations across Europe based on the pleasant and unpleasant weather conditions
- Accuracy after optimization was 95%.

- The dendrogram complete method above was used to cluster/group weather data (e.g. temperature, humidity, wind speed, etc. It helps to understand and simplify the complexity of weather patterns

Slide 2: Identifying and Tracking Unusual Weather Patterns in Europe

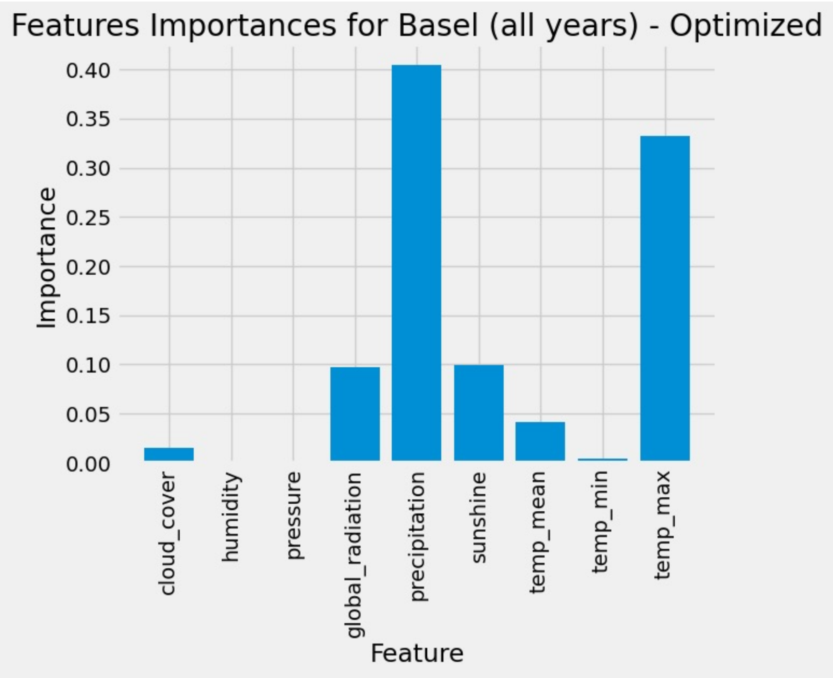
Link to ClimateWins Goals:
- Identify weather patterns outside the norm in Europe.
- Determine whether unusual weather patterns are increasing.

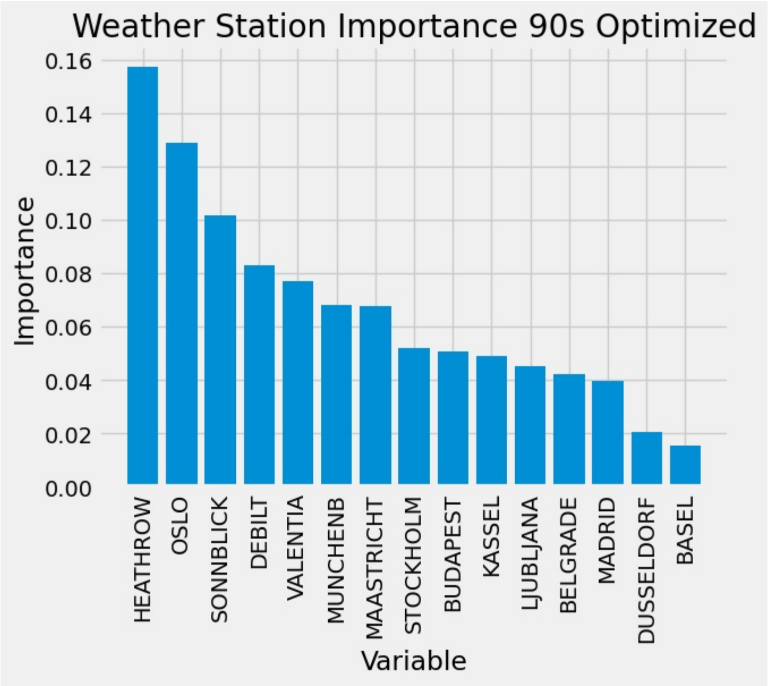
Machine Learning Methods:
- **CNN**: Process spatial and temporal climate data (e.g., satellite maps and weather grids) to detect European irregularity.
- **Random Forest**: Random forest is used to highlight which weather factors are the most prominent.

Data Needed:
- Historical and real-time European weather data (temperature, precipitation, wind speeds, pressure).

Outcome:
- Comprehensive maps highlighting unusual weather trends and frequency shifts across Europe, providing evidence of climatic changes.

Application





- This score shows the score of feature importance and its prominence in various.
- This was a 100% accuracy.

Slide 3: Future Weather Generation and Climate Safety Mapping

Link to ClimateWins Goals:

- Generate possibilities for future weather conditions over the next 25—50 years based on current trends.
- Determine the safest places for people to live in Europe.

Machine Learning Methods:

- **Generative Adversarial Networks (GAN)**: Generate realistic future weather scenarios by modelling changes based on historical trends and projections.
- **CNN**: Analyze simulated weather outputs to identify spatially safe or vulnerable areas.
- **Random Forest**: Evaluate the impact of generated scenarios on habitability and safety indicators.

Data Needed:

- Current and historical climate data (temperature, extreme weather events, emissions).
- Future projections: Industrial, demographic, and urbanization trends.
- Socioeconomic data for evaluating livability and vulnerability.

Outcome:

- Predictive maps and recommendations highlight safe zones for habitation based on minimized risk of extreme weather, aiding policymakers and planners.

Thought Experiment	Potential	Key Algorithm	Data Needed
1. Discovering New Patterns in Historical Weather	High: Provides a foundation for understanding long-term climate trends.	Random Forest, CNN	Historical weather data (1960–present), geographic data, emissions, and industrial activity trends.
2. Identifying and Tracking Unusual Weather	Medium-High: Critical for understanding and mitigating unusual climate phenomena.	CNN, Random Forest	Historical and real-time European weather data, anomaly criteria, and satellite imagery.
3. Generating Future Weather Scenarios and Safety	Very High: Directly informs policy and planning for future climate resilience.	GAN, CNN, Random Forest	Current and historical weather data, future projections, socioeconomic and demographic data.

Recommendations

1. Prioritize Thought Experiment 3:

- **Reason:** Offers actionable insights for future planning and aligns closely with ClimateWins' goals of future preparedness and safety.
- **Value:** Helps identify safe zones, aiding policy, urban planning, and climate adaptation strategies.

2. Support Thought Experiment 1 as a Foundational Study:

- **Reason:** Provides baseline insights into weather trends, critical for validating future projections.

3. Use Thought Experiment 2 for Short-Term Insights:

- **Reason:** Helps address current climate anomalies, which can provide immediate applications in mitigation strategies.

Thank You

[Github](#) [Email](#) [LinkedIn](#)