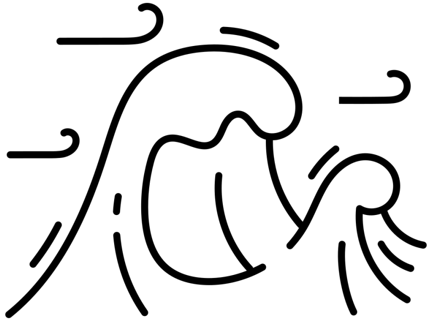
Weather Conditions and Climate Change with ClimateWins

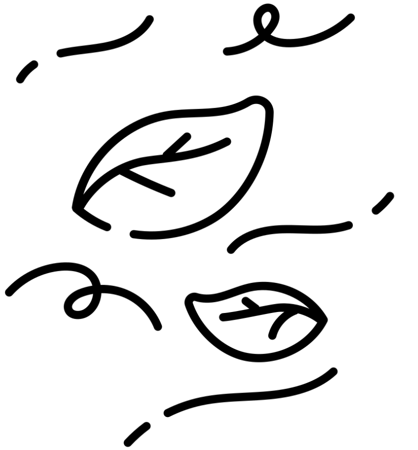
Rachel Grigiac



Objective

Using machine learning to help predict the consequences of climate change:

- Identifying weather patterns outside the regional norm in Europe.
- Determining if unusual weather patterns are increasing.
- Generating possibilities for futures weather conditions over the next 25 to 50 years based on current trends and determining the safest places for people to live in Europe.

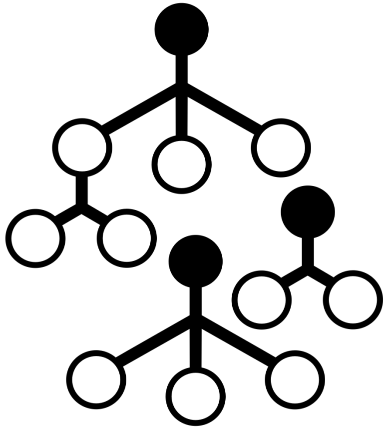


Machine Learning Options

<u>Random Forest:</u>

Builds multiple decision trees using random subsets of data and features, then combines their predictions to improve accuracy and reduce overfitting.

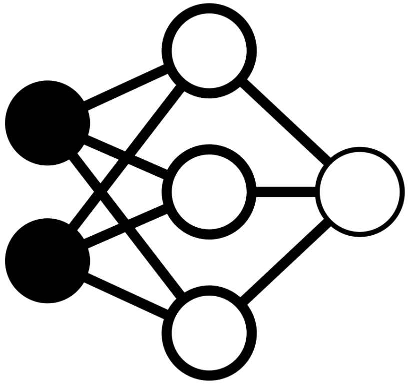
Can be used to predict weather outcomes, such as temperature or precipitation, by analyzing multiple environmental features and their interactions.



<u>CNNs & RNNs :</u>

CNNs detect spatial patterns in data, while RNNs handle sequences by tracking past inputs. Combined, they tackle tasks involving both spatial and sequential information, like video analysis.

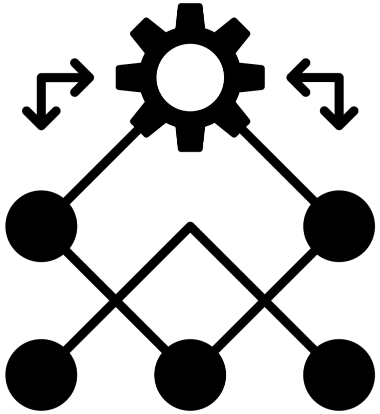
Can analyze satellite images for weather pattern recognition, while RNNs can model time series data to forecast future weather conditions based on historical trends.



<u>GANs (Generative Adversarial Networks):</u>

Involve two networks, a generator and a discriminator, competing to create realistic synthetic data.

Can generate realistic synthetic weather data to augment datasets for training models, improving the robustness of predictions in weather analysis.



Thought experiment #1
Detect and classify unusual weather events

Goal:

Detect and classify unusual weather events (e.g: storms, floods) using satellite imagery and historical weather data.

Approach:

- Use CNNs to analyze satellite images and identify visual patterns associated with severe weather events (e.g., cloud formations indicating storms).
- Extract weather features (temperature, humidity, etc.) from historical data and train a Random Forest model to classify weather events based on these features.
- Use RNNs to analyze historical time-series weather data to understand the sequences leading up to unusual weather events and improve classification accuracy.

Data Needed:

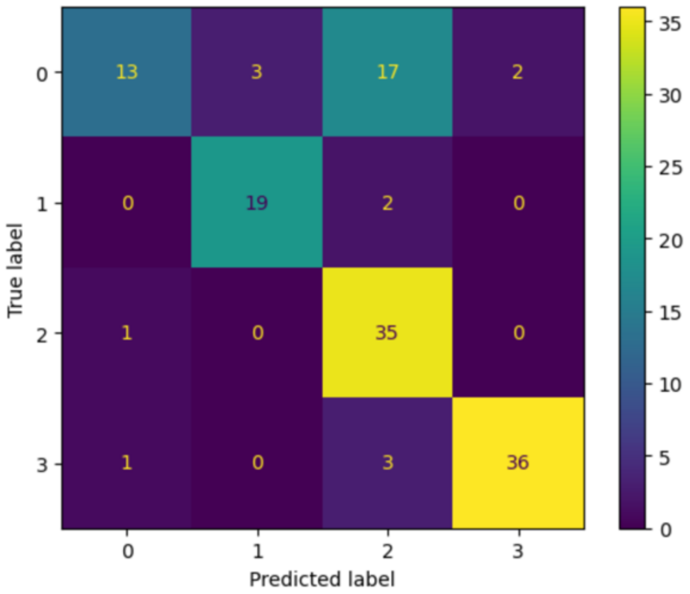
- High-resolution satellite images of weather phenomena.
- Historical weather datasets with labels for various weather events.
- Try to find specifical event leading up to unusual weather.

Data Bias:

Satellite imagery may not represent all geographic areas equally, resulting in underrepresentation of certain weather events and geographic phenomena.

A confusion matrix could help evaluate the CNN's and Random Forest's ability to correctly classify different weather events (storms, floods) by comparing predicted vs actual events. The RNN's predictions of sequences leading up to unusual events can be assessed through precision and recall to fine-tune detection.

Before-and-after image of a region during a storm



Confusion Matrix comparing predicted vs actual events

Thought experiment #2

Assess climate risk across different regions in Europe

Goal:

Assess climate risk across different regions in Europe by integrating image analysis, historical data, and predictive modeling.

Approach:

- Use CNNs to process satellite imagery for real-time identification of climate risks (e.g: drought areas, flood zones).
- Train a Random Forest model on regional historical weather data combined with CNN outputs to classify areas based on climate risk levels.
- Implement an RNN to predict future climate risks based on historical trends, considering the impact of temporal weather changes.

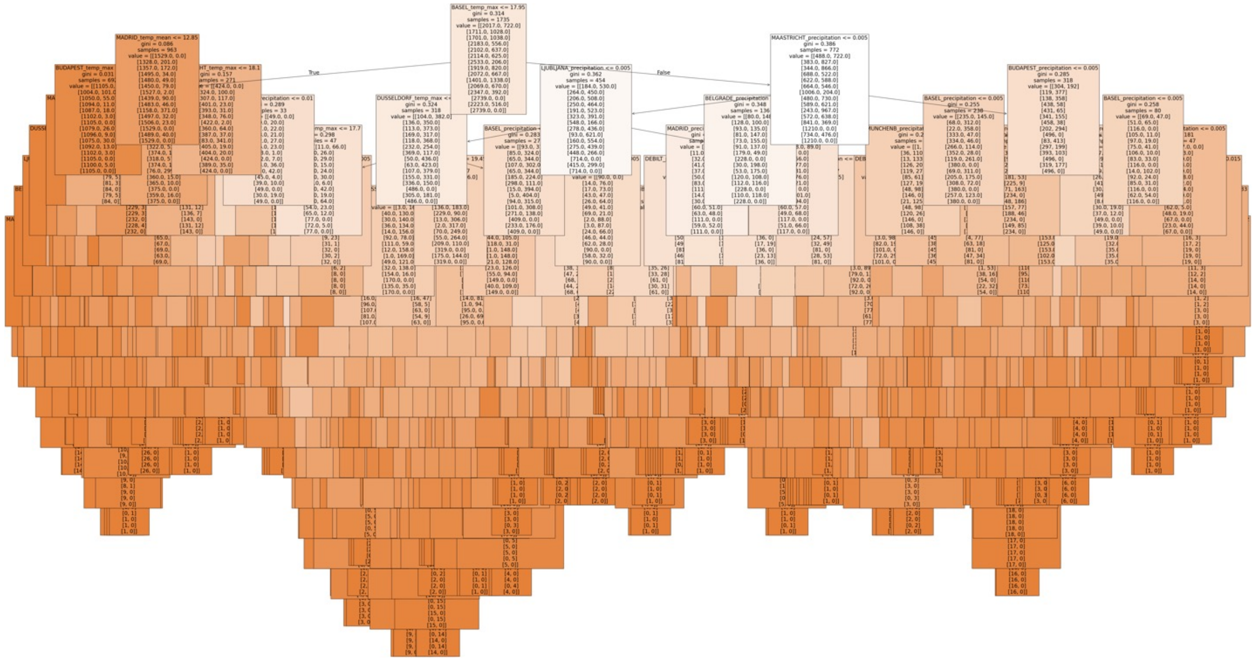
Data Needed:

- Satellite images showing land use and weather impacts.
- Historical weather datasets for different European regions.
- Climate risk assessments and impact data (e.g., past flood or drought occurrences).

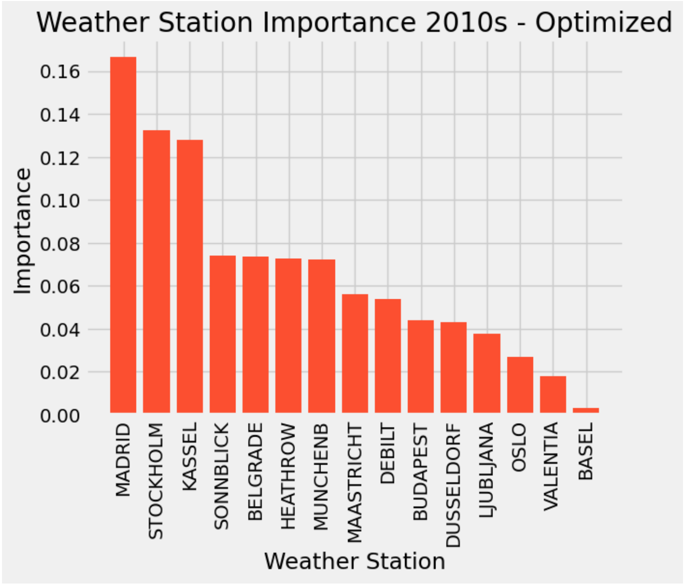
Data Bias:

Historical datasets may underrepresent certain geographic areas, particularly those with less research focus or funding, skewing climate risk evaluations.

The Random Forest model assesses climate risk levels, with CNNs analyzing satellite images to identify high-risk areas visually. RNN forecasts can be validated with metrics like RMSE, while PCA reduces data complexity by highlighting critical risk factors, such as droughts. Dendrograms cluster regions by similar vulnerabilities, aiding targeted adaptation strategies.



Random Forest Results: Classifies regions based on climatic characteristics



Feature Importance Bar Chart: Ranks weather stations by their importance in the classification model

Thought experiment #3
Simulate the impact of climate change

Goal:
Simulate the impact of climate change on weather patterns and assess future scenarios for adaptation strategies.

Approach:

- Use GANs to generate synthetic weather data under different climate change scenarios (e.g: increased temperatures, altered precipitation patterns).
- Train RNNs on both real and GAN-generated data to forecast future weather conditions, focusing on key variables affected by climate change.
- Analyze which generated scenarios most significantly affect weather patterns using Random Forest, identifying critical variables and thresholds for adaptation strategies.

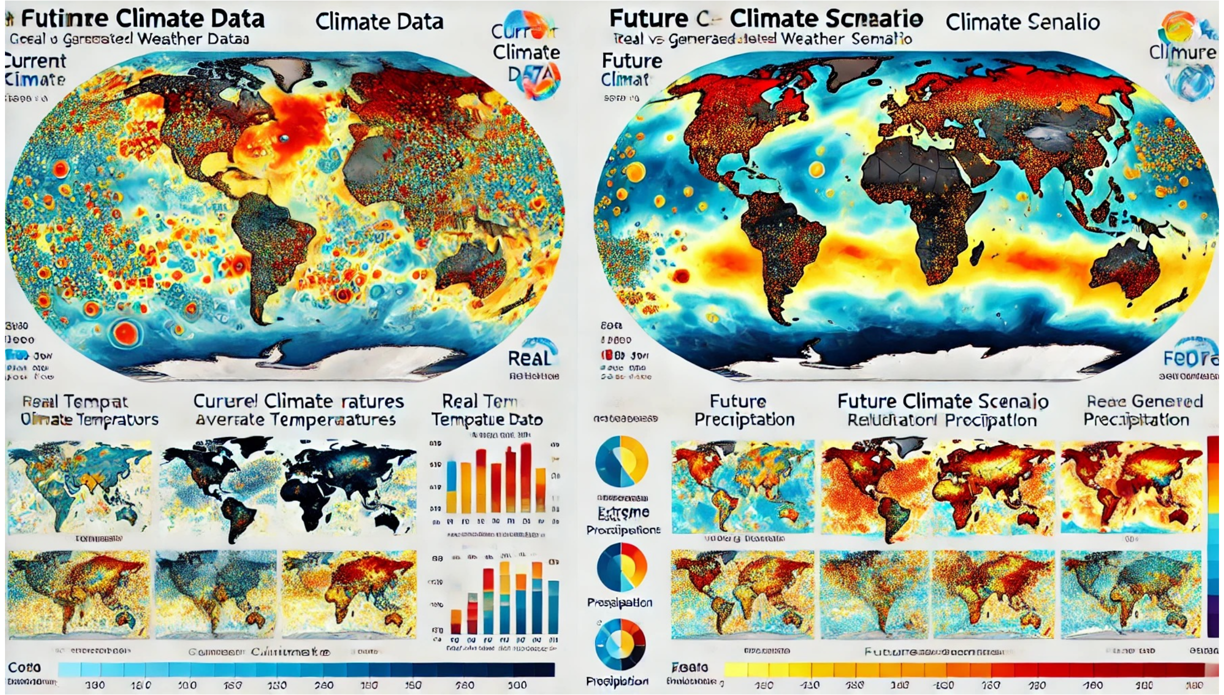
Data Needed:

- Comprehensive historical weather datasets.
- Climate models and projections on future conditions.
- Generated datasets representing different climate scenarios.

Data Bias:

The GAN-generated synthetic data may perpetuate biases present in historical datasets, resulting in unrealistic climate scenario simulations. Additionally, historical weather data may lack sufficient representation of extreme climate events, resulting in gaps in understanding future impacts and necessary adaptation strategies.

Generative Adversarial Networks (GANs) create synthetic datasets representing possible future climates, like increased temperature and extreme weather, which are then used to simulate various climate change scenarios. A linear regression model applied to both real and generated data can predict relationships between variables (e.g.: temperature rise and precipitation patterns), helping identify potential climate thresholds that could require adaptive strategies.



This image compares current climate data (left) with a synthetic future scenario generated by Generative Adversarial Networks (GANs) (right).
Note: This image was generated using ChatGPT as an illustrative example.

Scenario #1	Scenario #2	Scenario #3
✓ CNNs are highly effective for detecting visual patterns in satellite imagery, leading to accurate storm identification. ✓ Historical weather data integration improves predictive accuracy for classifying unusual weather events.	✓ Combining satellite imagery with historical data provides a comprehensive framework for accurate climate risk assessment across diverse regions. ✓ Real-time risk analysis using CNNs enables timely interventions in vulnerable areas.	✓ GANs allow for the generation of diverse synthetic datasets, exploring extreme climate scenarios that are not yet observable.
✗ The complexity of combining multiple models can lead to implementation challenges and high resource demands. ✗ Accessibility to high-resolution satellite imagery may limit the feasibility of the approach.	✗ Variability in data quality across regions can hinder model reliability significantly. ✗ The complexity of integrating different data types may complicate development and maintenance.	✗ The accuracy of synthetic data generated by GANs can be questionable, potentially undermining the reliability of simulations. ✗ The requirement for comprehensive historical datasets may limit feasibility if such data are not readily available. ✗ Training both real and synthetic data with RNNs adds complexity and requires substantial computational resources.

Conclusion:

Scenario #1 is strong in utilizing CNNs for accurate detection but faces feasibility issues due to the complexity of multiple models and data access.

Scenario #2 stands out for its robustness in risk assessment and real-time analysis, but it is limited by data variability across regions.

Scenario #3 offers innovative simulation capabilities through GANs but risks accuracy and feasibility if data quality is compromised.

Recommendation:

Scenario #2 appears the most feasible and comprehensive for assessing climate risks across Europe, balancing accuracy and implementation challenges effectively.

THANK YOU

rachelgrigiac@gmail.com