ClimateWins: Leveraging Machine Learning for Predicting and Understanding Climate Patterns

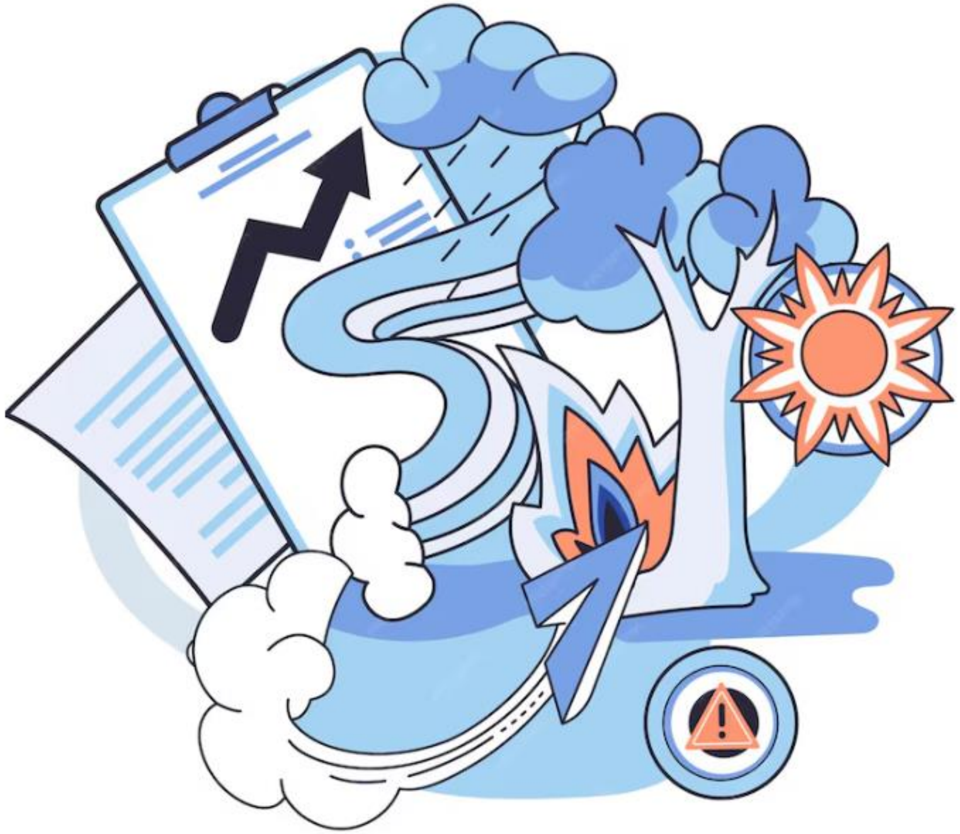
Asawer Maknoon
December 2024

AGENDA

- **Introduction:** Overview of project objectives and goals
- **Machine Learning Approaches:** Supervised, unsupervised, deep learning (RF, CNNs, RNNs and GANs)
- **Thought Experiment 1:** Predicting extreme weather events using historical data
- **Thought Experiment 2:** Image-Driven Weather Forecasting with Deep Learning
- **Thought Experiment 3:** Optimizing Weather Predictions with Synthesized Data using GANs
- **Recommendations & Next Steps:** Key insights and next steps

PROJECT OVERVIEW

ClimateWins is a small nonprofit organization with limited funds, offering hands on experience with advanced topics typically handled by data scientists.



ClimateWins aims to use machine learning (ML) to:

- Identify weather patterns outside regional norms in Europe.
- Assess if unusual weather patterns are increasing.
- Project future weather conditions over the next 20-50 years based on current trends.
- Determine the safest places to live in Europe for the next 25-50 years.

MACHINE LEARNING OPTIONS

Random Forest

- is an ensemble learning method that uses multiple decision trees to improve prediction accuracy and control overfitting.
- In weather and climate change prediction, it is used to analyze historical weather data, identify patterns, and predict future weather conditions and severe weather events. Its ability to handle large datasets and provide accurate predictions makes it a valuable tool in meteorology and climate science.

CNNs & RNNs

- are used in weather and climate change prediction to analyze spatial and temporal data, respectively.
- CNNs process data with a grid-like topology (e.g., satellite images), while RNNs handle sequential data (e.g., time series of temperature). Together, they improve the accuracy of weather forecasts and climate models by capturing complex patterns in large datasets.

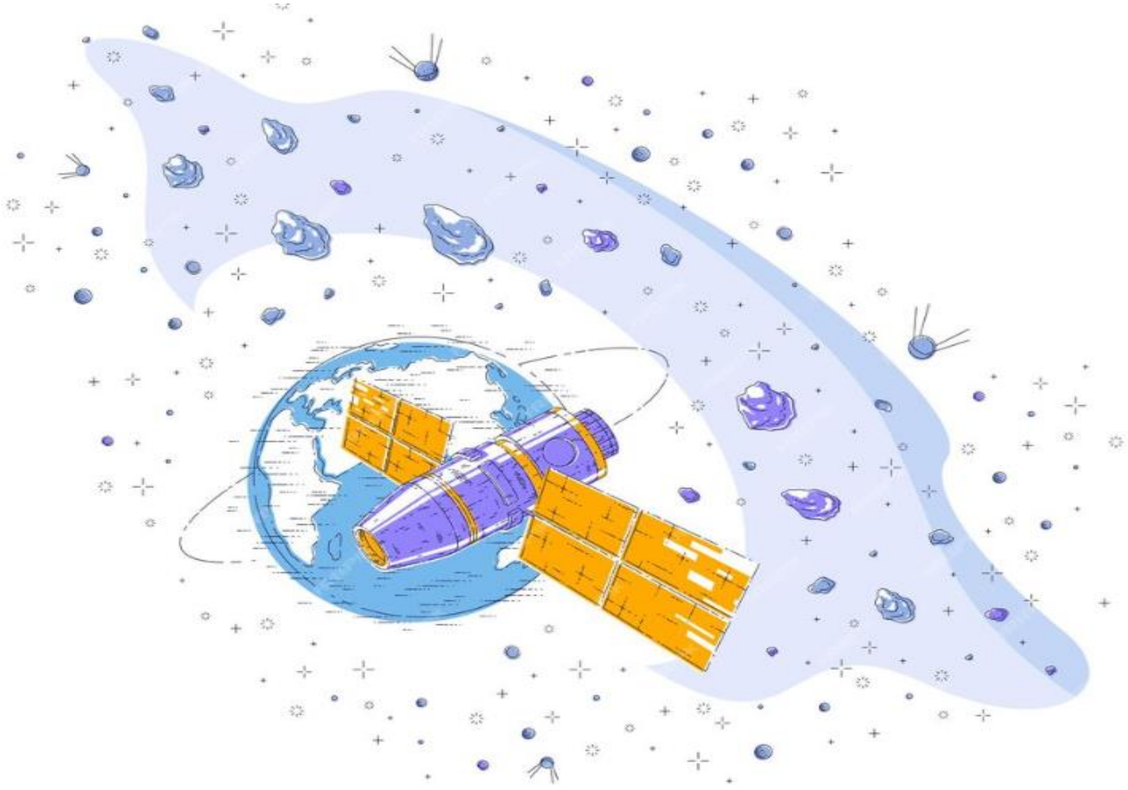
Generative Adversarial Networks (GANs)

- are used in weather and climate change prediction to generate realistic synthetic weather data by having two neural networks—the generator and the discriminator—compete against each other.
- This helps improve the accuracy of climate models and weather forecasts by augmenting real data with high-quality simulated data, capturing complex patterns and variability.

DATA REQUIREMENTS

Historical Weather Data

- Temperature, precipitation, wind speed, and humidity across Europe for past decades.
- Localized data from weather stations, satellites, and climate models.



Additional Data Needs

- **Geographical Data:** Terrain, elevation, population density, and urban/rural distinctions.
- **Socioeconomic Data:** Impact of extreme weather events on human health and infrastructure.
- **Climate Change Indicators:** Greenhouse gas emissions, sea-level rise, and ice cap melting data.
- **Remote Sensing Data:** Satellite imagery for detecting changes in vegetation, forest cover, and urban sprawl.

Thought Experiment 1: Predicting Extreme Weather Events Using Historical Data

Goal: Identify unusual weather patterns outside the regional norm in Europe.

Method:

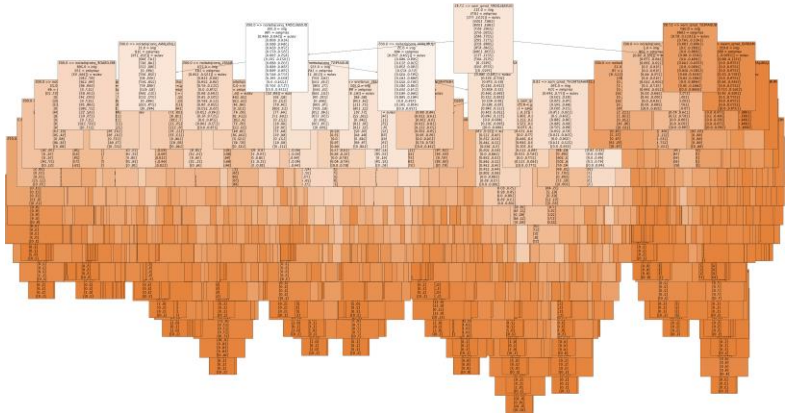
- Optimize the Random Forest model using Grid and Random Search.
- Compare the optimized model's performance to the unoptimized one with weather data analysis.

Data Required:

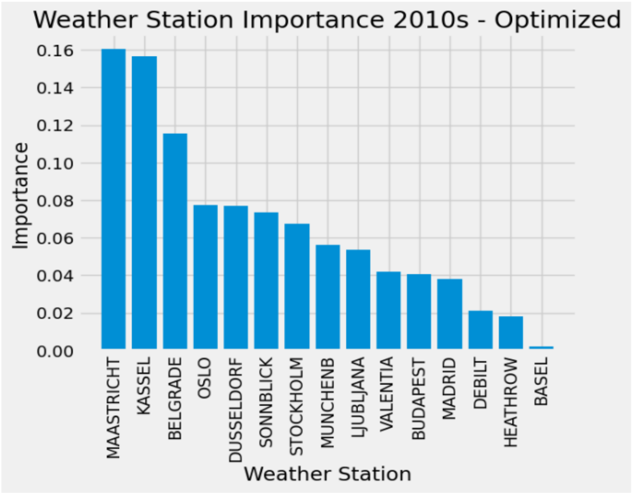
- Historical Weather Data for different European regions, including parameters like temperature, wind, rainfall, and pressure.
- Weather Event Databases: Records of historical extreme weather events, such as heatwaves, floods, and storms.

Results:

- After optimization, the accuracy improved from 60% to 70% across all stations.
- Maastricht, Kassel and Belgrade have the most influence on how the random forest divides up data.



Random Forest : Classifies regions based on climatic characteristics



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

Thought Experiment 2: Image-Driven Weather Forecasting with Deep Learning

Goal: Use deep learning models to analyze radar and satellite images for accurate weather prediction.

Methods:

- **CNNs:** Analyze satellite images to identify patterns linked to severe weather events.
- **RNNs:** Study time-series data to understand sequences leading to unusual weather events and enhance classification accuracy.

Data Required:

For CNNs:

- Satellite images capturing various weather phenomena.
- Visual data related to severe weather events like cloud formations, storms, etc.

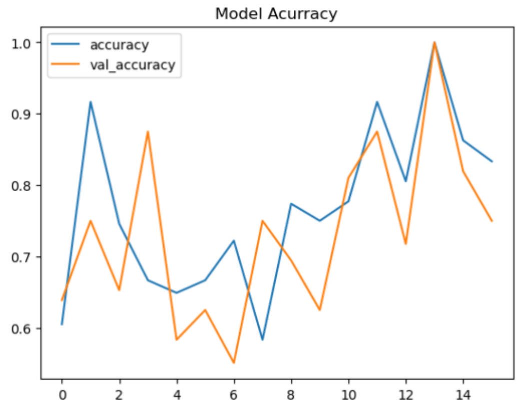
For RNNs:

- Historical time-series weather data, including temperature, humidity, wind speed, and atmospheric pressure.
- Sequential data that captures the progression of weather events over time.

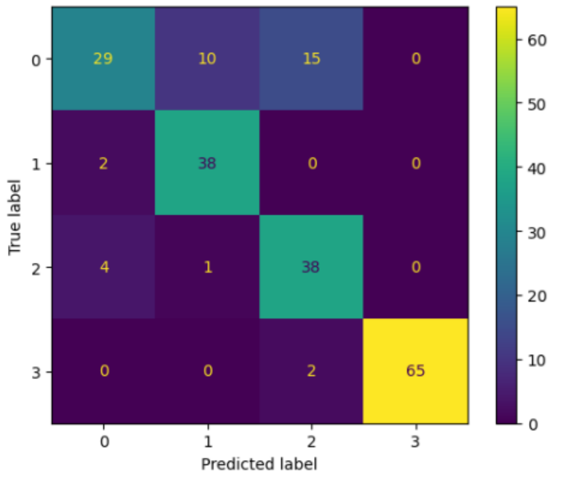
Results:

- **Training Accuracy:** 83.33% – The model performs well on training data.
- **Validation Accuracy:** 75% – Slightly worse on unseen data, indicating potential overfitting.
- **Training Loss:** 0.0473 – Low loss, indicating a good fit on training data.
- **Validation Loss:** 0.0901 – Higher than training loss, suggesting some overfitting, but the gap is not large.

Model loss over training epochs



Model accuracy over training epochs



Confusion matrix illustrating the model's prediction accuracy

Thought Experiment 3: Optimizing Weather Predictions with Synthesized Data using GANs

Goal: Enhance weather predictions and understand climate change impacts using GANs.

Methods:

- **Deep Learning Models:** Use GANs to generate realistic climate scenarios.

- **Climate Simulation:** Combine GANs with global climate models for refined regional predictions.

Data Needed:

- Historical Climate Data: Long-term weather data.
- Satellite Images: Visual data of weather phenomena.
- Climate Reports: IPCC and other global projections.

Results:

This approach not only provides valuable insights into the impacts of climate change but also equips stakeholders with the knowledge needed to plan effectively for extreme weather scenarios.

An example for a generative AI model visualizes how floods in Texas would look like in satellite imagery. The original photo is on the left, and the AI generated image is in on the right.

Credit: Pre-flood images from Maxar Open Data Program via Gupta et al., CVPR Workshop Proceedings. Generated images from Lütjen et al., IEEE TGRS.

RECOMMENDATIONS & NEXT STEPS

RECOMMENDATIONS

- ✓ Implement the optimized Random Forest model to identify unusual weather patterns more accurately.
- ✓ Integrate CNNs and RNNs into the existing weather prediction systems to enhance overall accuracy.
- ✓ Continue leveraging GANs to simulate realistic climate scenarios and study their impacts.

NEXT STEPS

- ✓ Combine data from all models to create a comprehensive tool for weather and climate prediction.
- ✓ Engage with stakeholders, including meteorologists, climate scientists, and policymakers, to implement the models effectively.
- ✓ Regularly update and validate models with new data to maintain accuracy and relevance.

THANK YOU

Asawer Maknoon
November, 2024