


Navigating Climate Change with Machine Learning


02.10.2024

GIADA GRISO

AGENDA

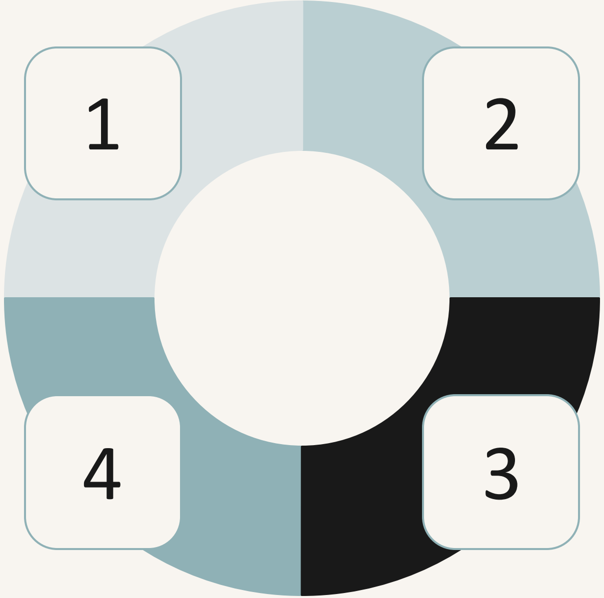
01 PROJECT OBJECTIVES

IDENTIFY

Unusual weather patterns in Europe.

DETERMINE

If such patterns are increasing.



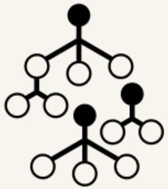
FIND

Safe places to live in Europe based on those trends.

PREDICT

Future weather trends (25 – 50 years).

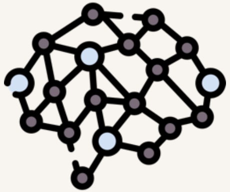
02 ML ALGORITHMS OVERVIEW



RANDOM FORESTS

How it works:
Multiple decision trees working together to make more accurate predictions by averaging the results from all the trees.

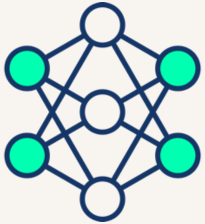
Applications:
Identify and predict key weather features at specific locations across Europe, recognise anomalies in weather patterns.



CNNs & RNNs – DEEP LEARNING

How it works:
They automatically learn to detect important patterns, like spatial or temporal features, through layers of convolutional filters.

Applications:
Recognise complex climate patterns and evolving trends, forecast long-term weather changes.



GANs

How it works:
They create synthetic data by training two neural networks (generator + discriminator) against each other.

Applications:
Simulate future climate changes, predict extreme weather events, find the safest places to live in Europe.

03 THOUGHT EXPERIMENTS

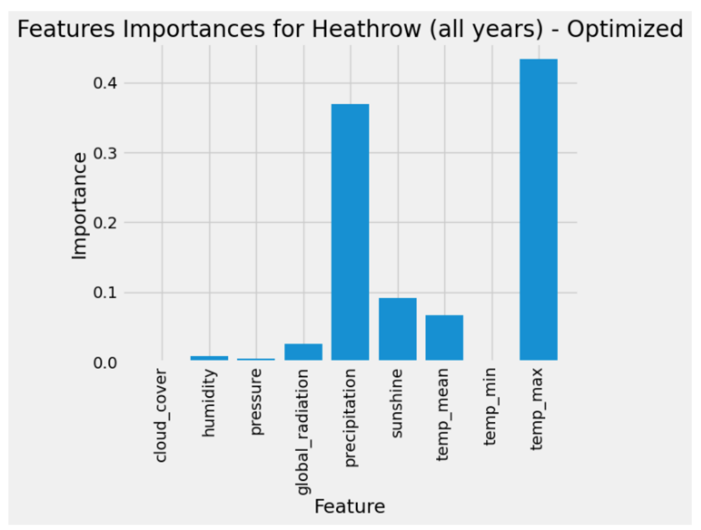
Scenario 1: Anomaly Detection for Unusual Weather Patterns

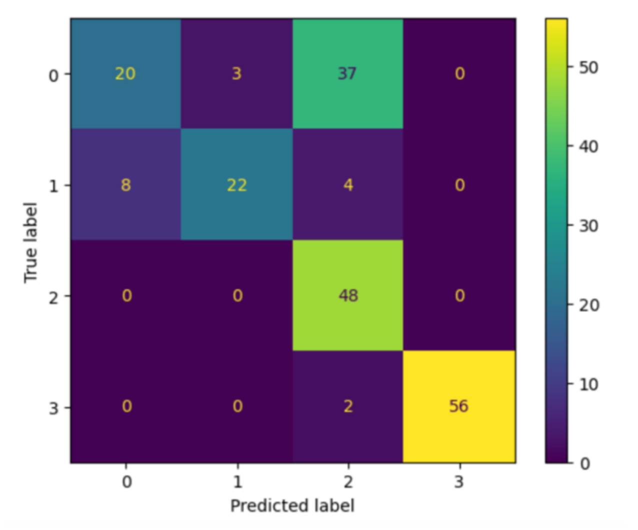
Goal: Combine **Random Forest** and **CNN** models to identify unusual weather patterns by detecting anomalies in both structured weather station data and unstructured satellite imagery.

Approach: Use the feature importance scores generated by the Random Forest to identify which weather factors (e.g., temperature, precipitation) are the primary drivers of anomalies. Use CNNs to analyse satellite images for identifying spatial anomalies like unexpected cloud formations or unusual precipitation patterns. Track anomalies over time to detect increasing trends in unusual weather patterns.

Data: historical weather data, satellite imagery.

1. How could this be applied practically?





- **Feature importance scores** obtained by training a random forest to distinguish between 'pleasant' and 'unpleasant' weather.
- **Accuracy rate** after optimization: **100%**.

- **Confusion matrix** derived from a CNN model model categorizing 4 different weather conditions (cloudy, rain, sunshine and sunrise).
- **Accuracy rate: 75%**.

Scenario 2: Future Climate Scenario Simulation

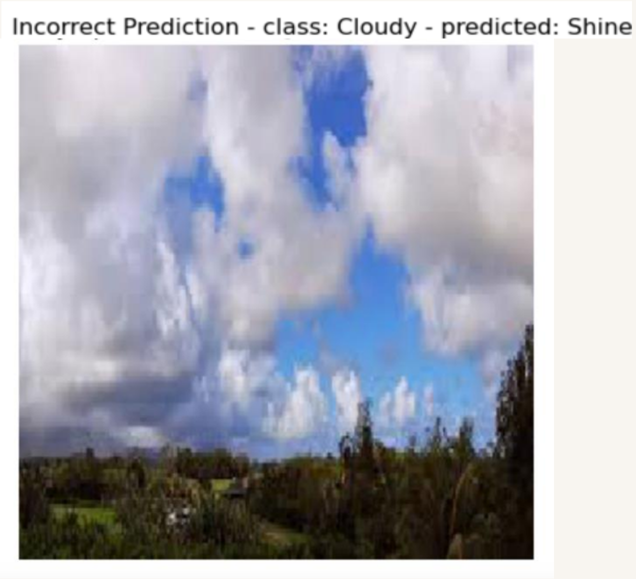
Goal: Use **GANs** to generate possible future climate scenarios and **CNNs** to validate and classify the severity of these scenarios.

Approach: Train GANs on historical weather data and satellite imagery to create synthetic future climate data (either numerical or visual), which represents plausible climate conditions for the next 25 to 50 years. The CNNs can then classify these simulated trends into categories like "mild", "severe", or "extreme" based on factors such as temperature increases or precipitation anomalies, helping to assess potential risks over time.

Data: historical weather data, satellite imagery, climate projections (models that that predict how the climate will respond to various factors, like CO_2 concentrations)

2. How could this be applied practically?

- ChatGPT simulated satellite image of Europe, produced by **GANs under the RCP 8.5 scenario,** showing potential extreme weather patterns.



- Example of incorrect prediction made by a **CNN** model model categorizing 4 different weather conditions.
- **Accuracy rate: 75%**.

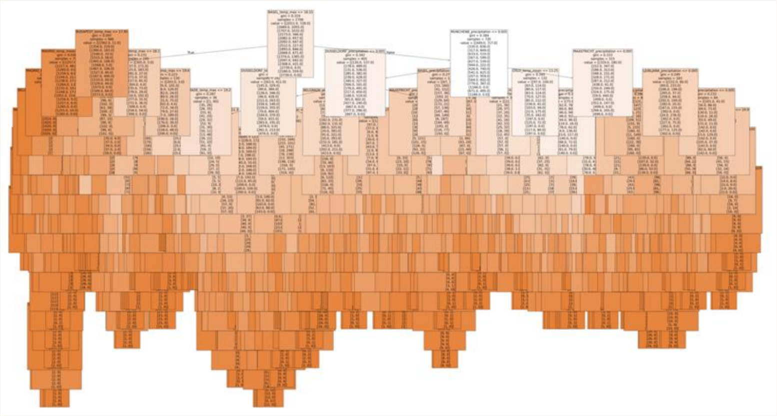
Scenario 3: Safe Areas Identification

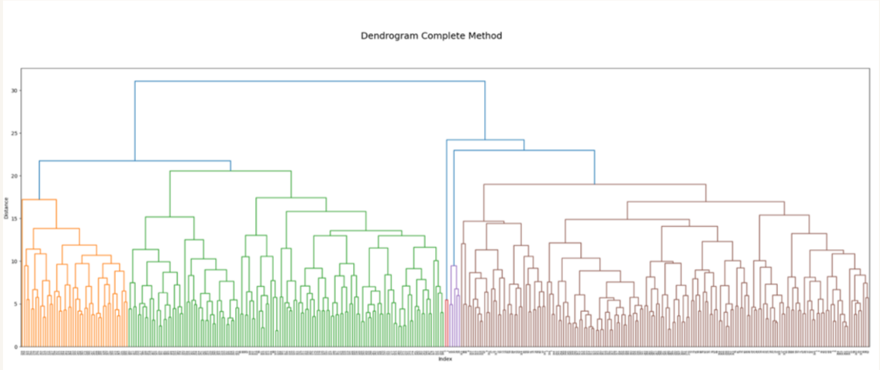
Goal: Identify the safest places to live in Europe in the next 25– 50 years by **clustering** regions based on weather patterns and using **Random Forest** to evaluate key factors like infrastructure resilience and historical weather conditions.

Approach: K-Means algorithm clusters regions across Europe based on weather data (e.g., temperature, precipitation) and climate risk factors (e.g., flooding potential). Each cluster is labelled with a safety score. Random Forest can then be used to predict which clusters are likely to remain stable and safe over time. PCA can be applied to reduce data complexity.

Data: historical and current weather data, geospatial data, infrastructure resilience data, demographic data, natural disasters data.

3. How could this be applied practically?





- **Random Forest** for 15 European weather stations categorizing a decade of pleasant and unpleasant weather data points.
- **Accuracy rate** after optimization**: 66,3%**.

- **Dendrogram Complete Method** after performing **PCA** to reduce analysis components.
- Clearer, more distinct clusters with less noise and overlap, indicating that the PCA has successfully highlighted the key patterns in the data.

04 SOCIAL & ETHICAL CONSIDERATIONS

EQUITY & CLIMATE JUSTICE

Across all experiments, the outputs must result in **equal access to safety, protection, or resources**.
ClimateWins should ensure that the most vulnerable communities are **included** in the models and benefit from the insights generated to avoid enhancing existing inequalities.

TRANSPARENCY & TRUST

Clear communication of how ML models work, their limitations, and the rationale behind their predictions is essential for **public trust** and **acceptance**.

RESPONSIBILITY & ACCOUNTABILITY

ClimateWins must be **accountable** for model outcomes and **adjust strategies** if negative impacts occur.
All models require ethical oversight and **continuous evaluation** to ensure responsible application and adaptability to feedback from impacted communities.

05 RESULTS & SUMMARY

Scenario #1 (Anomaly Detection)

PROS:
- Combines multiple data sources
- High accuracy
- Potential for real-time monitoring
- Transparency

CONS:
- System complexity
- Computational costs
- Overfitting risk (CNN)

Scenario #2 (Future Climate Scenario Simulation)

PROS:
- Forward-looking
- Versatile
- Multiple scenario exploration
- Risk assessment aid

CONS:
- Black box nature
- High computational demand
- Data dependency (biases)

Scenario #3 (Safe Areas Identification)

PROS:
- Transparent results
- Scalable
- Multi-factor analysis (holistic view of safety)
- Actionable insights

CONS:
- Limited focus
- Uncertain future conditions
- Tricky clusters interpretation

Which thought experiment has the most potential for ClimateWins?

Scenario #1 (Anomaly Detection) is the most achievable

- ✓ Valuable for detecting emerging climate anomalies in real-time, providing early warnings for short- to medium-term climate risks.
- ✓ High accuracy, transparency and lower computational costs compared to other scenarios.

Scenario #2 (Future Climate Scenario Simulation) is the most ambitious

- ✓ Complex future modelling, technically challenging (GANs) and high computational costs.

Scenario #3 (Safe Areas Identification) has the most political implications

- ✓ Highest influence on policy (e.g., regional investment, resource allocation), migration and infrastructure development.
- ✓ Profound political and societal consequences.
- ✓ Long-term uncertainty

06 FINAL RECOMMENDATIONS

1. Begin by deploying **Scenario #1 (Anomaly Detection)** as it is the most achievable and practical in the short term.

2. Allocate resources to develop the necessary infrastructure and expertise to develop **Scenario #2 (Future Climate Scenario Simulation)**. This one should be prioritized for long-term forecasting and strategic climate planning.

3. Integrate **Scenario #1** and **#2** for a more balanced climate strategy (short- and long-term).

4. Approach **Scenario #3 (Safe Areas Identification)** with careful stakeholder engagement and policy considerations, ensuring that the designation of "safe" and "at-risk" zones does not exacerbate social inequalities or fuel political tension.

THANKS!

DO YOU HAVE ANY QUESTIONS?

CREDITS: This presentation template was created by **Slidesgo**, and includes icons by **Flaticon** and infographics & images by **Freepik**