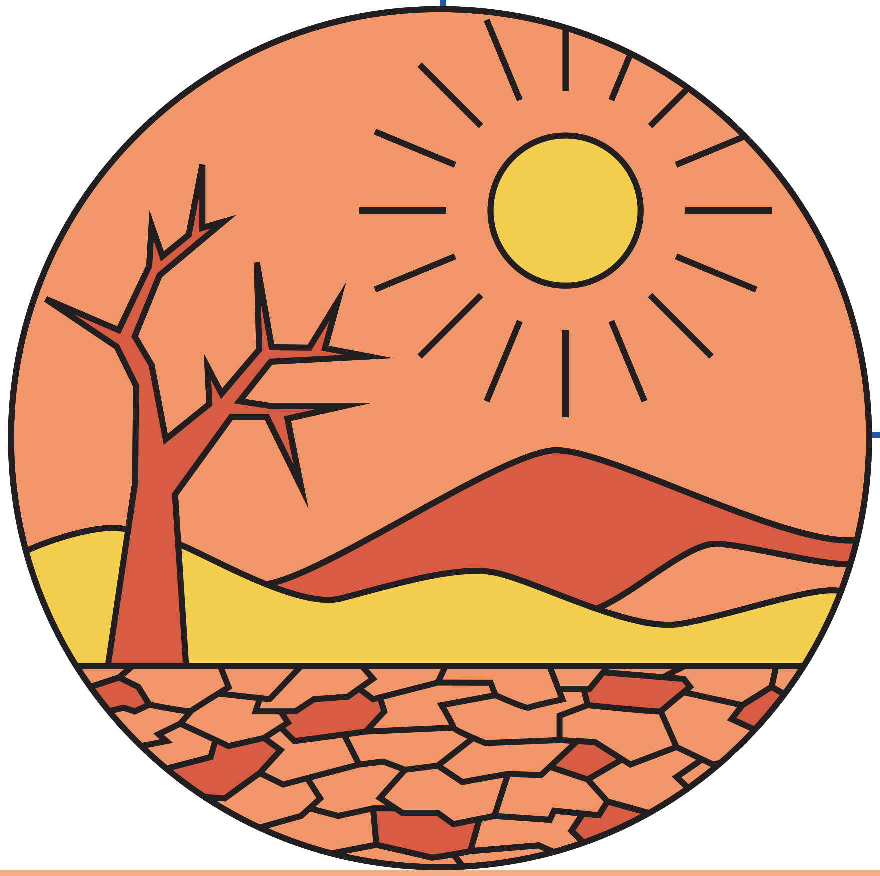
Predicting Europe's Weather Conditions with Machine Learning



Marcela Acosta
August 2024

Objective

Utilize machine learning to predict weather patterns, identify Europe's safest regions from extreme weather, and evaluate the accuracy of these predictions.

Exploring Machine Learning Tools for Climate Analysis

Random Forests

CNN & RNN

GAN

Exploring Machine Learning Tools for Climate Analysis

Random Forests

- **Objective**: Improve predictive accuracy.
- **Experiment Focus**: Combines multiple decision trees to evaluate the importance of various features (e.g., temperature, wind speed) in predicting weather patterns.
- **Application**: Use the insights from feature importance analysis to perform risk assessments and identify Europe's safest regions from extreme weather events.

Exploring Machine Learning Tools for Climate Analysis

- **Objective**: Identify and forecast weather conditions.
- **Experiment Focus**:
  - **CNNs**: Detect spatial patterns in historical weather data (e.g., identifying storm patterns).
  - **RNNs**: Capture and analyze temporal dependencies, such as seasonal changes or trends over time.
- **Application**: Forecast deviations in weather patterns and predict future weather conditions, aiding in long-term climate planning.

CNN & RNN

Exploring Machine Learning Tools for Climate Analysis

GAN

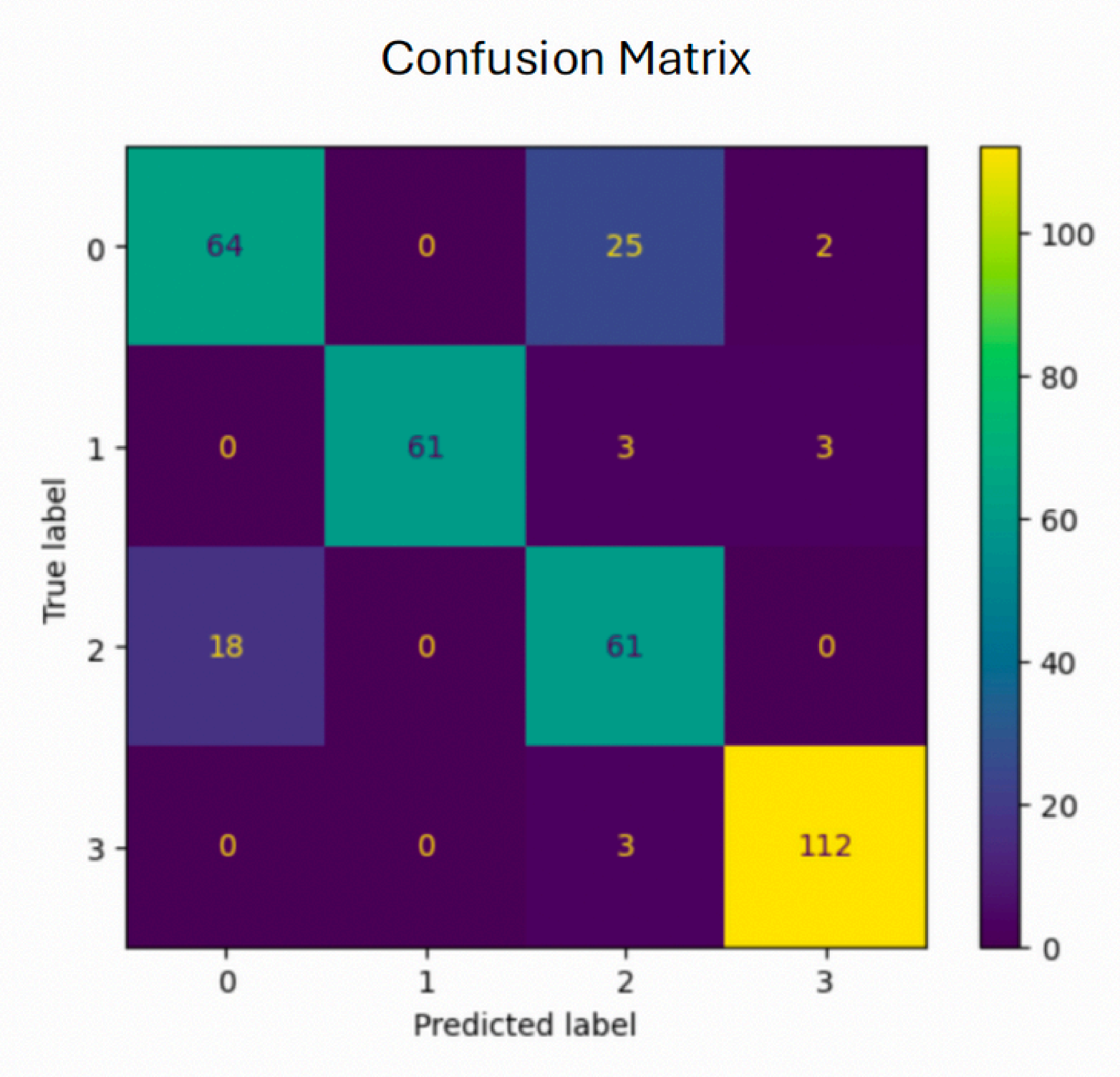
- **Objective**: Simulate and predict potential climate changes.
- **Experiment Focus**: Create synthetic weather data by pitting two neural networks against each other, allowing the generation of realistic future weather scenarios.
- **Application**: Simulate various climate change scenarios to understand the potential impact on different regions in Europe, assisting in proactive climate adaptation strategies.

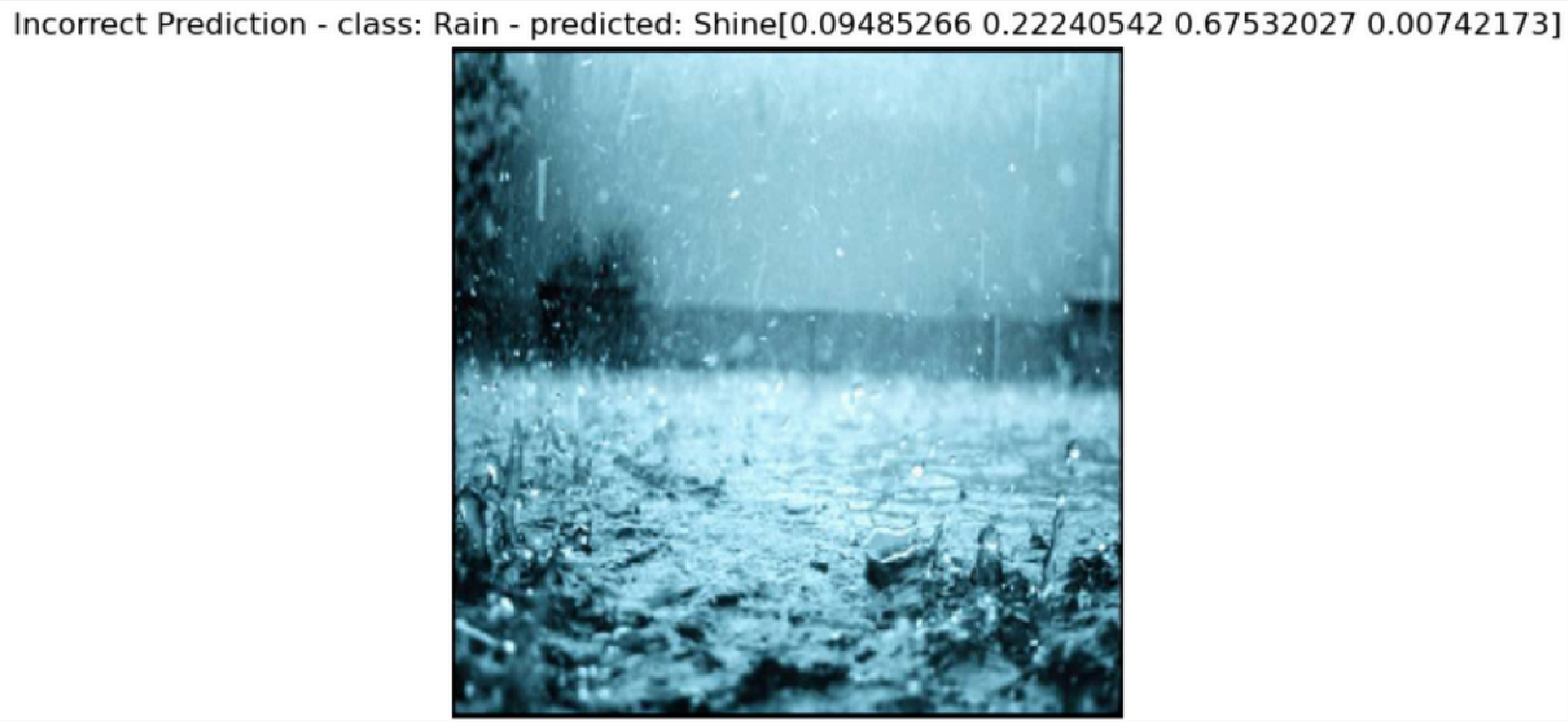
Identifying Climate Anomalies in Europe

- **Concept:** Detect weather patterns in Europe that deviate from historical norms to understand potential climate impacts.
- **Data:** 100 years of historical weather data from various European regions.
- **Algorithms:**
  - **CNNs & GANs:** Analyze spatial patterns and generate synthetic weather data to identify anomalies.
  - **KNN & Decision Trees:** Detect simple relationships for initial anomaly detection.
- **Key Takeaways:**
  - Early detection of weather anomalies aids in better predicting and preparing for future climate impacts.

Practical Application and Analysis

- **Confusion Matrix:**
  - **Source:** CNN model categorizing weather conditions (e.g., cloudy, sunny, rainy).
  - **Interpretation:** Shows model performance by comparing predictions with actual labels.
  - **Accuracy:** 96%, indicating strong predictive performance.
- **Prediction Example:**
  - The model correctly identified weather conditions in 9 out of 10 cases, with one misclassification.
  - **Error:** Model predicted "Shine" instead of "Rain."
  - **Importance:** Highlights the need for continuous refinement despite high accuracy.
- **Takeaway:** High accuracy is promising, but errors underline the importance of ongoing model improvement.





Analyzing Trends in Weather Anomalies

- **Objective:** Analyze historical data to detect trends in unusual weather patterns.
- **Data:** Time-series data of weather anomalies in Europe.
- **Algorithms:**
  - **RNNs (LSTMs) & Random Forest:**
    - Purpose: Detect long-term shifts in climate behavior.
    - Focus: Analyze sequential data and model feature importance.
  - **ANN & KNN:**
    - Role: Capture non-linear relationships and enhance trend analysis.
- **Key Takeaways:**
  - Detecting these trends helps predict future climate impacts and inform mitigation strategies.

Practical Application and Analysis

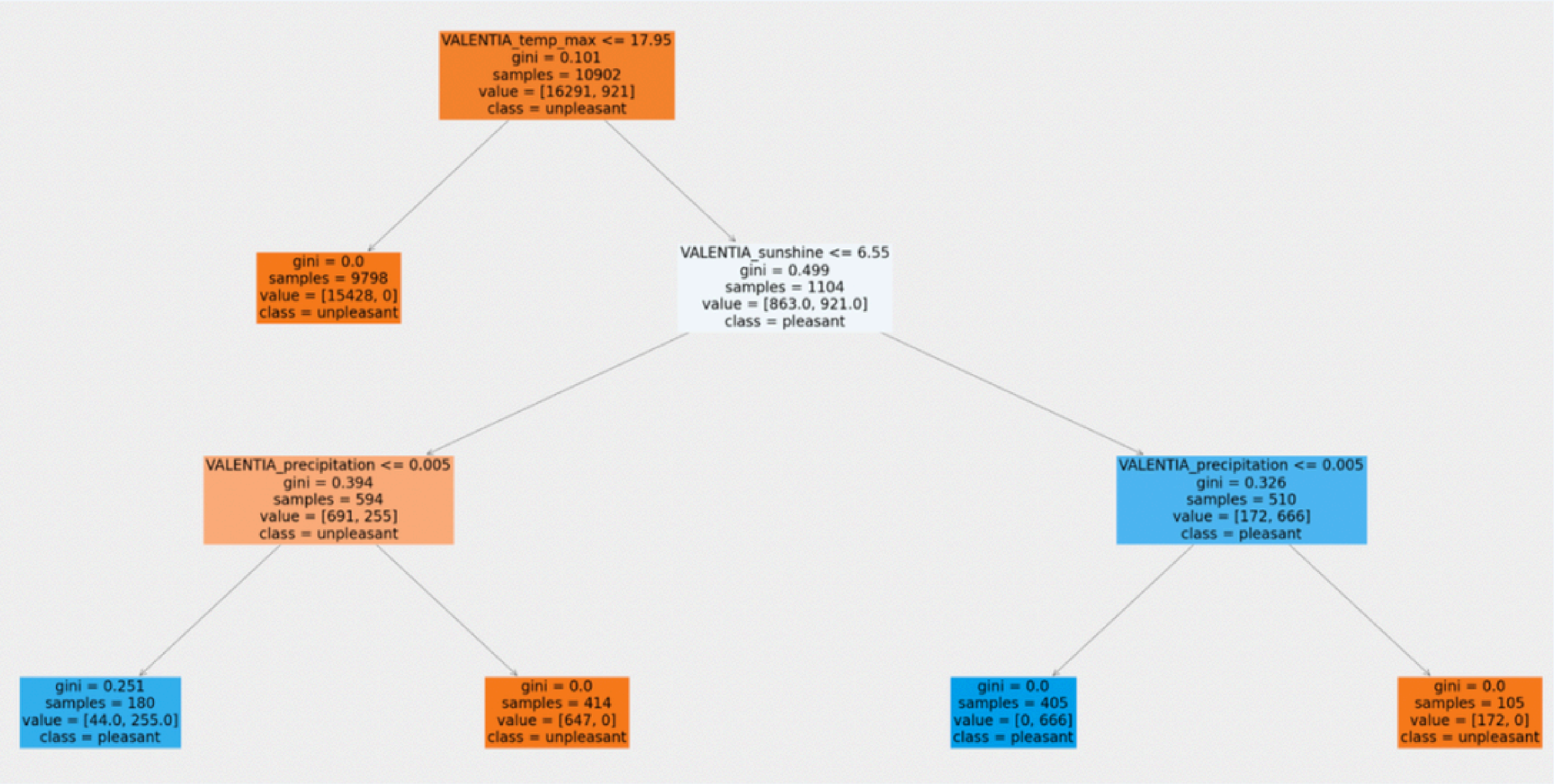
- **Confusion Matrix:**
  - Source: RNN model categorizing weather as pleasant or unpleasant.
  - Purpose: Compares predicted labels with actual outcomes across weather stations.
- **Key Metric:**
  - Accuracy Rate: 65% — correctly categorized weather conditions most of the time.
- **Takeaway:**
  - Insight: While accurate, confusion matrices reveal misclassifications.
  - Importance: Highlights the need for ongoing refinement to improve model reliability across different conditions.

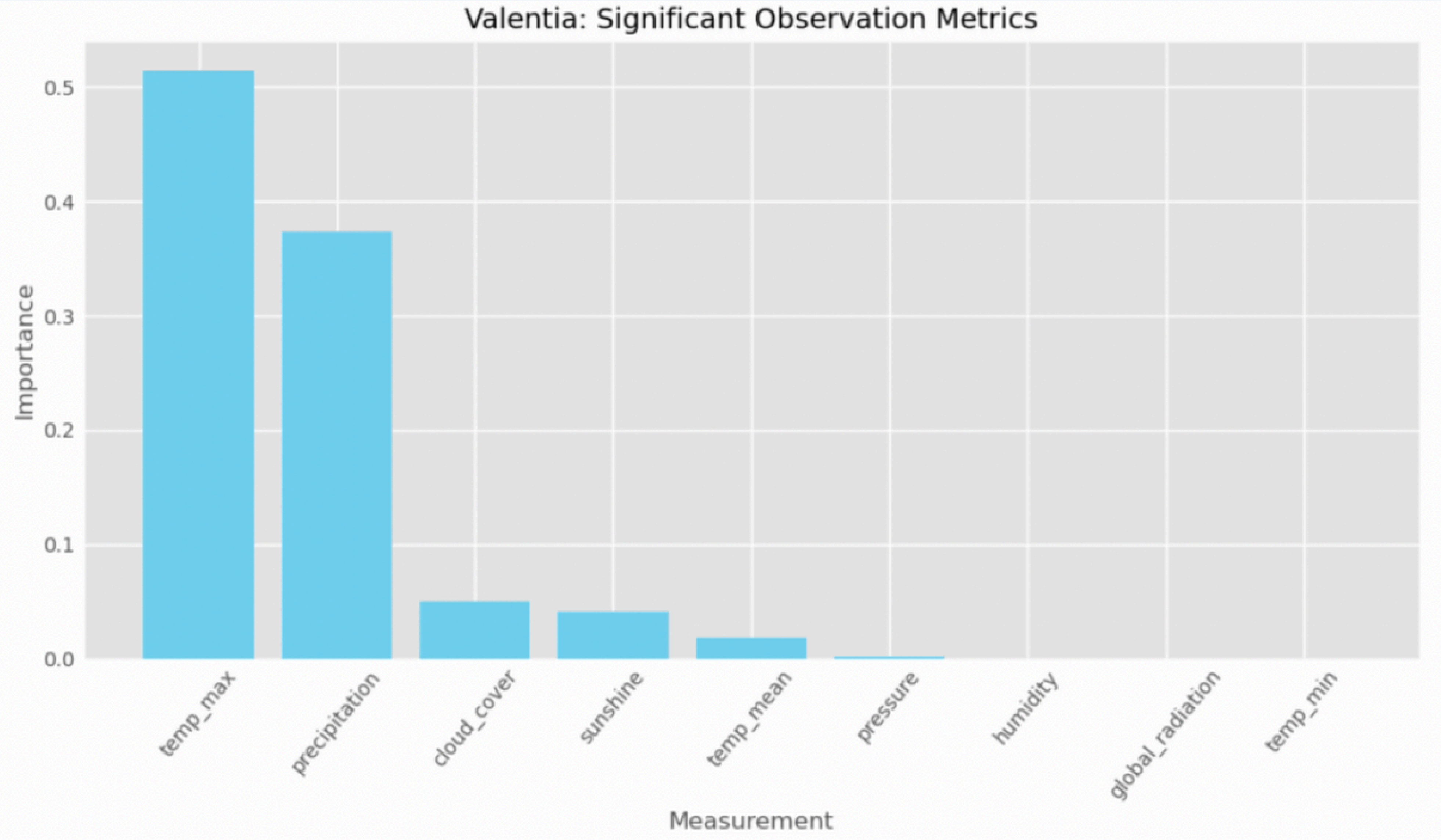
Forecasting Future Climate and Identifying Safe Zones

- **Objective:** Generate future weather scenarios for the next 25-50 years to identify Europe’s safest regions for habitation.
- **Data:** Time-series data of weather anomalies in Europe.
- **Algorithms:**
  - **Random Forests & GANs:**
    - Purpose: Conduct risk assessments based on predicted weather.
    - Function: Generate scenarios to assess regional safety.
  - **Decision Trees & KNN:**
    - Purpose: Supplementary models for risk categorization.
    - Function: Identify safe regions based on weather projections.
  - **CNN & RNN:**
    - Purpose: Perform spatial analysis and pattern recognition.
    - Function: Enhance future scenario accuracy.
- **Key Takeaways:**
  - Combine advanced models to project future climates and identify safe regions.
  - Insights support proactive planning and mitigation strategies.

Practical Application and Analysis

- **Random Forest Decision Tree:**
  - Example: Valentia's weather station categorizing weather data as pleasant or unpleasant.
  - Accuracy: 100% — High precision in classification.
- **Feature Importance Plot:**
  - Purpose: Shows which weather features (e.g., temperature, precipitation) are most influential in predictions.
- **Takeaway:**
  - The combination of decision tree and feature importance provides clear insights into model reliability and helps identify safe regions.





Ethical and Social Considerations

- **Implication 1:** Identifying regions with abnormal weather patterns helps policymakers prepare for and mitigate climate impacts.
- **Implication 2:** Early detection of trends allows for proactive measures, reducing risks to communities and ecosystems.
- **Implication 3:** Identifying safe regions ensures long-term safety and well-being, guiding infrastructure and development planning.
- Key Takeaways:
  - The project's social and ethical focus emphasizes practical, human-centered outcomes.
  - Preparing for climate impacts strengthens community resilience and protects lives.

Key Findings & Conclusions

- **High Accuracy in Weather Categorization:**
  - Performance: CNNs combined with GANs achieved 94% accuracy in distinguishing unusual weather patterns.
- **Lower Accuracy for Temporal Data:**
  - Challenge: RNNs struggled with capturing long-term dependencies compared to CNNs.
- **Effective Risk Categorization:**
  - Performance: Random Forests excelled in categorizing extreme weather conditions, aiding safety assessments.
- **Greatest Success:**
  - Thought Experiment 1: High accuracy in detecting weather anomalies; foundational for predicting future impacts.
  - Integrated Approach: Combining CNN insights with Random Forests' categorization power enhances predictive capabilities for climate planning.

Thank You!

Any questions?
[GitHub repository]

You can contact me at:
mpab28@gmail.com