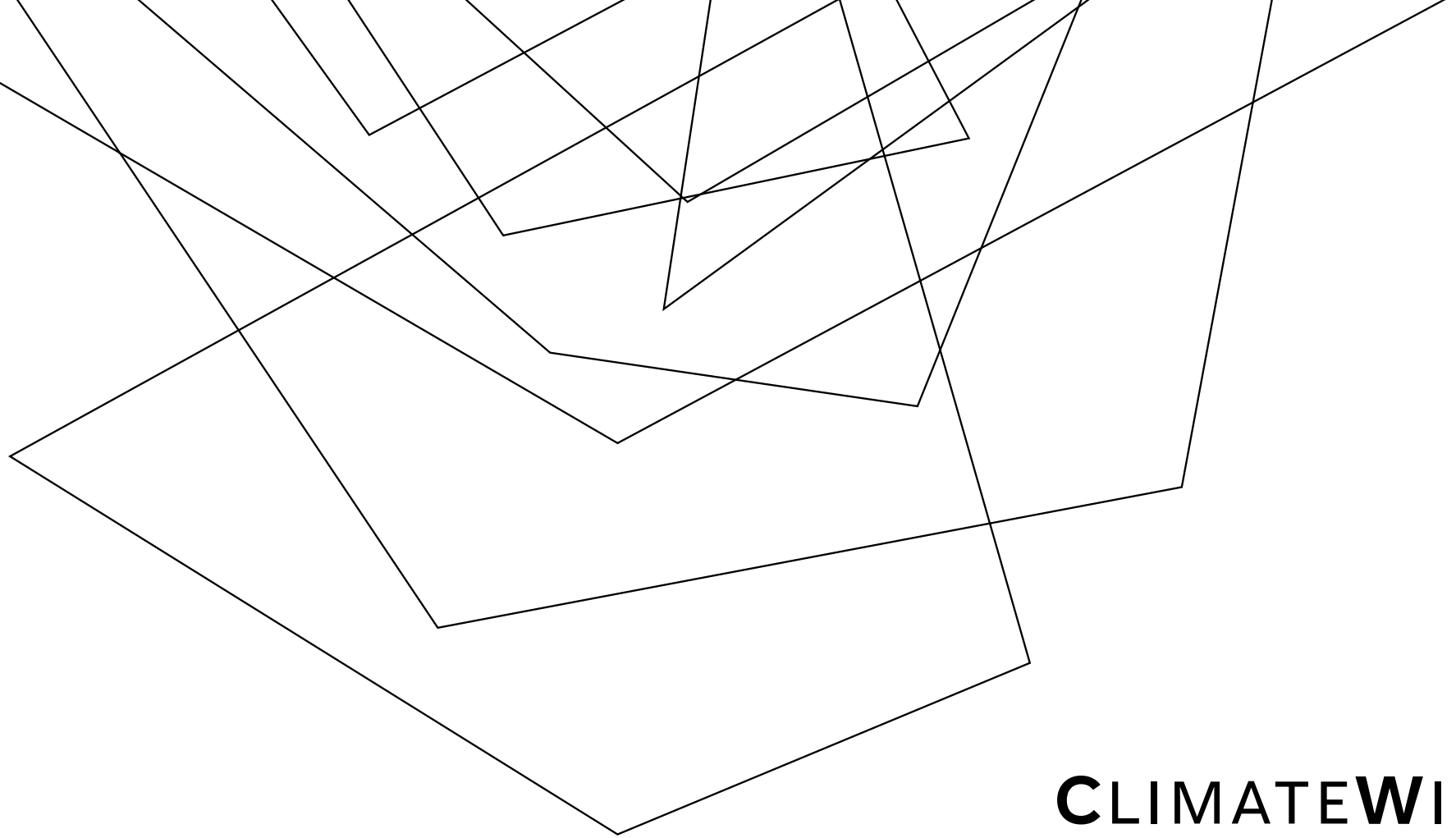


CLIMATEWINS

PREDICTING CLIMATE CHANGE WITH ML

RATUL DEBNATH

OCTOBER 25, 2025

AGENDA

Project Goals

Proposed Models for ML

Models used for ML

- Dendrograms
- Random Forest
- Visual

Summary

PROJECT GOALS

ClimateWins wants to assess what tools are available to categorize and predict the weather in mainland Europe. It's concerned with the increase in extreme weather events, especially in the past 10 to 20 years.

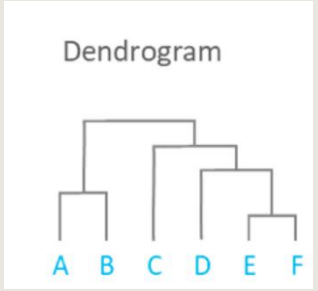
ClimateWins has a few areas it wants to cover:

1. Finding new patterns in weather changes over the last 60 years.
2. Identifying weather patterns outside the regional norm in Europe.
3. Determining whether unusual weather patterns are increasing.
4. Generating possibilities for future weather conditions over the next 25 to 50 years based on current trends.
5. Determining the safest places for people to live in Europe within the next 25 to 50 years.

PROPOSED MODELS FOR ML

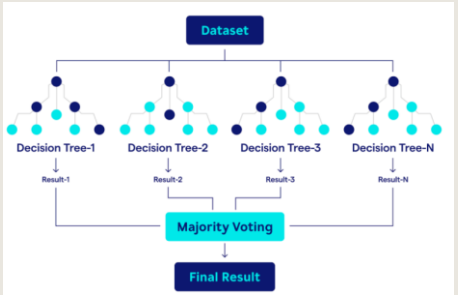
Dendrograms

Dendrograms are a powerful tool to group similar patterns together. This is extremely useful for creating segmented date for better accuracy as seasons and location are a big factor in reducing range and outliers



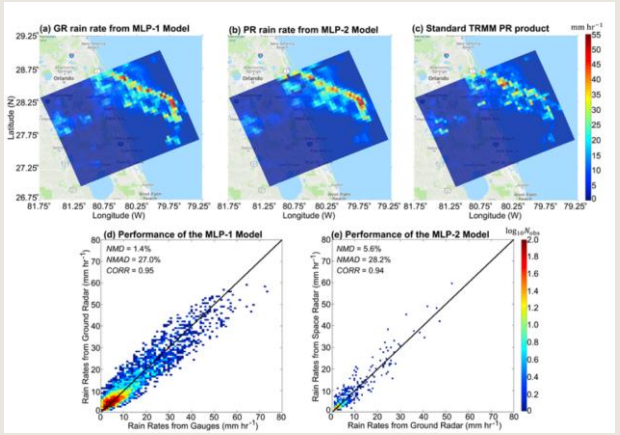
Random Forest

Random Forest is an ensemble learning method widely used in machine learning for both classification and regression tasks. It operates by constructing a multitude of decision trees during training and outputting the class that is the mode of the classes (for classification) or mean prediction (for regression) of the individual trees.



Visual ML Application

This is an extremely useful tool to predict short term changes in weather that is sometimes harder than predicting long term climate changes by utilizing hardware that already exist like thermal, surveillance and satellite cameras.



DENDROGRAMS

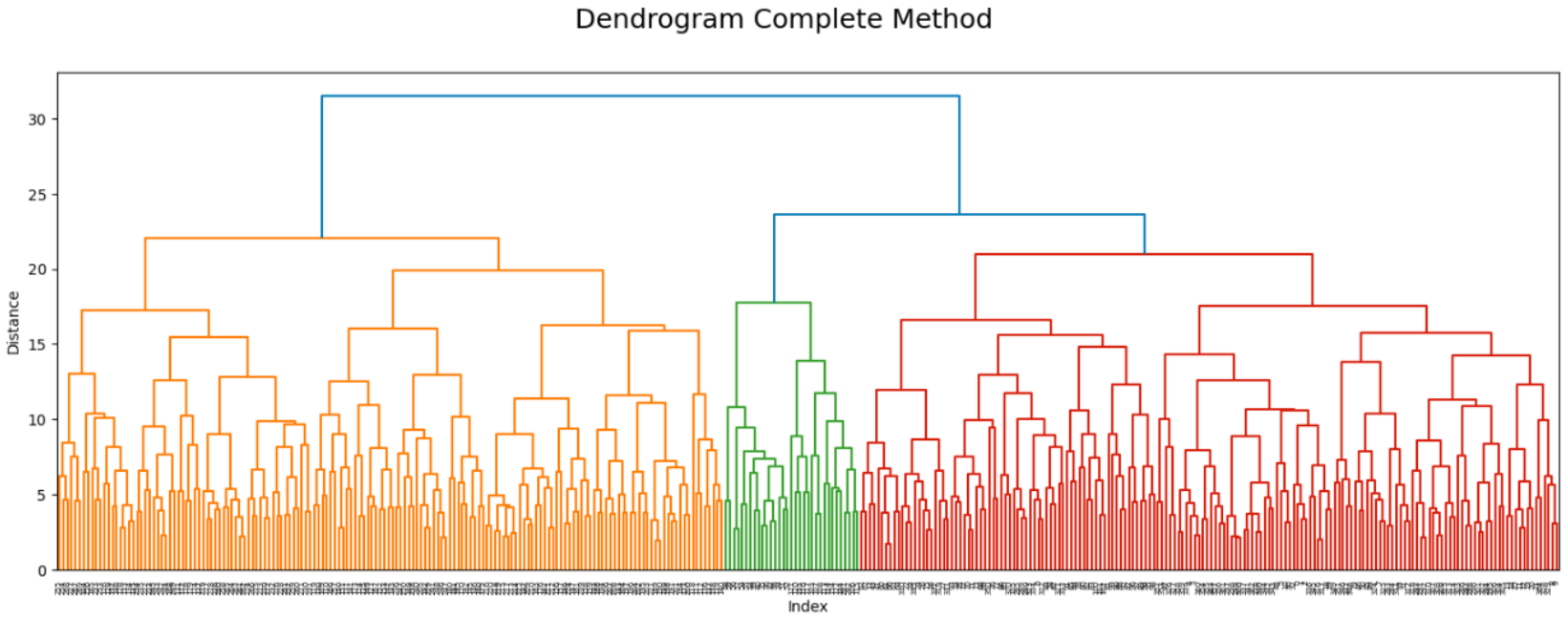
The method for calculating the distance between clusters is a critical component of the hierarchical clustering algorithm and affects the shape of the resulting dendrogram. The most common types of linkage criteria include:

- Single-linkage: Uses the distance between the two closest members of different clusters.

- Complete-linkage: Uses the distance between the two farthest members of different clusters.

- Average-linkage: Uses the average distance between all pairs of members in different clusters.

- Ward's method: Minimizes the total within-cluster variance when merging two clusters.



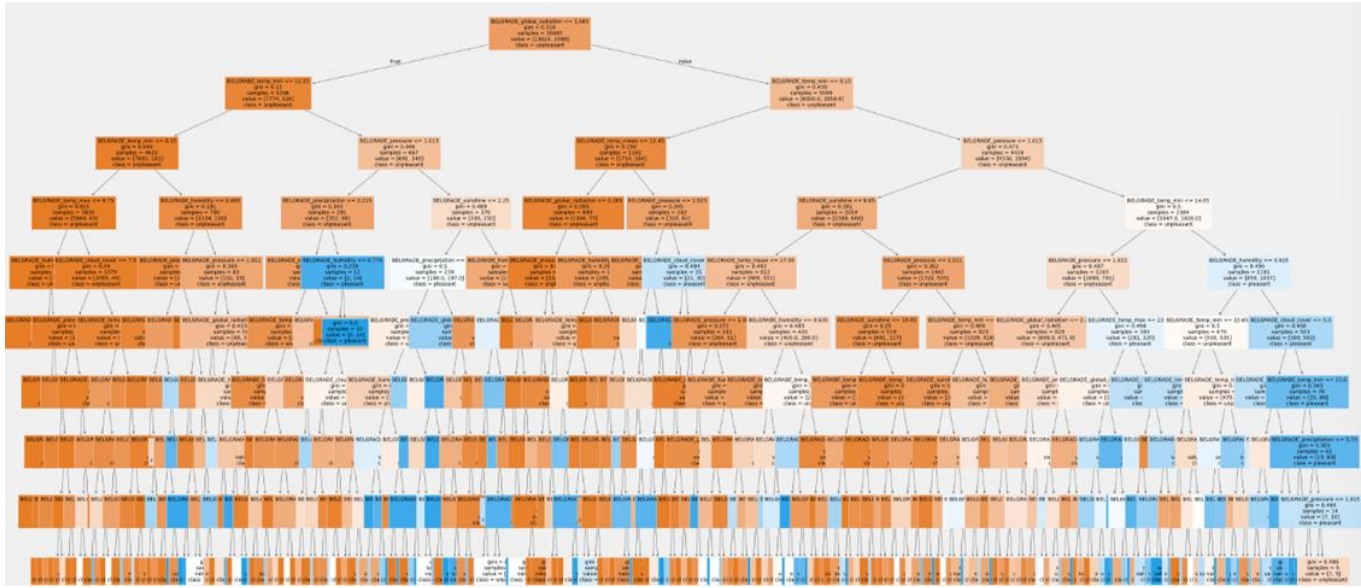
Observations with weather data:

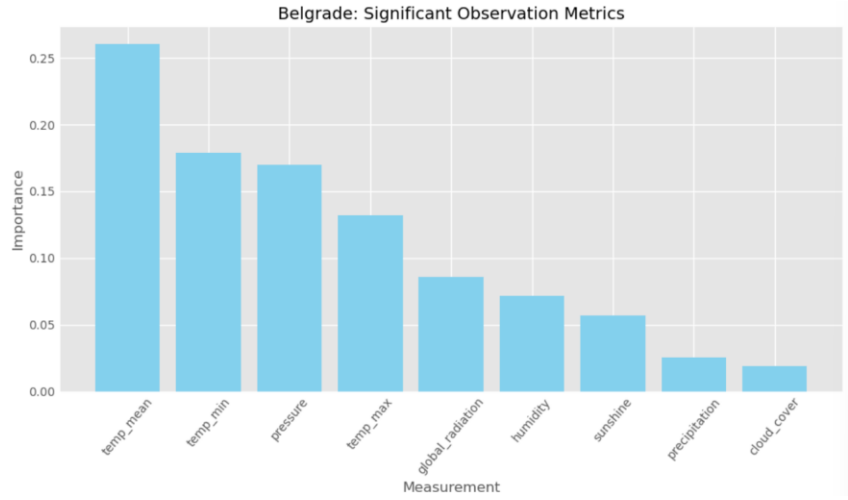
Single Method is not very useful as it makes one cluster dominate over the others.
Complete Method shows a good mix of clusters, which should be the case due to seasonal differences as well as differences between Madrid and Belgrade.
Average Method: Mostly shows three distinct categories. Possibly warm weather, cold weather and pleasant days cross over in the middle
Ward Method: This shows two clear distinct categories in both datasets. This along with complete method I think are the most valuable in this dataset.

RANDOM FOREST

Random forests operate by combining multiple decision trees to create a robust and accurate prediction model.

Using Random Forest method with optimization allowed us to get much better predictability of the dataset.





Observations with weather data:

With the help of ML optimization, we found that the importance table changed a fair bit.
The accuracy showed significant improvement after optimization.
Further improvements can be made by:

- Segmenting the data by hot/cold/dry/humid/windy weather by location.
- Creating seasonal segmentation.

VISUAL RECOGNITION

With long term weather data, we find the extremes and means, but this does not expose us to the complete picture. For example, short term spikes in weather are getting more common and so is unpredictability. These spikes cause extreme weather scenarios that can be harmful. Generative Adversarial Networks (GANs) can help us integrate climate data with weather events and help with short term forecasting.

```
Accuracy: 0.9235095381736755, Val_Accuracy: 0.875
Loss: 0.02913747727870941, Val_Loss: 0.05806592479348183
```

Proposal for the use of GANs in weather prediction:

- GANs could use traffic/surveillance/satellite imaging devices to record and interpret weather data
- GANs could use historic patterns and compare to current situation to predict the short term weather forecasts, which are much harder to do than the long term predictions.
- GANs could be trained to observe polar angular tilt shifts based on fixed camera observations of the position of the sun by comparing it to historical data.

SUMMARY

The three structured models can be used in sequence for a full featured implementation of weather prediction tool.

1

Segmentation

This step is important to analyze similar data from similar weather locations and similar seasons. This will only help reduce outliers and improve success rate.

2

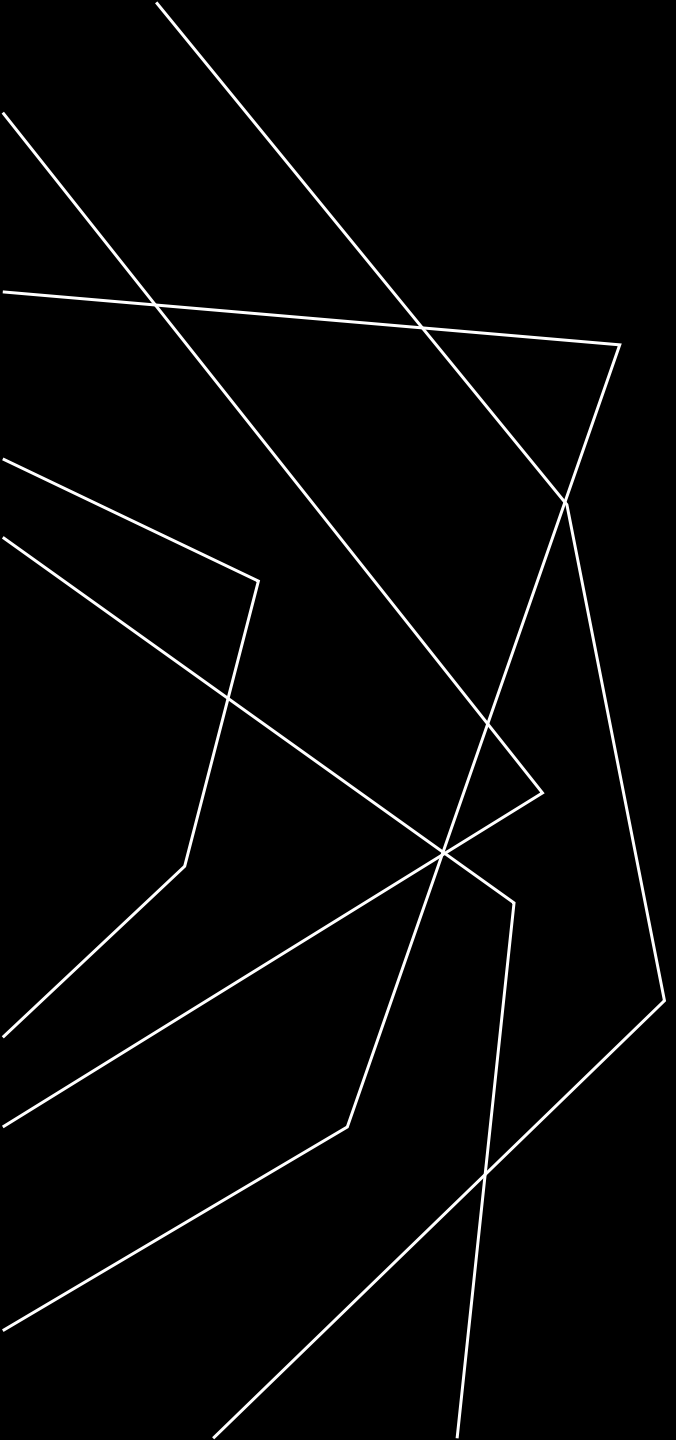
Prediction Model

Using CNN/RNN models with Random forest after ML optimization of the segmented data will help give accurate future predictions of general long-term forecasts.

3

Short term warnings

Using Visual recognition and GANs, we can use common tools like surveillance and satellite imaging to record and interpret short term weather but also large small shifts in climate that would affect the long term.



THANK YOU

Ratul Debnath

+1 705.970.0224

Ratul.Debnath@gmail.com

https://github.com/ratul-debnath