ClimateWins

Weather Conditions and Climate Change

Final Report, Joan Gandy, 02/04/2025

Objective: Use machine learning to predict the consequences of climate change

Recurrent Neural Network (RNN):

Designed for processing sequences of data. Use information from previous time steps to predict the current conditions. Suited for time series predictions, speech recognition and text generation.

Random Forest:

Can identify complex patterns from multiple variables and classify data into binary outcomes.

Generative Adversarial Network (GAN):

A generator and discriminator work together to classify data as either real or fake.

Thought Experiment #1:

Weather data from the European weather stations from 1960 to 2022 can be used to forecast the weather of these regions in upcoming years.

Thought Experiment #2:

Variables in weather data can be used by Random Forest models to predict whether a day is likely to experience severe or not severe weather.

Thought Experiment #3:

Weather data combined with images of the sky can be used by a GAN to identify when weather patterns suddenly shift from calm to dangerous.

Thought Experiment #1:

Weather data from the European weather stations from 1960 to 2022 can be used to forecast the weather of these regions for future years using the RNN model.

Final RNN parameters:

Epochs = 50

Batch_size = 32

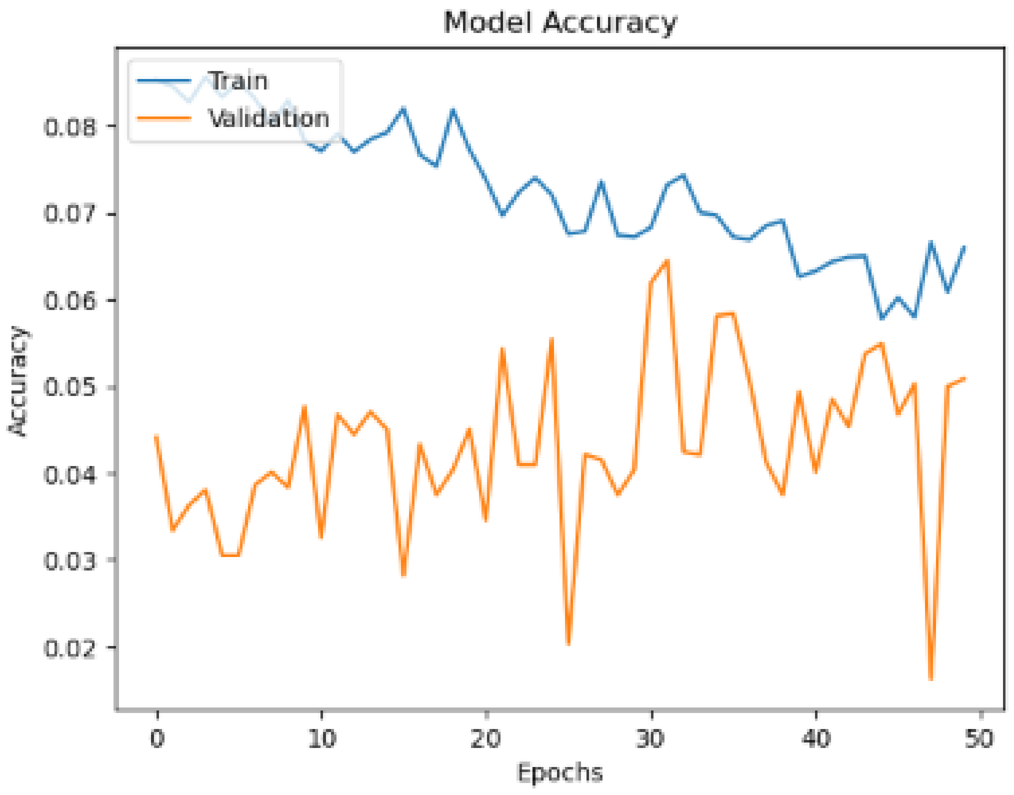
Hidden layers = 400

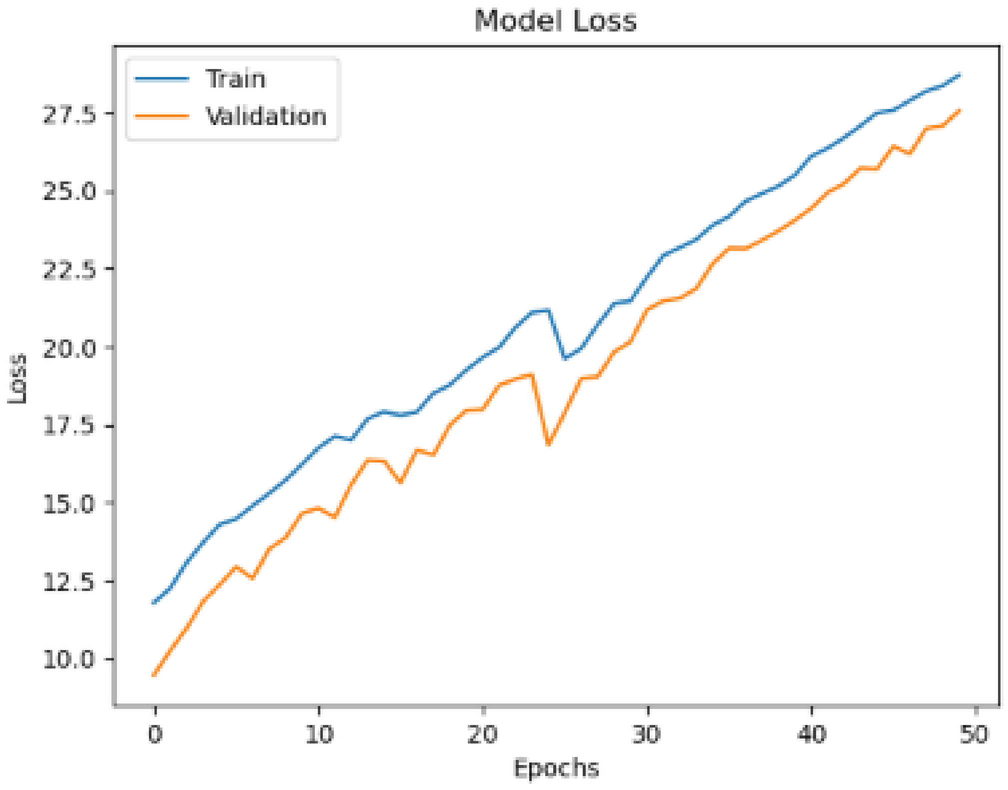
Attempted RNN parameters:

Epochs between 50 and 500 were tried. The increased number of epochs did not improve accuracy or decrease loss.

Batch sizes of 16 and 32 were tried. Only moderate improvements were made.

Hidden layers of 200, 300, and 400 were tried with little impact on the results.





Confusion matrix, along with accuracy, and loss outputs show this model did not perform well with these dataframes.

Thought Experiment #1 Conclusion

RNN models should be able to use time series data to make predictions about future weather patterns based on past weather data. This could improve the warning time before extreme weather events, i.e. floods, hurricanes, tornadoes, tsunamis, droughts, etc.

Recommendations:

- Source additional weather data from geographically disperse weather stations across the globe.
- Source data on weather patterns leading up to known extreme weather events.
- Evaluate the parameters used to determine the Pleasant Weather data set.

Thought Experiment #2:

Variables in the weather data can be used by Random Forest models to predict whether a day is likely to experience severe or not severe weather.

Final Random Forest parameters:

N_estimators = 300

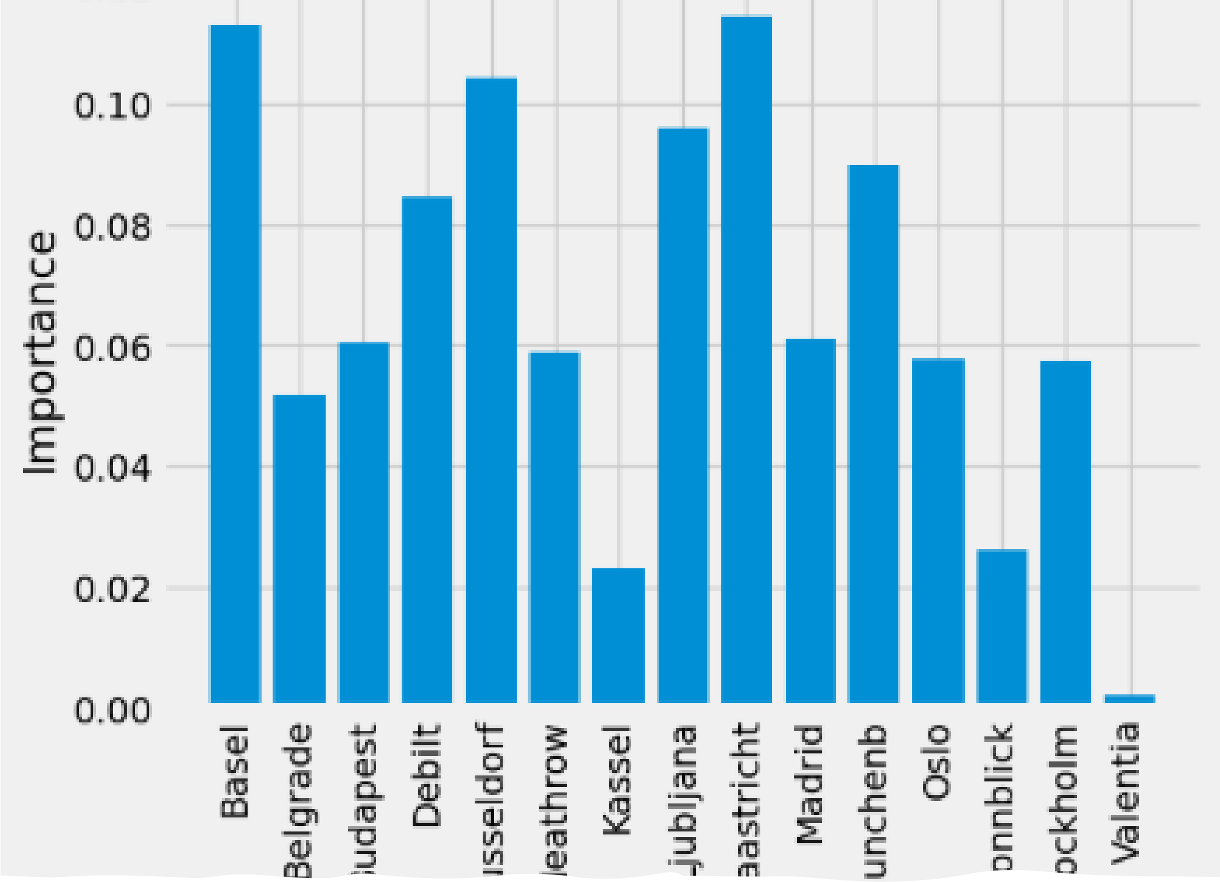
Max_depth = 10

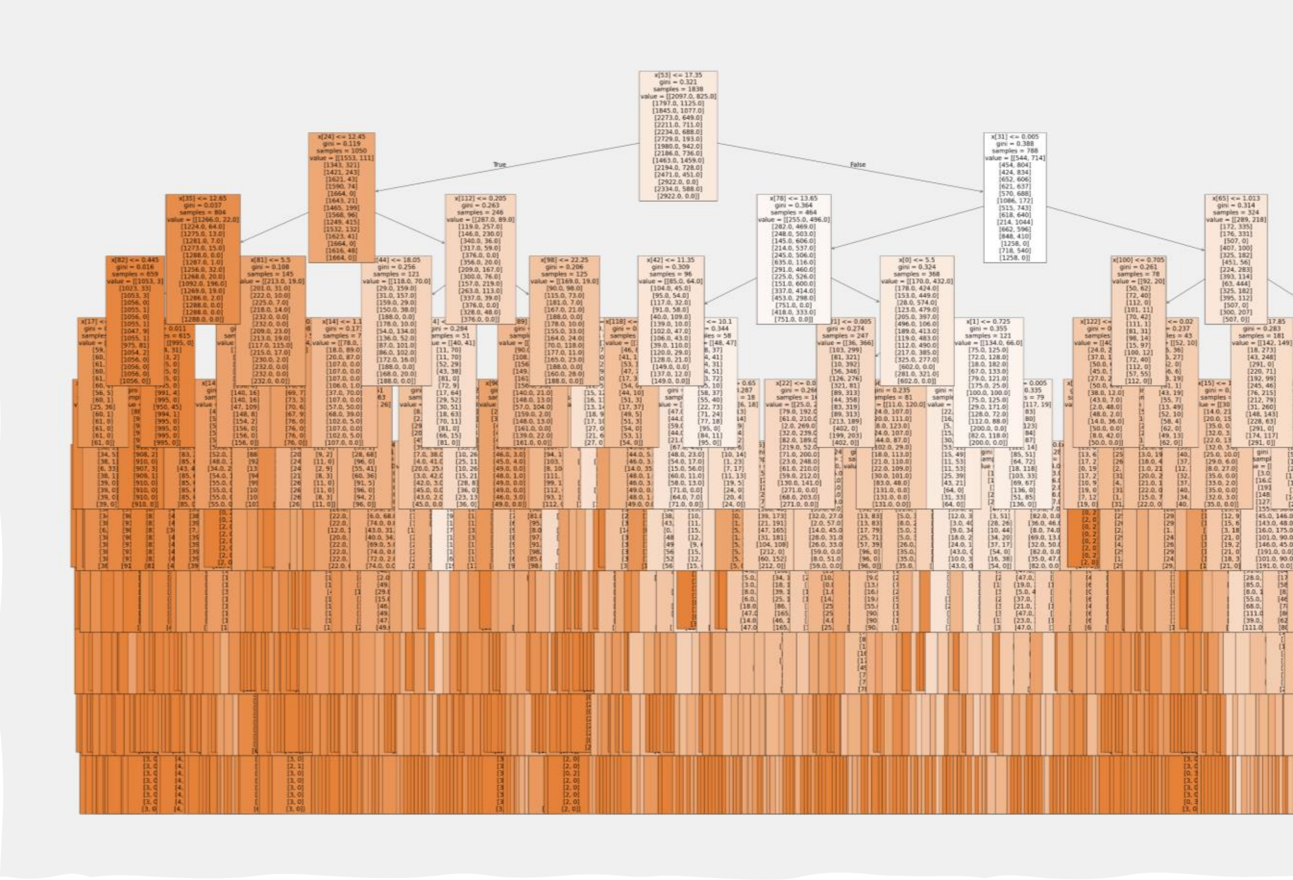
Accuracy = 58.2%

Attempted parameters:

N_estimators from 100 to 300 were tried.

Max_depths of 5 to 15 were used. Max depths of 5 and 15 produced less accurate results than 10.

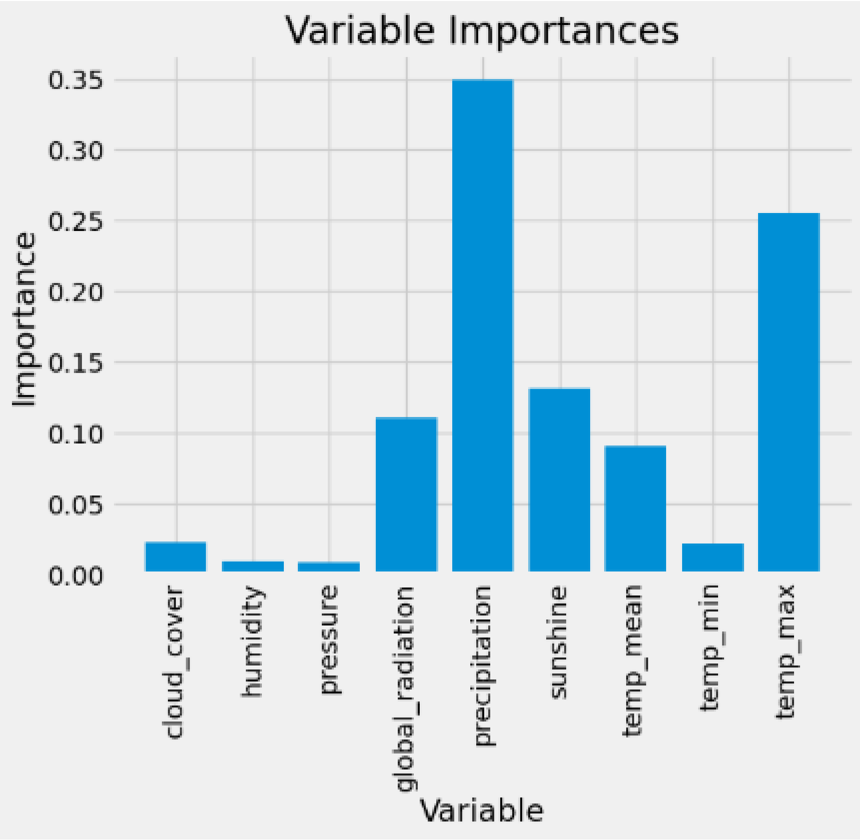




Weather stations at Maastricht, Basel, Dusseldorf, Ljubijana, Munchenb, and Debilt had the most influence on the random forest model. The stations could possibly be separated into groups of those whose importance was above and below .06.

Thought Experiment #2 Conclusion

Random Forest models can be used to separate data based on weather variables into binary categories including severe and not severe forecasts. This can be used to inform communities of the likelihood of dangerous weather.



Recommendations:

- Source additional weather data from geographically disperse weather stations across the globe and carefully select how the stations are combined into a data set.
- Define severe and not severe weather parameters.
- Evaluate to what degree the variables are influencing the Random Forest model, as shown for the Basel station in the graph to the right.

Thought Experiment #2:

Weather data combined with images of the sky can be used by a GAN to identify when weather patterns suddenly shift from calm to dangerous and alert the public to take precautions.

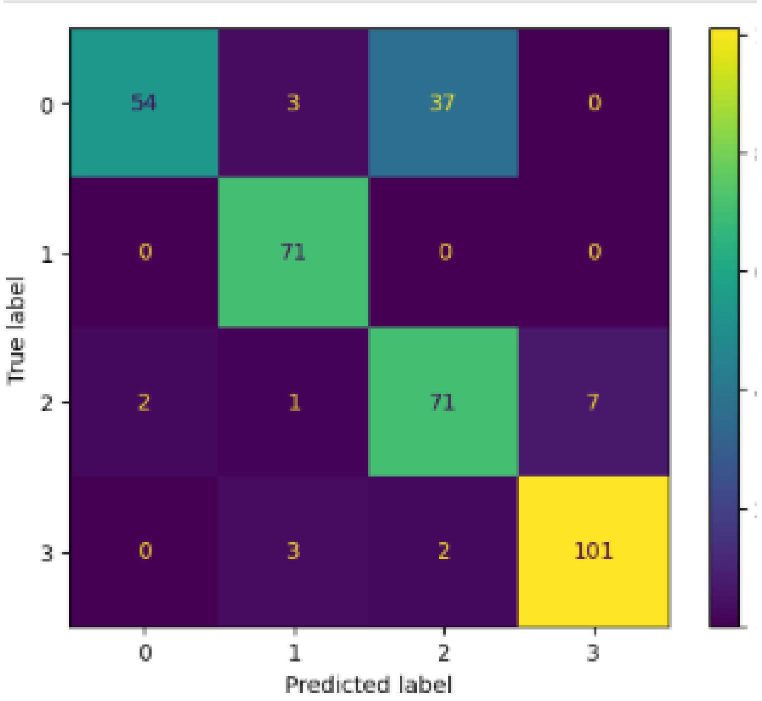
Final GAN parameters:

Epochs = 200

Attempted GAN parameters:

Increasingly large number of epochs were used to try and get the loss and accuracy of the training and testing models to converge.

Larger number of epochs should lead to better convergence.



- The confusion matrix for the GAN showed the model produced strong results with accuracy at 96.8%, value accuracy at 83.9%, loss at .015 and value loss at .068.

Thought Experiment #3 Conclusion

- GANs are able to accurately classify images of weather into categories such as cloudy or sunny. GANs may be able to quickly assess when weather is on the brink of turning dangerous in fast moving scenarios such as tornadoes and alert the public.

Recommendations:

In our research, Generative Adversarial Networks have shown the most potential for reaching ClimateWins' immediate goal of helping people who live in places where extreme weather may increase because of climate change.

Test the usefulness of this model more thoroughly by sourcing images of skies from areas prone to tornado activity, identifying how close in time the images occurred before a severe weather event.

Gather data about tornado activity from the same area as the images.

Continue to work on the convergence of the model, finding the right number of epochs to use with the minimum amount of computing power.

Test whether the model is able to classify images as pre-tornado or not.

Questions?

Joan Gandy

Cell: XXX-XXX-XXXX

Email: XXXXX@gmail.com