ClimateWins Analysis of Machine Learning Opportunities

Sydney Johnson

Dec’ 2024

Introduction

ClimateWins if machine learning can be used to help predict the consequences of climate change around Europe and, potentially, the world

The purpose of this project is to produce some potential best options for moving forward in this area

Objectives

- Identify weather patterns outside the regional norm in Europe.
- Find new patterns in weather changes over the last 60 years
- Determine if unusual weather patterns are increasing.
- Generate possibilities for future weather conditions over the next 25 to 50 years based on current trends.
- Determine the safest places for people to live in Europe over the next 25 to 50 years.

Thought Experiments

- Use random forest model to determine which places in Europe are the best to live over the next 25 to 50 years based on simple climatic criteria
- Use Long Short-Term Memory (LSTM) model to analyze weather conditions as time-series data for forecasting
- Generative Adversarial Network (GAN) synthesized weather data can be used to consider outcomes of weather systems not yet encountered

Machine Learning Model Definitions

- **Random Forest** is an ensemble machine learning algorithm that builds multiple decision trees during training and combines their outputs to improve accuracy, reduce overfitting, and enhance predictive performance
- A **recurrent neural network (RNN)** is a type of deep learning algorithm designed for sequential data that processes inputs through loops, allowing information to persist and enabling the modeling of temporal dependencies in data
  - A **Long Short-Term Memory (LSTM)** network is a type of RNN designed to effectively capture long-term dependencies in sequential data using specialized memory cells, gates, and mechanisms to regulate the flow of information
- **Generative Adversarial Network (GAN)** is a type of neural network architecture consisting of two models–a generator and a discriminator–that compete in a zero-sum game, where the generate creates realistic data samples and the discriminator tries to distinguish them from reality

Additional Data

- To accomplish ClimateWins' goals with sophisticated machine learning, more data will be needed on more existential climate threats, like sea level rise, water scarcity, temperature extremes, and increasingly devastating storms

Random Forest Concept

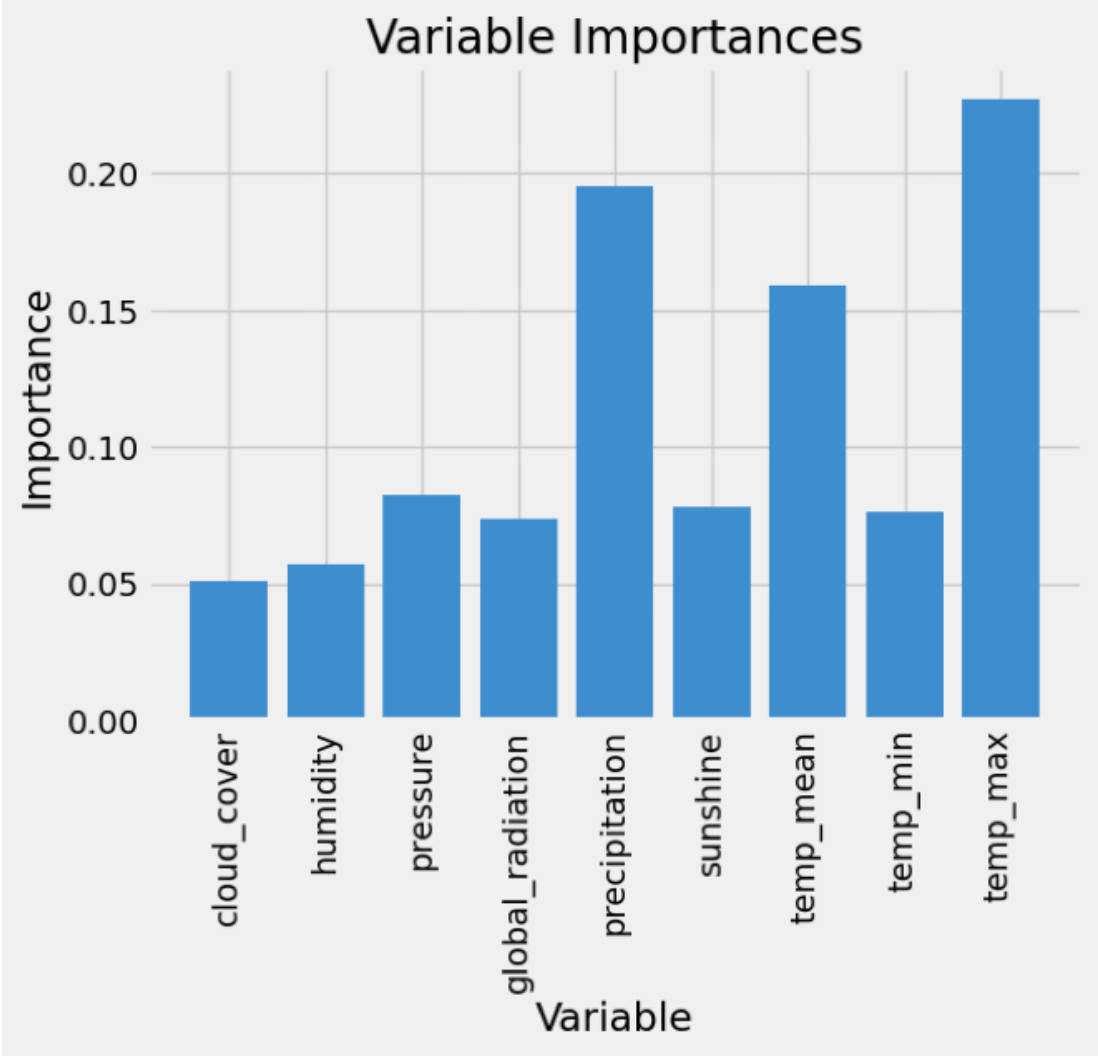
Thought Experiment: Use random forest model to determine which places in Europe are the best to live over the next 25 to 50 years based on simple climatic criteria

Why: The random forest model used in testing was highly effective at predicting which weather conditions were pleasant and which were not

When asked to distinguish between different weather stations based on 9 different variables, the model could do so with 88% accuracy

Considering the efficiency in which it took on that task, with the proper data on climate safety, it could sort through different places to live in Europe and help determine which are going to be the safest in the long run

Data that might help contribute to this might involve flood risk, relationship to sea level, positions in weather systems, access to clean water, and air quality



LSTM Concept

- Thought Experiment: Use LSTM to analyze weather conditions as time-series data for forecasting
- Why: The best deep learning model at our disposal for analyzing a large amount of dated data is the Long Short-Term Memory (LSTM) variation of a recurrent neural network (RNN) model
- We can anticipate what future weather conditions may look like by using an intelligent forecasting system based on past conditions

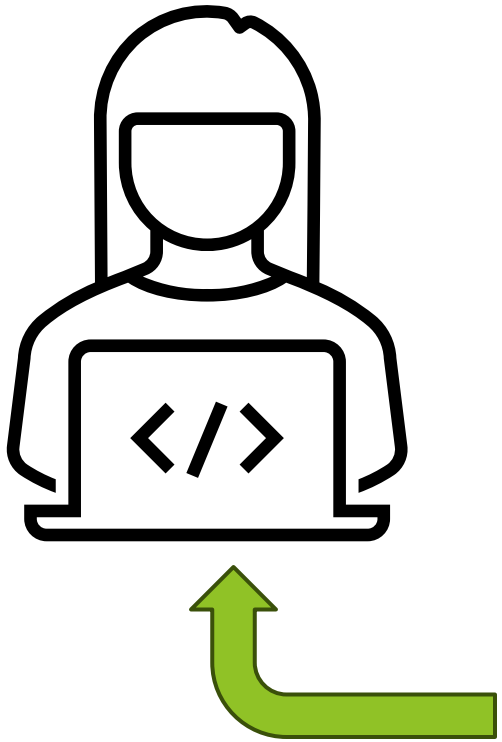
GAN Concept

- Thought Experiment: GAN synthesized weather data can be used to consider outcomes of weather systems not yet encountered
- Extreme weather events are becoming more and more common
- To prepare for the effects that these increasingly intense weather event might have on the broader climate, we can generate realistic synthetic data using generative adversarial networks (GANs) and examine their effects on related climates
- Generating synthetic climate data is already a possibility, as seen in this paper from the European Geosciences Union by Besombes, et al.

Recommendations

- While all these thought experiments have potential, the idea of using GANs to synthesize weather data has deeply valuable implications. We already have data on most weather events that we have experienced over the past 100 years, but as weather becomes more extreme, we're going to see unprecedented systems. Being able to prepare for the impact of those systems will be very valuable.

Thank You!



Please reach out with any questions regarding my analysis or conclusions:

sydneyruthj@gmail.com

Link to my GitHub Repository