Climate Predictions Using Machine Learning Approaches

Abuduaini Niyazi, Rexiati Dilimulati, Aihemaiti Maitituerdi

Overview

In contrast to modern climate models, which predict that precipitation will increase as temperatures rise, the Horn of Africa has experienced severe and recurring droughts over the past few decades. The region's agriculture-based economies have suffered greatly as a result of these droughts. Therefore, the quality of long-term weather prediction has become fundamentally important. In this project, we use multiple past climate proxy records to build a machine learning model to determine whether we can predict the future climate of the Horn of Africa.

Dataset

The dataset used in this project is sourced from previously published scientific research articles. Given the different sampling resolutions of input features, we adjusted them accordingly to match the resolution of the output feature through a combination of scaling and interpolation techniques.

Approaches

First, we trained a neural network using historical data. Following that, we forecasted the input features for the future time frame using both the baseline model and the ARIMA model. Finally, we used the trained neural network to forecast the future climate data.

Results

Our project encountered significant challenges, primarily concerning the characteristics of the dataset and the computational costs of training models. Our dataset comprises past climate indicator values averaged over every 200-year interval, with some data exhibiting significant fluctuations over the centuries since industrialization. In fact, These variations could potentially serve as more reliable indicators for future climate trends. Therefore, careful consideration is necessary when performing train-test splitting and selecting input features.

- For the input feature forecasting, we tested the linear trend model and the random walk with drift model. Finally, we determined the ARIMA model as the optimal approach.
- For neural network training, we noticed the model was overfitting on the test set. However, we implemented a range of optimization techniques and ultimately attained optimal results.

Future Work

- Seeking professional advice on feature selection to improve neural network performance
- Collecting modern-day data from instrumental records to develop better models
- Adopting a more sophisticated neural network model such as Recurrent Neural Network (RNNs) or LSTM optimized for time series data