ClimateWins Weather Report

Kuo Heng (Alex) Chao

4/22/2024

Introduction

About ClimateWins

ClimateWins is a European non-profit organization dedicated to mitigating the effects of climate change through advanced technology and scientific research. Due to a lack of extensive funds, innovative approaches such as machine learning are required to make the best use of available data.

Objective

Use machine learning techniques to predict and understand the effects of climate change on weather patterns across Europe, resulting in better preparedness and response strategies.

Hypothesis

1. Using historical weather data, machine learning can accurately predict the increasing frequency of extreme weather events in Europe.

2. Supervised learning models are particularly effective for forecasting specific weather conditions, such as extreme temperatures or heavy precipitation, on given days.

3. By applying optimization techniques, we can significantly enhance the feature selection process, thereby improving the accuracy and reliability of our weather predictions.

Data Source & Integrity

Data Source

The European Climate Assessment & Data Set project provided the primary dataset, which includes comprehensive weather observations from 18 weather stations across Europe, meticulously recorded from the late 1800s to 2022.

Data Integrity

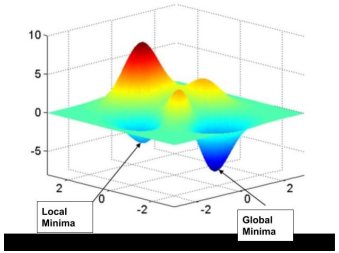
- In the Data, the weather stations are not dispersed equally. While certain regions—especially those that are rural or difficult to access—may have sparse data, others may have dense coverage, offering comprehensive local data. Regional biases may result from this unequal distribution, with models performing well in places with lots of data and badly in those with less.

- The techniques and tools used to measure weather conditions have changed over the years. Previous records may have been manually logged, which could have resulted in inconsistent data recording or human error. More accuracy and consistency are available with modern automated systems, but correctly assessing trends requires a grasp of how different approaches change from one another.

Optimization Techniques

Gradient Descent

It is a basic machine learning optimization approach that minimizes a function. Gradient descent aids in optimizing the model's weights for your project by reducing the discrepancy between the expected results and the actual data. By computing the gradient of the loss function, which counts prediction errors, with respect to each weight, this method iteratively modifies the weights.



- **Batch Gradient Descent**: Calculates the gradient of the cost function based on the parameters θ and θ throughout the entire dataset. This is computationally expensive and unfeasible for really large datasets.
- **Stochastic Gradient Descent (SGD)**: Unlike batch gradient descent, SGD adjusts the parameters based on a single data point (or small batch). This makes updates coarser, but substantially faster when working with huge datasets.
- **Mini-Batch Gradient Descent**: This is a hybrid of batch and stochastic versions. The gradient and parameters are computed and updated using a small section of the dataset. This approach minimizes the variance of parameter updates, potentially leading to more steady convergence than SGD.

Supervised Learning

Artificial Neural Networks (ANN)

ANN is essential for modeling complicated, non-linear interactions in huge datasets, including different climatic data. This adaptability enables ANNs to forecast a wide range of climate impacts, from extreme weather occurrences to the spatial spread of climate change.

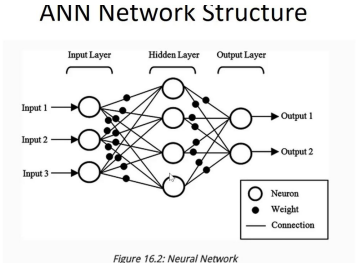
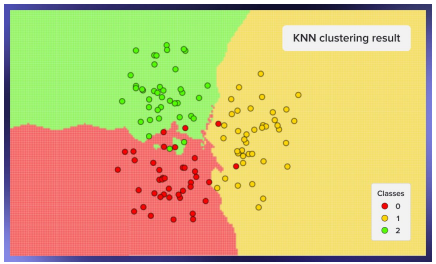


Figure 16.2: Neural Network

K-Nearest Neighbors (KNN)

KNN adjusts to changes in the input data, improving our ability to anticipate. It works especially well for localized weather forecasting, using past data from nearby geographic locations to provide precise and timely predictions about the weather.



Summary & Future Steps

Summary

In this study, we employed machine learning models to evaluate our hypothesis that supervised learning models are effective in forecasting specific weather conditions, such as extreme temperatures or heavy precipitation. We utilized Linear Regression, K-Nearest Neighbors (KNN), and Artificial Neural Networks (ANN) due to their strengths in handling complex climate data. Our findings so far confirm that these models can accurately predict various weather events, affirming our hypothesis about the potential of machine learning to enhance our understanding and prediction of weather patterns influenced by climate change.

Exploring Unsupervised Learning Models

- Our future goal is to use unsupervised learning techniques to identify previously uncategorized patterns and abnormalities in weather data. This method will aid in finding new weather patterns and, potentially, new climatic phenomena caused by global warming.
- Techniques such as cluster analysis and principal component analysis will be critical in segmenting data into meaningful groupings without prior labeling, bringing fresh insights into complicated climatic processes.

Comprehensive Climate Modeling

- The ultimate goal is to create a complete climate model that predicts both specific weather events and wider climate trends. This model would combine supervised and unsupervised learning results to provide a comprehensive understanding of climate impacts.
- Collaboration with worldwide meteorological and climate organizations will also be investigated to improve model inputs and validation, ensuring that our predictions are reliable and globally relevant.

Thank You For Your Time!

Any Questions Please Feel Free to Contact me via Email