CLIMATEWINS

Weather Conditions and Climate Change

Esther Howard
May 29th 2024

OBJECTIVE

Use Machine Learning to help predict the consequences of climate change

HYPHOTHESES

Higher levels of cloud cover are associated with lower maximum temperatures, as clouds block incoming solar radiation.

Precipitation events are more likely to occur on days with higher humidity levels and lower atmospheric pressure.

Sunshine duration is positively correlated with temperature, with longer periods of sunshine associated with higher temperatures.

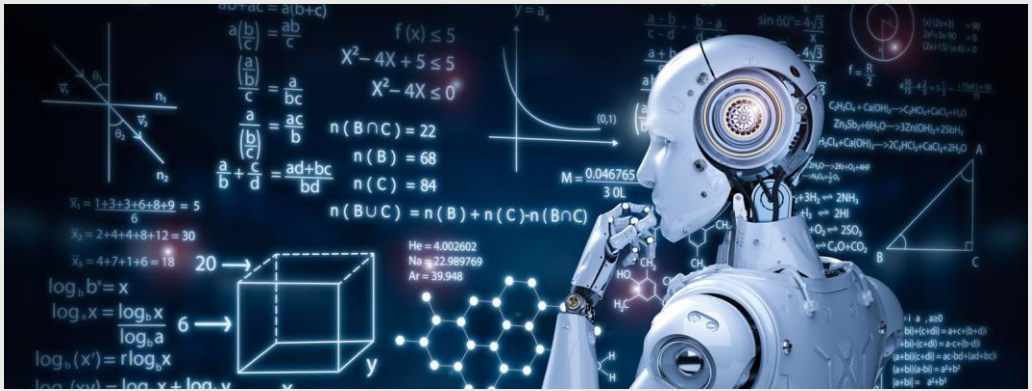
DATASET

- The dataset used in this project was collected by the European Climate Assessment & Data Set project (ECA&D) and can be found here. Since it is an official source, the dataset can be considered accurate and trusthworthy.
- The dataset includes weather data collected from 18 distinct weather stations spanning across Europe, covering the time period from 1960 to 2022. This dataset comprises various meteorological parameters including:
  - Temperature (min, max and mean)
  - Humidity
  - Cloud cover
  - Wind speed
  - Pressure
  - Global radiation
  - Precipitation
  - Sunshine
  - Snow depth

DATA BIAS

Collection Bias	Sampling Bias	Data Processing Bias
Changes in instrumentation, measurement methods, or station locations over time can introduce biases.	Data might be collected from specific regions or stations, potentially skewing the representation of climate patterns.	Biases can be introduced during data processing steps, such as data cleaning or interpolation, which may inadvertently alter the original climate signal.

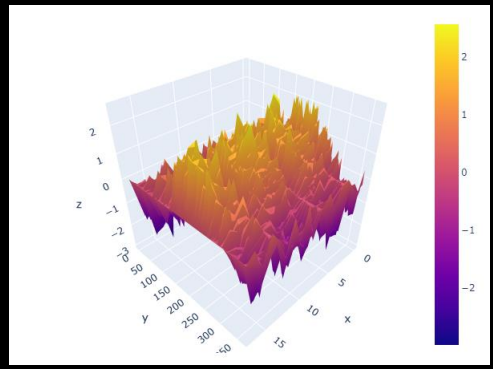
SUPERVISED MACHINE LEARNING



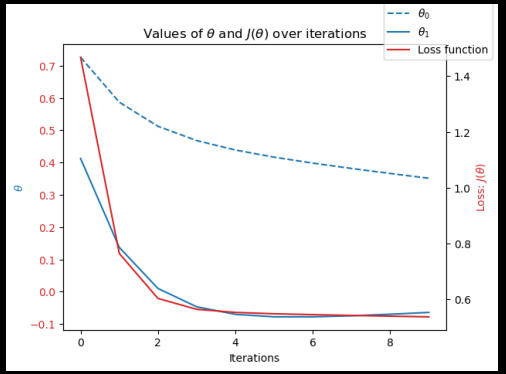
DATA OPTIMIZATION

Data Optimization employs a range of techniques and strategies to enhance data management across various aspects, thereby boosting efficiency, reliability, accessibility, and overall utility, aiming to extract maximum value from existing resources

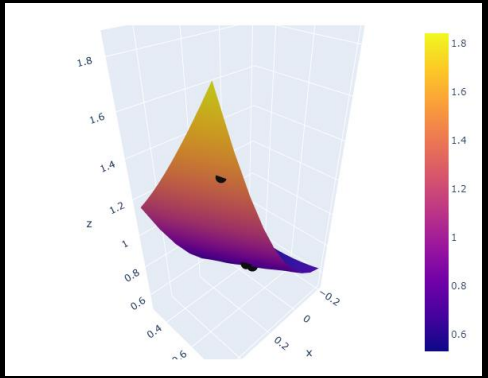
OPTIMIZATION TECHNIQUES



3D visualizations were used to map the weather data for all stations throughout a year.

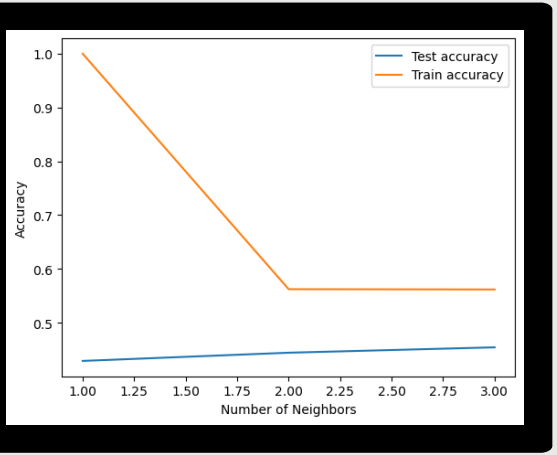


Loss function was employed to assess the deviation between predicted and actual weather data.



Gradient descent was applied to pinpoint the local minimum in the dataset, optimizing the model's performance towards an optimal solution.

SUPERVISED LEARNING MODELS



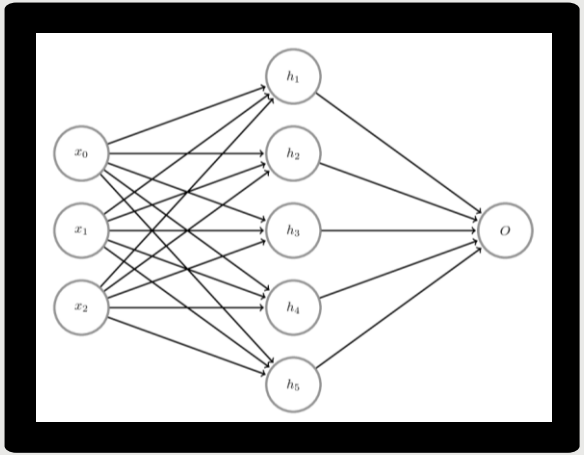
KNN was used to classify data points into groups by examining the categories of their closest neighbors.

Confusion matrix were used to provide visual insights and showcasing instances where classed were misclassified (confused).

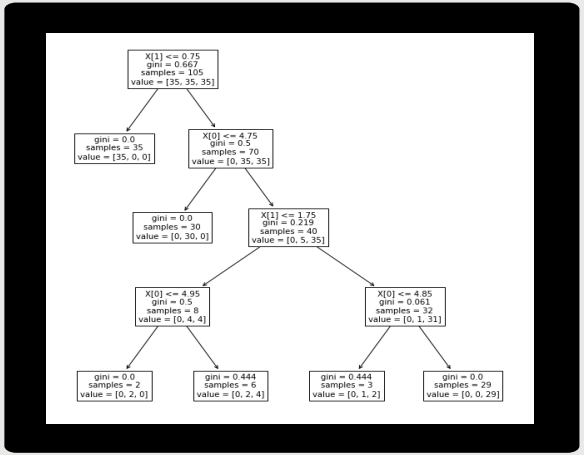
	precision	Recall	f1-score	support
0	0.68	0.69	0.68	1099
1	0.63	0.80	0.71	1561
2	0.65	0.85	0.74	1461
3	0.74	0.73	0.73	878
4	0.72	0.72	0.72	975
5	0.63	0.61	0.62	935
6	0.64	0.55	0.59	743
7	0.62	0.81	0.70	1229
8	0.73	0.74	0.74	933
9	0.74	0.74	0.79	2033
10	0.70	0.61	0.65	952
11	0.64	0.48	0.55	675
12	0.00	0.00	0.00	0
13	0.63	0.46	0.54	788
14	0.00	0.00	0.00	228
micro avg	0.68	0.71	0.69	14490
macro avg	0.58	0.59	0.58	14490
weighted avg	0.67	0.71	0.68	14490
samples avg	0.32	0.33	0.31	14490

Classification reports were used to evaluate the model's performance across various weather stations in the dataset.

SUPERVISED LEARNING MODELS



Artificial neural networks (ANN) were used enabling complex computations for predictions and classifications.



Decision trees were employed to progressively refine solutions effectively narrowing down optimal outcomes

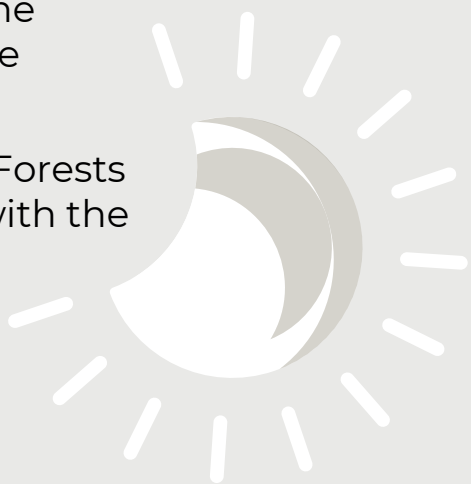
SUMMARY

- A dataset containing various weather parameters such as temperature, humidity, cloud cover, precipitation, sunshine, etc. was analyzed with the goal of predicting climate change trends using a variety of optimization techniques and machine learning algorithms.

- ANN was the ML model that had the highest accuracy for predicting weather data.

RECOMMENDATIONS AND NEXT STEPS

- Use Artificial Neural Network (ANN) for predicting weather data as it had the highest accuracy rate of the algorithms employed.
- Further evaluate the weather of all the stations in the dataset over the 60 years' worth of data to determine if Machine Learning can help predict the consequences of climate change.
- Use other Machine Learning algorithms like Random Forests and Gradient Boosting Machines to compare results with the already used models.



Do you have any questions?

howardesther13@gmail.com

THANKS!