WEATHER CONDITIONS AND CLIMATE CHANGE WITH CLIMATEWINS

Theo Liang
11/15/24

OBJECTIVE

ClimateWins is interested in using machine learning to help predict the consequences of climate change around Europe and, potentially, the world.

The organization is particularly interested in extreme weather events over the past 10 to 20 years.

HYPOTHESES

1

The frequency of extreme weather events, such as heatwaves, storms, and floods, has significantly increased in mainland Europe over the past 20 years due to climate change.

2

Historical weather data from the past century can significantly improve the accuracy of future extreme weather predictions in mainland Europe when used to train machine learning models.

3

Implementing machine learning models to predict extreme weather events will improve emergency preparedness and reduce damage in at-risk areas.

DATA SOURCE

The dataset from the European Climate Assessment & Dataset project includes daily weather data from 18 weather stations across Europe, covering the period from 1960 to 2022.

The data includes information on temperature, wind speed, snow, and other weather factors.

DATA BIAS

Collection Bias: The dataset only includes data from 18 weather stations in Europe, covering the years from 1960 to 2022. This is a small subset compared to the 23,789 stations across Europe and the Mediterranean, potentially limiting the data's scope and diversity.

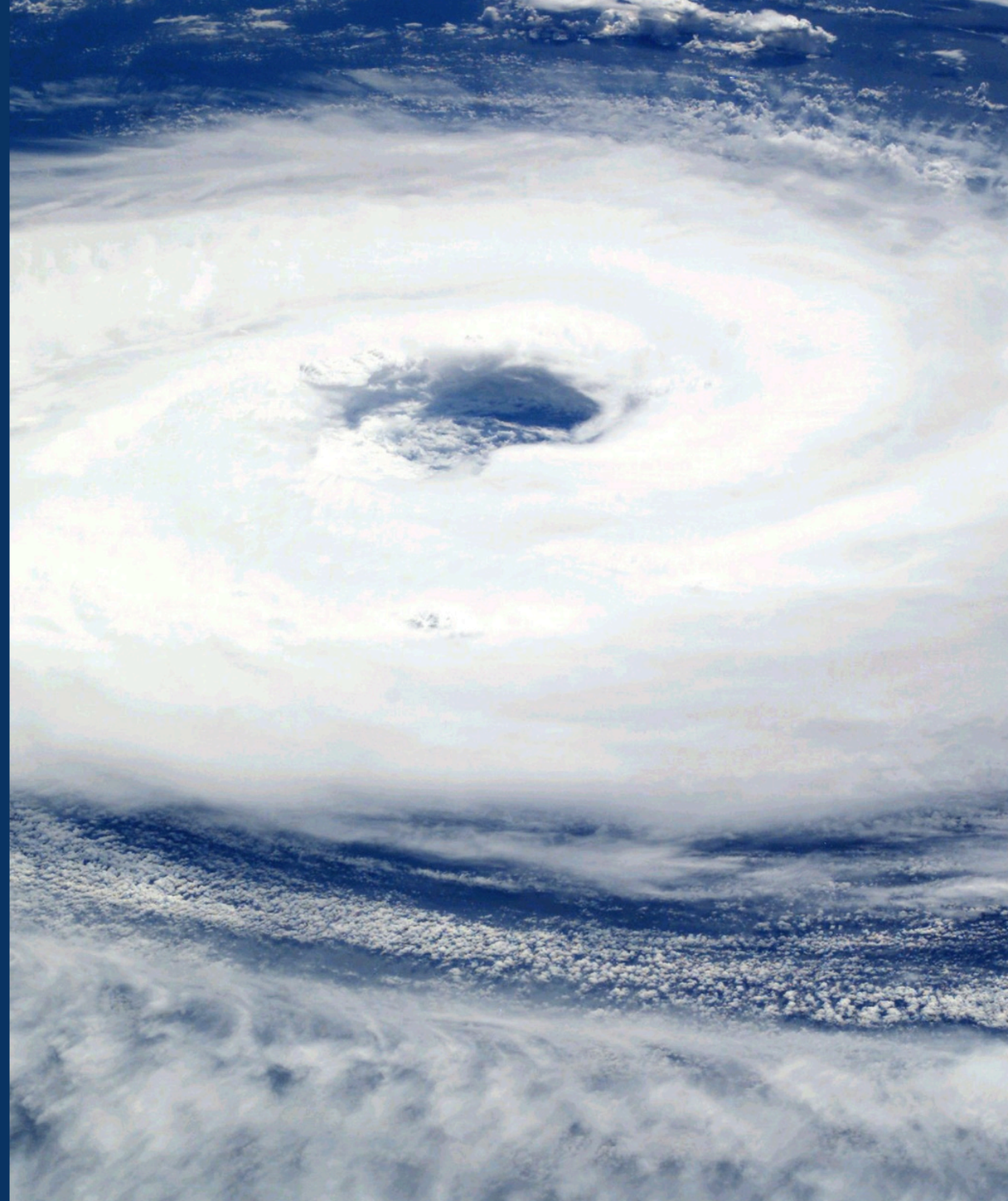
Sampling Bias: All 18 weather stations in the dataset are located in urban cities, which may result in a bias toward urban climate patterns. This could skew predictions about weather or climate changes, as it doesn't fully represent rural or less densely populated areas.

Historical Bias: Older data may have been collected with less accurate instruments or fewer measurement points, leading to inconsistencies when combined with more precise, modern data.

DATA OPTIMIZATION

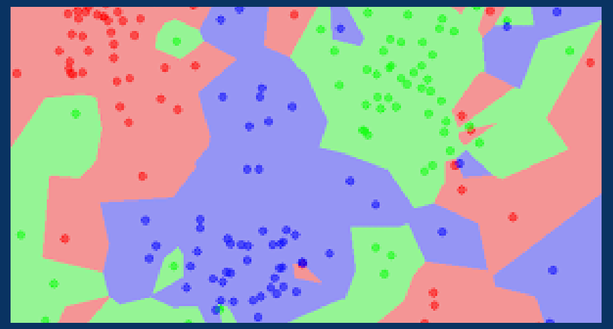
A method called **gradient descent** optimization was used to see how well temperature patterns could be predicted, helping ClimateWins work toward more accurate weather forecasts.

A model was used to predict future temperatures based on historical data to adjust its settings and reduce the gap between the predicted temperature and the actual temperature, improving its accuracy over time.



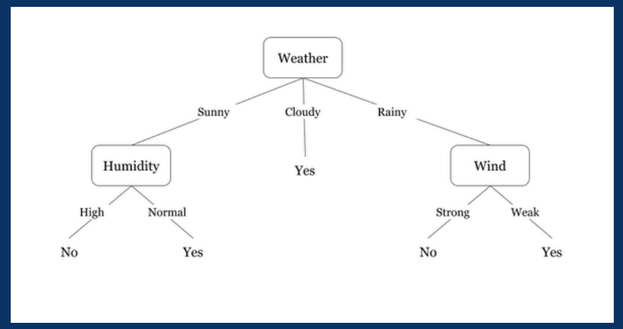
SUPERVISED MACHINE LEARNING

K-Nearest Neighbor (KNN)



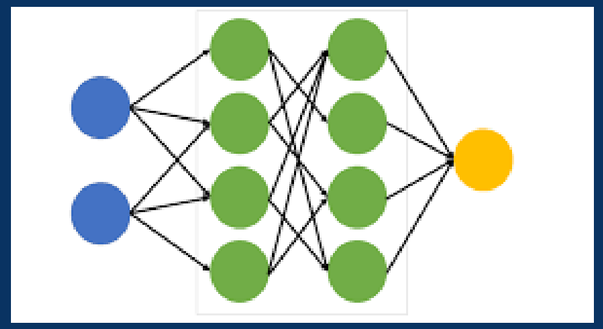
The KNN algorithm looked at the temperatures recorded at the K closest stations and predicted the temperature based on those nearby values.

Decision Tree Model



The decision tree predicted pleasant weather by splitting the data based on temperature values and classifying the day as pleasant or not.

Artificial Neural Network (ANN)



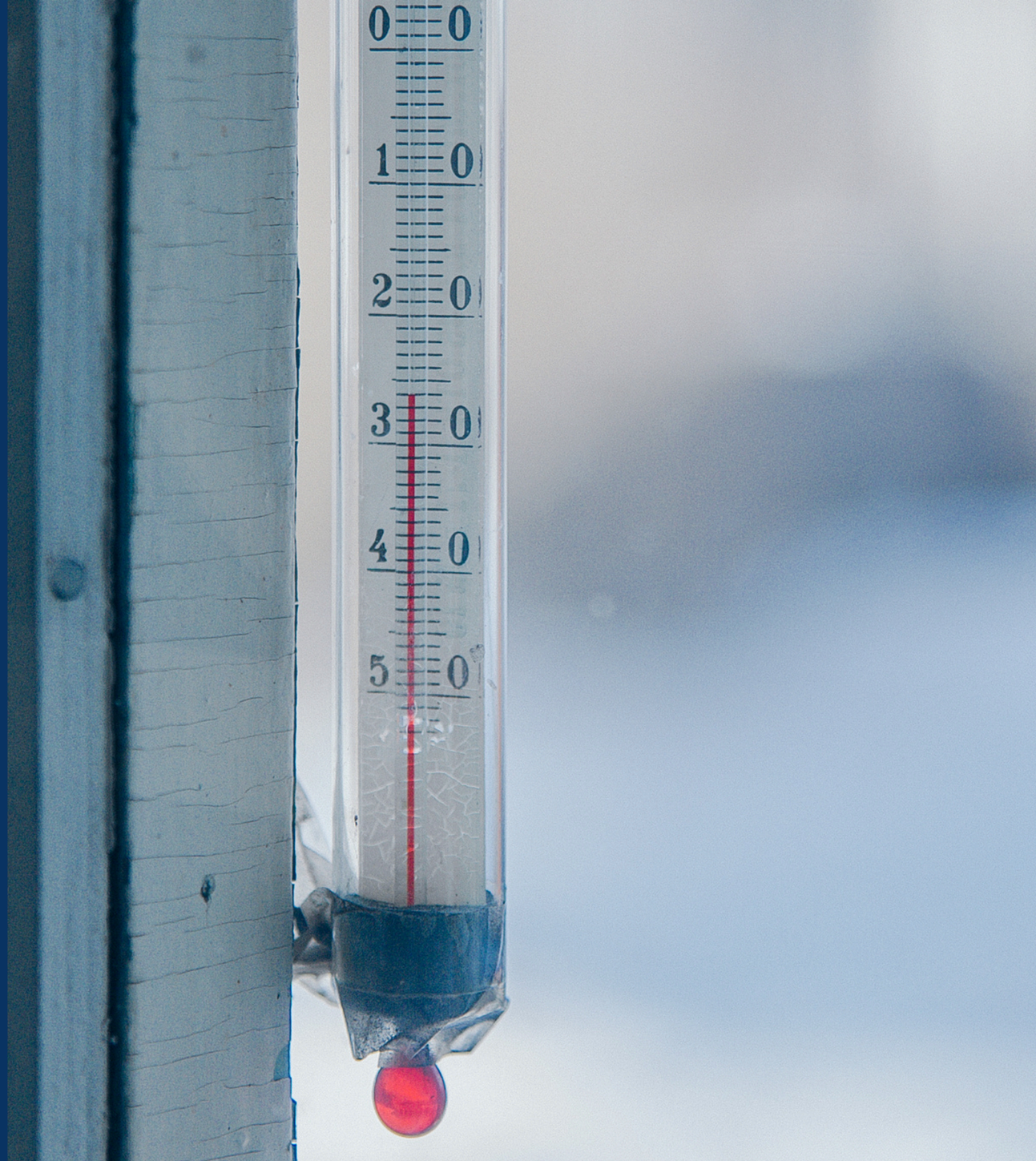
The ANN predicted pleasant weather by processing temperature values through layers of nodes to classify the day as pleasant or not.

FINDINGS

The KNN algorithm achieved high accuracy, with most weather stations reaching 84% or above, averaging 88%.

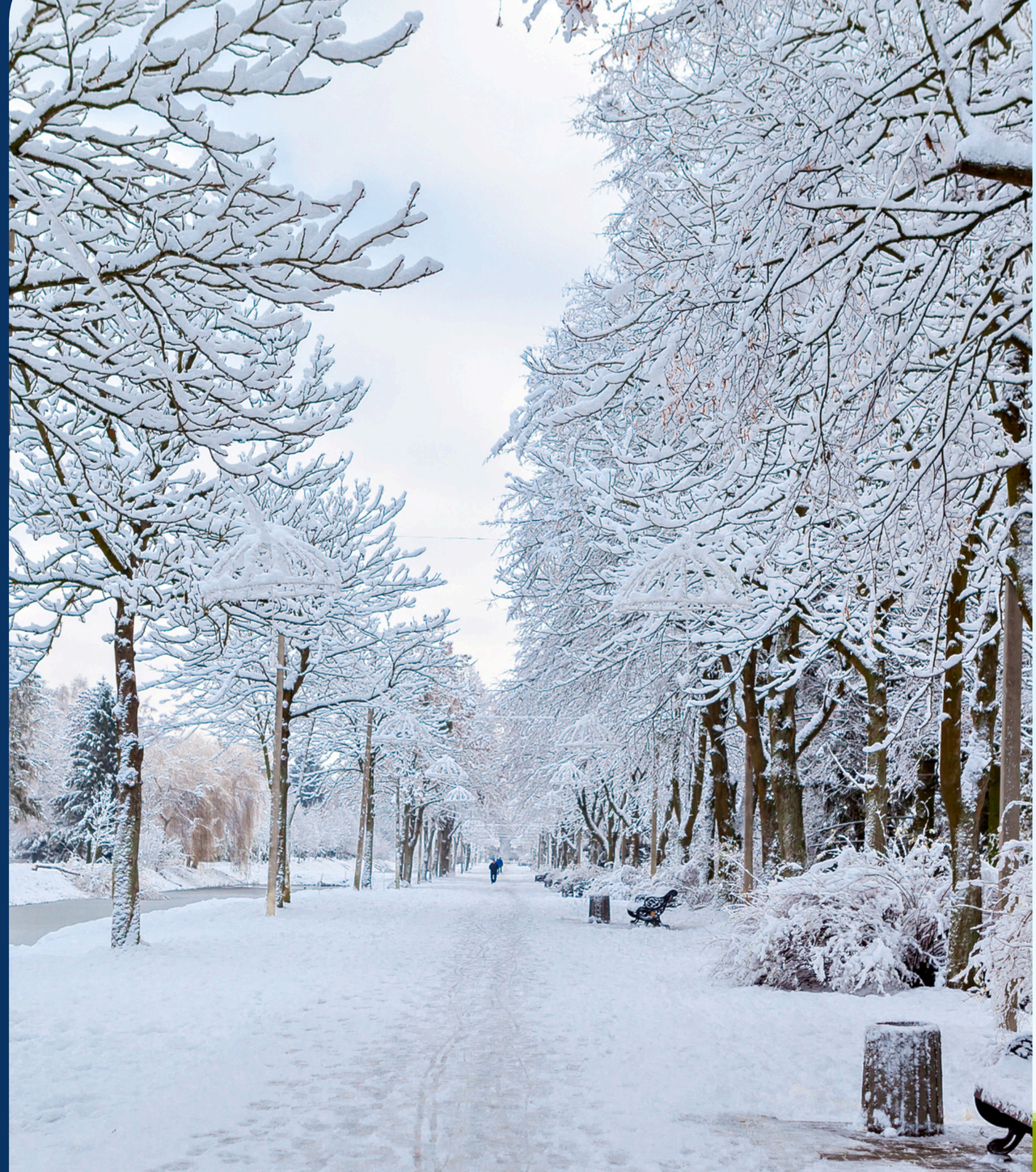
The Decision Tree model had perfect training accuracy (100%) but struggled with testing accuracy (63.04%), indicating overfitting.

The ANN model improved over three iterations, from 45.65% to 54.26% accuracy, but also showed signs of overfitting in the final model.



CONCLUSION

ClimateWins' goal is to make accurate weather predictions, and for this, the KNN algorithm is the best choice, with an average accuracy of 88%. The Decision Tree model overfitted with 100% training accuracy but only 63.04% testing accuracy, while the ANN showed improvement but still faced overfitting.



NEXT STEPS

To improve upon the models and make more accurate weather predictions, ClimateWins could consider:

- pruning the Decision Tree Model to reduce overfitting
- adjusting parameters such as number of neighbors and layers in the KNN algorithm and ANN
- applying unsupervised learning methods such as clustering

THANK YOU!

GitHub

in

tableau

For more information or to connect, visit the links above or reach out via email at theodore.liang@gmail.com.