Forest Fire Prediction

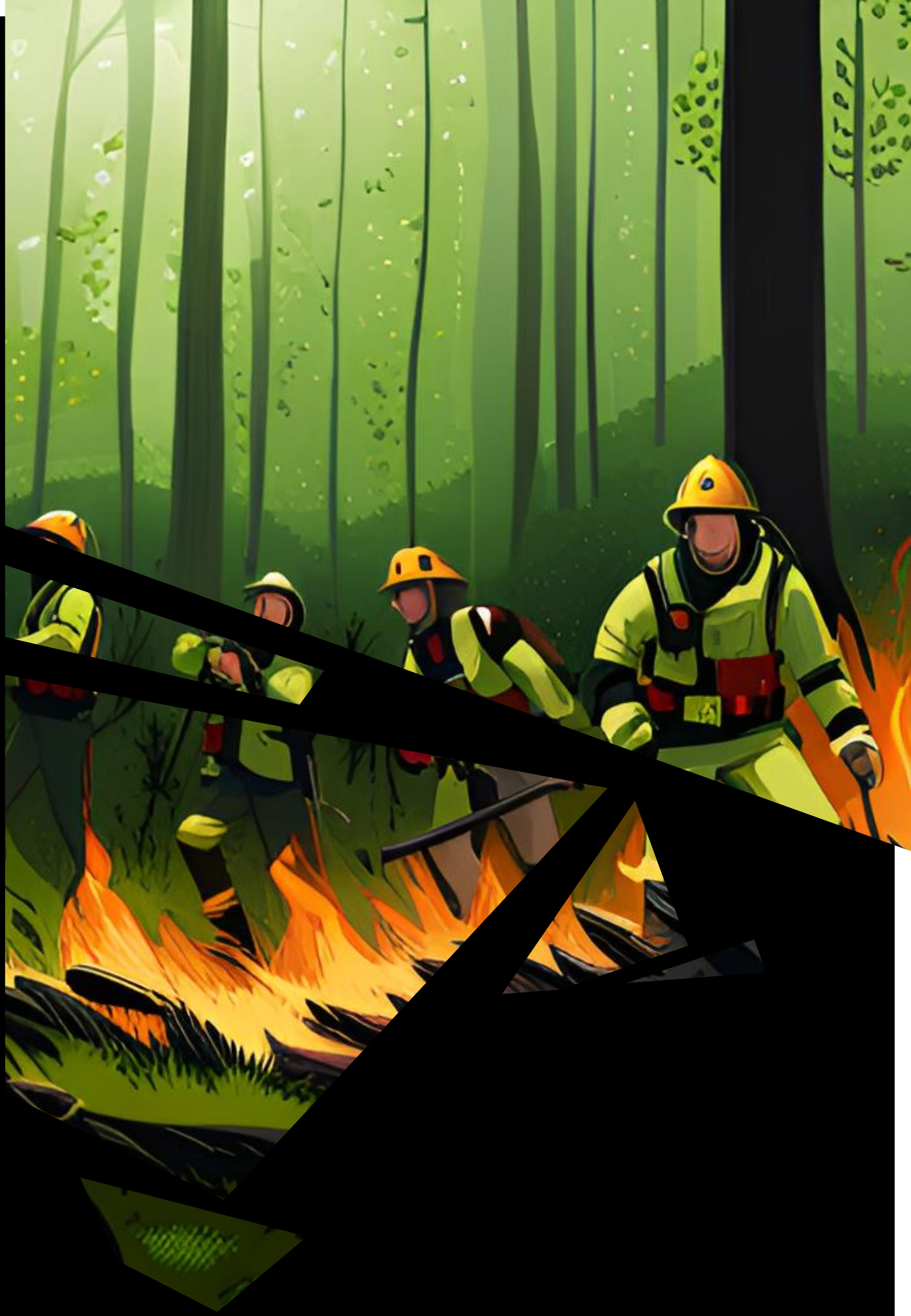
by Yevgenia Zalkind

Introduction

Every year, wildfires devastate vast areas of forests, causing significant damage to both the environment and human life. These fires are often unpredictable and can spread rapidly, making them difficult to control. Early detection and prediction of these fires are crucial in mitigating their impact and preventing further damage.

In order to deal with and take control of such a disaster, there is a need for money, the use of relevant human resources, and, of course, treatment as quickly as possible in order to reduce the damage as much as possible. The goal is to enable better resource allocation with faster response time.

Hence, using the meteorological data in combination with machine learning, we can identify the centers of the fires together with their possible location.



Data sources and acquisition

The acquisition of data for forest fire prediction is a complex process that involves multiple sources.

One important source is the National Interagency Fire Center website, which provides Point Locations for all wildland fires in the United States reported to the IRWIN system.

Another important source of data is the Open-Mateo historical weather API which provides weather data from 1940 until now.

By combining data from multiple sources, it is possible to build a comprehensive picture of the environment and identify areas at risk.

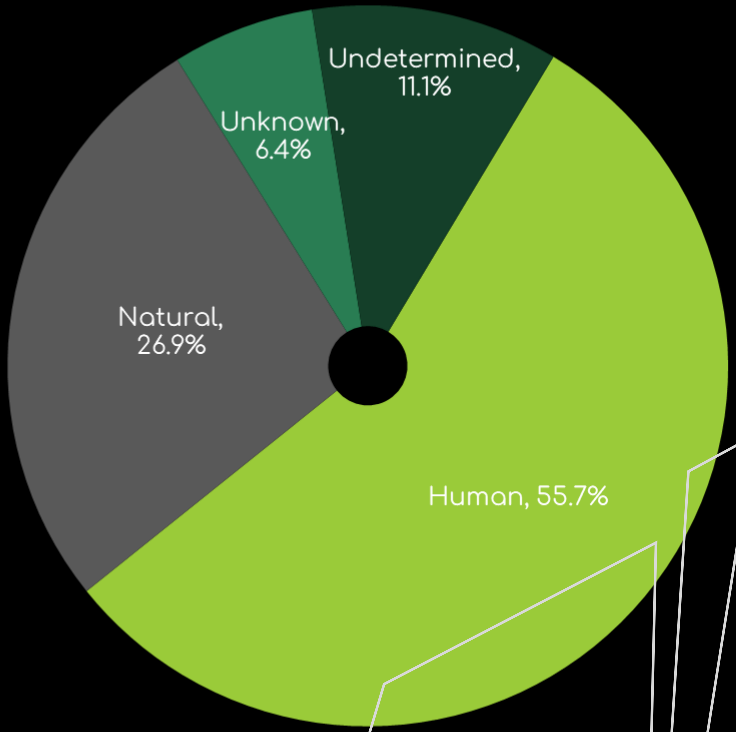
Initial analysis and data optimization

In order to accurately predict forest fires, we must first analyze and optimize the data that we have collected.

This involves removing any outliers or errors in the data, as well as identifying any missing values or inconsistencies.

Once the data has been cleaned, we can then begin to explore patterns and trends within the data. By doing so, we can identify key variables that are most predictive of forest fires and use them to build our predictive models.

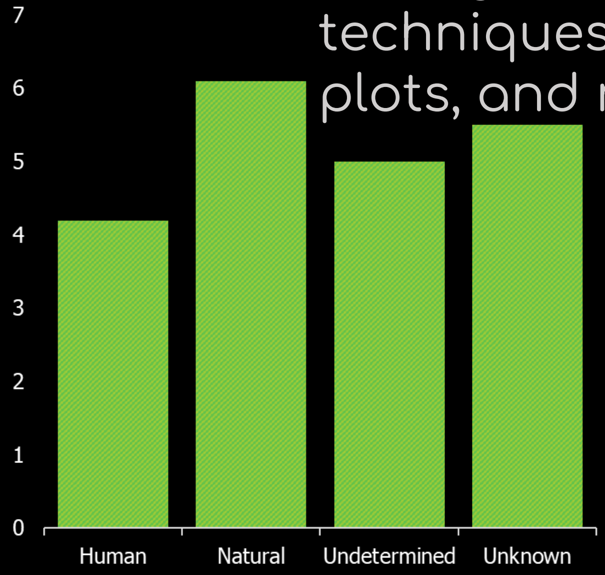
Visualization and EDA



Visualization and exploratory data analysis (EDA) are crucial steps in understanding the patterns and relationships within our forest fire dataset.

By creating visual representations of our data, we can identify trends, outliers, and potential correlations that may not be apparent through numerical analysis alone.

During this stage, we used various visualization techniques such as scatter plots, geo plots, pie plots, and more to gain insights into our data.



Implementation and machine learning

To ensure the accuracy and reliability of our predictions, we implemented a machine-learning model using a random forest algorithm.

The model was trained on a large dataset of historical forest fire incidents, taking into account various factors such as minimum and maximum temperature, wind speed, precipitation, and more.

Our model achieved an impressive overall accuracy rate of 88%, indicating its effectiveness in predicting naturally caused fires and an 84% accuracy rate in predicting locations.

However, we need to continue to fine-tune and optimize the model to further improve its performance.

Summary and conclusions

In conclusion, our analysis has shown that predicting forest fires is a complex but crucial task.

By utilizing various data sources and advanced analytical techniques, we were able to optimize our models for maximum performance.

Our findings suggest that implementing these models can greatly improve the accuracy of fire prediction and ultimately aid in preventing catastrophic events.

Furthermore, our visualizations have highlighted key patterns and trends in the data, providing valuable insights for future research.