Institute of Engineering & Technology

JK Lakshmipat University, Jaipur

Report Title:

California Wildfires Analysis & Prediction Model using Machine Learning

Submitted By:

Pranshu Kumar Mishra

November 2024

Table of Contents

Abstract

California's wildfire seasons are growing increasingly severe, fuelled by climate change, prolonged drought, and urban development in high-risk areas. Once confined to summer and fall, fire seasons now extend nearly year-round, with wildfires rising in both frequency and intensity. This leads to devastating impacts: millions of acres are burned, thousands of homes are destroyed, and infrastructure and ecosystems suffer immense disruption. Economic losses reach billions annually, while environmental consequences include degraded air quality and habitat destruction. Despite enhanced firefighting resources and strategies like prescribed burns, California's wildfire threats continue to escalate, requiring innovative solutions. Machine learning offers critical advancements by enabling more precise fire risk prediction through analysis of weather patterns, vegetation health, and historical data. This approach supports more efficient resource allocation, early warnings, and targeted responses, helping to protect lives, property, and ecosystems in an era of increasing wildfire risk.

Introduction

What is a Wildfire?

> Wildfire is a Naturally / Man-Made occurring disaster that takes place in a region having sufficient moist climate, green or dry vegetation and hot weather that promotes burning to a longer extent.
>
> Many areas across the globe experience wildfire, which include –
>
> - India – forests of Odisha, Maharashtra, Jharkhand, Chhattisgarh
> - United States of America (West)
> - Australia
> - Amazon Rainforest

What are the possible causes of Wildfire?

- Natural
  - Lightning
  - Volcanic Eruptions
  - Sparks from rock falls
  - Spontaneous combustion

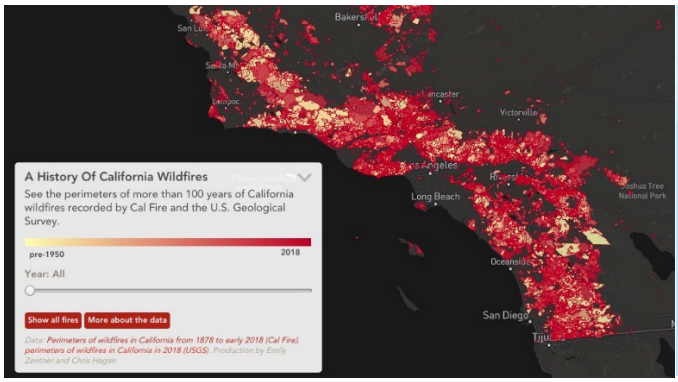
- Man-Made
  - Deforestation
  - Burning fossil fuels
  - Pollution
  - Discarded cigarettes

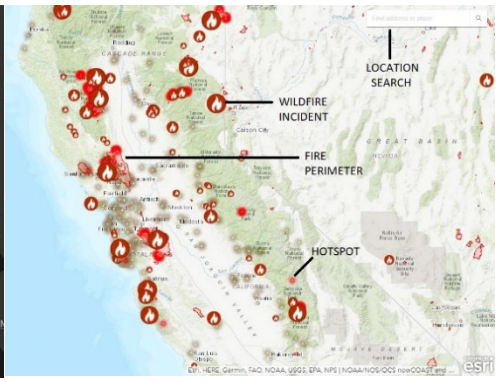
The *California* Wildfires

California wildfires have long been a significant environmental challenge, driven by a mix of natural and human factors. The state's dry climate, frequent winds, and abundant vegetation create the ideal conditions for fires to ignite and spread rapidly.


History

Wildfires have been a natural occurrence in California for millennia, but the frequency and intensity have surged in recent decades due to climate change, urban expansion, and poor land management. Some of the most destructive wildfires have occurred in recent years, including the 2020 fire season which burned over 4 million acres. California's fire season used to peak in the late summer and fall, but recent years have shown year-round fire activity, reflecting the changing climate patterns.





The Aftermath of those Wildfires


Role of Technology in Disaster

Q. How can Machine Learning help in these Wildfires?

Early Detection and Prediction of Wildfires

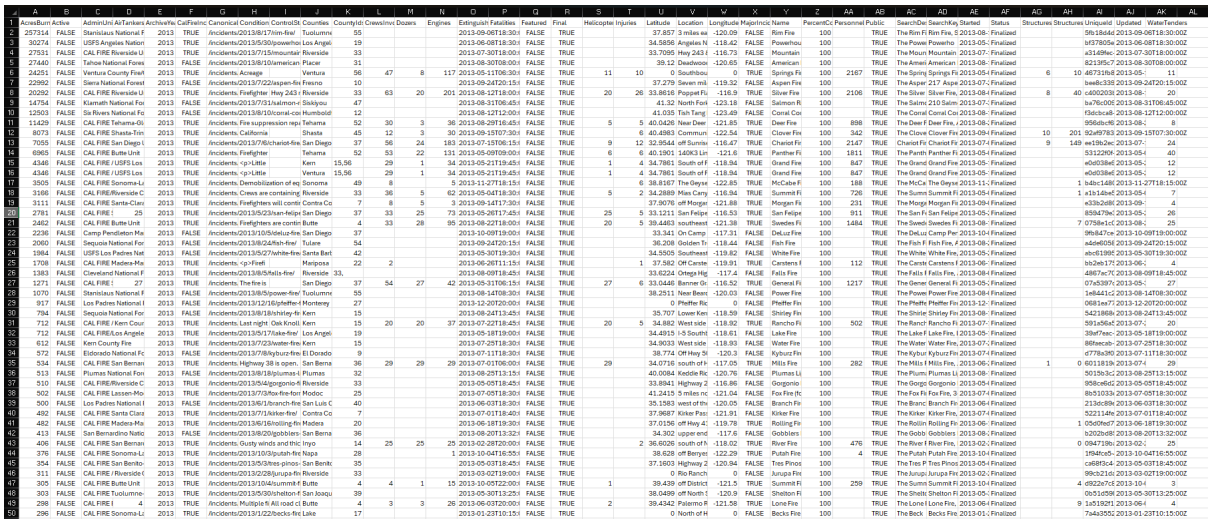
Fire Risk Prediction Models - ML models use historical fire data/weather patterns/vegetation density, or other environmental factors to predict the likelihood of fires. By identifying high-risk areas, authorities can deploy resources more effectively.

The steps include -

1. **Data Collection and Integration**

   Predictive models use vast amounts of historical and real-time data to make accurate forecasts. Key types of data include:

   - **Historical Fire Data**: Information on past wildfire incidents, including their location, duration, and cause. This data helps identify patterns and trends over time.
   - **Weather Patterns**: Weather conditions like temperature, wind speed, humidity, and precipitation have a significant impact on fire risk.
   - **Vegetation Density and Type**: Certain types of vegetation are more flammable, especially during dry seasons.
   - **Topographical Data**: Terrain data, such as elevation, slope, and aspect (the direction the slope faces), also impact fire behaviour, as fires can spread more quickly uphill and in certain terrain configurations.
   - **Human Activity Data**: Human actions, like camping, burning trash, or operating machinery near dry areas, are common causes of wildfires.

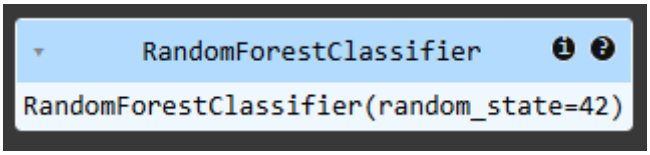


2. Model Development and Training

Different types of ML algorithms can be used, depending on the complexity of the data and the desired accuracy. Examples include:

- **Logistic Regression and Decision Trees**: These models can provide quick, interpretable results on which factors are most influential in fire occurrence.
- **Random Forests and Gradient Boosting Machines (GBM)**: These models capture complex interactions among multiple features, improving predictive accuracy by aggregating predictions from multiple decision trees.
- **Deep Learning**: Neural networks and recurrent neural networks (RNNs) can handle large-scale datasets, especially those with temporal (time-based) components, like weather patterns over time.

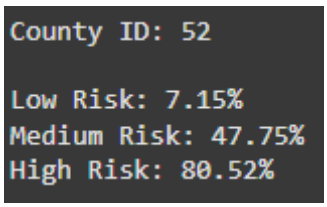
The model is trained on historical data, learning the relationships between the input features and the target variable (fire occurrence or fire risk level).



3. Real-Time Prediction and Early Warning Systems

Once trained, ML models can be used to make real-time predictions, providing early warnings for regions at high risk of wildfire:

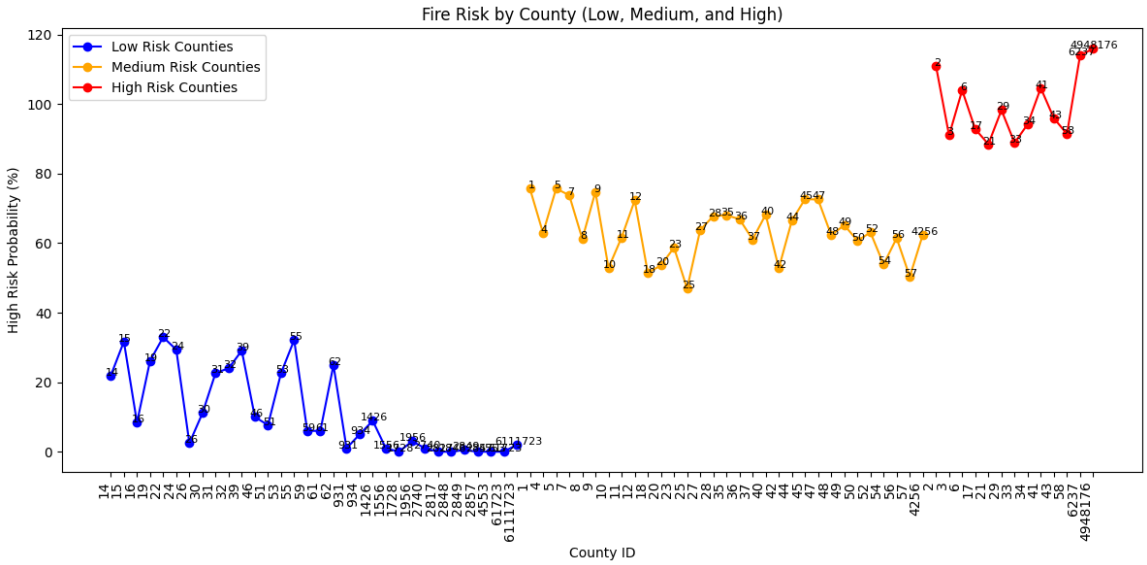
- **Risk Scoring**: For each region, the model outputs a probability score or risk level (e.g., low, medium, high) indicating the likelihood of a wildfire occurring within a certain timeframe.
- **Alert Systems**: Automated alerts can be sent to relevant authorities or residents in high-risk areas, allowing them to prepare for or mitigate the fire risk.



4. Deployment of Resources

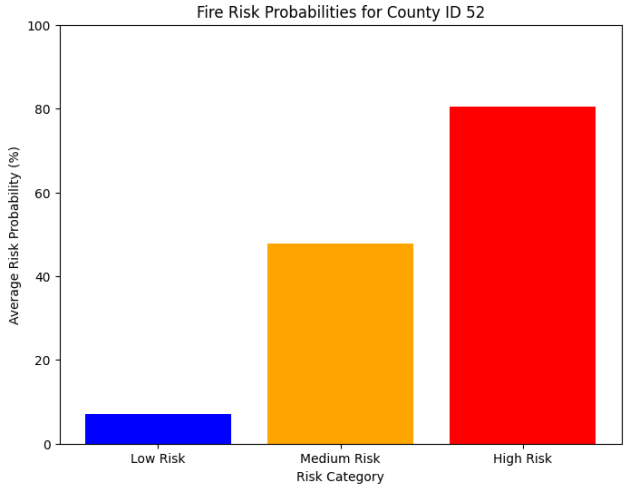
By identifying high-risk areas in advance, authorities can deploy resources more effectively:

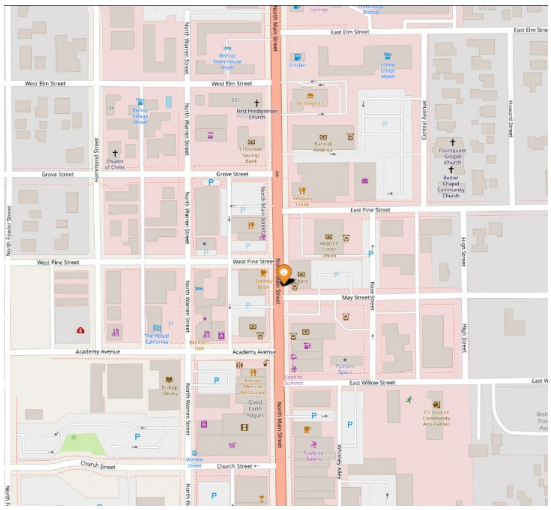
- **Pre-Positioning Firefighting Resources**: Firefighters, equipment, and aerial support can be strategically stationed near predicted high-risk zones, reducing response time if a fire does occur.
- **Public Awareness Campaigns**: For areas at high risk, public warnings can be issued, advising residents to take precautions, such as clearing brush around homes and avoiding open flames or machinery that may spark fires.
- **Controlled Burns and Vegetation Management**: Knowing which areas are at high risk enables authorities to proactively reduce fuel loads through controlled burns or vegetation clearance, lowering the likelihood of large, unmanageable fires.



Conclusion

California's escalating wildfire season, fuelled by climate change, drought, and urban expansion, poses a persistent threat to homes, infrastructure, and ecosystems. Despite advancements in firefighting and proactive strategies, the frequency and intensity of wildfires continue to surge, highlighting the urgent need for enhanced risk management. Predictive models can play a crucial role in this effort by assessing and forecasting wildfire risk levels across different regions. By analysing variables like weather patterns, vegetation conditions, and human activity, these models provide valuable insights that help allocate resources, prioritize high-risk areas, and inform early warning systems. Through accurate risk evaluation, predictive modelling empowers communities and authorities to better protect lives, property, and California's natural landscapes from the growing wildfire threat.





Areas of Improvement

Satellite imagery enhances wildfire prediction by providing real-time, high-resolution data on vegetation, soil moisture, and topography—all key factors influencing fire behaviour. Through continuous monitoring, satellites help track vegetation health and fuel load, identifying high-risk areas before fires start. Thermal imaging detects hot spots early and monitors fire spread, allowing rapid response to emerging threats. Combined with machine learning, satellite data can analyse historical fire patterns and environmental conditions, improving risk assessment accuracy. This satellite-driven insight enables more proactive, targeted wildfire predictions, supporting efficient resource allocation, better protection for vulnerable communities, and timely firefighting efforts. Also, an up-to-date Weather API can help in monitoring the climate in real time to take actions instantly.

References

- https://www.kaggle.com/datasets/ananthu017/california-wildfire-incidents-20132020
- https://www.kaggle.com/datasets/braddarrow/23-million-wildfires
- https://www.kaggle.com/code/skpatel12/california-wildfire-analysis
- https://github.com/Qunlexie/A-Comprehensive-Classification-Model-for-Predicting-Wildfires-with-Uncertainty
- https://python-visualization.github.io/folium/latest/
- https://www.kaggle.com/datasets/chelseazaloumis/cimis-dataset-with-fire-target