**Wildfire Prediction Based on Environmental Factors**

Kaiyi Chen：[kennychen@vt.edu](mailto:kennychen@vt.edu)

Guangkai Chen: [guangkai@vt.edu](mailto:guangkai@vt.edu)

**Introduction:**  
 Wildfires, often exacerbated by climate change and human activities, represent a critical challenge for environmental management and public safety. These fires can rapidly devastate vast areas, causing ecological imbalance, destroying wildlife habitats, and posing serious risks to human lives and property. The need for advanced, predictive solutions in this domain has never been more urgent.

Our project taps into the power of data analytics and artificial intelligence to develop a predictive model capable of forecasting wildfire risks with higher accuracy. By integrating diverse environmental data points such as temperature variations, precipitation patterns, soil moisture levels, and historical fire occurrences, we aim to build a comprehensive model that offers an insightful understanding of the conditions leading to wildfires.This initiative not only aims to enhance the predictive capabilities but also to empower decision-makers with actionable insights. The model can serve as an early warning system, enabling forest departments and emergency response teams to deploy resources more effectively, plan evacuation routes, and initiate fire containment strategies promptly. Moreover, it can aid in long-term environmental planning and policy-making, helping to design sustainable practices that reduce the risk of future wildfires.

By leveraging machine learning algorithms and sophisticated data processing techniques, we intend to push the boundaries of predictive modeling in natural disaster management. Our commitment goes beyond the technical aspects; we envision our work as a step towards safeguarding natural ecosystems, protecting wildlife, and ensuring the safety and well-being of communities at risk from these devastating events.

**Project Problem Statement:**

Wildfires are complex and dynamic natural disasters influenced by a variety of environmental factors. The unpredictability and potential for significant destruction associated with these fires pose a formidable challenge to communities, ecological systems, and emergency response strategies. This project aims to predict the likelihood of wildfires with an emphasis on leveraging data analytics to analyze key environmental factors such as temperature, rainfall, humidity, and sunlight.

Stakeholders including local communities, forest departments, insurance companies, government agencies, and environmentalists are deeply affected by the impact of wildfires. For local communities and forest departments, the primary concern lies in safeguarding lives and preserving the natural environment. Insurance companies, facing financial risks, need accurate predictions to adjust premiums and prepare for claims. Government agencies require this data to allocate resources effectively and implement preventive measures, while environmentalists can use this information to advocate for policies that address the root causes of wildfires.

The justification for a data-driven approach stems from the sheer scale and devastation of wildfires, which can lead to loss of life, destruction of property, and significant ecological damage. Traditional methods of predicting wildfires often rely on limited empirical data and localized observations, which may not capture the complex interplay of factors that lead to such events. A data-driven model, powered by historical data and current environmental patterns, would provide a more systematic, comprehensive, and reliable method to anticipate wildfire occurrences. This model could help authorities in planning and executing effective preventive and mitigation strategies, ultimately saving lives, protecting properties, and preserving the environment.

**Data Set：**

1. <https://www.kaggle.com/datasets/rtatman/188-million-us-wildfires>

**Preprocessing steps:**