Classifying Australian Bushfires Severity Using Machine Learning

Author: Arya Raj Khadka

Abstract:

Bushfires are one of the most devastating natural disasters, with the 2012-2020 Black Summer fires highlighting the gaps in national preparedness. This study develops a machine learning framework to classify detected bushfires into three severity categories- Low, Medium, High- using VIIRS satellite data from 2020-2024. It uses 200,000 stratified fire detection data from 5.3 million MODIS/VIIRS observations. Logistic Regression achieved an accuracy of 0.64 (f1_score:0.69), whereas random forest model reached an accuracy of 0.81(f1_score:079), demonstrating its superior predictive power. These findings highlight that machine learning can be used to enhance bushfire severity and support proactive emergency management in Australia.

1. Introduction:

1.1 Problem Context and Motivation:

Australia experiences sever bushfires seasons annually, with fire causing catastrophic impacts on ecosystems, society and economy. One prime example is the 2019-2020 Black Summer fire which cost the life of animals and humans and inversely affected the economy. Fundamentally, a bushfire is caused by the combination of three conditions – fuel, oxygen, ignition source- also known as fire triangle. Though these conditions might cause bushfires, other environmental factors like high fuel load (fallen bark, litter, twigs, and branches), dry fuel moisture, strong wind speed, and high temperature enhance the spread and intensity of it.

While bushfires cannot completely be prevented, we can mitigate the impact by early detection and classification systems. This research aims to develop a Machine Learning (ML) model that can classify Australian bushfire severity to inform emergency response planning.

1.2 Research Objective and Hypothesis:

Primary Hypothesis: Machine learning algorithms can classify Australian bushfire by integrating meteorological, topological, and vegetation data.

Secondary Hypothesis:

a) Random Forest will outperform traditional fire weather indices in classification accuracy.
b) Combining multiple ensemble algorithms will provide superior performance to single models.
c) High Temperature, low rainfall, low humidity, and strong winds significantly increase fire severity.

1.3 Limitations and Scope:

1. Vegetation Data Absence: NDVI/EVI vegetation indices not integrated despite collection, missing critical fuel conditions.
2. Computational Constraints: Analysis limited to 200,000 stratified samples due to processing limitations.
3. Temporal Scope: Focused of VIIRS data from 2020-2024; historical data excluded to ensure temporal consistency