ABSTRACT

This project offers a machine learning-based methodology for early warning and prediction of forest fires in India's ecologically rich Nallamala Forest region. Employing a combination of satellite-based Normalized Difference Vegetation Index (NDVI) data and synthetically generated weather data from 2012 to 2025, the research constructs a strong model to classify fire events. The pipeline combines MODIS HDF-format NDVI time series with historical temperature and humidity patterns, supplemented by engineered lag features. Ground truth fire events are obtained from MODIS and VIIRS fire archive data sets. For class imbalance in fire event data, the Synthetic Minority Over-sampling Technique (SMOTE) is used. The ultimate predictive model utilizes an ensemble of XGBoost and LightGBM classifiers within a voting approach, with strong potential for operational deployment in forest fire alert systems. This work emphasizes the need for a combination of remote sensing and ML methods for proactive forest management and climate resilience.