Tempest FWI Predictor

Project Overview

The Tempest FWI Predictor is a machine learning project designed to forecast the Fire Weather Index (FWI), a numerical rating of fire danger developed from weather variables. The goal is to build an accurate predictive model that uses meteorological data to estimate the likelihood and intensity of forest fires, aiding early warning and resource allocation.

Data Acquisition and Structure

The dataset used is the Bejaia Region Dataset containing daily weather observations and calculated fire danger indices, including temperature, relative humidity, wind speed (Ws), rainfall, and several FWI system components such as FFMC, DMC, DC, ISI, BUI, and the target variable FWI itself. Date information is present but limited; notably, the "Year" column is constant (2012) and excluded from modeling due to lack of variance.

Data Preprocessing

- The dataset was loaded using pandas, ensuring accurate import and inspecting data types and completeness.
- Leading and trailing spaces in column headers were stripped using `df.columns.str.strip()` to avoid access errors.
- The categorical "Classes" column, indicating fire occurrence, was mapped to numeric (0 for "not fire", 1 for "fire") to enable quantitative analysis.
- Constant columns, such as "Year", were detected and removed automatically since they provide no predictive value or correlation information.

Exploratory Data Analysis

Histograms

Histograms were plotted for the FWI and all relevant numerical features to understand their distributions. This visual analysis informs the shape, skewness, and presence of outliers in each variable, which is critical for appropriate preprocessing and model design. For example, a skewed FWI distribution might require transformation or special modeling strategies to balance rare high-risk cases.

Correlation Heatmap

A correlation matrix heatmap was generated to explore linear relationships among the features. This revealed how variables relate to one another and to the target FWI, guiding feature selection and engineering. The "Year" column was blank in the heatmap, reflecting its constant value and exclusion from analysis.

Summary

This project’s initial phases establish a strong foundation in data integrity, exploration, and prep for effective machine learning. The exploratory steps are vital in understanding the Fire Weather Index’s behavior in historical data, ensuring models built are robust, interpretable, and actionable for wildfire risk prediction.