Project Documentation

Title: Fire Weather Index (FWI) Prediction System Using Machine Learning

1. Project Overview

The **FWI Prediction System** is a machine learning project developed to predict the **Fire Weather Index (FWI)** based on meteorological and environmental factors. The system helps estimate wildfire risk using weather data, supporting early warning and disaster prevention mechanisms.

This project uses **Linear Regression** and **Ridge Regression** models to analyze and predict FWI values accurately.

2. Objectives

The objectives of this project are:

- To analyze weather-related fire data.
- To perform data preprocessing and feature engineering.
- To build and train regression models.
- To evaluate model performance using standard metrics.
- To optimize the model using Ridge Regression and alpha tuning.
- To visualize model performance using graphs.

3. Dataset Description

Dataset Used:

Algerian_forest_fires_dataset.csv

Important Columns:

Column Name	Description
Temperature	Ambient atmospheric temperature
RH	Relative Humidity
Ws	Wind Speed
Rain	Rainfall
FFMC	Fine Fuel Moisture Code

Column Name	Description
DMC	Duff Moisture Code
DC	Drought Code
ISI	Initial Spread Index
BUI	Buildup Index
FWI	Fire Weather Index (Target variable)
Classes	Fire / Not Fire
Region	Geographical region

4. Technologies Used

Component	Technology
Programming Language	Python
IDE	VS Code
Libraries	Pandas, NumPy, Matplotlib, Seaborn, Scikit-learn, Joblib

5. System Architecture

Flow of the system:

1. Load dataset
2. Data cleaning and preprocessing
3. Feature selection
4. Feature scaling
5. Train-Test split
6. Linear Regression training
7. Ridge Regression training
8. Model evaluation
9. Graph visualization
10. Save trained model

6. Data Preprocessing

Steps performed:

- Removed duplicate rows
- Handled missing values
- Converted categorical columns into numeric values
- Generated statistical summaries

After preprocessing, the cleaned dataset is saved as:

cleaned_fwi_dataset.csv

7. Exploratory Data Analysis (EDA)

The following visualizations were generated:

- **Histogram** – to analyze data distribution
- **Correlation Heatmap** – to identify relationships between features

These helped in selecting the most influential features for FWI prediction.

8. Feature Engineering

Selected input features:

Temperature, RH, Ws, Rain,FFMC, DMC, DC, ISI, BUI

Target variable:

FWI

9. Model Training

9.1 Linear Regression

Linear Regression was used as the baseline model to learn the relationship between weather parameters and FWI.

9.2 Ridge Regression

Ridge Regression was used to handle multicollinearity and reduce overfitting by adding regularization.

Different alpha values tested:

0.01, 0.1, 1, 10, 100

10. Model Evaluation

The following metrics were used:

Metric	Description
MAE	Mean Absolute Error
MSE	Mean Squared Error
RMSE	Root Mean Squared Error
R^2	Coefficient of Determination

Performance Summary:

Linear Regression showed excellent results with:

- High R^2 score (≈ 0.98)
- Low error values

Ridge Regression performance depended on alpha:

- Best alpha = **0.01**
- Higher alpha values caused underfitting

11. Overfitting and Underfitting Analysis

Criteria used:

- If Training MSE ≈ Testing MSE → Well-fit model
- If Training MSE << Testing MSE → Overfitting
- If both Training and Testing MSE are high → Underfitting

Observation:

- Low alpha → Good fit
- High alpha → Underfitting

12. Model Visualization

A graph was plotted:

MSE vs Alpha (Ridge Regression)

This helped identify the optimal alpha value visually.

13. Model Saving

The trained models were saved using joblib:

File Name	Purpose
scaler.pkl	Saved StandardScaler
ridge_model.pkl	Saved Ridge Regression model

14. Conclusion

The Fire Weather Index Prediction System was successfully developed and evaluated. The Linear Regression model achieved high accuracy, and Ridge Regression helped optimize model performance. The system demonstrated strong generalization and can be used as an effective predictive tool for wildfire risk assessment.