Project Title: Machine Learning Based Forest Fire Analysis

Dataset Source: The dataset was obtained from the Machine Learning Repository at the University of California, Irvine.
http://archive.ics.uci.edu/ml/machine-learning-databases/forest-fires/forestfires.csv

Description: There are 13 attributes and **a total of 517 rows** in the above dataset, each attribute corresponds to a particular trait associated with forest fires in the Montesinho park.

Attribute Descriptions:

1. *X and Y (Spatial Coordinates):*
   i) X: The spatial coordinate along the x-axis in the Montesinho park map, which ranges from 1 to 9.
   ii) Y: The spatial coordinate along the y-axis in the Montesinho park map, which ranges from 2 to 9. Every data point in the park has its position indicated by these coordinates.
2. *Month:* This indicates the month of the year the data was collected; January through December are denoted by the letters "jan" through "dec." The seasonality of forest fires is captured by this feature.
3. *Day:* Marked as 'mon' for Monday through 'sun' for Sunday, this field indicates the day of the week when the data was collected. This characteristic aids in comprehending how weekdays affect the frequency of forest fires.
4. *FFMC (Fine Fuel Moisture Code):* The Fire Weather Index (FWI) system's FFMC index, which has a value range of 18.7 to 96.20. It measures the moisture content of fine fuels, which are materials that ignite easily, such as little twigs, grass, and leaves. Higher numbers signify drier weather, which raises the possibility of a fire.
5. *DMC (Duff Moisture Code):* The FWI system's DMC index, which has a value range of 1.1 to 291.3. It gauges the amount of moisture in the organic layer underneath the forest floor that is breaking down. Higher DMC readings indicate organic material that is dryer, which may intensify the fire.
6. *DC (Drought Code):* The FWI system's DC index, which has a value range of 7.9 to 860.6. It depicts the deficiency of moisture in the soil's deeper layers. Increased soil dryness, which might influence fire behavior, is shown by higher DC values.
7. *ISI (Initial Spread Index):* The FWI system's ISI index, which has values between 0.0 and 56.10. The potential rate of fire spread is measured by ISI. Higher values under certain conditions imply a faster beginning spread of the fire.
8. *Temp (Temperature):* This indicates the temperature in degrees Celsius, with a range of 2.2 to 33.30. It offers data about the temperature of the surrounding area, which may have an impact on how easily a fire may start and behave.

9. *RH (Relative Humidity):* This indicates the percentage of relative humidity, with a range of 15.0 to 100. Lower values of relative humidity indicate drier air and a higher fire danger, which is one of the major factors influencing fire behavior.
10. *Wind (Wind Speed):* This variable ranges in value from 0.40 to 9.40 and represents wind speed in kilometers per hour (km/h). Higher winds can expedite a fire, hence wind speed is important in determining how quickly a fire spreads.
11. *Rain (Rainfall):* A measurement of the amount of precipitation outside expressed in millimeters per square meter (mm/m2), with a range of 0.0 to 6.4. Rainfall can reduce the chance of a fire by making the air more wet.
12. *Area (Burned Area):* This variable, which ranges from 0.00 to 1090.84, indicates the burned area of the forest in hectares (ha). The goal variable for this characteristic is the amount of forest destroyed during a fire occurrence.

The numerous environmental elements and variables that may have an impact on forest fires and their propagation are explained in detail by these attribute descriptions.

Sample Rows from the Dataset:

X,Y,month,day,FFMC,DMC,DC,ISI,temp,RH,wind,rain,area
1, 2, 'mar', 'fri', 86.2, 26.2, 94.3, 5.1, 8.2, 51, 6.7, 0.0, 0.00
4, 4, 'oct', 'tue', 90.6, 35.4, 669.1, 6.7, 18.0, 33, 0.9, 0.0, 0.00
7, 4, 'aug', 'sat', 90.2, 110.9, 559.7, 6.0, 18.7, 45, 4.0, 0.0, 0.00
7, 4, 'aug', 'sun', 92.3, 88.9, 495.6, 8.5, 24.1, 29, 1.3, 0.0, 0.00
8, 6, 'aug', 'sun', 91.5, 145.4, 608.2, 10.7, 8.8, 30, 2.8, 0.0, 0.00

There are **a total of 517 rows** in the dataset.

Goal:
Developing a predictive model with machine learning approaches to comprehend and forecast the magnitude of forest fires based on environmental conditions is the main goal of this investigation. Our goal is to obtain an understanding of the correlations between temperature, humidity, wind speed, and other variables, and the resulting burned area measured in hectares, by utilizing the qualities that have been provided. By assisting authorities and stakeholders in making well-informed decisions to lessen the impact of forest fires in wooded areas, this information can be extremely valuable for fire prevention and management measures.