A COMPARATIVE STUDY OF FOREST FIRE PREDICTION USING MACHINE LEARNING MODELS

by

Mr. Sachin Malego (st125171) & Mr. Sila N. Mahoot (st125127)

A Project Proposal Submitted in Partial Fulfillment of the Requirements for the Degree of Master of Science in Data Science & AI

Submitted to: Dr. Chantri Polprasert
Assistant Professor
CSIM, Department of ICT

Nationality: Nepalese and Thai

Asian Institute of Technology

CSIM, Department of Data Science & AI

Thailand

October 2024

ABSTRACT

Forest fires are one of the most pressing incidents that have a significant impact to both the environment and human life, making early prediction is crucial for minimizing the damage. This proposal presents a comparative study of various machine learning models for forest fire prediction leveraging the historical data from various dataset. Techniques such as XGBoost, KNN, Decision Tree, Random Forest, and Linear and Logistic Regression are explored, with models evaluated based on accuracy and other performance metrics. The study highlights the strength and weakness of each approach with geographic maps of the country, demonstrating that machine learning can significantly enhance predictive capabilities, providing a valuable tool for fire prevention and disaster management. This project does not determine if a forest fire will take place or not, however we are predicting the confidence of the forest fire based on some attributes.

CONTENTS

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

With the high incident rates of forest fires, its prediction holds a significant environmental and scientific importance.[1] In recent years, forest fires have emerged as a critical concern, posing a recurring threat to vast forested areas worldwide. Observing the data from Nepal alone, it has experienced 965 forest fire incidents in the last one year (08-10-2023 to 07-10-2024) with an estimated economic loss of around USD. 644,696. Such is not only the scenario of Nepal but at a global level. Here in this project, we review such dataset of forest fire from Forest fire dataset of Algeria, and Portugal and run different models to make the best possible prediction.

1.1 Background

Forest fire forecasting is critical for mitigating the environmental, economic, and public health impacts of wildfires. According to existing literature, forest fire prediction methods can be broadly categorized into three main approaches: physics-based models, statistical models, and machine learning models.[2] Physics-based models rely on simulating fire behavior based on physical laws, such as fluid dynamics and thermodynamics. Statistical models focus on identifying correlations between past fire events and contributing factors, often using regression techniques. However, machine learning (ML) models have recently gained prominence due to their ability to learn complex, non-linear relationships from large datasets, making them highly suitable for dynamic and uncertain phenomena like forest fires.

This project leverages data-driven algorithms in the machine learning domain to predict the likelihood of forest fire occurrences. Our goal is to train ML models on historical weather and environmental data and use these models to predict fire risks based on current conditions. The data, sourced from regions in Portugal and Algeria, contains key variables such as temperature, relative humidity, wind speed, and climate events like El Niño. These variables are integral to capturing the complex interplay of factors that contribute to fire ignition and spread.

Before delving into the technicalities of forest fire prediction, it's crucial to explore the climatic conditions that exacerbate wildfire risks. El Niño is a particularly significant

factor, characterized by the abnormal warming of sea surface temperatures in the central and eastern Pacific Ocean. Although El Niño is a Pacific-based event, it has far-reaching implications for global weather patterns, leading to drier or wetter conditions in various regions, including Europe and North Africa. In both Portugal and Algeria, El Niño can disrupt typical weather patterns, resulting in reduced rainfall, prolonged dry spells, and increased temperatures - conditions that are conducive to the ignition and spread of forest fires.

In addition to El Niño, several other environmental and climatic variables influence fire risk, including rainfall patterns, wind speed, relative humidity, fine fuel moisture, duff moisture, and drought conditions. These factors, when integrated into machine learning models, can provide a powerful predictive tool for early fire detection and risk assessment.
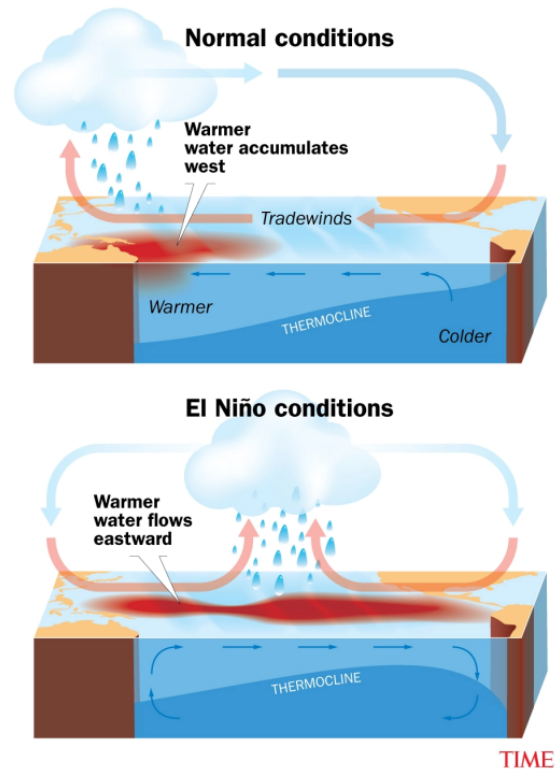

Figure 1: Understanding El Nino

1.2 Objective

The objective part in these answers the why part of need of forest fire prediction. The objective of forest fire prediction is to improve early detection and risk assessment, a crucial need globally. Despite the experience of forest departments, human limitations in processing multiple variables hinder accurate predictions. Hence in such situations Machine learning can analyze numerous factors simultaneously, making fire prediction more efficient and effective. The major objectives of this project are:

i.   Evaluate Model Performance: Compare the accuracy, precision, recall, F1 score and overall predictive performance of various machine learning models in predicting forest fires.

ii.   Identify Key Features: Analyze the importance of various environmental and weather-related factors (e.g., temperature, humidity, wind speed, El Niño) across different models to identify which features contribute most to fire prediction accuracy.

2

iii. Assess Generalization Capability: Investigate how well each machine learning model generalizes to unseen data, particularly when predicting forest fires in different geographic regions or under varying climate conditions.

iv. Model Interpretability and Usability: Evaluate the interpretability of each model, determining how easily stakeholders (e.g., forest management agencies) can understand and utilize the model's predictions for decision-making.

v. Predictive Alerts for Health Precautions: Evaluate how machine learning models can provide early fire risk alerts, enabling timely health interventions, such as air quality warnings or evacuation notices, to protect vulnerable populations (e.g., children, elderly, people with pre-existing conditions).

vi. Air Quality and Environmental Health: Investigate how machine learning-driven forest fire predictions can contribute to improving air quality by preventing large-scale fires that release pollutants like carbon monoxide and nitrogen oxides, which have harmful effects on the environment.

vii. Support Sustainable Forest Management: Explore how machine learning predictions can aid in sustainable forest management by allowing for preemptive actions like controlled burns or forest thinning to mitigate the severity of future fires, reducing long-term environmental damage.

1.3 Business Understanding

With climate change intensifying conditions such as droughts, and unpredictable weather patterns, forest fires are becoming more frequent and destructive. Effective prediction is crucial but challenging due to the complexity of interacting factors.

This project will focus on forest fire prediction in regions prone to fires, specifically Portugal and Algeria. It will utilize weather and environmental data from these regions and compare multiple machine learning models to assess their performance in predicting fire occurrences. The project will evaluate the impact of different variables such as temperature, humidity, wind, and the presence of El Niño on fire risk prediction.

1.4 Impact

The proposed study on forest fire prediction using machine learning models has the potential to generate significant and far-reaching impacts across multiple sectors. These impacts can be categorized as environmental, social, economic, technological, policy-

related, and scientific, all contributing to a more informed and effective approach to forest fire management.

The comparative study of forest fire prediction using machine learning models has the potential to significantly impact environmental conservation, public safety, economic stability, technological progress, policy development, and scientific advancement. By addressing a critical global challenge, this research will contribute to a safer and more resilient world in the face of increasing forest fire risks due to climate change.

CHAPTER 2
PROBLEM STATEMENT

The increasing frequency and intensity of forest fires due to climate change pose significant threats to ecosystems, human health, and economies worldwide. To address this pressing issue, there is a need to develop accurate predictive models that can forecast forest fire occurrences based on historical data and environmental conditions. This study aims to investigate the relationships between various factors contributing to forest fires, including:

- Temporal Trends: Analyzing seasonal, and monthly trends in forest fire occurrences to identify patterns and peak risk periods.
- Correlations Between Environmental Parameters: Examining the correlations between key weather parameters (e.g., temperature, humidity, wind speed) and the incidence of forest fires to understand how these factors influence fire risk.
- Impact of Climate Phenomena: Assessing the effects of significant climate phenomena, such as El Niño, on forest fire occurrences, providing insights into how broader climatic trends affect fire behavior.
- Comparative Analysis of Machine Learning Models: Evaluating the performance of different machine learning algorithms in predicting forest fire risk, determining which model offers the best accuracy and reliability for real-time decision-making.
- Visualization: Visualize impact areas in maps for prediction and effective decision making.

By comprehensively analyzing these elements, the project seeks to enhance forest fire prediction capabilities, enabling more effective prevention and response strategies to mitigate the environmental and societal impacts of forest fires. The main goal of this project is to create a reliable machine learning model that can predict the chances of forest fires based on weather and environmental data. By using data from Portugal and Algeria, this project aims to solve the problems with traditional prediction methods and provide a tool to help authorities make better decisions for preventing and managing

fires. The challenge is to make a model that can handle different weather patterns and conditions, ultimately helping reduce the damage caused by forest fires.

CHAPTER 3
RELATED WORKS

The project goal is to train the model to be able to predict the forest fire. Here the focus is on the use of different factors along with climate change phenomenon of El Niño to generalize and predict the fire occurrence scenario for better decision making.

Jing et al., in the research paper "***Toward a more resilient Thailand: Developing a machine learning-powered forest fire warning system***" describes a machine learning-based forest fire warning system in Thailand. It uses satellite data and gas measurements to predict forest fires, with models like linear classifiers, gradient boosting classifiers, and neural networks.[3] The XGBoost model had the best performance with an accuracy of 99.6%. Our project aims to improve this approach by adding climate factors like El Niño and testing more models across different regions.

Abdelhamid Zaidi in the paper "***Predicting wildfires in Algerian forests using machine learning models***" focuses on developing a predictive system for forest fires in Algeria, a region that has seen an increase in fire occurrences over recent years.[4] The study aims to create an accessible, low-budget system for predicting wildfires based on climatic data and machine learning algorithms. The research highlights the challenges faced in forest fire prediction, such as the complexity of fire behavior, the need for substantial and accurate datasets, and the necessity for models to be adaptable to various conditions. The author utilizes machine learning algorithms, to analyze weather data and predict potential fire occurrences. The study concludes that a well-designed machine learning-based system could significantly benefit local authorities and the general public by enabling early fire detection and facilitating timely interventions.

In another paper from IEEE Viswa et al., "***Comparison of Forest Fire Prediction System using Machine Learning Algorithms***" presents a comparative analysis of various machine learning algorithms for predicting forest fires. It emphasizes the importance of early detection in minimizing the devastating effects of wildfires on ecosystems and human settlements.[5] The study evaluates the performance of different algorithms based on accuracy, precision, recall, and other relevant metrics. The conclusion iterates the importance of machine learning in forest fire prediction and

suggests future research directions, including the integration of additional data sources and advanced algorithms.

CHAPTER 4
DATASETS

4.1 Description

In this project, we are using two main datasets from Kaggle and one weather data from another public source via web scraping. These datasets represent key information about forest fire occurrences and influencing factors. The datasets are in structured format in EXCEL spreadsheet.

- Forest Fires Dataset – Algeria
- Forest Fires Dataset – Portugal
- Weather (Meteorological) Dataset

Along with this dataset if feasible we are also trying to study El Niño if it has any significane in forest fire occurrence although the period for this is very low.

Figure 2: Area of Interest

4.2 Features

The dataset contains following features:

1) *Algerian Forest Fires Dataset*: This dataset will be used for **training** the model. The dataset includes 244 instances that regroup a data of two regions of Algeria, namely the Bejaia region located in the northeast of Algeria and the Sidi Bel-abbes region located in the northwest of Algeria. 122 instances for each region. The period from June 2012 to September 2012. The dataset includes 11 attributes and 1 output attribute (class). The 244 instances have been classified into fire (138 classes) and not fire (106 classes) classes.

   a) Date: The date of the observation.

   b) Temperature: Daily average temperature in degrees Celsius.

   c) Humidity: Daily average relative humidity as a percentage.

   d) Wind Speed: Daily average wind speed in km/h.

   e) Rainfall: Daily rainfall in mm.

   f) Region: The specific region in Algeria where the data was collected.

   g) Fire Occurrence: Binary indicator (Yes/No) for whether a fire occurred.

2) *Forest Fires Dataset - Portugal*: This dataset will be used for **testing** the model. The dataset includes 517 instances. The dataset includes 13 attributes and 1 output attribute (class).

   a) Date: The date of the observation.

   b) Temperature: Daily average temperature in degrees Celsius.

   c) Humidity: Daily average relative humidity as a percentage.

   d) Wind Speed: Daily average wind speed in km/h.

   e) Rainfall: Daily rainfall in mm.

   f) FFMC (Fine Fuel Moisture Code): Index representing the moisture content of surface litter.

   g) DMC (Duff Moisture Code): Index indicating the moisture content of loosely compacted organic layers.

   h) DC (Drought Code): Index indicating long-term drying conditions.

   i) ISI (Initial Spread Index): Index for predicting the rate of fire spread.

   j) Fire Occurrence: Binary indicator (Yes/No) for whether a fire occurred.

3) *Weather and Environmental Data*: These are the additional factors which may have some role in the fire. This dataset will be mapped with the other two on basis of timestamp and location.

   a) Temperature: Daily average temperature in degrees Celsius.

   b) Humidity: Daily average relative humidity as a percentage.

   c) Wind Speed: Daily average wind speed in km/h.

   d) Precipitation: Daily rainfall in mm.

4) *El Niño Data (Sea Surface Temperature):* This dataset is used only for observing if it has any significance or not. This is not a required dataset however we wanted to have a look in to this as well to see its significance if any.

   a) Niño 3.4 Index: Represents Sea surface temperature anomalies in the central equatorial Pacific, which is a critical indicator of El Niño events.

The figure below depicts the El Nino occurrence and average temperature of the 3 months of data from the Algerian dataset, it tries to map the El Nino occurrence with temperature observed in that period.
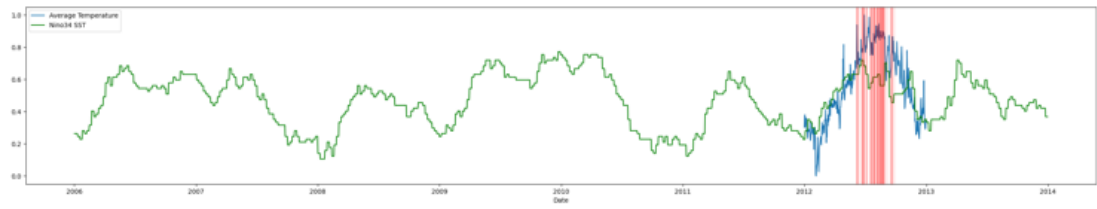
Figure 3: El Nino occurrence and Average Temperature

CHAPTER 5
METHODOLOGY

5.1 Data Acquisition

The data for this project was collected from publicly available sources on Kaggle and other weather data platforms. The main datasets are the Algerian Forest Fires Dataset and the Forest Fires Dataset - Portugal, which have valuable information on past forest fires and weather conditions. The data is provided in CSV format, making it easy to load and work with using Python libraries like Pandas.

We also got weather data for Algeria and Portugal from Meteostat and Weather Underground. Meteostat provides historical weather metrics, while Weather Underground allows us to gather more recent daily weather data using web scraping. This weather data is very important for understanding the factors that contribute to forest fires.

To understand the role of climate events like El Niño, we also used El Niño Sea surface temperature data from the Climate Prediction Center (CPC). This helps us see how larger climate patterns might influence forest fire occurrences. All the data was downloaded or accessed through APIs, then integrated for further exploration and analysis.

5.2 Exploratory Data Analysis

Exploratory Data Analysis (EDA) helps us understand the patterns and relationships in our data. For this project, we are doing EDA on the Algerian and Portugal datasets, as well as weather and El Niño data, to find trends that might predict forest fires.

The following steps and best practices will be applied during the EDA process:
1) Data Cleaning

   We will check for missing values, mistakes, and inconsistencies. Missing data will either be filled in or removed based on its importance.

2) Data Transformation

We will turn date features into more useful information, like year or month, to help find seasonal trends.

3) Data Visualization

We will use different types of plots to understand the data better:

 a) Line Plots to show trends over time, like changes in temperature and humidity.

 b) Histograms show how individual features are distributed.

 c) Scatter Plots to look at relationships between key features, like temperature and fire occurrences.

 d) Correlation Heat Maps to find which features are most important for predicting fires.

 e) Box Plots to find and deal with outliers in the data.

EDA will give us important insights that will help us prepare the data, choose features, and build better models.

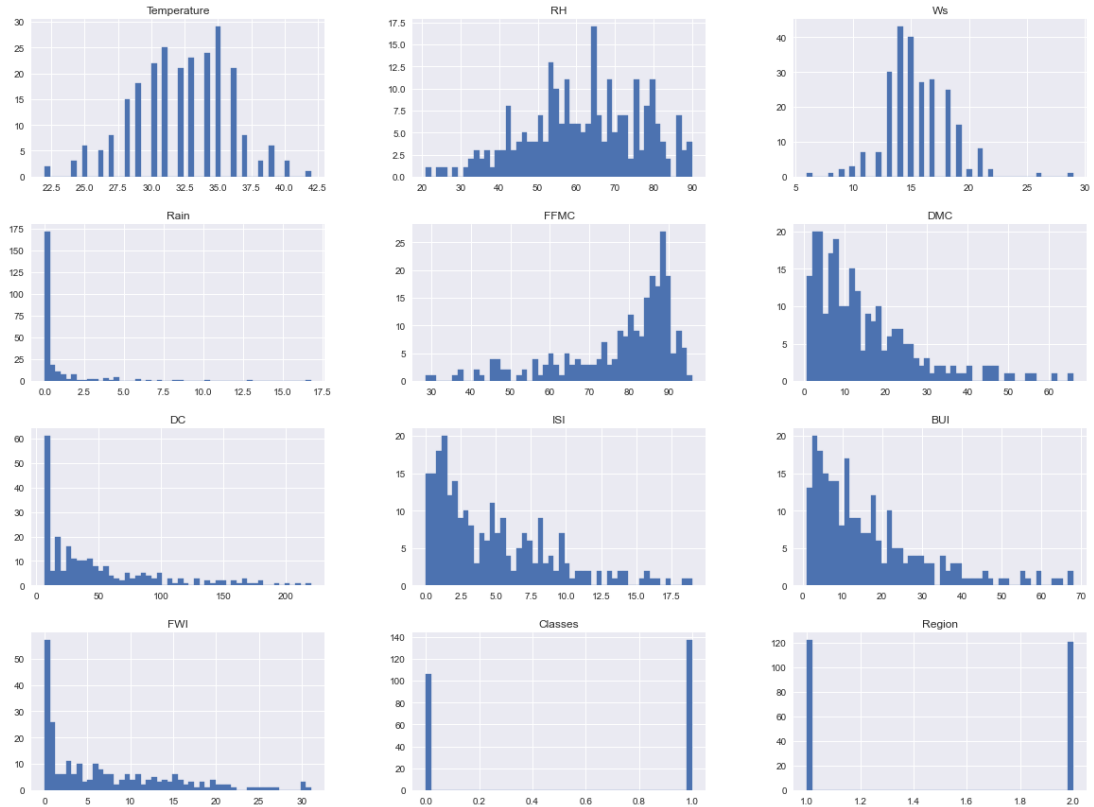The histogram provides us with some idea on the Algerian fire dataset.

Figure 4: Histogram plot of Algerian fire dataset

a) Temperature: The distribution of temperature values appears roughly symmetric, with the most frequent values between 30 and 35 degrees Celsius.

b) RH (Relative Humidity): The distribution is more spread out, with the highest frequency around 50-60% relative humidity. There are still some values between 30 and 90%.

c) Ws (Wind Speed): This has a peak around 15, meaning wind speed typically clusters around this value, and fewer data points appear at lower and higher speeds.

d) Rain: The data for rain is highly skewed to the left, with most values close to zero, indicating very few instances of significant rainfall.

e) FFMC (Fine Fuel Moisture Code): The distribution shows that most values are concentrated around 80 to 90, meaning most of the observations have high FFMC values.

f) DMC (Duff Moisture Code): The distribution is skewed to the left with a peak around 0 to 10, meaning the majority of the observations have low DMC values, with fewer instances having higher values.

g) DC (Drought Code): The DC data also exhibits a skewed left distribution, with most values clustering around the lower end of the scale but extending up to 200.

h) ISI (Initial Spread Index): This distribution is left-skewed, with the majority of values between 0 and 5. There are few instances of higher ISI values.

i) BUI (Buildup Index): This shows a similar left-skewed distribution, with a peak in lower values (0 to 20), indicating lower BUI is common.

j) FWI (Fire Weather Index): The distribution shows most values are concentrated between 0 and 10, with a steep drop-off after that.

k) Classes: This likely represents some form of classification (e.g., fire severity), and the histogram shows that one class (possibly 0) dominates the data, while other classes have fewer occurrences.

l) Region: This variable has only two distinct categories, with one region (possibly represented by 1.0) dominating the data.

Key Observations in the histograms:

- Many of the features (Rain, DC, ISI, BUI, etc.) show a left-skewed distribution, meaning most values are low, with a few high outliers.

- Temperature and wind speed (Ws) show more symmetric distributions.
- The dominance of one class in the "Classes" and "Region" histograms could imply an imbalanced dataset for those categorical variables.

The **heatmap** of multicollinearity check in the Algerian fire dataset provides some insight information in the relationship between features in the dataset.

Key Observations in the heatmap:

- Strong positive correlations between fire-related indices (e.g., BUI, DC, FWI, ISI) suggest that these factors increase together, affecting fire severity.

Figure 5: Multicollinearity check in Algerian fire dataset

- Negative correlations between relative humidity and fire indices (FWI, ISI, FFMC) highlight the role of moisture in reducing fire spread risk.
- Temperature and certain moisture indices (DMC, FFMC) show moderate correlations, indicating a relationship between weather conditions and fire risk.
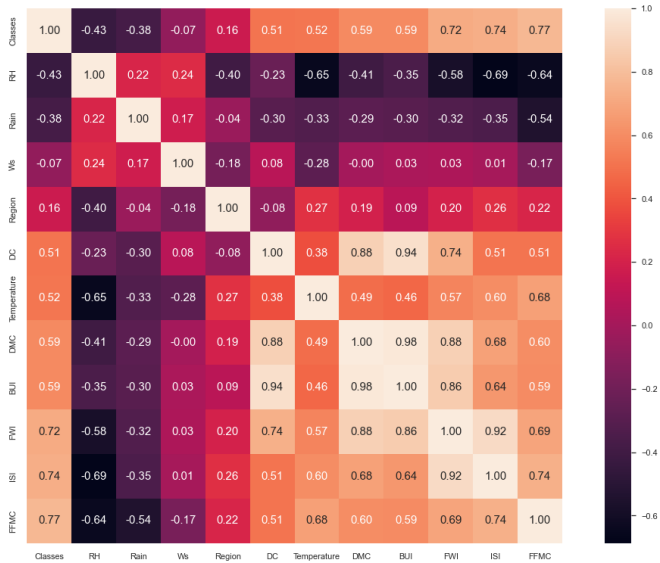
This **bar plot** displays a month-wise fire analysis for the Bejaia region, comparing the occurrence of fire and non-fire incidents across the months of June, July, August, and September. Observation in the bar graph:
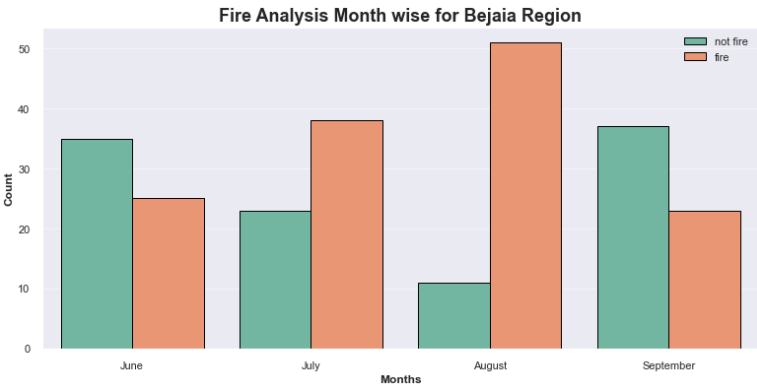

Figure 6: Monthly fire analysis of Algerian fire dataset

- August stands out as the peak month for fire incidents in the Bejaia region, with significantly more fire events compared to non-fire events.
- In contrast, June and September have more non-fire events, while July shows a similar count for both fire and non-fire events.

This plot helps understand the seasonal variation of fire occurrences, showing that fire incidents are more likely in the late summer (July and August) in the Bejaia region. This could indicate that environmental factors, like temperature and dry conditions, are more conducive to fires during these months.

In this **bar graph** there is a clear increase in fire occurrences as temperatures rise, peaking between 34°C and 35°C. After this peak, the number of fires decreases, indicating that the highest fire risk occurs at these temperature levels. From this we can conclude
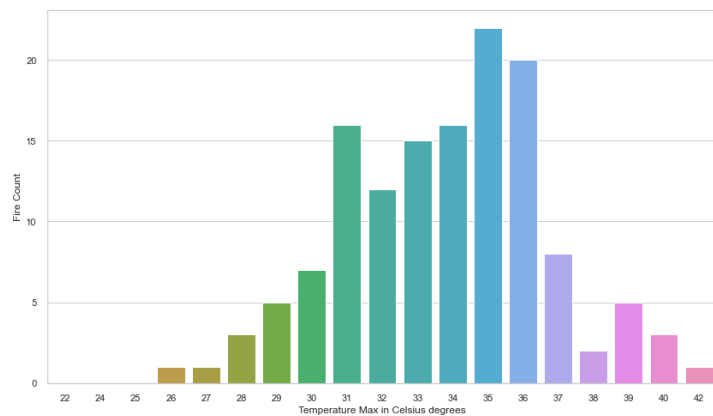

Figure 7: Temperature data representation of Algerian fire dataset

that, high temperatures, especially around 34-35°C, significantly contribute to fire incidents, likely due to the drying out of vegetation and increased flammability.

We can observer that almost all fires occur when there is **very little to no rainfall** (below 0.2 mm). After this point, the fire count drops dramatically, with almost no fires occurring as rainfall increases. Hence, it is conclusive that drier conditions (low rainfall) are the most critical
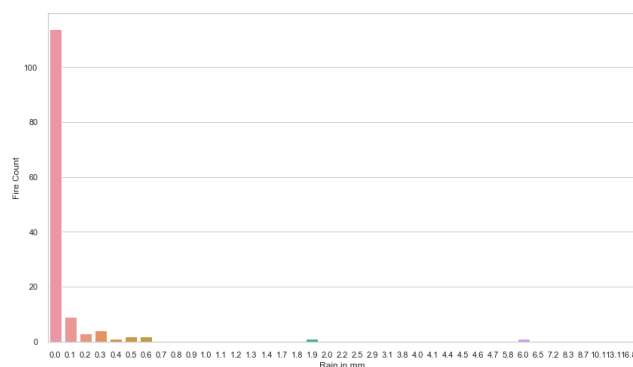

Figure 8: Rainfall in mm in Algerian fire dataset

factor for fire outbreaks. Rainfall appears to have a strong dampening effect on the number of fires, as more moisture makes ignition less likely.

Here, fire activity peaks when **wind speeds** are between 13 km/h and 16 km/h. Beyond this range, the fire count decreases. Moderate wind speeds seem to enhance fire spread by providing enough airflow to fuel the flames, while lo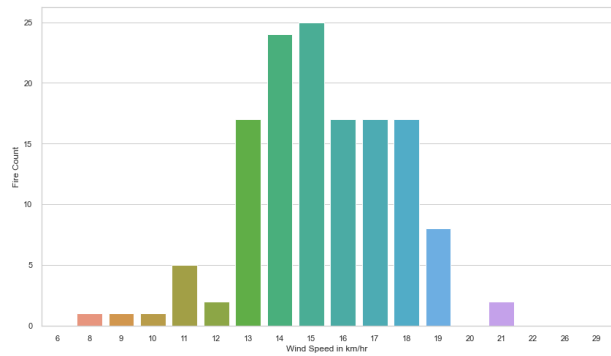wer and higher wind speeds may limit fire occurrences. Very high winds might disperse fire sources, while lower winds may not provide enough oxygen for the fire to spread rapidly.

Figure 9: Wind speed

The graph suggests that relative humidity is a significant factor in determining the number of fires. There seems to be a general inverse relationship 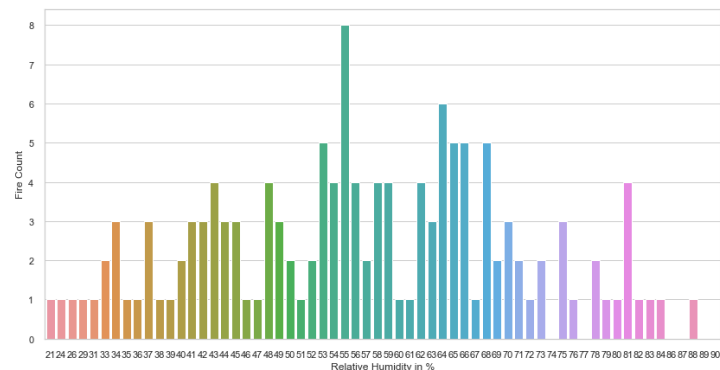between relative humidity and the number of fires. As the relative humidity increases, the number of fires tends to decrease. This suggests that higher humidity levels may have a suppressing effect on fires. There are noticeable clusters of fires at certain relative humidity levels, such as around 30-35% and 60-70%. This might indicate that specific weather conditions or factors associated with these humidity levels are more conducive to fires. The bar corresponding to 55% relative humidity stands out with an exceptionally high number of fires. This could be due to a specific event or circumstance that caused a significant spike in fires at that particular humidity level.

Figure 10: Relative humidity

5.3 Pre-processing

Data preprocessing is important for making sure our datasets are clean and ready for training. Here are the steps we will take:

1) Handling Missing Values: Missing values will be filled in or removed, depending on their impact.

2) Encoding Categorical Variables: We will turn categorical features, like region, into numerical data using one-hot or label encoding.

3) Feature Scaling: We will normalize or standardize features like temperature and wind speed to make sure they're on a similar scale.

4) Date Feature Extraction: Dates will be transformed to get useful information like month or day of the week.

5) Handling Outliers: We will find and handle outliers so they don't affect our model badly.

6) Data Integration: We will combine all datasets—fire data, weather data, and El Niño data—into one comprehensive dataset for modeling.

7) Feature Engineering: We will create new features that could improve the model's performance, like combining temperature, humidity, and wind speed.

8) Data Balancing: Since forest fires are rare, we will balance the dataset so the model can learn well from both fire and non-fire cases.

9) Splitting Data: The data will be split into training, validation, and test sets to make sure our model performs well on new data.

These preprocessing steps will help make sure our data is suitable for training a reliable machine learning model.

5.4 Modeling

To find the best model for predicting forest fires, we will use **cross-validation** and grid search to compare different machine learning algorithms and tune their settings. Cross-validation

Figure 11: K-Fold cross validation

will help make sure the models can work well on new data, and grid search will help find the best settings for each model.

1) Random Forest and XGBoost are good options because they handle many features and complex relationships well.

2) Gradient Boosting is also strong for modeling interactions between features.

18

3) Regression models will be used as a simple baseline model to compare with the others.

We will compare these models based on accuracy, precision, recall, and F1 score, then choose the best one for predicting forest fires.

5.5 Training

Training the model means using the training data to teach the model how to predict forest fires. We will use the scikit-learn library to train our models. The model will learn by finding the relationships between input features, like temperature and wind speed, and the target feature, fire occurrence.

Steps in Training:
1) Model Initialization: Start the model with the best settings found in grid search.
2) Fitting the Model: Train the model with the .fit() method.
3) Monitoring Performance: Keep track of how well the model is learning by checking the loss function and metrics like accuracy.
4) Addressing Overfitting: Use methods like early stopping to prevent the model from overfitting.

After training, we will evaluate the model with the validation set. If needed, we will adjust the settings to improve performance, then test the final model with new data.

5.6 Evaluation

We will evaluate the model to make sure it accurately predicts forest fires. We will use these metrics:

- Accuracy: How often the model makes the correct prediction.
- Precision: How often the model correctly predicts a fire when it says there is one.
- Recall: How many actual fires the model can predict.
- F1 Score: A balance between precision and recall.

These metrics will help us understand how well the model works and what areas need improvement.

5.7 Deployment

The forest fire prediction model will be deployed using Azure Machine Learning for hosting and Dash for the web-based interface. Azure Machine Learning provides a safe and scalable way to host machine learning models.
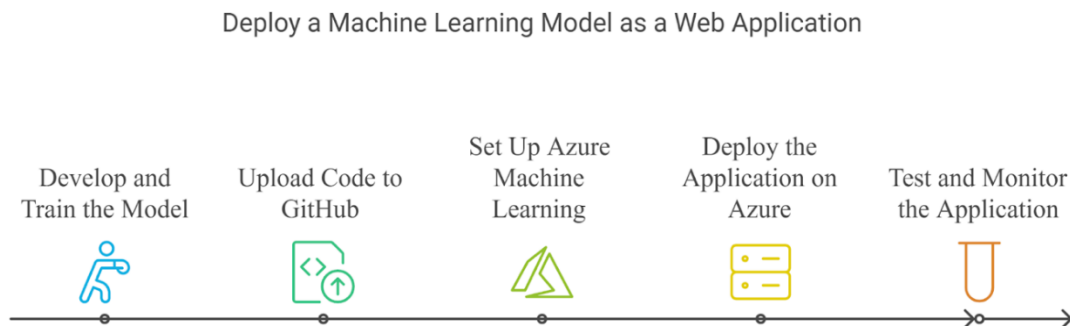

Figure 12: Web Application Deployment

Here are the steps for deployment:

1) Create Prediction Model and Web Application Locally

   a) Develop ML Model (model.py): Build the model in a Python script named model.py.

   b) Convert Model to Deployable Format: Save the trained model using pickle or joblib to prepare it for deployment.

   c) Create Dash Web Application (app.py): Make a Dash web app that lets users input data and see predictions, including maps showing fire risk.

2) Upload Code to GitHub: The relevant codes, including the model.py and app.py scripts, along with dependencies, will be uploaded to GitHub for version control and collaboration.

3) Azure Machine Learning Setup

   a) Create Azure Account: Use Azure to access Machine Learning services.

   b) Register Model on Azure: Register the model in Azure to manage it easily.

   c) Create an Inference Endpoint: Create an endpoint that the Dash app can use to get predictions from the model.

   d) Containerize the Model: Use Docker to containerize the model, making it easier to deploy.

4) Deploy Dash Application on Azure Web App

   a) Create Azure Web App Service: Use Azure App Service to host the Dash app.

b) Configure Azure App Service: Connect the web app to the model endpoint.

c) Clone Repository and Deploy: Clone the GitHub repository to Azure and deploy the app.

5) Testing and Monitoring

a) Test the Deployment: Make sure the deployed app works well and interacts properly with the model.

b) Monitor the Application: Use Azure's monitoring tools to track performance and fix any issues.

By using Azure Machine Learning and Dash, we will make the forest fire prediction system accessible online. This system will help users input data and see predictions in real time, including maps that show areas at risk of forest fires.

CHAPTER 6
PRELIMINARY RESULTS

This chapter presents preliminary results of the analysis conducted on various environmental factors that influence forest fire occurrence. The findings provide insights into the conditions under which forest fires are most likely to occur, helping to identify critical thresholds for different parameters that contribute to fire risk.

- Temperature: The highest fire counts occurred between 30-37 degrees Celsius.

- Rain: The highest fire counts occurred when there was no rain or very little rain, i.e., between 0.0 to 0.3.

- Wind Speed: The highest fire counts occurred when the wind speed was between 13 to 19 km/hr.

- Relative Humidity: The highest fire counts occurred when the relative humidity (RH) was between 50 to 80%.

- Fine Fuel Moisture Code (FFMC) index, which ranges between 28.6 to 92.5, shows that above 75 there is a higher chance of forest fires.

- Duff Moisture Code (DMC) index, which ranges between 1.1 to 65.9, indicates that values between 1.1-10 have a lower chance of forest fires, whereas above 10-30 DMC has very high evidence of forest fires in the past.

- Drought Code (DC) index, which ranges between 7 to 220.4, shows that values between 0-25 are safe and have a lower chance of forest fires, whereas values above 25 DC have a higher chance of forest fires.

- Initial Spread Index (ISI) index, which ranges between 0 to 18, shows that values between 0-3 have lower fire chances, and values above 3 ISI have a higher chance of forest fires.

- Buildup Index (BUI) index, which ranges between 1.1 to 68, shows that values between 1.1 to 10 have a lower fire chance, and values above 10 BUI have a higher chance of forest fires.

- Fire Weather Index (FWI) index, which ranges between 1 to 31.1, shows that values between 0-3 have a lower chance of forest fires, and values between 3-25 FWI have a higher chance of forest fires.

REFERENCES

[1]   Sanjeev Sharma, Puskar Khanal 2024. Forest Fire Prediction: A Spatial Machine Learning and Neural Network Approach. Fire, Available at: https://www.mdpi.com/2571-6255/7/6/205 [Accessed 7 Oct. 2024].

[2]   Abdelhamid Zaidi, 2023. Predicting wildfires in Algerian forests using machine learning models. Heliyon. Available at: https://www.sciencedirect.com/science/article/pii/S2405844023052726 [Accessed 7 Oct. 2024]

[3]   Tang, J. et al. (2024). Toward a more resilient Thailand: Developing a machine learning-powered forest fire warning system', Heliyon. Available at: https://www.sciencedirect.com/science/article/pii/S2405844024100527?via%3 Dihub [Accessed 7 Oct. 2024]

[4]   Zaidi, A. (2023). Predicting wildfires in Algerian forests using machine learning models. Heliyon. Available at: https://www.sciencedirect.com/science/article/pii/S2405844023052726?via%3 Dihub [Accessed 7 Oct. 2024]

[5]   Bharathi, V. and Pedda Reddy, C. (2023). Comparison of forest fire prediction system using machine learning algorithms, in Proceedings of the 2023 3rd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), Greater Noida, India. Available at: https://ieeexplore.ieee.org/document/10182818 [Accessed 7 Oct. 2024]

[6]   Datasets:

A.   Algerian Forest Fires Dataset: https://www.kaggle.com/datasets/nitinchoudhary012/algerian-forest-fires-dataset

B.   Forest fires Dataset Portugal: https://www.kaggle.com/datasets/ishandutta/forest-fires-data-set-portugal