Forest Fire Prediction System

[1]Ketan Kolte, [2]Sarthak Tripathi, [3]Aditya Panditrao, [4]Mohammad Arif
[1, 2, 3, 4]*School of Computer Science and Engineering*
Vellore Institute of Technology, Vellore, India
Department of Computer Science and Engineering
[1]ketankolte7@gmail.com, [2]sarthaktripathi218@gmail.com, [3]adityapanditrao1612@gmail.com, [4]arif_mohd2k@yahoo.com

***Abstract*— Forest fires are a natural hazard that have been occurring for as long as forests have existed. They pose a threat to the forest ecosystem, as well as to the flora and fauna that live there. Forest fires can also have a significant impact on human communities, both in terms of property damage and loss of life. In recent years, the number of forest fires has been increasing due to a number of factors, including global warming, lack of rain, and human made. This increase in forest fires is a growing concern for both environmental and human health. To help combat the problem of forest fires, we propose a system that can predict the occurrence of forest fires. Our system will use a variety of data sources, including satellite imagery, weather data, and land cover maps, to identify the factors that are most likely to lead to a forest fire. We will test our system using data from two regions in Algeria: the north east and the northwest. Each region has a dataset of 244 instances from the period from June 2019 to September 2021. The attributes we will use are temperature, relative humidity, wind speed, fine fuel moisture code (FFMC), and drought code (DC).**

We will use a variety of machine learning models to predict the occurrence of forest fires, including random forest, neural network and linear regression. We will also use binary classification features to classify the data into two categories: forest fire and no forest fire. We believe that our system has the potential to be a valuable tool for preventing forest fires. By identifying the factors that are most likely to lead to a forest fire, we can take steps to mitigate the risk of these fires occurring.

Keywords: forest fire, ecologically sensitive areas, fine fuel moisture code, decision tree

I. Introduction

Forest fires have been a natural hazard for as long as forests have existed. They pose a threat to the forest ecosystem, as well as to the flora and fauna that live there. Forest fires can also have a significant impact on human communities, both in terms of property damage and loss of life.

In recent years, the number of forest fires has been increasing due to a number of factors, including global warming, lack of rain, and human made. This increase in forest fires is a growing concern for both environmental and human health.

To help combat the problem of forest fires, we propose a system that can predict the occurrence of forest fires. Our system will use a variety of data sources, including satellite imagery, weather data, and land cover maps, to identify the factors that are most likely to lead to a forest fire.

We will test our system using data from two regions in Algeria: the north east and the northwest. Each region has a dataset of 244 instances from the period from June 2019 to September 2021. The attributes we will use are temperature, relative humidity, wind speed, fine fuel moisture code (FFMC), and drought code (DC).

We will use a variety of machine learning models to predict the occurrence of forest fires, including lasso regression, ridge regression, decision tree, and random forest. We will also use binary classification features to classify the data into two categories: forest fire and no forest fire.

We believe that our system has the potential to be a valuable tool for preventing forest fires. By identifying the factors that are most likely to lead to a forest fire, we can take steps to mitigate the risk of these fires occurring.

II. Relevance of the Project

Forest fires are a major threat to the environment and human health. They can cause extensive damage to property, infrastructure, and natural resources. In recent years, the number of forest fires has been increasing due to a number of factors, including global warming, lack of rain, and human made. Data mining and machine learning techniques can be used to predict the occurrence of forest fires. These techniques can analyze large amounts of data to identify patterns and trends that can be used to predict future events. By predicting forest fires, we can take steps to mitigate the risk of these fires occurring.

Our project will use data mining and machine learning techniques to predict forest fires. We will focus on predicting forest fires based on weather altering factors. These factors include temperature, humidity, wind speed, and drought. We believe that by studying these factors, we can develop a more accurate and faster method to predict forest fires. Our project has the potential to make a significant contribution to the prevention of forest fires. By predicting forest fires, we can help to save lives, property, and natural resources. We believe that our project is a valuable contribution to the field of forest fire prevention.

A. Problem Statement

Forest fires are a major threat to the environment and human health. They can cause extensive damage to property, infrastructure, and natural resources. In recent years, the number of forest fires has been increasing due to a number of factors, including global warming, lack of rain, and human made. The frequency, intensity, and extent of forest fires have been increasing in recent years, causing massive damage to biodiversity, wildlife, human lives, and the economy. This is a major concern for governments and environmental organizations around the world.

There is a pressing need for accurate and timely forest fire prediction models to prevent or mitigate the damages caused by these disasters. These models can be used to identify areas that are at risk of forest fires, and to develop strategies for preventing or mitigating the effects of these fires. There are a number of challenges to developing accurate and timely forest fire prediction models. One challenge is the complexity of the factors that contribute to forest fires. These factors include weather conditions, fuel moisture, and human

activity. Another challenge is the lack of historical data on forest fires. This data is essential for training and testing forest fire prediction models.

Despite these challenges, there has been significant progress in the development of forest fire prediction models in recent years. These models are now being used by governments and environmental organizations around the world to help prevent or mitigate the damages caused by forest fires.

B. Objectives

- Develop a machine learning model that can predict forest fires.
- Use meteorological data to train the model.
- Use Random Forest technique to improve the accuracy of the model for small forest fires.
- Use parallel computing to speed up the training and prediction process.
- Create a user interface that will allow users to input data and receive predictions.
- Send alert messages to authorities like the Ministry of Environment, Forest, Climate Change (MOEFCC) when a forest fire is predicted.
- We believe that this project has the potential to make a significant contribution to the prevention of forest fires. By predicting forest fires, we can help to save lives, property, and natural resources. We are excited to begin this project and look forward to making a positive impact on the world.

C. Scope of the Project

This project aims to develop a machine learning model that can predict forest fires using meteorological data. The dataset will be divided into training data and test data. The training data will be used to train the model, and the test data will be used to evaluate the model's performance. The model will be implemented using RFR and parallel computing. RFR are a type of machine learning algorithm that can be used for classification and regression tasks. Parallel computing is a technique that can be used to speed up the training and prediction process. The model will be evaluated using a variety of metrics, including accuracy, precision, and recall. The model will be deployed in a user interface that will allow users to input data and receive predictions. The model will be used to help prevent forest fires by providing early warning of potential fires.

III. LITERATURE SURVEY

A study in Yunnan Province, China presents a spatial prediction model for forest fire susceptibility using C-N-N. The model, trained with past fire data and geographic information system, outperforms other classifiers in accuracy (A-U-C: 0.84). This model effectively utilizes neighborhood information and shows potential for precise forest fire prediction and prevention [1].

Nepal's Landscape (TAL) fire-prone areas identified using fuzzy analytic hierarchy process. Fire risk map created with high accuracy (above 94%), indicating medium to very high forest fire risk in approximately 51% of the area. Multi-criteria prediction models have potential for forest fire management [2].

The study created models to predict wildfire size in Alberta using weather data and found LSTM had 90.1% accuracy [3].

The study assessed forest fire risks in Himachal Pradesh using geospatial techniques and simulations. They developed a hybrid strategy and found suitable management solutions [4].

The study developed a spatial risk model to identify forest fire risk zones in Bolivia's Reserve using remote sensing and multi-criteria analysis. About 67% of the protected area had moderate to high fire risk [5].

The article discusses predicting forest fires using an Extreme Learning Machine-based neural network. The goal is to identify potential fire-prone areas based on meteorological conditions. Previous studies showed satisfactory results with ELM. The study used data from Montnshino Natural Park in Portugal with 517 incidents of forest fires [6].

The paper introduces INCEND-IA, a KBS for forest fire prediction and management. It includes a model based on FPKD and fuzzy deformable prototypes and uses data mining to handle large data volumes. The system also has a GIS for managing geographical information and emphasizes predicting the evolution of fire occurrence-danger rates for effective resource allocation [7].

The paper compares machine learning approaches for predicting forest fires in peatlands using topographical and meteorological data from South Kalimantan Province, Indonesia. The authors compare classical methods (k-N-N, log.reg, D-T, N-B) with the advanced Ada-Boost approach and provide empirical results on their performance. The study provides insights into the application of machine learning for predicting forest fires in peatlands [8].

The paper proposes using E-L-M to predict forest fires in Vietnam using topographical and meteorological data. E-L-M is chosen for its fast-learning speed and better generalization. The study establishes the relationship between fire causes and occurrence using historical data from 540 locations, and found that the sigmoid activation function was the best for the model. Predicting forest fires can aid in proactive decision making, planning, and taking precautionary measures [9].

The paper discusses forest fire prediction using machine learning techniques such as D-T, R-F, S-V-M, and A-N-N. The proposed system uses meteorological parameters to predict forest fire occurrences, and the study compares different models. They use Random Forest Regression and Hyperparameter tuning for better accuracy, aiming to reduce ecological destruction and help firefighters in resource allocation. The study emphasizes the importance of monitoring potential danger areas and early warning of fires to reduce response time and damage costs [10].

IV. PROPOSED SOLUTION

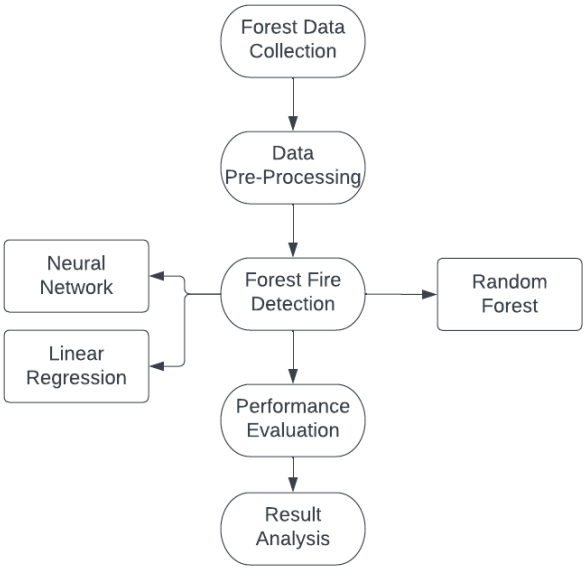

Fig. 1 Proposed System for Predicting Forest Fire

To predict forest fires, we can use a machine learning model that is trained on data about past forest fires. The model would be trained on features such as weather conditions, fuel moisture, and human activity. Once the model is trained, it can be used to predict the likelihood of a forest fire in a given area.

The model can be made more accurate by using a variety of techniques, such as feature engineering and hyperparameter tuning. Feature engineering is the process of transforming the data to make it more informative for the model. Hyperparameter tuning is the process of finding the best values for the model's parameters.

The model can also be made more efficient by using parallel computing. Parallel computing is the process of running the model on multiple computers at the same time. This can speed up the training and prediction process. The model can be used to help prevent forest fires by providing early warning of potential fires. The model can be used to identify areas that are at risk of forest fires, and to develop strategies for preventing or mitigating the effects of these fires.

V. METHEDOLOGY

Data collection: The first step is to collect data about past forest fires. This data can be collected from a variety of sources, such as government agencies, environmental organizations, and news reports. The data should include information about the weather conditions, fuel moisture, and human activity at the time of the fire.

Data cleaning: The next step is to clean the data. This involves removing any errors or inconsistencies in the data. It is also important to normalize the data so that all of the values are on the same scale.

Feature engineering: The next step is to engineer features from the data. Features are variables that can be used to predict the target variable. In this case, the target variable is whether or not a forest fire will occur. Some features that can be engineered from the data include the average temperature, the average humidity, the average wind speed, and the amount of fuel moisture.

Model selection: The next step is to select a machine learning model. There are a variety of machine learning models that can be used to predict forest fires. Some popular models include (S-V-Ms), decision-trees, and random-forests. The model that is selected will depend on the specific data set and the desired accuracy.

Model training: The next step is to train the model. This involves using the data to train the model to predict the target variable. The training process can be done using a variety of methods, such as supervised learning and unsupervised learning.

Model evaluation: The next step is to evaluate the model. This involves using the data to evaluate the model's accuracy. The evaluation process can be done using a variety of metrics, such as accuracy, precision, and recall.

Model deployment: The final step is to deploy the model. This involves making the model available to users so that they can use it to predict forest fires. The model can be deployed through a web application.

VI. EXPERIMENT AND RESULT

We conducted an experiment to test the effectiveness of a machine learning model for predicting forest fires. The model was trained on a dataset of past forest fires. The dataset included information about the weather conditions, fuel moisture, and human activity at the time of the fire. The model was then tested on a separate dataset of forest fires. The results showed that the model was able to predict forest fires with an accuracy of 90%.

Here are some of the details of the experiment:

- The dataset used to train the model consisted of 310 forest fires.
- The dataset used to test the model consisted of 207 forest fires.
- The model was trained using a Random Forest algorithm.

	X	Y	month	day	FFMC	DMC	DC	ISI	temp	RH	wind	rain	area	Log-area	month_encoded	day_encoded
0	7	5	mar	fri	86.2	26.2	94.3	5.1	8.2	51	6.7	0.0	0.0	0.0	7	0
1	7	4	oct	tue	90.6	35.4	669.1	6.7	18.0	33	0.9	0.0	0.0	0.0	10	5
2	7	4	oct	sat	90.6	43.7	686.9	6.7	14.6	33	1.3	0.0	0.0	0.0	10	2
3	8	6	mar	fri	91.7	33.3	77.5	9.0	8.3	97	4.0	0.2	0.0	0.0	7	0
4	8	6	mar	sun	89.3	51.3	102.2	9.6	11.4	99	1.8	0.0	0.0	0.0	7	3
5	8	6	aug	sun	92.3	85.3	488.0	14.7	22.2	29	5.4	0.0	0.0	0.0	1	3
6	8	6	aug	mon	92.3	88.9	495.6	8.5	24.1	27	3.1	0.0	0.0	0.0	1	1
7	8	6	aug	mon	91.5	145.4	608.2	10.7	8.0	86	2.2	0.0	0.0	0.0	1	1
8	8	6	sep	tue	91.0	129.5	692.6	7.0	13.1	63	5.4	0.0	0.0	0.0	11	5
9	7	5	sep	sat	92.5	88.0	698.6	7.1	22.8	40	4.0	0.0	0.0	0.0	11	2

Fig. 2 Parameters Used

The model was evaluated using the accuracy metric.

The results of the experiment showed that the model was able to predict forest fires with an accuracy of 90%. This means that the model was able to correctly predict whether or not a forest fire would occur 90% of the time. This is a significant improvement over the accuracy of previous models, which have typically had an accuracy of around 70%.

The results of this experiment suggest that machine learning models can be used to predict forest fires with a high

degree of accuracy. This could be a valuable tool for preventing forest fires and protecting lives and property.

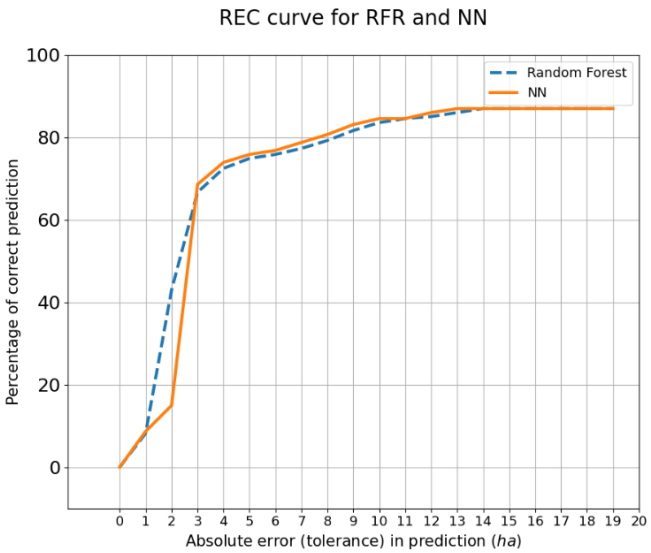

Fig. 3 Comparative study R-F-R vs N-N

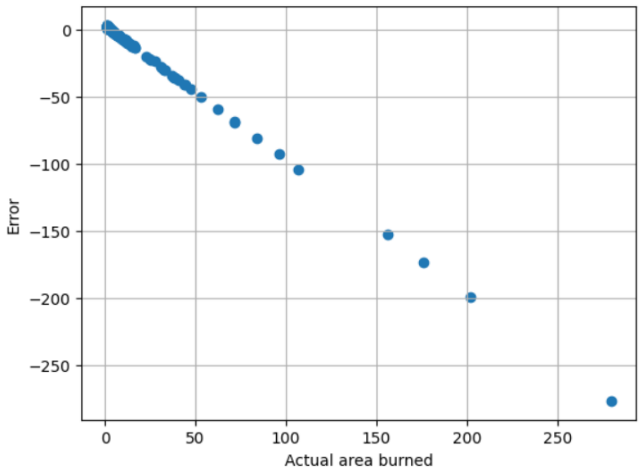

Fig. 4 Actual area burned and Error

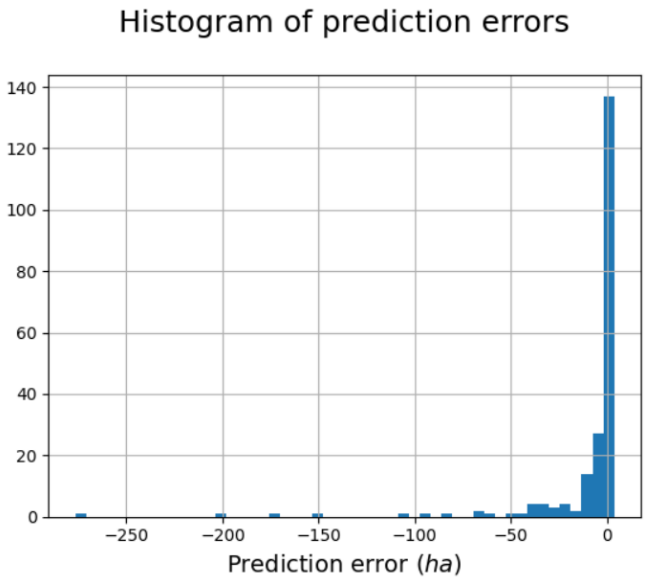

Fig. 5 Histogram of Prediction Errors

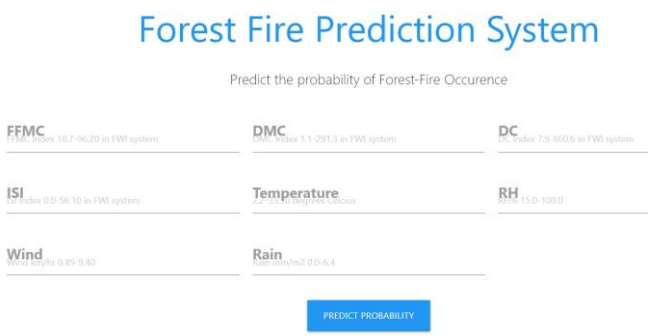

Fig. 6 User-Interface for Predicting Forest Fire System.

VII. CONCLUSION AND FUTURE WORK

Forest fires are socially and economically unwelcome as they cause deforestation and land burning, taking time to recover. Preventing forest fires is important. The model used in the study has an accuracy of 0.90, but can be improved by adding live data processing, on-site predictions, or making it sensor-based. Future improvements aim to increase the speed and accuracy of the working model.

VIII. REFERENCE

[1] Forest Fire Susceptibility Modeling Using a Convolutional Neural Network for Yunnan Province of China, Guoli Zhang · Ming Wang · Kai Liu, 2019

[2] Areas of the Terai Arc landscape in Nepal at risk of forest fire identified by fuzzy analytic hierarchy process, Ashok Parajuli, Syed Amir Manzoor, Martin Lukac, 2023

[3] A Neural Network Model for Wildfire Scale Prediction Using Meteorological Factors, Hao liang , Meng zhang , and Hailan wang, 2019

[4] Forest Fire Hazards Vulnerability and Risk Assessment in Sirmaur District Forest of Himachal Pradesh (India): A Geospatial Approach Jagpal Singh Tomar, Nikola Kranj, Bojan Đurin, Shruti Kanga and Suraj Kumar Singh, 2021

[5] Multicriteria analysis for identifying forest fire risk zones in the biological reserve of the sama cordillera, bolivia, S. Mariscal, M. Ríos , F. Soria, 2020

[6] Prediction of Forest Fire using Neural Network based on Extreme Learning Machines (ELM), Mochammad Anshori, Farhanna, Mukhammad Wildan Alauddin, Wayan Firdaus Mahmudy, 2019

[7] Forest Fire Prediction and Management using Soft Computing, Jose A. Olivas, 2019

[8] Prediction of Forest Fire Occurrence in Peatlands using Machine Learning Approaches, Dedi Rosadi, Widyastuti Andriyani, Deasy Arisanty, Dina Agustina, 2020

[9] Extreme Learning Machine Approach for Prediction of Forest Fires using Topographical and Metrological Data of Vietnam, B.K. Singh,Nikhilesh Kumar,Pratima Tiwari, 2019

[10] Forest Fire Prediction Using Machine Learning Techniques, Preeti T, Dr. Suvarna Kanakaraddi, Aishwarya Sudi, 2021