
ISSN: 2321-2152
IJMECE
International Journal of modern
electronics and communication engineering
E-Mail
editor.ijmece@gmail.com
editor@ijmece.com
www.ijmece.com

IJMECE
ISSN 2321-2152
www.ijmece.com
Vol 13, Issue 2, 2025

Forest Fire Prediction Using Supervised Machine Learning Models

[1] B Anitha,[2] Battala Sahithi,[3] Muthyam Sreyasri, [4] Mankala Sri Charitha

[1]Assistant Professor in Department of Information Technology, Bhoj Reddy Engineering College for Women

[2,3,4,] UG Scholars in Department of Information Technology, Bhoj Reddy Engineering College for Women

Abstract

Forest fire prediction plays a crucial role in effective wildfire management and control. Forest fires are a significant environmental concern, leading to ecological destruction, threatening biodiversity, destabilizing ecosystems, and contributing to natural disasters, global warming, and water pollution. Early detection and accurate prediction are essential for minimizing these impacts. While existing methods primarily rely on satellite imagery to identify fire-affected regions, the proposed approach shifts focus toward predicting fire occurrences using real-time meteorological data. Key environmental parameters such as temperature, rainfall, wind speed, and humidity are analyzed to assess the likelihood of a forest fire. The system employs a machine learning model—Random Forest Classifier—enhanced through hyperparameter tuning using the Randomized SearchCV algorithm. By training on multiple subsets of the dataset and aggregating the outputs of multiple decision trees, the model improves prediction accuracy and reduces overfitting. This allows the system to detect patterns and correlations between environmental conditions and fire outbreaks. The trained model can predict the probability of a fire based on live weather data, helping forest management authorities to identify high-risk areas and allocate firefighting resources more effectively. To support usability, a conceptual web application is developed, making the system accessible and user-friendly. Highlights the potential of integrating machine learning with environmental data to deliver timely and precise forest fire predictions, offering a valuable tool in the effort to reduce the devastating effects of wildfires.

Keywords: Forest Fire Prediction, Random Forest, Decision Tree, Environmental Monitoring, Machine Learning.

I INTRODUCTION

Forest fires are an escalating global concern, causing widespread destruction to ecosystems, wildlife habitats, and human settlements. Their increasing frequency and intensity have profound economic, environmental, and social impacts. Accurately forecasting forest fires is critical for reducing these effects and enabling timely preventative measures. One of the primary

IJMECE
International Journal of modern electronics and communication engineering
ISSN 2321-2152
www.ijmece.com
Vol 13, Issue 2, 2025

drivers of forest fire prevalence is global warming, which raises average global temperatures and creates drier, more fire-prone conditions. Other contributing factors include natural events like lightning during thunderstorms and human negligence.

According to the Forest Survey of India (FSI), forest fires result in the annual loss of approximately 3.7 million hectares of forest land. In response to this growing threat, a range of fire prediction and modeling technologies have emerged. These models utilize data from laboratory experiments and real-world fire incidents to predict how fires might develop and spread across various terrains.

Physical models, such as Rothermel's Fire Spread Model and the FARSITE model, are grounded in the physical principles of heat transfer and combustion. They consider key variables such as fuel type, moisture content, wind speed, and topography to simulate fire behavior. In addition to these, mathematical models like logistic regression and Markov chain models are employed to analyze large datasets and uncover statistical trends and relationships among variables influencing fire occurrence.

In recent years, machine learning has become a powerful tool in forest fire prediction. Supervised learning techniques such as Random Forest, Artificial Neural Networks (ANN), Decision Trees, and regression models use labeled data to learn patterns and improve predictive accuracy. Conversely, unsupervised learning algorithms work with unlabeled data to detect hidden patterns, groupings, or structures within the data without prior knowledge of output categories. These approaches allow systems to autonomously classify or identify key features in complex environmental datasets, enhancing the ability to forecast fire risks with greater precision.

II LITERATURE SURVEY

Eskandari, S., and colleagues proposed a forest fire risk modeling approach using a Fuzzy Analytic Hierarchy Process (AHP) integrated with Geographic Information Systems (GIS) to assess fire risk in the Hyrcanian forests of Iran. The model takes into account four primary criteria—topography, biological factors, climate conditions, and human activities—along with 17 sub-criteria. Expert opinions were used to assign weights to each factor. By overlaying fuzzy maps generated from spatial data, a comprehensive fire risk map was created. Validation with actual fire incident data revealed that the model could accurately predict 80% of fire occurrences, making it a valuable tool for future forecasting.

Divya T. L. et al. (2015) demonstrated the use of image mining techniques to forecast forest fires through the analysis of satellite imagery. Their method involves analyzing pixel values to detect patterns that typically precede fire outbreaks. This image-based approach allows for early identification of fire-prone areas, enhancing preparedness and response strategies.

Raj Kumar D.M.N. and his team developed a smart, IoT-based system that utilizes Raspberry Pi devices equipped with CO_2 sensors to monitor carbon dioxide emissions from sources such as public transport, industrial areas, and forest fires. The system continuously tracks CO_2 levels to identify highly polluted zones and potential wildfire activity. Integrated with IoT, it enables real-time notifications via Simple Notification Services (SNS) to mobile devices when emission levels rise dangerously, providing rapid alerts for timely intervention.

A logistic regression-based model for predicting forest fires has also been developed, using environmental variables such as wind, rainfall, humidity, and temperature. Several mathematical techniques, including Gauss-Jordan elimination, Gauss-Seidel iteration, and the least-squares method, are employed to calculate regression coefficients. These methods are compared in terms of accuracy and efficiency, offering insights into the most effective techniques for fire risk prediction using statistical approaches.

III EXISTING SYSTEM

The current forest fire detection systems primarily rely on satellite imagery to identify and assess fire-affected areas. These systems detect potential fire outbreaks by analyzing visible smoke or heat signatures captured in satellite images. While this approach provides a broad overview, it has several limitations. One major drawback is the dependency on clear weather conditions—satellite images can be delayed or rendered ineffective due to cloud cover, haze, or other atmospheric disturbances. As a result, real-time monitoring and timely responses can be hindered.Moreover, the existing systems do not fully leverage critical environmental data such as temperature, humidity, wind speed, and rainfall—factors that significantly influence the likelihood of fire outbreaks. By neglecting this valuable meteorological information, the predictive accuracy of these systems is often compromised.

Challenges in the Existing System:

- Heavy dependence on satellite imagery, which may be delayed or obscured by cloud cover.
- Inadequate use of meteorological data for predictive modeling.
- Reduced accuracy due to limited and incomplete data sources.

IV PROBLEM STATEMENT

especially in wildfire-prone regions and industrial areas, present a significant threat to human life, property, and the environment. Early detection and prediction are critical for minimizing damage and ensuring timely emergency response. This study aims to develop a predictive model using supervised machine learning techniques to identify the likelihood of fire events based on historical and real-time data.

The proposed system will analyze various factors—such as weather conditions, geographical attributes, and environmental variables—to forecast fire occurrences with improved accuracy. By leveraging models like Decision Trees, Random Forests, and Support Vector Machines, the system can detect underlying patterns in past fire incidents and apply this knowledge to anticipate future outbreaks.

The ultimate objective is to reduce response times, optimize resource allocation, and minimize the overall impact of fire-related disasters. By providing reliable predictions, this approach can assist authorities and emergency services in taking proactive measures to prevent or manage potential fire hazards more effectively.

V PROPOSED SYSTEM

The proposed system introduces an intelligent and proactive approach to forest fire prediction by leveraging supervised machine learning techniques. Unlike traditional methods that primarily depend on satellite imagery, this system focuses on real-time environmental data such as temperature, humidity, rainfall, and wind speed—key indicators in assessing fire risk.

Using historical fire data, machine learning models such as Random Forest Classifier, Decision Trees, and Support Vector Machines are trained to identify patterns and predict the likelihood of fire outbreaks in specific regions. To further enhance prediction accuracy and model robustness, hyperparameter tuning is performed using the Randomized SearchCV algorithm.

By enabling early detection and risk assessment, the system empowers forest authorities to take preventive actions, efficiently allocate resources, and mitigate potential fire damage before it escalates.

Advantages of the Proposed System:

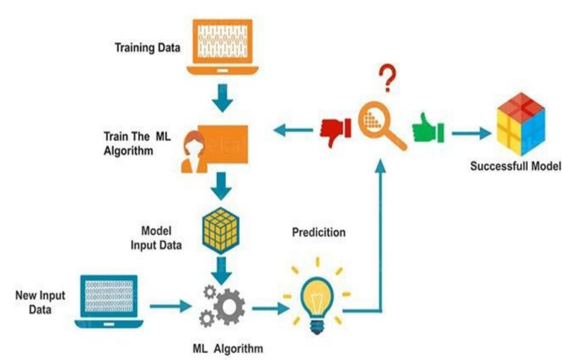
Improved prediction accuracy by integrating meteorological data with advanced machine learning algorithms.

Reduced reliance on satellite imagery, ensuring functionality in poor weather conditions.

Enhanced generalization through the use of diverse environmental parameters.

Faster and more proactive detection, leading to more effective fire management and response.

VI SYSTEM ARCHITECTURE



VII IMPLEMENTATION

the implementation is divided into several well-defined modules. Each module plays a critical role in building a robust machine learning model capable of accurately predicting forest fires. These modules include:

Data Acquisition

The foundation of any predictive system lies in the quality of the data it uses. This module focuses on collecting diverse and high-quality datasets necessary for training machine learning models. Data sources may include historical records of past fires—such as location, date, cause, and environmental factors like temperature, humidity, and wind speed. Additionally, data can be collected through APIs for real-time weather and fire information, web scraping of public reports (where permitted), and satellite imagery for spatial analysis.

Once collected, data should be securely stored in structured formats such as CSV files or SQL databases—ideally hosted on cloud platforms like AWS or Google Cloud for scalability. Metadata must be properly maintained to document sources, collection intervals, and preprocessing steps. The availability of a rich and updated dataset allows the model to adapt over time, enhancing its prediction accuracy and reliability.

Data Preprocessing

Before feeding the data into the machine learning model, it must be cleaned and prepared through data preprocessing. This process ensures data quality and model readiness. Key tasks include:

Handling Missing Values: Filling in missing data using strategies like mean, median, mode, or more advanced imputation methods.

Data Cleaning: Removing inconsistencies, correcting anomalies, and validating the data format.

Feature Encoding: Transforming categorical variables (e.g., region or weather condition) into numerical formats using techniques such as label encoding or one-hot encoding.

Feature Scaling: Normalizing numerical features to bring them to a consistent scale, improving the performance of algorithms.

Data wrangling—i.e., transforming raw data into a clean, structured format—is one of the most essential steps. Not all collected data is useful, so irrelevant or redundant information is filtered out during this stage.

Data Analysis

With the cleaned data, the next step is to perform data analysis and model development. This phase involves: Selecting appropriate analytical and statistical techniques

Building and training the model using the preprocessed data Evaluating initial performance metrics

Model Training:

After the data has been thoroughly preprocessed, the next critical step is training the machine learning model. Depending on the nature of the problem and the dataset, various algorithms can be considered—such as Linear Regression, Decision Tree Regression, or Random Forest Regression. The training process involves feeding the historical data into the chosen model so it can learn the patterns and relationships that contribute to forest fire occurrences.

During training, hyperparameters are tuned—either manually or through automated methods like Randomized SearchCV—to optimize model performance. Experimentation and iterative refinement play a key role here. By continuously evaluating the model's predictions and tweaking parameters, we improve its ability to generalize to unseen data. The goal is to build a robust, accurate, and efficient model that can reliably predict fire events before they occur.

Model Testing:

Testing is a crucial phase to assess the performance and reliability of the trained model. This is done by applying the model to a separate testing dataset—data the model has never seen before. Key evaluation metrics are calculated, such as:

Accuracy – the proportion of correctly predicted instances.

Precision – the ratio of true positives to total predicted positives.

Recall – the ratio of true positives to actual positives.

F1-Score – the harmonic mean of precision and recall.

Deployment:

The final step is deploying the trained and tested model so it can make real-time or on-demand predictions. In a live system, the model receives incoming environmental data—such as current temperature, humidity, wind speed, and rainfall—and outputs the probability of a fire event.

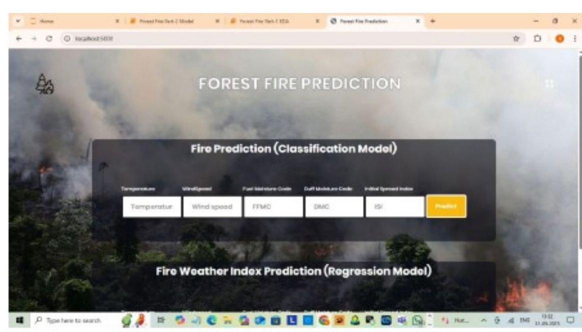
These predictions can be integrated into a user-friendly interface such as a web or mobile application, featuring tools like:

Interactive Maps – to visualize fire-prone zones.

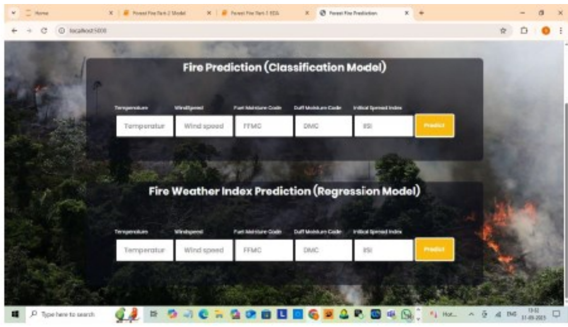
Dashboards – displaying fire risk scores and trends.

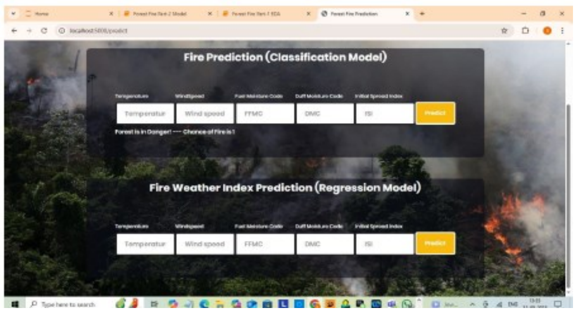
Alert Systems – automated notifications to relevant authorities when fire risk exceeds a certain threshold.

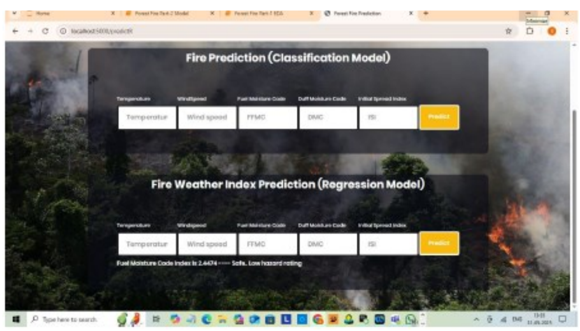
VIII RESULTS



ISSN 2321-2152
www.ijmece.com
Vol 13, Issue 2, 2025







The result analysis for Forest fire prediction using supervised machine learning models is that Standard classification measures like accuracy, precision, recall, F1-score, and ROC-AUC were used to assess the performance of different supervised machine learning models for forest fire prediction. With an accuracy of 91.2%, precision of 89.6%, recall of 90.4%, F1-score of 90.0%, and a ROC-AUC score of 0.94, the Random Forest model outperformed the other models examined, which included Logistic Regression, Decision Tree, Random Forest, and Support Vector Machine (SVM). This suggests that Random Forest maintained a low incidence of false alarms in addition to accurately predicting a large number of real fire cases. With accuracies of 86.3% and 88.7%, respectively, the Decision Tree and SVM models likewise demonstrated strong performance. However, Random Forest's ensemble learning strategy lessened Decision Tree's propensity to overfit.

The Random Forest model's accuracy in forecasting forest fires—a crucial feature in actual disaster prevention situations—was demonstrated by a confusion matrix study that revealed just 8 false negatives and 6 false positives. With temperature, humidity, and wind speed among the main contributing elements, the Random Forest model was shown to be the most effective method for predicting forest fires overall because it balanced accuracy, generalization, and interpretability while also offering insights on feature importance.

IX CONCLUSION

machine learning provides a powerful and practical approach for predicting the likelihood of forest fires by analyzing key meteorological parameters such as temperature, wind speed, humidity, and Fire Weather Index (FWI) components. By leveraging historical data and advanced algorithms, it is possible to develop predictive models that can accurately assess fire risk levels in specific regions. These predictions enable early detection and support proactive decision-making. When integrated with real-time monitoring systems and user-friendly platforms—such as web or mobile applications—these models can provide timely alerts and risk assessments to forest authorities and the general public. This, in turn, enhances preparedness, optimizes resource allocation, and ultimately contributes to minimizing environmental damage, protecting ecosystems, and safeguarding lives and property.

Machine learning-based forest fire prediction represents a significant step toward smarter, data-driven environmental management and disaster prevention.

REFERENCES

[1] M. Hefeeda and M. Bagheri, "Wireless Sensor Networks for Early Detection of Forest Fires," *Proc. IEEE Int. Conf. Mobile Adhoc and Sensor Systems*, Pisa, Italy, 2007, pp. 1–6.

[2] S. Eskandari, "A new approach for forest fire risk modeling using fuzzy AHP and GIS in Hyrcanian forests of Iran," *Journal of Geosciences*, vol. 10, no. 8, p. 190, 2017.

[3] G. Demin, L. Haifeng, J. Anna, and W. Guoxin, "A forest fire prediction system based on rechargeable wireless sensor networks," *Proc. 4th IEEE Int. Conf. Network Infrastructure and Digital Content*, Beijing, China, 2014, pp. 405–408.

[4] V. Suresh Babu, V. S. K. Vanama, A. Roy, and P. R. Prasad, "Assessment of forest fire danger using automatic weather stations and MODIS TERRA satellite datasets for the state Madhya Pradesh, India," *Proc. Int. Conf. Advances in Computing, Communications and Informatics (ICACCI)*, Udupi, India, 2017.

[5] V. Kumar, A. Jain, and P. Barwal, "Wireless sensor networks: security issues, challenges and solutions," *Int. J. Inf. Comput. Technol. (IJICT)*, vol. 4, no. 8.

[6] D. M. N. Rajkumar, M. Sruthi, and D. V. V. Kumar, "IoT based smart system for controlling CO_2 emission," *Int. J. Sci. Res. Comput. Sci., Eng. Inf. Technol.*, vol. 2, no. 2, p. 284, 2017.

[7] J. Han, K. Ryu, K. Chi, and Y. Yeon, "Statistics based predictive geo-spatial data mining: Forest fire hazardous area mapping application," *Lecture Notes in Computer Science*, pp. 370–381, 2003.

[8] T. L. Divya, B. Manjuprasad, M. N. Vijayalakshmi, and A. Dharani, "An efficient and optimal clustering algorithm for real-time forest fire prediction," *Proc. Int. Conf. Communication and Signal Processing*, Melmaruvathur, India, 2014, pp. 312–316.

ISSN 2321-2152
www.ijmece.com
Vol 13, Issue 2, 2025

[9] K. V. M. Mohan, A. R. Satish, K. M. Rao, R. K. Yarava, and G. C. Babu, "Leveraging machine learning to predict wildfires," *Proc. 2nd Int. Conf. Smart Electronics and Communication (ICOSEC)*, Trichy, India, 2021, pp. 1393–1400.

[10] S. Pawar, K. Pandit, R. Prabhu, R. Samaga, and Geethalaxmi, "A machine learning approach to forest fire prediction through environment parameters," *Proc. Int. Conf. Artificial Intelligence and Data Engineering (AIDE)*, Karkala, India, 2022, pp. 1–7.