



International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.699

Volume 14, Issue 4, April 2025

A Data Mining Approach for Prediction of Forest Fire Using Machine Learning Algorithm

Kadraka Jayavardhan, Darmaraju Anil Santosh, Nethala Kavya Sree, Nanduri Raja Sathvik,
Yeduresapu Priyanka

UG Scholar, Dept. of Computer Science Engineering (Artificial Intelligence and Data Science), Satya Institute of
Technology and Management, Vizianagaram, India

Mogali Madhu Babu

Associate Professor, Dept. of Computer Science Engineering (Artificial Intelligence and Data Science), Satya Institute
of Technology and Management, Vizianagaram, India

ABSTRACT: Every year, millions of hectares of forests are destroyed by forest fires, which are a common and significant natural calamity that seriously endangers both human life and property. For the purpose of developing quick risk management plans and putting into practice efficient firefighting techniques, precise quantitative forecasting of forest fire spread is crucial. The Forest Fire Spread Behaviour Prediction (FFSBP) model, which includes two essential components—the Forest Fire Spread Process Prediction (FFSPP) model and the Forest Fire Spread Results Prediction (FFSRP) model—is presented in this work in answer to this requirement. Cellular Automata and Wang Zhengfei models are combined to create the FFSPP model, which predicts the direction and rate of forest fire propagation. The FFSRP model, on the other hand, employs machine learning techniques to anticipate the size of the burnt area. An actual case study is conducted utilising the "3.29 Forest Fire" disaster in China to verify the effectiveness of the suggested models. This instance is used to verify the FFSPP model, while an actual fire dataset from Portugal's Montesinho National Forest Park is used to evaluate the FFSRP model. The validation process's findings show that the FFSPP model forecasts a burnt area of 286.81 hm² with a corresponding relative error of 28.94% during the "3.29 Forest Fire's" natural growth phase. Compared to the Farsite and Prometheus fire behaviour simulation models, this relative inaccuracy is noticeably lower. Furthermore, the FFSRP model has excellent prediction performance, especially when considering small and medium-sized fire situations. These results highlight the FFSBP model's potential as a useful instrument for improving the accuracy of forest fire predictions and enabling more effective risk reduction and suppression tactics.

KEYWORDS: Fire prediction, Cellular Automata, Machine learning

I. INTRODUCTION

Forest represents a material foundation that is important for a country's sustainable development, and it is also an important renewable resource in the construction of the economy and ecological environments [1]. Forests provide habitat for wildlife and protect biodiversity [2]. However, forest ecosystems are often subjected not only to natural disturbances but also to manmade damage, including forest diseases, pests, and forest fires [3–5]. Every year, millions of hectares of forests are destroyed by wildfires that seriously threaten human life and property safety [6,7]. When a forest fire is very serious, it will destroy natural ecological species and ecological structure, even threatening human safety and affecting the regional carbon cycle [8–10]. Suppressing wildfires to preserve life and properties has become one of the most critical tasks in preserving natural resources [11,12].

The accurate modeling of forest fire spread facilitates the rational allocation of fire prevention resources and the development of timely and effective fire suppression strategies [13]. In general, forest fire spread models can be classified into three types [14]: remote sensing, physical, and algorithmic. Remote sensing technology can provide the necessary environmental information required to monitor and forecast fire spreading [15,16]. In remote sensing models, a GIS (geographic information system) is usually used to explore the spatial distribution pattern of forest fires [17]. GIS data have been used to analyze the spatial and temporal distribution characteristics of factors that affect forest fires in

different regions [18]. It has been shown that GIS is available for constructing forest fire simulation and rescue systems in formulating corresponding fire prevention measures for local forestry bureaus [19].

To create a user-friendly and self-organizing model, this study integrates the Wang Zhengfei model with the Cellular Automata (CA) model. This integrated approach aims to predict the direction and speed of forest fire spread, ultimately achieving a visually intuitive representation. Furthermore, as computer computing power has advanced and Machine Learning (ML) has evolved, ML has garnered attention for its ability to discern nonlinear relationships among diverse input parameters. Researchers have applied ML to forest fire prediction, utilizing methods such as backpropagation neural networks, random forests (RFs), deep learning, and ensemble learning. While various ML techniques have demonstrated superior results in fire risk prediction compared to probability and statistical methods, the emphasis has primarily been on predicting the likelihood of forest fire occurrence rather than the burned area. In recent decades, ensemble learning has gained prominence in the ML field due to its efficacy in addressing practical application challenges. Consequently, this article opts for ensemble learning to forecast the outcome of forest fire occurrences, specifically, the burned area. This choice aims to enhance the prediction of the fire situation and facilitate timely response measures.

II. METHODOLOGY

In the proposed context, the focus of this paper is on investigating the prediction of both the forest fire spread process and its outcomes, leading to the establishment of a Forest Fire Spread Behavior Prediction (FFSBP) model. This model comprises two integral components: the Forest Fire Spread Process Prediction (FFSPP) model and the Forest Fire Spread Results Prediction (FFSRP) model. The FFSPP model involves combining the Wang Zhengfei model with the Cellular Automata (CA) model to predict the direction and speed of forest fire spread, while the FFSRP model employs ensemble learning methods to predict the burned area. To validate the forest fire spread prediction model, the “3.29” forest fire in Anning, Southwest China, serves as an illustrative example. Additionally, the burned area prediction model is verified using a real fire dataset from January 2000 to December 2003 obtained from Montesinho National Forest Park in Portugal. The research outcomes hold significant practical implications for forest fire response:

- Predicting and visualizing the direction and speed of forest fire spread offers valuable insights for deploying firefighting resources;
- Forecasting the burned area provides guidance for aggregating total firefighting resources;
- The proposed method’s straightforward calculation process, coupled with its adaptability to different datasets, facilitates easy application and optimization of the model.

A. FFSPP Model

Main Settings of CA

The forest fire spread model based on CA is a discrete event model that divides forest land into numerous discrete cells, simulating the fire spread process according to a defined set of rules. Each cell denotes a location on the forest land, and the model simulates the fire propagation by updating the status of each cell. The key settings for a forest fire spread model based on CA include:

- (1) State definition: Each cell can possess distinct states. This paper adopts four states: “unburned,” “partially burned,” “fully burned,” and “end burned.” Initially, the fire source status is assigned to one or more cells based on the actual situation.
- (2) Proximity definition: The adjacent cells of each cell are defined. Typically, four or eight neighborhoods represent cells adjacent to each other, with this paper opting for eight neighborhoods to accurately represent units adjacent at the top, bottom, left, right, or diagonal.
- (3) Update rules: Rules for updating the status of each cell are defined based on the status of adjacent cells and environmental factors like wind direction, wind speed, fuel humidity, and terrain. For instance, when a cell is in a burning state, surrounding unburned cells can be ignited.
- (4) Iterative update: The spread of fire is simulated by iteratively updating the status of each cell. Discrete time steps control the frequency of updates. Each update considers the current state and the state of surrounding cells, updating according to predefined rules.
- (5) Boundary condition handling: Special handling of boundary conditions is required for boundary cells to prevent the fire from spreading beyond the simulation range.

Improvement of CA for Forest Fire Spread

The fundamental components of CA encompass four elements: cell, state, neighborhood, and rule. Presently, research predominantly focuses on one-dimensional and two-dimensional CA. This study opts for a two-dimensional CA, akin to a regular grid. Each grid serves as a cell, possessing its state at every moment. The transition of cell states is contingent upon dynamic rules: functions reliant on the current state of the cell and its neighboring cells to ascertain the subsequent state. Throughout the simulation process, dynamic iterations and calculations are executed, incorporating changes in the neighborhood based on the transformation rules.

B. FFSRP MODEL

ML allows for constructing statistical models by providing sufficient sample data and employing a series of algorithms to carry out predictions. It is found that ML models are advantageous in high prediction accuracy and flexible model structure. In recent decades, the ensemble learning has attracted much attention in the field of ML because it can efficiently solve practical application problems. The basic process is: 1) generate a series of different learners; 2) use a certain integration method to combine learners to improve the prediction ability and generalization ability of the model. Boosting, Bagging, Stacking, and Blending are the most commonly used methods in ensemble learning. Among them, Boosting and Bagging use similar learners for integration. Stacking and Blending are hierarchical structures, which can realize the integration of heterogeneous learners. Moreover, stacking uses K-fold cross validation, and learners are trained with all data. When the data set is not large, the performance of Stacking is more robust compared with Blending. Therefore, this paper uses Stacking to integrate multiple learners to establish the burned area of forest fire prediction model. Stacking is the abbreviation of stacked generalization proposed by Wolpert in 1992. Its core idea is to carry out cross validation training on the base learner, form a training set based on the output of the base learner to train the meta learner, and the prediction result is formed by the output of the meta learner. The base learner and meta learner are the core of the Stacking model, and the selection and combination of learners are the key to the integration of the Stacking model. If the multi-layer stacking learner framework is adopted, it is very difficult to determine the best stacking method due to the large number of learners available and combinations. And the performance improvement after the integration of multi-layer stacking models is limited. Therefore, this paper selects a two-layer learner framework based on Stacking, chooses appropriate learners and their combinations, and constructs a prediction model to improve the prediction ability and generalization ability of the model.

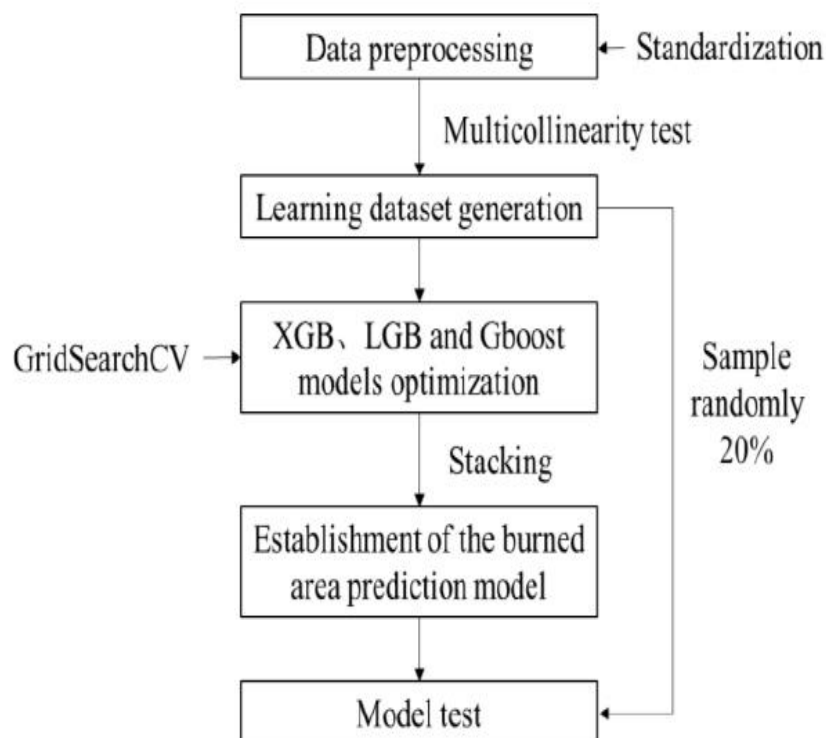


Figure1. Flow chart for establishment of the burned area prediction model

III. RESULTS AND DISCUSSION

A. System Environment

The proposed model is developed using said hardware system configuration and software required to implement the design is shown below.

Hardware System Configuration:

- Processor: Intel i5
- RAM : 8 GB
- Hard Disk : 256 GB
- Key Board : Standard Windows Keyboard
- Mouse : Two or Three Button Mouse
- Monitor : SVGA

Software Requirements:

- Operating system : Windows 7 Ultimate.
- Coding Language: Python.
- Front-End : Python.
- Back-End : Django-ORM
- Designing : Html, css, javascript.
- Data Base : MySQL (WAMP Server).

B. Experimental findings

The processed results and user interface results are shown in below figures.

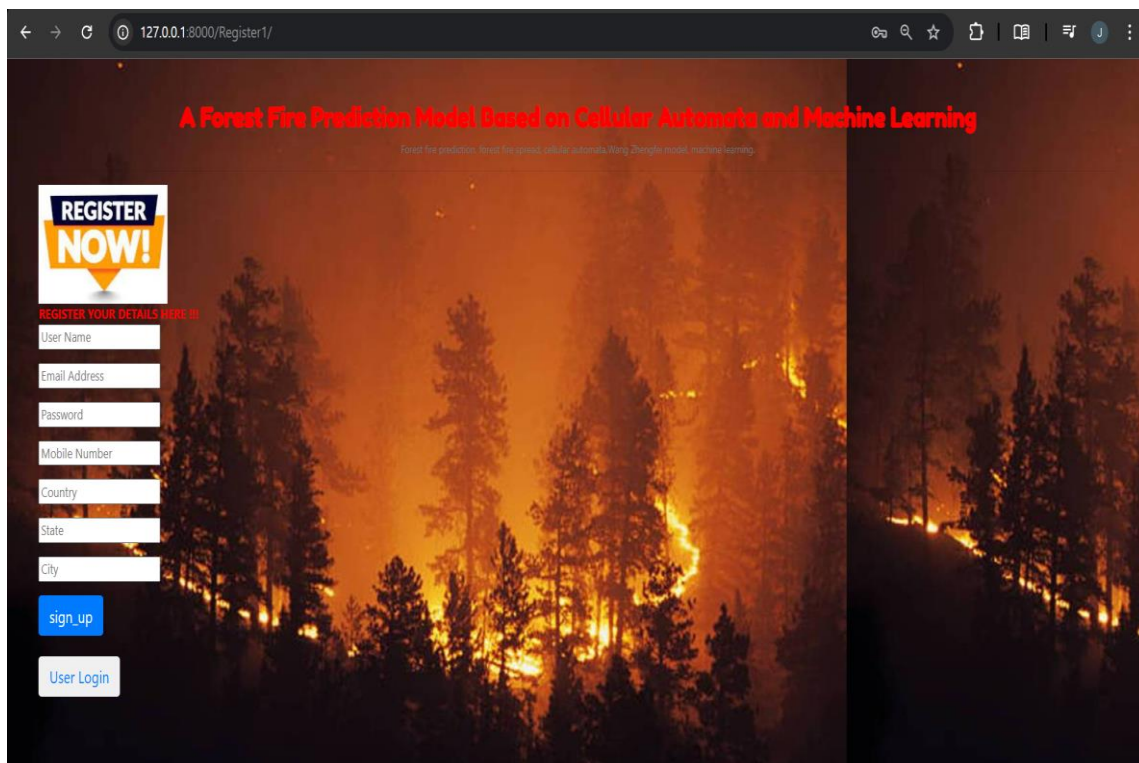


Figure 2. Registration of user

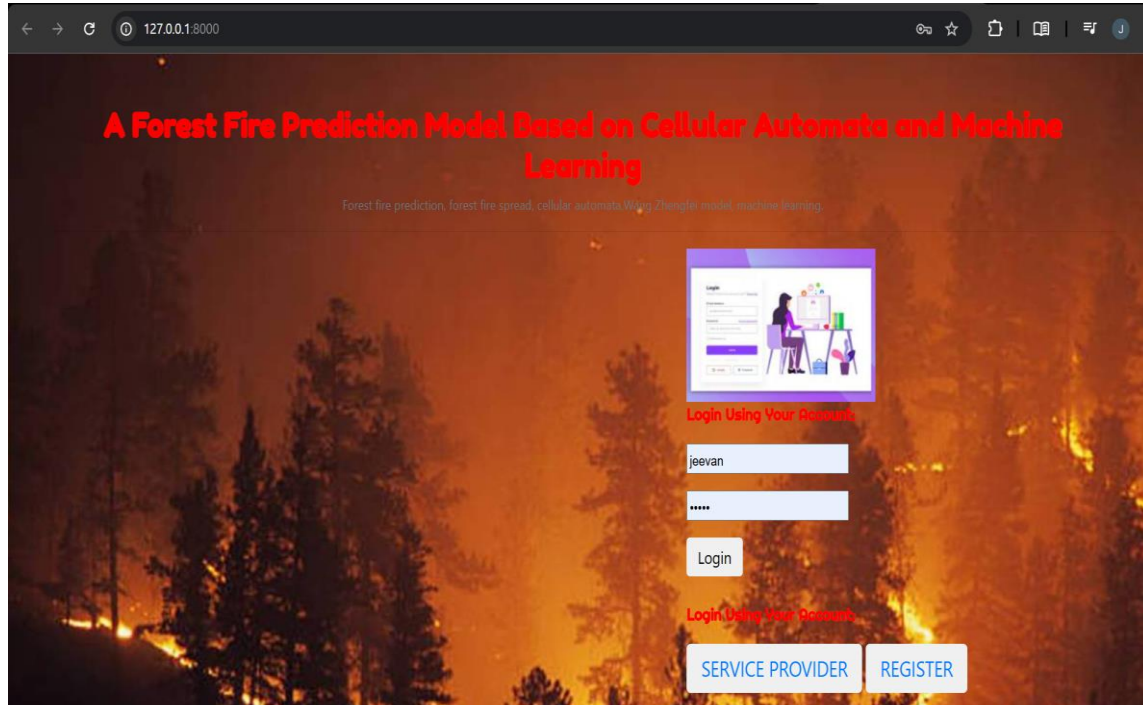


Figure 3. Login interface

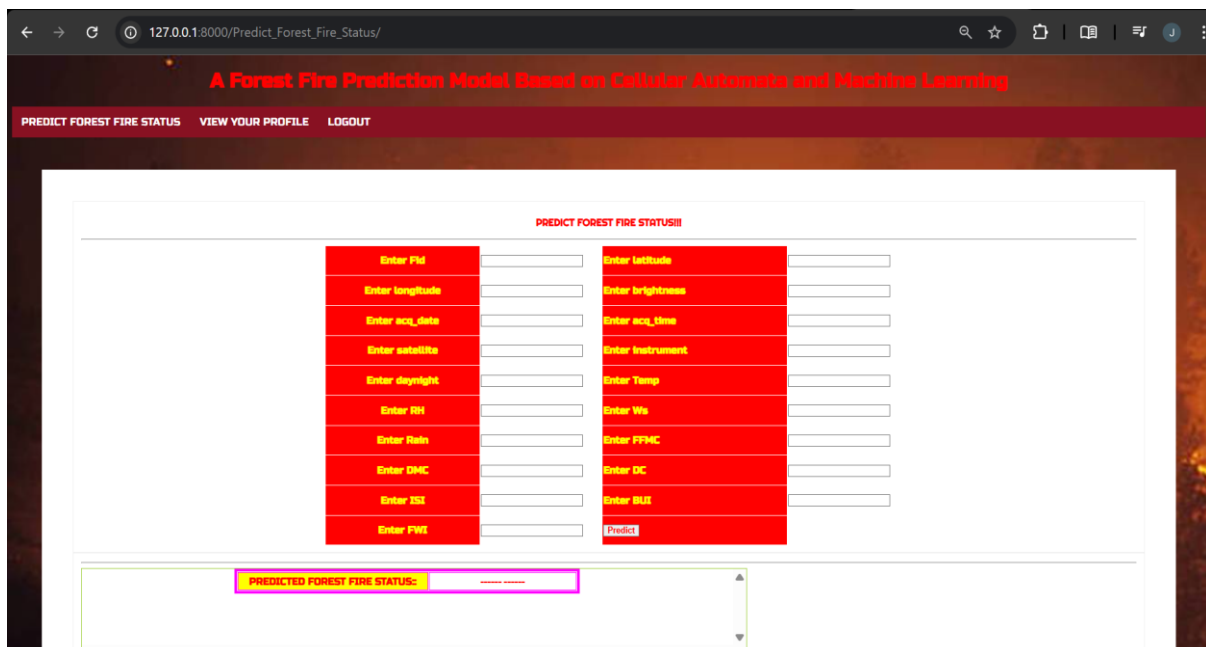
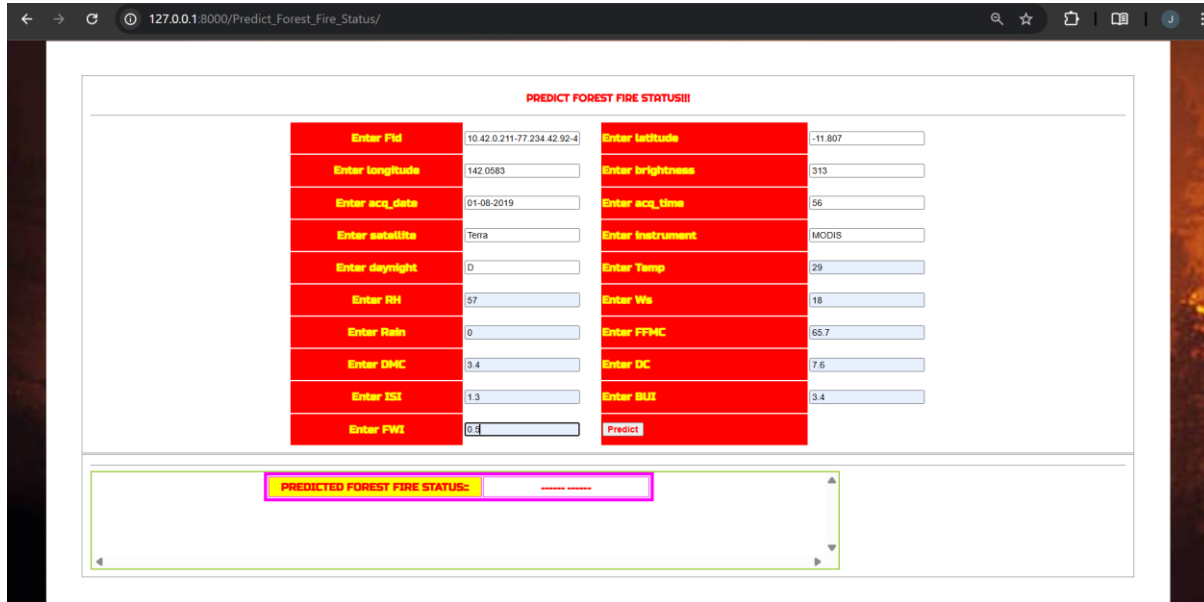


Figure 4. Search location interface



PREDICT FOREST FIRE STATUS!!!

Enter Fld	10.42 0.211-77.234.42.92-4	Enter latitude	11.807
Enter longitude	142.0583	Enter brightness	313
Enter acq_date	01-08-2019	Enter acq_time	56
Enter satellite	Terra	Enter instrument	MODIS
Enter daylight	0	Enter Temp	29
Enter RH	57	Enter Ws	18
Enter Rain	0	Enter FFMC	65.7
Enter DMC	3.4	Enter DC	7.6
Enter ISI	1.3	Enter BUI	3.4
Enter FWI	0.4	Predict	

PREDICTED FOREST FIRE STATUS:- -----

Figure 5. Input search details



PREDICTED FOREST FIRE STATUS:- -----Fire-----

Figure 6. Predicted output

For the given input details as shown in figure 5, the machine learning algorithm classifies the data and helps to predict the output. For the given input details the output predicted is fire which is shown in figure 6.

IV. CONCLUSION

A FFSRP model following an unintentional forest fire is presented in this research. First, the combustible state, climatic elements, and topographical conditions have been identified as the forest fire affecting components that are readily available during the practical implementation of the model, based on pertinent literature and real-world scenarios. The CA model and Wang Zhengfei model are then combined to create the FFSPP model. Furthermore, using machine learning techniques, the FFSRP model is suggested to forecast the burnt area. Lastly, the real fire dataset from Montesinho National Forest Park in Portuga and the "3.29 Forest Fire" in China are used as examples to validate the suggested model and methodology. The suggested FFSPP model's relative error is determined to be lower than that of the Prometheus and Far site fire behaviour simulation models. Additionally, for small and medium-sized fire situations, the suggested FFSRP model can perform well in terms of prediction.

REFERENCES

1. Li, Q.; Cui, J.; Jiang, W.; Jiao, Q.; Gong, L.; Zhang, J.; Shen, X. Monitoring of the fire in Muli County on 28 March 2020, based on high temporal-spatial resolution remote sensing techniques. *Nat. Hazards Res.* **2021**, 1, 20–31.
2. Kariyeva, J.; van Leeuwen, W.J.; Woodhouse, C.A. Impacts of climate gradients on the vegetation phenology of major land use types in Central Asia (1981–2008). *Front. Earth Sci.* **2012**, 6, 206–225.
3. Islam, S.; Bhuiyan, M.A.H. Sundarbans mangrove forest of Bangladesh: Causes of degradation and sustainable management options. *Environ. Sustain.* **2018**, 1, 113–131.

4. Khan, A.; Gupta, S.; Gupta, S.K. Multi-hazard disaster studies: Monitoring, detection, recovery, and management, based on emerging technologies and optimal techniques. *Int. J. Disaster Risk Reduct.* **2020**, *47*, 101642.
5. Kulakowski, D.; Seidl, R.; Holeksa, J.; Kuuluvainen, T.; Nagel, T.A.; Panayotov, M.; Svoboda, M.; Thorn, S.; Vacchiano, G.; Whitlock, C.; et al. A walk on the wild side: Disturbance dynamics and the conservation and management of European mountain forest ecosystems. *For. Ecol. Manag.* **2017**, *388*, 120–131.
6. Collins, B.M.; Kelly, M.; Van Wagtendonk, J.W.; Stephens, S.L. Spatial patterns of large natural fires in Sierra Nevada wilderness areas. *Landsc. Ecol.* **2007**, *22*, 545–557.
7. Neary, D.G.; Robichaud, P.R.; Beyers, J.L. Burned area emergency watershed rehabilitation: Program goals, techniques, effectiveness, and future directions in the 21 century. *Proc. RMRS* **1998**, *375*, 13.
8. Sileshi, G.; Mafongoya, P. The short-term impact of forest fire on soil invertebrates in the miombo. *Biodivers. Conserv.* **2006**, *15*, 3153–3160.
9. Pham, B.T.; Jaafari, A.; Avand, M.; Al-Ansari, N.; Dinh Du, T.; Yen, H.P.H.; Phong, T.V.; Nguyen, D.H.; Le, H.V.; Mafi-Gholami, D.; et al. Performance evaluation of machine learning methods for forest fire modeling and prediction. *Symmetry* **2020**, *12*, 1022.
10. Matin, M.A.; Chitale, V.S.; Murthy, M.S.; Uddin, K.; Bajracharya, B.; Pradhan, S. Understanding forest fire patterns and risk in Nepal using remote sensing, geographic information system and historical fire data. *Int. J. Wildland Fire* **2017**, *26*, 276–286.
11. Van Leeuwen, W.J. Monitoring the effects of forest restoration treatments on post-fire vegetation recovery with MODIS multitemporal data. *Sensors* **2008**, *8*, 2017–2042.
12. Van Leeuwen, W.J.; Casady, G.M.; Neary, D.G.; Bautista, S.; Alloza, J.A.; Carmel, Y.; Wittenberg, L.; Malkinson, D.; Orr, B.J. Monitoring post-wildfire vegetation response with remotely sensed time-series data in Spain, USA and Israel. *Int. J. Wildland Fire* **2010**, *19*, 75–93.
13. Zhenyang, X.; Lin, H.; Wang, F. A small target forest fire detection model based on YOLOv5 Improvement. *Forests* **2022**, *13*, 1332.
14. Li, H.; Long, Z.; Yang, Z.; Xu, Z.; Li, Y. Analysis of forest fire risk in Sichuan Liangshan based on logistic model. *J. Saf. Environ.* **2021**, *21*, 498–505.
15. Huang, Z.; Huang, X.; Fan, J.; Eichhorn, M.P.; An, F.; Chen, B.; Cao, L.; Zhu, Z.; Yun, T. Retrieval of aerodynamic parameters in rubber tree forests based on the computer simulation technique and terrestrial laser scanning data. *Remote Sens.* **2020**, *12*, 1318.
16. Jiecheng, Z. Application of remote sensing technology in forest fire-proof technique. *Inf. Agric. Sci. Technol.* **2019**, *8*, 67–68.
17. Vemulapalli, G. (2024). AI-Driven Predictive Models Strategies to Reduce Customer Churn. *International Numeric Journal of Machine Learning and Robots*, 8(8), 1-13.
18. Chu, C.C.; Zhang, G.; Sun, Y.R. Trend forecast of forest fire in Hunan province based on Kriging interpolation model. *J. Cent. South Univ. For. Technol.* **2014**, *34*, 66–70.
19. Jiachang, L.; Bin, T.; Yuan, Z. GIS-Based spatial and temporal distribution characteristics and factor analysis of forest fires—Taking California, USA as an example. *J. Northeast. For. Univ.* **2020**, *7*, 70–74.
20. Meilin, W.; Linlin, J.; Xiaohong, W.; Bin, W.; Xingxing, X. Forest fire simulation and rescue system based on Geographic Information System. *Inf. Rec. Mater.* **2019**, *9*, 144–145.



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details