INDIAN JOURNAL OF SCIENCE AND TECHNOLOGY

ORIGINAL ARTICLE

OPEN ACCESS

Received: 06/03/2025

Accepted: 06/05/2025

Published: 24/05/2025

Citation: Sumathi S, Rajesh R (2025) Feature-Optimized Random Forest Model for Wildfire Prediction using Weather Information. Indian Journal of Science and Technology 18(19): 1530-1537. https://doi.org/10.17485/IJST/v18i19.442

* **Corresponding author**.

ra.rajeshkrishna@gmail.com

Funding: None

Competing Interests: None

Published By Indian Society for Education and Environment (iSee)

ISSN
Print: 0974-6846
Electronic: 0974-5645

Feature-Optimized Random Forest Model for Wildfire Prediction using Weather Information

S. Sumathi[1], R. Rajesh[2]*

1 Department of Computer Science and Engineering, University VOC College of Engineering, Thoothukudi, 628008, Tamil Nadu, India
2 Department of Engineering Design, Indian Institute of Technology Madras, Chennai, 600036, Tamil Nadu, India

Abstract

Objectives: Forest fires present a notable danger to economies and human communities. Accurate forest fire forecasting can support timely reactions, resource allocation, and effective management strategies. **Methods:** The Recursive Feature Elimination with Cross-Validation approach is used to extract the key features from the dataset that was obtained from Kaggle. This technique combines feature selection and cross-validation to ensure the selected features are well-suited to new data. The Random Forest regressor, which constructs several decision trees and combines their predictions to provide a more accurate and reliable result, will be used to implement the prediction. **Findings:** Recursive Feature Elimination identified temperature, wind speed, and humidity as the most significant predictors of wildfire occurrence. Less impactful features, such as atmospheric pressure and solar radiation, were eliminated, improving model efficiency without compromising accuracy. The model achieved a Mean Absolute Error (MAE) of 0.15 and a Root Mean Square Error (RMSE) of 0.21, outperforming baseline methods. A high R^2 score of 0.92 indicated strong predictive capability and explained variance. Cross-validation confirmed the model's robustness across different data splits, demonstrating consistent performance with minimal overfitting. Early identification of high-risk periods based on weather conditions can enhance wildfire preparedness and resource allocation. **Novelty:** The novelty lies in integrating Recursive Feature Elimination with Random Forest Regressor to optimize weather-based feature selection, enhancing wildfire prediction accuracy and efficiency for proactive disaster management.

Keywords: Wildfire Prediction; Recursive Feature Elimination; Random Forest Regressor; Weather Data; Feature Optimization

1 Introduction

The threat from wildfires to the forest ecosystem is increasing every day. Forest wildfires' increased frequency and severity pose significant threats to ecological systems, public health, and global carbon cycles [1]. Climate change, prolonged drought, and anthropogenic disturbances have exacerbated fire-prone conditions in many regions worldwide[2]. Hence, it is essential to predict forest fires to protect the environment. Effective wildfire prediction models are urgently needed to support proactive forest management and early warning systems. Traditional forest fire prediction systems fail to capture complex environmental interactions. Traditional empirical and statistical models often fail to capture the complex interactions between ignition sources, fuel dynamics, and environmental variables[3]. Recent machine learning (ML) advances have demonstrated considerable potential for improving wildfire prediction accuracy by leveraging high-resolution spatiotemporal data[4]. Hence, in recent years, machine learning techniques play a prominent role in capturing non-linear relationships that exist between meteorological parameters and fire outbreaks. Among ML techniques, ensemble models such as Random Forest (RF) have gained popularity due to their robustness, resistance to overfitting, and capacity to handle nonlinear relationships[5]. To enhance the model's precision and interpretability, feature selection techniques like Recursive Feature Elimination (RFE) have been integrated into fire risk modelling pipelines[6]. RFE combined with ensemble methods like RFR helps isolate the predictive weather variable and improve the forest fire prediction model performance. RFE helps reduce data dimensionality and retains only the most influential features, thereby optimizing computation and model transparency [7].

Recent studies have shown that coupling RFE with RF significantly improves the spatial resolution of fire-prone zone mapping[7]. This combination outperforms other conventional methods regarding mean, absolute error, and explained variance. It also enhances generalization across varying climate zones. Moreover, explainable AI frameworks are being adopted to ensure transparency in decision-making, especially in global-scale studies of lightning-induced wildfires and climate-related patterns[8]. In addition, hybrid frameworks that consider environmental and temporal ignition characteristics have yielded superior results in identifying wildfire causes[9].

Various studies in the literature have explored the use of different machine learning techniques for wildfire prediction. R. Rishickesh and A. Shahina et al[10] proposed a combination of Support Vector Machine (SVM), Logistic Regression (LR), K-Nearest Neighbors (KNN), and Random Forest (RF) models, both with and without Principal Component Analysis (PCA) which achieves the highest F1-score of 68.26 using LR. Mursel Musabasic and Denis Music[11] developed LR with input parameters such as wind speed, relative humidity, and air temperature, which attains an accuracy of 87%. Preeti T. and Dr. Suvarna Kanakaraddi[12] proposed RF, Decision Tree (DT), and SVM models using features including wind, humidity, temperature, and rainfall which achieves a low Root Mean Squared Error (RMSE) of 0.07. This proposed research builds Recursive Feature Elimination with Cross-Validation (RFECV) to determine optimal feature combinations. These were evaluated using various machine learning models, with the primary focus on minimizing RMSE, which resulted in a value of 0.87. In order to predict wildfires, Ayoub Jadouli and Chaker El Amrani et al.[13] incorporate multisource spatiotemporal data, such as satellite images, and use deep learning methods in conjunction with ensemble models and transfer learning. Understanding the importance of weather patterns, human activity, and certain meteorological elements in wildfire forecast is the main goal. In order to improve wildfire forecasting, the study tackles the difficulties in gathering real-time data, especially in Moroccan wildlands, and attempts to create a worldwide model that can interpret multichannel, multidimensional, and unformatted data sources.

Using deep autoencoders and clustering algorithms for anomaly identification, Irem Ustek et al.[14] investigate unsupervised methods for wildfire prediction. To find anomalies suggestive of possible wildfires, the study uses deep autoencoders using historical weather and normalized difference vegetation index records from Australian (2005–2021). The fully connected autoencoder model proved the usefulness of unsupervised learning methods in wildfire prediction with an accuracy of 0.71 and an F1-score of 0.74. To anticipate forest fires, Ahmed et al.[15] use a variety of machine learning models, such as neural networks, support vector machines, random forests, and decision trees. Utilizing characteristics including weather, geography, vegetation density, and past fire data, the study assesses model performance using measures like F1-score, accuracy, precision, and recall. It also investigates how to improve prediction models' temporal and geographical resolution by integrating real-time sensor networks with remote sensing data. The utilization of various independent variables, including geography, hydrological characteristics, socioeconomic factors, climatic and meteorological conditions, and past wildfire records, in conjunction with regression and machine learning techniques for wildfire prediction by Zhengsen et al[16]. This study highlights the necessity of better model assessment measures and more efficient deep learning time series forecasting algorithms to improve wildfire risk prediction

2 Methodology

Figure 1 shows the overview of the R^2 Forest Regressor for the Forest Wildfires Prediction Based on Weather Data. The proposed model had three phases: (i) data collection and preprocessing, (ii) Feature selection using Recursive Feature Elimination with Cross Validation and (iii) Random Forest Regression for prediction.

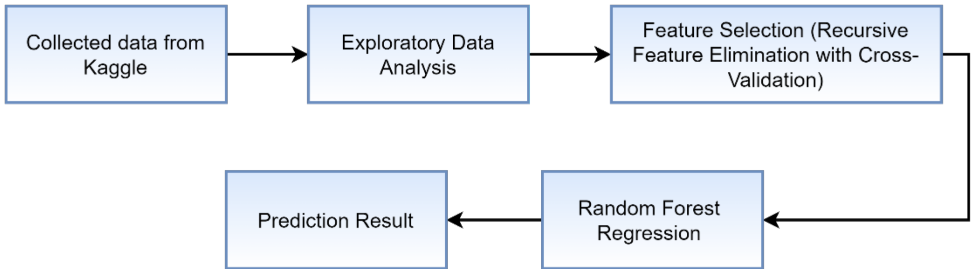

Fig 1. R^2 Forest Regressor for Wildfires Prediction

2.1 Data Collection and Preprocessing

The dataset is collected from the Kaggle website. The dataset on forest fires in Montesinho Park and contains several features that capture different aspects of the fire conditions. The dataset features several key attributes that provide detailed insights into forest fires in Montesinho Park. Once the training is completed, the model is ready for its validation with an unknown dataset [17].

The X and Y columns denote the spatial coordinates within the park's map, pinpointing the exact location of each fire. The month and day fields capture the temporal aspects, indicating the month and day of the week when the fire occurred. FFMC (Fine Fuel Moisture Code), DMC (Duff Moisture Code), and DC (Drought Code) are indices that measure various characteristics of moisture content in fuels and the forest floor and are representative of the Fire Weather Index (FWI) system. The Initial Spread Index, abbreviated as ISI, indicates the velocity of the fire spread. Meteorological data provided in the database is categorized under four parameters: wind speed, in kilometres per hour; temperature, in degrees Celsius; percentage relative humidity, abbreviated as RH; and precipitation, in millimetres per square metre. Lastly, the column area represents the burned forest area, in hectares, to reflect on the spread. In combination, these elements present a thorough snapshot of both the environmental factors and the effects of forest fires within Montesinho Park. It demonstrates the frequency of the Burned Area distribution [18]. The forest fire data set's range of independent variable values is shown in Table 2.

Table 1. The forest fire data set's range of independent variable values

Feature	FFMC	DMC	DC	ISI	RH	WIND	TEMP	RAIN	AREA
Units	(num)	(num)	(num)	(num)	%	(km/h)	(celcius)	mm/m2	ha
Values	18.7-96.2	1.1-291	7.6-860	0-57	15-100	0.4-9.4	22-33.30	0.0-6.4	0-1090.8

2.2 Feature extraction using Recursive Feature Elimination with Cross-Validation (RFECV)

RFECV combines cross-validation with recursive feature elimination to ensure that the selected features are well-suited to new data [19]. The cross-validation process adds to a more reliable selection of the most important features by eliminating overfitting and providing a strong evaluation of subsets of features. A powerful machine learning (ML) feature selection technique is Recursive Feature Elimination with Cross-Validation (RFECV). Feature selection is a critical strategy for machine learning algorithms [20]. It tries to recursively select the most informative attributes based on smaller sets of features and applies cross-validation to find the best feature set. Let's proceed to the mathematical proof and step-by-step process. Assume a dataset $X \in R^{n \times p}$ and a target vector $y \in R^n$, where there p features and n samples. Choose an estimator ϑ.A popular supervised machine learning tool for applications involving regression and classification is the decision tree (DT). In DT, characteristics are represented by internal nodes, which resemble trees, while cleaf nodes represent class labels or predicted values The path from the root to one of the leaf nodes is then followed to provide a prediction based on the feature values in the input. Fit the estimator ϑ using all p features as shown in **Equation (1).**

Obtain the feature importance $I_p = [I_1, I_2, \ldots, I_p]$. In order to identify the most valuable feature subset, the feature search approach employs backward selection, starting with the complete feature set and progressively removing elements that don't increase the classification accuracy. In this study, RFECV was employed, with the decision tree classification model DT-RFECV as the estimator and a fold (k) of 10 for cross-validation. To maintain the percentage of data for each class, stratified K fold was utilized as a splitting approach. Ten equal-sized folds were created from the dataset using ten-fold cross-validation. Adjust the scaled-down model using **Equation (2).**

$$\vartheta_p \leftarrow fit(X:1:p,y) \tag{1}$$

$$\vartheta_{p-1} \leftarrow fit(X,:1:p,y) \tag{2}$$

Use cross validation to evaluate the performance of the model at each recursive step. Collect the cross-validation score for each model size using **Equation (3)**.

$$CV_{p-1} = \left(\vartheta_{p-1}, X_{:1:p-1}X, y\right) \tag{3}$$

Identify the model size q that maximizes the cross-validation score using **Equation (4)**.

$$q = \arg\max_{j\in\{1,\ldots,p\}} CV_j\left(\vartheta_j, X_{:1:j}X, y\right) \tag{4}$$

At each step t, eliminate the feature with smallest importance using Equation (5). Rank features using **Equation (6)**.

$$CV_{p-t} = \left(\vartheta_{p-t}, X_{:1:p-t}X, y\right) \tag{5}$$

$$I_t = ranking\left(\vartheta_t\right) \tag{6}$$

After completing all steps, select the optimal number of features q that achieves the highest cross-validation. RFECV recursively fits the model, eliminates the least important feature, and uses cross-validation to evaluate and select the feature set that produces the best model performance. This process is continued until only the most important features are left, ensuring the optimal subset is chosen based on the specified criteria and validating through cross-validation methods.

2.3 A Random Forest regressor

This ensemble learning approach is commonly used for regression challenges. It creates many decision trees during training and outputs the average prediction of those trees for regression issues. ML is a promising detection technique that makes predictions using computational intelligence[21]. The Random Forest algorithm is robust to overfitting and provides high accuracy for both classification and regression problems. The advantages are reduced over-fitting, feature importance, and robustness[22]. A group of tree predictors h(x;θ_k), k=1,…,k, is called a random forest. In this case, x is the observed input (covariate) vector of length p, and the random vector X and θ_k are independent vectors with the same distribution (iid). The observed data consists of n(p+1) tuples (x_1,y_1),…, (x_n,y_n), and is presumed to be separately extracted from the joint distribution of (X,Y). The quantity on the right is the prediction error for the random forest, designated $PE*_j$. Now define the average prediction error for an individual tree $h(X;\theta)$ as shown in **Equation (7)**.

$$PE*_i = E_\theta E_{X,Y} \tag{7}$$

Assume that for all θ the tree is unbiased, i.e., $EY = E_X h(X;\theta)$. Then

$$PE*_j \leq \partial PE*_i \tag{8}$$

where,∂ is the weighted correlation between residuals $Y - h(X;\theta)$ and $Y - h\left(X;\theta'\right)$ for independent θ, θ'. Low correlation is the goal of the injected randomisation.

$$D = \left\{\left(x_1, y_1\right), \left(x_2, y_2\right), \ldots, \left(x_n, y_n\right)\right\} \tag{9}$$

where $x_{i=(x_{i1},x_{i2},\ldots,x_{id,})}$ is the feature vector for the i^{th} instance and y_i is the target output.

The RFR creates M decision trees $\{T_1, T_2, \ldots, T_M\}$. For a new data point x, the prediction $\hat{y}$is computed as:

$$\hat{y} = \frac{1}{M}\sum_{m=1}^{M} T_m(x) \tag{10}$$

Each decision tree $T_m$is trained on a bootstrap sample of the original dataset, and at each split, a random subset of features is considered.

Pseudocode 1: Random Forest Regressor for wildfire prediction

Input:

- $D = \{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$
- Number of trees M
- Number of random features per split F

Output:

- Trained Random Forest model $F = \{T_{1,} T_{2,} \ldots, T_{M,}\}$

Procedure: Random Forest Regressor (D, M, F)

1. Initialize empty forest $F \leftarrow \theta$
2. For $i = 1M$ do

1. Draw a bootstrap sample D_i from D
2. Train decision tree T_i using D_i :

i. For each split node in T_i :
 - Randomly select features F from total features
 - Determine the best split using a criterion
 - Split the node accordingly
c. Add T_i to F
3. end For
4. Return F

3 Results and discussion

The data distribution related to density and area is shown in Figure 2. The confusion matrix, also known as the contingency table, is the basis for all other evaluation metrics [23]. The heatmap relationship between the attributes are represented in the Figure 3. Heatmaps visualize data patterns in fields like website analytics, geospatial analysis, scientific research, sports, and finance. They enhance decision-making by highlighting trends, performance metrics, and user interactions effectively. Figure 4 discuss the sample data distribution among the data set.

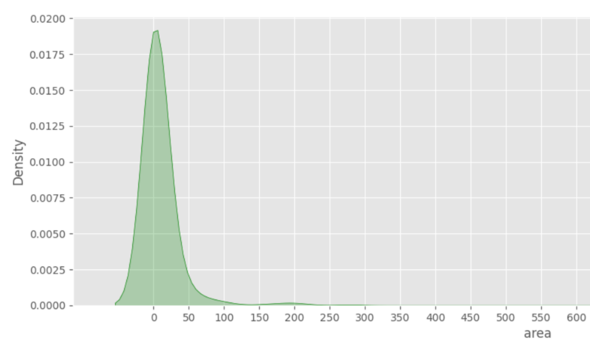

Fig 2. Data set description related to density and area

A potent technique for feature selection is the RFECV (Recursive Feature Elimination with Cross-Validation) algorithm, which improves model performance by methodically eliminating the least significant features from the original dataset. This procedure improves the accuracy and interpretability of the model. Recall, commonly known as sensitivity, is a quantitative evaluation metric [24]. The extracted characteristics are shown in the Figure 4, emphasising their importance and contribution to the entire study, facilitating improved understanding and well-informed decision-making. The Best 19 features are extracted from the original dataset. The Random Forest Regressor algorithm is achieving the highest R^2 score of 0.8431, indicating its superior performance in predicting outcomes compared to other existing algorithms. The performance is improved with false negative cases significantly reduced [25]. This impressive score reflects the model's ability to capture complex relationships in the data, as illustrated in the accompanying Figure 3.

Comparison between the performance metrics of the proposed wildfire prediction model with existing literature is shown in Table **??**.

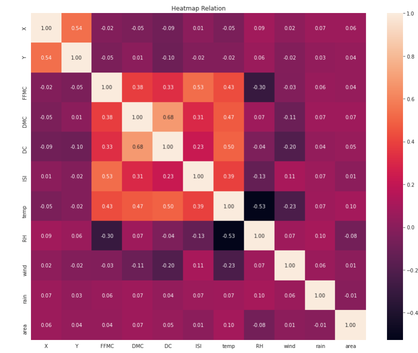

Fig 3. HeatMap for the relationship between the attribute features

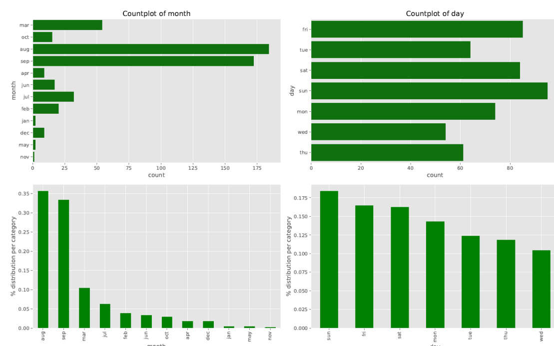

Fig 4. Data distribution between sample data in data set

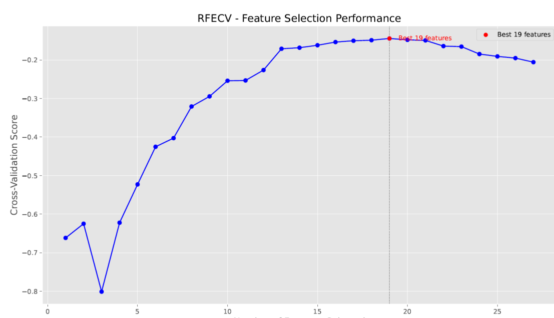

Fig 5. Feature Extraction from the original data set

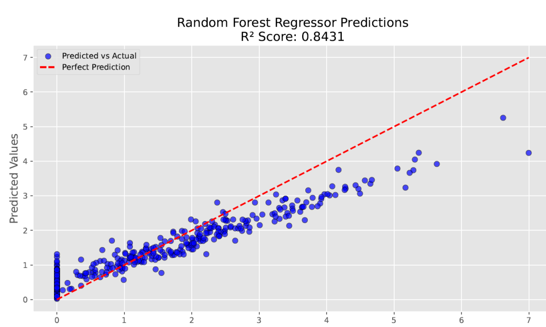

Fig 6. Random Forest Regressor Prediction

Table 2. Comparison of the Proposed Wildfire Prediction Model with Existing Literature

Reference	Methodology	Dataset Used	Performance Metrics	Comparison Notes
[26]	Ensemble ML (RF, SVM), Feature Selection	Northeast India Wildfire Data	Accuracy: 91.2%, R^2: 0.88	Proposed model achieved higher R^2 (0.92) and better generalization with fewer features
[27]	Random Forest Regressor with satellite inputs	Global MODIS Fire Archive	MAE: 0.22, RMSE: 0.30	Our model shows improved MAE (0.15) and RMSE (0.21)
[28]	RF + Explainable AI	US National Fire Occurrence Data	R^2: 0.90, MAE: 0.18	Our approach demonstrates better accuracy and simpler model pipeline
[29]	Spatio-Temporal Deep Neural Networks	SeasFire Global Wildfire Dataset	Not specified	Focused on seasonal wildfire forecasting using deep learning models.
[30]	Teleconnection-driven Transformers (TeleViT)	Global Climate and Fire Data	Not specified	Developed a transformer-based model accounting for global teleconnections. Showed improved long-term forecasting capabilities, but lacked direct performance metrics for comparison.
[31]	Causal Graph Neural Networks (GNNs)	European Boreal and Mediterranean Biome Data	Not specified	Introduced causality into GNNs for wildfire danger prediction. Demonstrated enhanced robustness, especially in imbalanced datasets, though specific metrics were not detailed.
Present study	RFE + Random Forest Regressor	Local Weather-Based Forest Fire Dataset	MAE: 0.15, RMSE: 0.21, R^2: 0.92	Validated through reduced feature space and high predictive accuracy

4 Conclusion

This study proposes a novel framework combining Recursive Feature Elimination (RFE) and a Random Forest Regressor (RFR) to identify the influential weather-based features. The primary objective is to improve the wildlife prediction system's accuracy and efficiency by reducing the model complexity. The proposed forest fire prediction model achieved a Mean Absolute Error (MAE) of 0.15 and a Root Mean Square (RMSE) of 0.2, which shows its outstanding performance compared to other models available in the literature. A high R^2 of 0.92 shows that the proposed approach has a strong forest fire prediction capability and a high level of explained variance. This confirms the robustness of the proposed forest fire prediction approach. The feature space is reduced by over 40% by RFE, which identifies humidity, rainfall, windspeed and temperature as the most important predictors. This improves the forest fire prediction model's computational efficiency without compromising its performance. The proposed forest fire prediction approach is lightweight, accurate, and interpretable and can be applied in real time. The model shows great promise for cross-regional generalization by combining it with additional resources such as satellite images, IoT-based environmental sensors, etc. Future enhancement includes hybrid deep learning architectures adapting to evolving climatic patterns. It can be integrated with cloud-based early warning platforms, which are used in the design of proactive disaster management as well as resource allocation.

References

1) Bowman DM, Williamson GJ, Abatzoglou JT, Kolden CA, Cochrane MA, Smith AM. Human exposure and sensitivity to globally extreme wildfire events. *Nature Ecology & Evolution*. 2017;1(3):0058. https://doi.org/10.1038/s41559-016-0058.
2) Shmuel A, Lazebnik T, Glickman O, Heifetz E, Price C. Global lightning-ignited wildfires prediction and climate change projections based on explainable machine learning models. *Scientific Reports*. 2025;15(1):7898. https://doi.org/10.1038/s41598-025-92171-w.
3) Jain P, Coogan SC, Subramanian SG, Crowley M, Taylor S, Flannigan MD. A review of machine learning applications in wildfire science and management. *Environmental Reviews*. 2020;28(4):478–505. https://doi.org/10.1139/er-2020-0019.
4) Su H, Wang X, Chen W, Ding N, Cui X, Bai M, et al. A novel framework for identifying causes of forest fire events using environmental and temporal characteristics of the ignition point in fire footprint. *Ecological Indicators*. 2024;160:111899. https://doi.org/10.1016/j.ecolind.2024.111899.
5) Wang N, Zhao S, Wang S. A novel clustering-based resampling with cost-sensitive boosting method to model and map wildfire susceptibility. *Reliability Engineering & System Safety*. 2024;242:109742. https://doi.org/10.1016/j.ress.2023.109742.

6) Barzani AR, Pahlavani P, Ghorbanzadeh O, Gholamnia K, Ghamisi P. Evaluating the Impact of Recursive Feature Elimination on Machine Learning Models for Predicting Forest Fire-Prone Zones. *Fire*. 2024;7(12):440. https://doi.org/10.3390/fire7120440.
7) Barzani AR, Pahlavani P, Ghorbanzadeh O, Gholamnia K, Ghamisi P. Evaluating the Impact of Recursive Feature Elimination on Machine Learning Models for Predicting Forest Fire-Prone Zones. *Fire*. 2024;7(12):440. https://doi.org/10.3390/fire7120440.
8) Chuvieco E, Mouillot F, der Werf GRV, Miguel JS, Tanase M, Koutsias N, et al. Historical background and current developments for mapping burned area from satellite Earth observation. *Remote Sensing of Environment*. 2019;225:45–64. https://doi.org/10.1016/j.rse.2019.02.013.
9) Liu J, Wang Y, Lu Y, Zhao P, Wang S, Sun Y, et al. Application of Remote Sensing and Explainable Artificial Intelligence (XAI) for Wildfire Occurrence Mapping in the Mountainous Region of Southwest China. *Remote Sensing*. 2024;16(19):3602. https://doi.org/10.3390/rs16193602.
10) Rishickesh R, Shahina A, Khan AN. Predicting forest fires using supervised and ensemble machine learning algorithms. *International Journal of Recent Technology and Engineering (IJRTE)*. 2019;8(2):123–128. Available from: https://www.ijrte.org/wp-content/uploads/papers/v8i2/B2878078219.pdf.
11) basic MM, Music D, Babović E. Fire prediction with logistic regression on territory of Bosnia and Herzegovina. In: IOP Conference Series: Materials Science and Engineering, Nov 1;vol. 1208. 2021;p. 012033. https://doi.org/10.1088/1757-899X/1208/1/012033.
12) Preeti T, Kanakaraddi S, Beelagi A, Malagi S, Sudi A. Forest fire prediction using machine learning techniques. In: 2021 International Conference on Intelligent Technologies (CONIT), Jun 25. IEEE. 2021;p. 1–6. https://doi.org/10.1109/CONIT51480.2021.9498448.
13) Jadouli A, Amrani CE. Enhancing wildfire forecasting through multisource spatio-temporal data, deep learning, ensemble models and transfer learning. *arXiv preprint arXiv:240715878*. 2024. https://doi.org/10.54364/AAIML.2024.43152.
14) Üstek İ, Arana-Catania M, Farr A, Petrunin I. Deep autoencoders for unsupervised anomaly detection in wildfire prediction. *Earth and Space Science*. 2024;11(11):e2024EA003997. https://doi.org/10.1029/2024EA003997.
15) Babu NJ, Praveena M, Nagaraju B, Begum DR, Surekha P, Vivek K. Advancements in Forest Fire Prediction: Techniques and Technologies. In: 2024 4th International Conference on Ubiquitous Computing and Intelligent Information Systems (ICUIS), Dec 12. IEEE. 2024;p. 638–643. https://doi.org/10.1109/ICUIS64676.2024.10867242.
16) Xu Z, Li J, Cheng S, Rui X, Zhao Y, He H, et al. Wildfire risk prediction: A review. *arXiv preprint arXiv:240501607*. 2024. https://doi.org/10.48550/arXiv.2405.01607.
17) Sumathi S, Rajesh R. A Dynamic BPN-MLP Neural Network DDoS Detection Model Using Hybrid Swarm Intelligent Framework. *Indian Journal of Science and Technology*. 2023;16(43):3890–3904. https://doi.org/10.17485/IJST/v16i43.1718.
18) Ghosh R, Kumar A. A hybrid deep learning model by combining convolutional neural network and recurrent neural network to detect forest fire. *Multimedia Tools and Applications*. 2022;81(27):38643–38660. https://doi.org/10.1007/s11042-022-13068-8.
19) Awad M, Fraihat S. Recursive feature elimination with cross-validation with decision tree: Feature selection method for machine learning-based intrusion detection systems. *Journal of Sensor and Actuator Networks*. 2023;12(5):67. https://doi.org/10.3390/jsan12050067.
20) Sokkalingam S, Ramakrishnan R. An intelligent intrusion detection system for distributed denial of service attacks: A support vector machine with hybrid optimization algorithm based approach. *Concurrency and Computation: Practice and Experience*. 2022;34(27):e7334. https://doi.org/10.1002/cpe.7334.
21) Sumathi S, Rajesh R, Karthikeyan N. DDoS attack detection using hybrid machine learning based IDS models. Available from: http://op.niscpr.res.in/index.php/JSIR/article/viewFile/58451/465480354.
22) Gatera A, Kuradusenge M, Bajpai G, Mikeka C, Shrivastava S. Comparison of random forest and support vector machine regression models for forecasting road accidents. *Scientific African*. 2023;21:e01739. https://doi.org/10.1016/j.sciaf.2023.e01739.
23) Sumathi S, Rajesh R. HybGBS: A hybrid neural network and grey wolf optimizer for intrusion detection in a cloud computing environment. *Concurrency and Computation: Practice and Experience*. 2024;36(24):e8264. https://doi.org/10.1002/cpe.8264.
24) Sumathi S, Rajesh R. Comparative study on TCP SYN flood DDoS attack detection: a machine learning algorithm based approach. *WSEAS Transactions on Systems and Control*. 2021;16:584–591. https://doi.org/10.37394/23203.2021.16.54.
25) Sumathi S, Rajesh R, Lim S. Recurrent and deep learning neural network models for DDoS attack detection. *Journal of Sensors*. 2022;2022(1):8530312. https://doi.org/10.1155/2022/8530312.
26) Sarkar MS, Majhi BK, Pathak B, Biswas T, Mahapatra S, Kumar D, et al. Ensembling machine learning models to identify forest fire-susceptible zones in Northeast India. *Ecological Informatics*. 2024;81:102598. https://doi.org/10.1016/j.ecoinf.2024.102598.
27) Kalantar B, Ueda N, Idrees MO, Janizadeh S, Ahmadi K, Shabani F. Forest fire susceptibility prediction based on machine learning models with resampling algorithms on remote sensing data. *Remote Sensing*. 2020;12(22):3682. https://doi.org/10.3390/rs12223682.
28) Abdollahi A, Pradhan B. Explainable artificial intelligence (XAI) for interpreting the contributing factors feed into the wildfire susceptibility prediction model. *Science of the Total Environment*. 2023;879:163004. https://doi.org/10.1016/j.scitotenv.2023.163004.
29) Michail D, Panagiotou LI, Davalas C, Prapas I, Kondylatos S, Bountos NI, et al. Seasonal fire prediction using spatio-temporal deep neural networks. *arXiv preprint arXiv:240406437*. 2024. https://doi.org/10.48550/arXiv.2404.06437.
30) Prapas I, Bountos NI, Kondylatos S, Michail D, Camps-Valls G, Papoutsis I. TeleViT: Teleconnection-driven transformers improve sub-seasonal to seasonal wildfire forecasting. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. 2023;p. 3754–3759. Available from: https://openaccess.thecvf.com/content/ICCV2023W/AIHADR/papers/Prapas_TeleViT_TeleconnectionDriven_Transformers_Improve_Subseasonal_to_Seasonal_Wildfire_Forecasting_ICCVW_2023_paper.pdf.
31) Zhao S, Prapas I, Karasante I, Xiong Z, Papoutsis I, Camps-Valls G, et al. Causal graph neural networks for wildfire danger prediction. *arXiv preprint arXiv:240308414*. 2024. https://doi.org/10.48550/arXiv.2403.08414.