The Report is Generated by DrillBit Plagiarism Detection Software

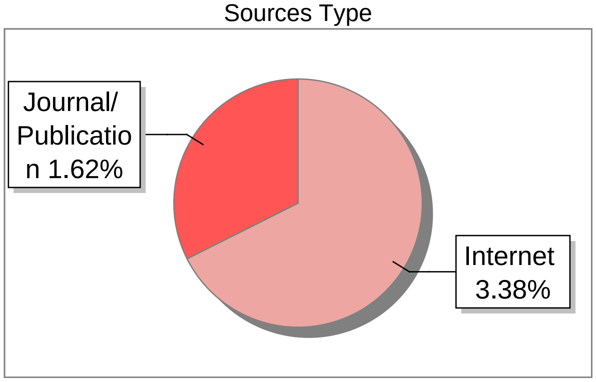
Submission Information

Author Name	Keerthana C N
Title	Forest Fire Risk Assessment
Paper/Submission ID	3594344
Submitted by	charankumarba@gmail.com
Submission Date	2025-05-08 13:47:25
Total Pages, Total Words	6, 2591
Document type	Research Paper

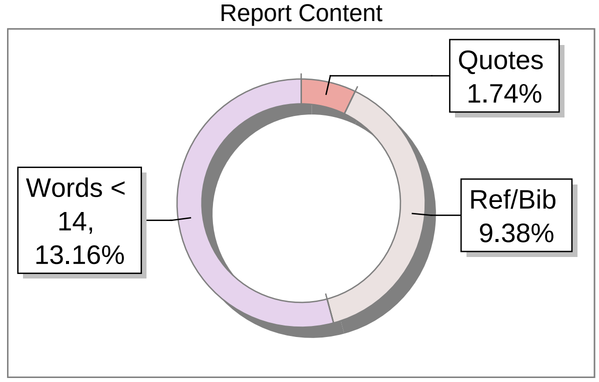
Result Information

Similarity **5 %**

Sources Type



Report Content



Exclude Information

Quotes	Excluded
References/Bibliography	Excluded
Source: Excluded < 14 Words	Excluded
Excluded Source	**0 %**
Excluded Phrases	Not Excluded

Database Selection

Language	English
Student Papers	Yes
Journals & publishers	Yes
Internet or Web	Yes
Institution Repository	Yes

A Unique QR Code use to View/Download/Share Pdf File

DrillBit Similarity Report

5	6	A	A-Satisfactory (0-10%) B-Upgrade (11-40%) C-Poor (41-60%) D-Unacceptable (61-100%)
SIMILARITY %	MATCHED SOURCES	GRADE	

LOCATION	MATCHED DOMAIN	%	SOURCE TYPE
1	link.springer.com	<1	Internet Data
2	journals.vilniustech.lt	1	Internet Data
3	frontiersin.org	1	Internet Data
5	Thesis Submitted to Shodhganga Repository	1	Publication
6	moam.info	1	Internet Data
7	www.dx.doi.org	1	Publication

Forest Fire Risk Assessment

Keerthana C N, Harshitha V M

Student, Department of Computer Science and Engineering, The Oxford College Of Engineering, Bangalore-560068, Karnataka, India

keerthugowda12e@gmail.com, harshithavm4@gmail.com

Abstract — The biosphere has been seriously threatened by wildfires in recent years, which frequently cause financial and environmental harm. For the purpose of taking preventative action, wildfire probabilities must be accurately and promptly predicted. The Artificial Neural Network (ANN) method for forecasting the likelihood of herbage, topographical, meterological, and fire weather index data is reviewed in this research. To evaluate the model, performance metrics are employed. Machine learning algorithms are a means of making predictions. The best course of action is to pre-process the data by normalising and separating it into training and testing data.

Keywords–Topographical, Meterological, Fire Weather Index, ANN, ReLu Activation

I. Introduction

Wildfires and bushfires are other names for forest fires. It is an unintentional fire in a foilage-ignitable area. A wildfire can be described more precisely as a bushfire, grass fire, desert fire, hill fire, preat fire, or vegetation fire, depending on the kind of vegetation that is present. Wildfires are categorised based on their physical characteristics, meteorological impact, flammable material content, and cause of ignition[1].

These kinds of disasters are frequent in the US, Canada, Australia,Russia, Washington, and California. Particularly vulnerable are regions with Mediterranean or taiga biomes. When flammable material exposed to heat and an igniting source come into contact, wildfires occur.

In general, denser forest patches provide more shade, which lowers temperatures and increases humidity, making them less vulnerable to flames. Since they contain less water than denser materials like branches and trunks, less dense materials like grasses and leaves are simpler to burn. Technology like controlled burns and WFU (wildland fire use) may be incorporated into wildfire prevention. The suggested system aims to leverage cutting-edge machine learning algorithms that make predicting and detecting wildfires much easier[3].

II. Literature Review

The publications under review all used a range of modeling techniques and algorithms to address forest cover change, post-fire recovery, mercury cycling, and wildfire detection. Landsat satellite data was used to classify land use and land cover using machine learning techniques like Random Forest (RF) and Support Vector Machine (SVM). SVM performed well in high-dimensional spaces with small sample sizes for classification, whereas RF provided good accuracy, noise resistance, and was well suited for large datasets[5].

Deep learning models like AlexNet and custom Convolutional Neural Networks (CNN) were used to classify images that had a fire versus those that did not. The accuracy of CNN was higher than that of AlexNet (88.19%).

Space–Time Scan Statistics Permutation (STSSP) models were used to detect spatiotemporal fire clusters and establish a relationship between fire frequency and biodiversity risks. We examined the relationships among species features, post-fire tree regeneration densities, and latitude using linear mixed models that included both fixed and random factors.

A dynamic multimedia system model was created following wildfires to incorporate complex biogeochemical interactions and simulate the movement of mercury (Hg) through various environmental media. Maps of individual trees, clumps, and canopy gaps require very accurate 3D reconstructions of forest structures, which were also produced by drone-based Structure from Motion(SfM)[9].

III. Proposed Methodology

Decision Tree : A supervised learning method for regression and classification models is the decision tree. Using characteristic values to divide the provided data into branches, it may make decisions at the leaf nodes. Every leaf node symbolizes a final product, internal nodes show the choice, and branches show the result of the choice[7].

Often, the decision tree approach starts with the entire dataset. Then, utilizing information gain (for classification) or variance reduction (for regression), the optimal feature is chosen to separate the data. Create two or more subgroups from the provided data, then repeat the process recursively for each subset. This process ends when all the samples are in the same class, the maximum depth is reached, or there are no more possible splits.

ANN(Artificial Neural Network): The biological neural networks present in animal brains serve as the model for an artificial neural network (ANN), a form of computer system. It interprets input data by employing layers of interconnected "neurones" to detect complex patterns. It learns by changing weights when it makes mistakes.
It is common for ANNs to initialize random weights and biases. It calculates losses as well as forward and backward passes. Data is passed across layers using the forward pass method to help calculate the result. Compare the expected and actual values to determine the loss. In the backward pass, weights are updated using gradients to minimize loss[6].

Clustering: With clustering, an unsupervised learning method, data points are grouped into clusters based on how similar they are. In this instance, the algorithm finds structure in the data on its own without labels. As initial cluster centers, clustering usually starts with the selection of K randomly selected points known as centroids. The centroids are then updated by figuring out the mean of all the points added to that cluster. Each point is then assigned to the nearest centroid using a distance, like the Euclidean distance. Until the centroids show minimal change, this process will be repeated.

SARSA(State-Action-Reward-State-Action): SARSA is an on-policy reinforcement learning approach. The update includes both current and planned activities, adjusting the Q-value based on the action the agent actually performed.

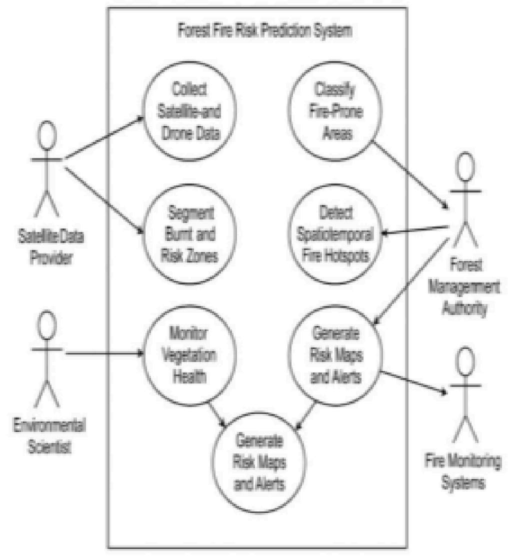


Fig1: Use case diagram for Forest Fire Prediction

IV. Implementation

A. Data Acquisition Reliable sources such as NOAA weather databases, NASA's MODIS, and forest fire datasets are used to gather data. Typically, the data comprises vegetation kinds, topographical parameters (elevation, slope), weather conditions (temperature, wind speed, humidity), and historical fire events together with their location and date.

B. Data Exploration and Visualization Data is gathered from dependable sources, including NASA's MODIS, NOAA meteorological databases, and forest fire information. Vegetation types, topographical features (slope, elevation), meteorological factors (temperature, wind speed, humidity), and past fire incidents along with their location and date are typically included in the data.

C. Data Cleaning Reliable sources of data are used, such as NOAA weather databases, NASA's MODIS, and data on forest fires. Data usually include vegetation kinds, topographical parameters (slope, elevation), weather conditions (temperature, wind speed, humidity), and previous fire incidences, along with the date and location of the incident.

D. Feature Engineering The performance of the model is enhanced by generating new characteristics from pre-existing ones. Temperature, humidity, and wind, for example, can be used to calculate a fire weather index. Repetitive or unnecessary features are eliminated, and features are scaled (normalized) to guarantee consistent input to the ANN.

E. Model Selection The ability of a feedforward Artificial Neural Network (ANN) to capture intricate nonlinear interactions is the reason for its selection. The architecture is made up of an output layer that uses sigmoid activation for probability prediction, hidden layers using ReLU activation, and an input layer that matches the amount of features.

F. Model Training Subsets of the dataset are separated for testing, validation, and training. With backpropagation and an optimizer such as Adam, the ANN is trained using the training data. To avoid overfitting, early halting or dropout is utilized. The number of neurons and learning rate are examples of hyperparameters that are adjusted.

G. Model Evaluation The test set is used to assess the trained model using metrics like ROC-AUC, F1-score, recall, accuracy, and precision. The visualization of true positives and false negatives, which are essential in wildfire forecast, is aided by a confusion matrix. Model robustness is ensured by cross-validation.

H. Interpretation and Analysis The decision-making process of the model is analyzed using interpretability techniques such as SHAP or LIME. Users and decision-makers can trust and act upon the model's findings by using these to emphasize which parameters (such as temperature and wind speed) have the most influence on the wildfire probability estimate.

I. Visualization of Results Visualization of the Findings Using bar charts, maps, and heatmaps, final forecasts and their probability are displayed. These aid in successfully presenting high-risk areas. GIS maps or interactive dashboards can be used to show real-time fire risk forecasts and support in-the-moment wildfire management decision-making.

V. Results

The dataset can have blank columns, extraneous rows, or data that isn't formatted correctly. The first row, which could be a repeating header or a description, is thus skipped when using pd.read_csv().

To avoid mistakes, empty or unidentified columns are eliminated, rows with missing data are deleted, and column names are cleared of spaces. In order to make sure that numerical columns such as Temperature, RH (Relative Humidity), FFMC, and FWI may be statistically evaluated, we next transform them into the appropriate numeric types. The Classes column, which displays "fire" or "not fire," is finally changed to a new column called Fire, which has values of 1 or 0. This makes it simpler to visualize and train on.

Bar Charts – Class Distributions

A bar chart (or bar graph) is a graphical representation of data using rectangular bars to show the frequency, count, or value of different categories. It helps you easily compare data across different groups.

To illustrate how many records fall into the "fire" and "not fire" classes, bar charts are created independently for the training and testing sets. With the aid of these charts, you can rapidly determine whether the dataset is unbalanced. For example, if there are a lot more "not fire" records, any machine learning model may start to predict "not fire" more frequently. For a meaningful evaluation, the distributions of the training and test datasets must be comparable.

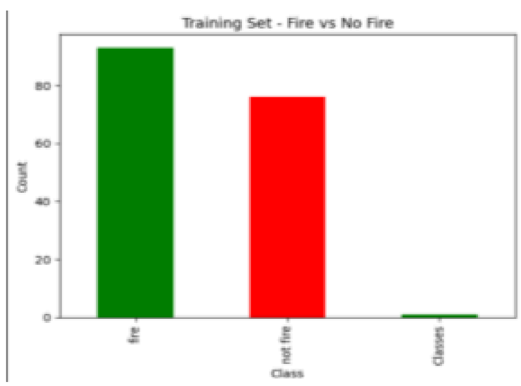


Fig2: Bar chart showing Training set

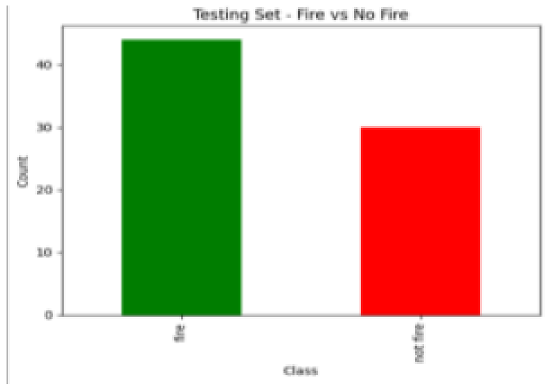


Fig3: Bar chart showing Testing set

Heatmap – Correlation Matrix

A heatmap is a matrix with colors that indicates the degree of correlation between each numerical feature and other features, such as the Fire label. Values of correlation vary from -1 to 1: A significant positive correlation is shown by a value near +1 (for example, higher temperatures are associated with a higher danger of fire). There is a substantial negative correlation when the value is around -1. There is no significant association if the value is close to 0. Temperature and the Fire Weather Index (FWI) have strong positive relationships with fire, which makes them crucial characteristics for comprehending or forecasting forest fires, as you may find out from the heatmap.

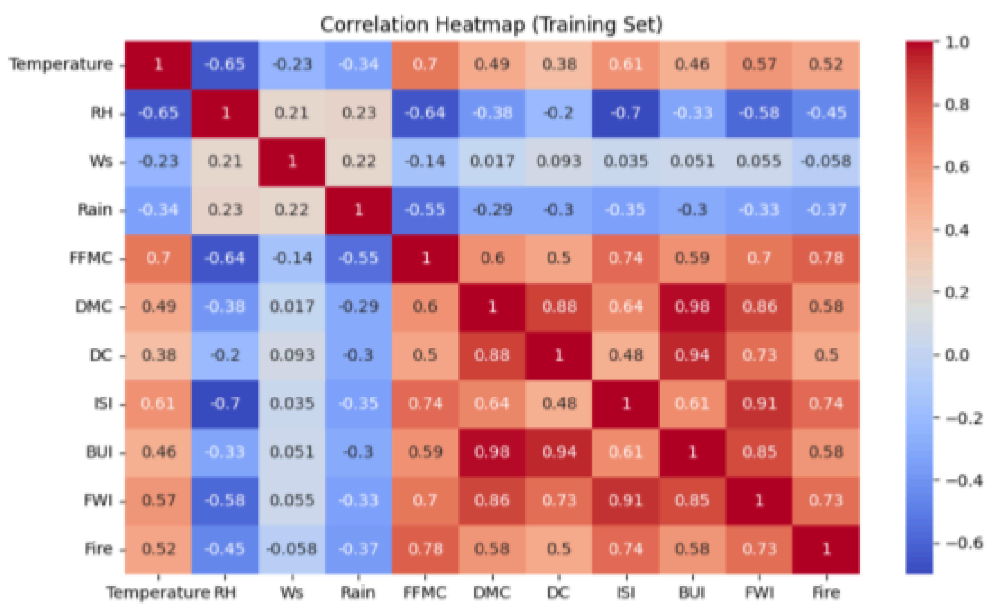


Fig4: Representation of Correlation Heatmaps

VI. Discussions

Wildfire damage has increased as a result of climate change, human encroachment, and environmental negligence. Early identification and prediction methods are critical to minimizing damage to woodlands, biodiversity, and human settlements. The focus of this study has been developing a wildfire prediction model based on Artificial Neural Networks (ANN) using the Algerian Forest Fires Database. It has also looked into how to enhance future forecasting capabilities by integrating satellite-based data from the MODIS 2017 India dataset.

Training an artificial neural network (ANN) with structured meteorological data to predict the chance and risk score of forest fires was the primary objective of the study. Because of its abundant and labeled data, which contained several fire hazard indices such as FFMC, DMC, DC, and ISI, as well as significant environmental factors like temperature, relative humidity (RH), wind speed (Ws), and rainfall, the Algerian dataset was selected. These features have been discovered to play a significant role in the occurrence and spread of forest fires.

As demonstrated in the study's conclusion, when trained on well-structured environmental datasets, Artificial Neural Networks are capable of effectively classifying and assessing the risk of forest fires. A potentially helpful tool for decision assistance is provided by the model to forest departments, emergency responders, and environmental agencies.

The potential integration with satellite datasets like MODIS further enhances the real-time applicability of the technologies and enables a scalable and automated wildfire detection system. Future studies should look at the use of ensemble techniques, real-time data streaming, and more thorough geographic analysis to improve the accuracy and adaptability of the model in a range of ecosystems and climate zones.

Pairplot – Feature Relationships by Class

For each pair of numerical variables, a pairplot generates a scatter plot with color labels to identify "fire" and "not fire" instances. This is particularly helpful for class separation and pattern recognition. Green (not fire) points steer clear of areas where red (fire) points are concentrated, indicating that those traits are highly suggestive of fire risk. The majority of fires, for instance, may occur at high temperatures and low humidity, according to a pairplot of temperature vs relative humidity.

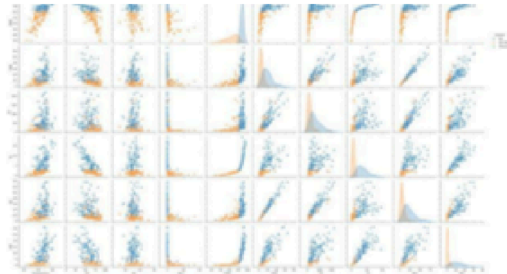


Fig5: Representation of Pair Plots

Scatter Plot – Temperature vs FWI

Two important features are highlighted in this scatter plot: temperature and FWI. Typically, fire instances (shown in red) occur when both numbers are higher. This lends credence to the theory that fire is more likely to occur when hot, dry weather and high FWI values are present.

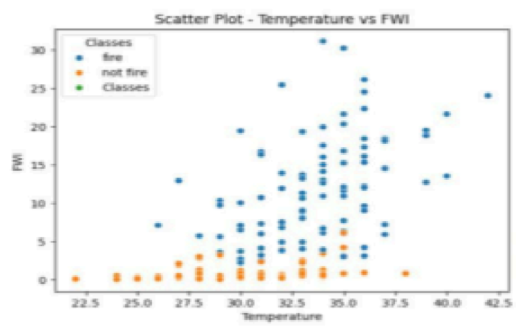


Fig6: Representation of Scatter Plots

VII. Conclusion

Due to environmental neglect, human encroachment, and climate change, wildfires have become more destructive. To reduce harm to forests, biodiversity, and human settlements, early detection and prediction tools are essential. Utilising the Algerian Forest Fires dataset, this project has concentrated on creating an Artificial Neural Network (ANN)-based wildfire prediction model. Additionally, it has examined the possibility of incorporating satellite-based data from the MODIS 2017 India dataset to improve future forecast capabilities.

The study concludes by showing that Artificial Neural Networks can efficiently categorise and evaluate the danger of forest fires when trained on well-structured environmental datasets. The model offers environmental agencies, emergency responders, and forest departments a potentially useful tool for decision support. A scalable and automated wildfire detection system is made possible by the potential integration with satellite datasets like MODIS, which further improves the systems' real-time applicability. Future research should investigate the application of ensemble approaches, real-time data streaming, and more in-depth geographic analysis to enhance the model's accuracy and flexibility across a variety of climate zones and ecosystems.

References

[1] Naderpour, M., Rizeei, H. M., & Ramezani, F. (2021). Forest fire risk prediction: A spatial deep neural network-based framework. *Remote Sensing*, *13*(13), 2513.

[2] Naderpour, Mohsen, Hossein Mojaddadi Rizeei, and Fahimeh Ramezani. "Forest fire risk prediction: A spatial deep neural network-based framework." *Remote Sensing* 13, no. 13 (2021): 2513.

[3] Alonso-Betanzos, A., Fontenla-Romero, O., Guijarro-Berdiñas, B., Hernández-Pereira, E., Andrade, M. I. P., Jiménez, E., ... & Carballas, T. (2003). An intelligent system for forest fire risk prediction and fire fighting management in Galicia. *Expert systems with applications*, *25*(4), 545-554.

[4] Salehi, Mahsa, and Lida Rashidi. "A Survey on Anomaly detection in Evolving Data: [with Application to Forest Fire Risk Prediction]." *ACM SIGKDD Explorations Newsletter* 20, no. 1 (2018): 13-23.

[5] Salehi, M., & Rashidi, L. (2018). A Survey on Anomaly detection in Evolving Data: [with Application to Forest Fire Risk Prediction]. *ACM SIGKDD Explorations Newsletter*, *20*(1), 13-23.

[6] Sakr, George E., Imad H. Elhajj, George Mitri, and Uchechukwu C. Wejinya. "Artificial intelligence for forest fire prediction." In *2010 IEEE/ASME international conference on advanced intelligent mechatronics*, pp. 1311-1316. IEEE, 2010.

[7] Sakr, G. E., Elhajj, I. H., Mitri, G., & Wejinya, U. C. (2010, July). Artificial intelligence for forest fire prediction. In *2010 IEEE/ASME international conference on advanced intelligent mechatronics* (pp. 1311-1316). IEEE.

[8] Zheng, Z., Gao, Y., Yang, Q., Zou, B., Xu, Y., Chen, Y., ... & Wang, Z. (2020). Predicting forest fire risk based on mining rules with ant-miner algorithm in cloud-rich areas. *Ecological Indicators*, *118*, 106772.

[9] Zheng, Z., Gao, Y., Yang, Q., Zou, B., Xu, Y., Chen, Y., ... & Wang, Z. (2020). Predicting forest fire risk based on mining rules with ant-miner algorithm in cloud-rich areas. *Ecological Indicators*, *118*, 106772.