iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 9 **Issue:** VI **Month of publication:** June 2021

DOI: https://doi.org/10.22214/ijraset.2021.35258

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VI Jun 2021- Available at www.ijraset.com

A Neural Network Model for Wildfire Scale Prediction using Meteorological Factors

Lavanya I[1], Meghana C V[2], Sanju S[3], Shreyas R[4], Gagana S[5]
[1, 2, 3, 4]B.E Student, [5]Assistant Professor, Dept. of Information Science and Engineering, VVCE, Mysuru

Abstract: Forest fires are natural hazards defined as movements of fire through unregulated and uncontrolled forested areas. They pose a permanent risk of loss of forest and forest land. The ability to reliably forecast the region that could be involved in a forest fire incident will help to optimize fire prevention efforts. It appears that Portugal may theoretically make better use of the wildfire risk assessment. More than any other region in Europe, it is a country overrun by wildfires. It has a large amount of forest. Forest fires have a long-term impact on the climate because they contribute to deforestation and global warming, which is one of the main causes of the phenomenon. This research employs Back Propagation Neural Network (BPNN) and Recurrent Neural Network (RNN) models with meteorological parameters as inputs to anticipate forest fires as a means of safeguarding forest biodiversity. The results indicate that using meteorological data, it is possible to anticipate the severity of a forest fire at the beginning.
Keywords: forest fire, BPNN, RNN, meteorological factors, Prediction

I. INTRODUCTION

In recent years, forest fires have been one of the most common disasters. Most of these forest fires occur by climate, human activities, and ecosystem processes. These are particularly susceptible in dry regions of the forest and are often seen during the summer season. Valuable forest resources, including carbon stored in biomass, are lost each year, in part owing to forest fires, disrupting the aspects of the supply chain from forests. Forest fires have devastating ecological implications for the people, communities, flora, and fauna directly affected.

Forest fires have a significant impact on our lives and surroundings. The capacity to properly estimate the area affected by forest fires will aid in the optimization of firefighting efforts. However, fire forecasting is difficult. Fire hazard is highly influenced by humidity and temperature over the previous several weeks. The longer it takes to dry and stay damp through the season, the thicker or deeper the material is. For low-pass fire hazards, this functions as a filter.

Several real-time forest fire detection and forecast technologies have been implemented, including video-based, sensor-based, and satellite-based systems. The standard procedure for dealing with wildfires is to record satellite photos, notify the authorities, and then take action to put out the fire. However, this procedure will only begin after the wildfires, and even before taking the first step, a few acres of forest area will have already been lost. The entire method should be thoroughly understood and finished ahead of time. However, the idea of putting this in place and then maintaining it is too expensive. Not only should the process be time-efficient, but it should also be cost-effective [2].

We present a solution to analyze forest fire in the Montesinho National Park, which is located in the north-eastern region of Tras-os-Montes, Portugal, and comprises 517 forest fire occurrences, demonstrating the applicability of the proposed approach for burnt area forecasting. To evaluate prediction models, a Back Propagation Neural Network (BPNN) and Recurrent Neural Network (RNN) will be used, with meteorological factors as input values. Our prediction model assists fire rescuers in taking appropriate actions depending on the projected size of a wildfire's destruction in its early phases. The findings will indicate that by utilizing meteorological data at the time of occurrence, it is possible to anticipate the size of a forest wildfire.

II. LITERATURE SURVEY

Arnida L. Latifah et al [1] aim to test the random forest model based on environmental variables and satellite data of the burned area in the prediction of forest fires. Forest fire forecasts were anticipated to reduce the impact of forest fires in the future. The random forest model with all selected climatic variables described the event of forest fires over Borneo based on analysis of spatial and annual variability.

Kshama Shalini et al [2] aim to educate the local fire authorities by exploring different real-time forest fire detection and prediction approaches. The geometry and characteristics of the images generated were considered by the algorithms used in fire detection.

Gao Demin et al [3] introduce a fuzzification method to determine the risk of fire in the field of study and establish a future fire risk scheme quantitatively. The technology of the wireless sensor network was used to continuously collect 24-hour weather data, providing a great opportunity to accurately represent the status of the forest climate. Depending on the method, they obtained days that were highly likely to pose a risk of forest fires and paid particular attention to forest guard forest fire prevention.

Mochammad Anshori et al [4] introduced a neural network training technique based on Extreme Learning Machines (ELM). Several experiments were conducted in this study to increase the effectiveness of the ELM approach, resulting in the best forecasts possible.

Osama Elsarrar et al [5] developed a Dynamic Cell Structure (DCS) neural network that was used to predict forest fire data and determine which environmental elements cause fires. An intuitive rule extraction approach was used to derive understandable rules for this prediction. By comparing the results to the raw data, the results were confirmed.

Tarun Dadlani et al [6] used machine learning to try to anticipate forest fires utilizing operational monitoring across a region and different sensors to detect changes in climate. The Wildfire Prediction System (WiPreSy) tracked and documented changes in meteorological conditions and forecasted the intensity of forest fires using real-time data, preventing significant forest fire losses.

Mikhail D. Molovtsev et al [7] aim to analyze ML algorithms and the possibility of their application to forest fire data.

Martina PETKOVIĆ et al [8] aim to discuss the preventive measures against forest fires as well as the forest monitoring approach as a way of conserving the biodiversity of forests. Using modern information technology, such as geographic information systems (GIS), they enhanced strategic and tactical fire prevention preparation and behavior. The paper demonstrates the use of a selected information system as forest fire protection in Serbia.

III. METHODOLOGY

A. Proposed Solution

The main aims of this work are to develop a system using advanced algorithms that make wildfire prediction and detection extremely simple, to suggest a new concept of the wildfire scale, defined by the combination of the length of a fire and the size of the region it burns, to study the level of neural network models (i.e., Backpropagation Neural Network (BPNN), and RNN models) to predict the wildfire scales, and to provide an accurate estimate of the wildfires in the future.

B. Data Collection and Data Preprocessing

The raw data comes from the Forest Fire data set in the UCI Machine Learning Repository, which was developed by Paulo Cortez and Anbal Morais from the Montesinho natural park in Portugal's northeast area, Tras-os-Montes [14]. The flora and fauna in this park are extremely diverse. In the supra-Mediterranean region, the average yearly temperature ranges from 8 to 12 degrees Celsius.

The data for this experiment was collected between January 2000 and December 2003, and it was collected using two sources. The initial database was obtained by the investigator in charge of the Montesinho fire. The time, date, geographical location within a 9x9 grid (x and y-axis), the kind of vegetation involved, the six components of the FWI system, and the total area burned if a forest fire occurred were all reported on a regular basis.

Attribute	Description
X	x-axis coordinate (from 1 to 9)
Y	y-axis coordinate (from 1 to 9)
month	The Month of the year (January to December)
day	Day of the week (Monday to Sunday)
FFMC	FFMC code
DMC	DMC code
DC	DC code
ISI	ISI index
temp	Outside temperature (in ◦C)
RH	Outside relative humidity (in %)
wind	Outside wind speed (in km/h)
rain	Outside rain (in mm/m2)
area	Total burned area (in ha)

Table 1: The preprocessed dataset attributes

Table 1 describes the attributes present in the dataset. The data from the data set includes 517 instances which have 12 attributes as inputs and 1 output.

C. Prediction Modeling

1) *Backpropagation Neural Network (BPNN):* It is a supervised learning algorithm, for training Multi-layer perceptrons (Artificial Neural Networks). The configuration of the network will be calculated by the number of input layer nodes, hidden layer, and output layer in the BPNN. It is an error backpropagation algorithm-trained multi-layer feed-forward network and is one of the most commonly used neural network models. The BP network can be used to learn and store a lot of input-output model mapping relationships, and the mathematical equation that defines these mapping relationships does not need to be revealed in advance. The learning rule is to follow the steepest descent technique in which the backpropagation is used to control the network's weight value and threshold value to obtain the minimum square error number. [11].
2) *Recurrent neural network (RNN):* It is a type of artificial neural network that utilizes data from sequential data or time series. Recurrent neural networks use training data to learn, including feed-forward and convolutional neural networks (CNNs). As they take information from previous inputs to affect the current input and output, they are differentiated by their' memory'. While traditional deep neural networks assume that inputs and outputs are independent, the performance of recurrent neural networks depends on the previous elements of the series [12].

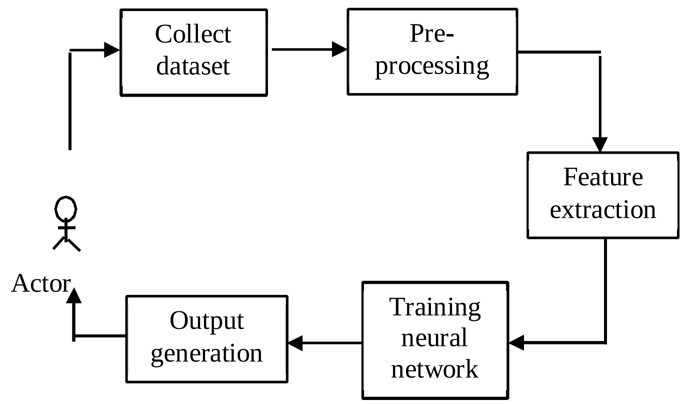


Figure 1: System Architecture

Figure 1 represents the system architecture. The required database is first collected. The datasets are then preprocessed and training samples are generated followed by the feature selection. The model is then properly trained using BPNN and RNN algorithms, followed by the generation of outputs.

IV. COMPARISON

Several analyses have been carried out for forest fire prediction using data mining algorithms such as Logistic regression, Gradient boosting, SVM, and Random forest both with and without principal component analysis (PCA) [9].

In two different ways, the dataset was modeled, one where PCA was added and the other without PCA. After applying PCA to the dataset, it was split into training and test data, resulting in a table with five essential components, with test data accounting for 20% of the dataset.The ML methods for modeling were then applied and the findings are tabulated.

A. Logistic Regression

Logistic regression is a model that predicts the probability of occurrence of a certain event by adjusting the data to the logistic curve [7]. It is a type of regression model where the dependent variable is categorical.

B. Gradient Boosting

Gradient boosting is a technique for training multiple models in an additive, sequential, and measurable way. This algorithm determines the disadvantages of weak learners by using gradients in the loss function. The loss function is a statistic that shows how effective the model's coefficients are at fitting data. In the gradient boost function [10], users have the option of optimising the cost function instead of the loss function.

C. Support Vector Machines (SVM)

For both regression and classification assignments, SVM can be used. But, in classification targets, it is commonly used. The support vector machine algorithm aims to find a hyperplane that separately classifies the data points in N-dimensional space (N-the number of features).

D. Random Forest

A random forest is a compositional algorithm that is based on the concept of ensemble learning. The prediction is made by aggregating the ensemble's predictions. In general, Random Forest demonstrates a major increase in efficiency over the single tree classifier.

Algorithm	Accuracy(%)
Logistic Regression	68.26
Gradient Boosting	54.87
SVM	42.26
Random forest	51.14

Table 2: Accuracies of algorithms after performing PCA

Table 2 shows the prediction accuracies of algorithms after performing PCA. For Logistic Regression, it gave an accuracy value of 68.26. It was then modeled with gradient boosting, giving a value of 54.87. An accuracy of 51.14 was given by the Random Forest model. Among the models that were implemented with PCA, the Support Vector Machine gave the least value. The value was 42.66 for the SVM model.

Algorithm	Accuracy(%)
Logistic Regression	66.69
Gradient Boosting	68.38
SVM	66.33
Random Forest	50.13

Table 3: Accuracies of algorithms without performing PCA

The second scenario was to model the dataset without using Principal Component Analysis. The dataset was partitioned into test and train data, with test data accounting for 20% of the initial dataset. After splitting, the test and train data were both scaled within the range of 0-1 using Min-Max scaling. After the dividing was completed and the findings were gathered, the models were created.

The accuracy for Logistic Regression was 66.69. The accuracy for Gradient Boosting was 68.38, the highest value in the second case among the models. An accuracy of 50.13 was provided by the Random Forest model. A value of 66.33 was given by the Support Vector Machine.

In the first scenario, logistic regression gave the highest value of 68.26, while SVM gave the lowest value of 42.66. In the second scenario, the Gradient Boosting classifier provided the highest value of 68.38, which proved to be the greatest accuracy of all the models considered (both Case-1 and Case-2). The Random Forest classifier was assigned the lowest value in the study, with a score of 50.13 in the second scenario. Machine Learning Techniques like BPNN and RNN can improve prediction accuracy even more.

V. CONCLUSION

In this paper, a neural-networks-based approach to the problem of predicting forest fires is developed. The UCI machine learning repository's forest fire dataset is used to forecast the occurrence of a wildfire. The size of the data set is 517 instances, and some attributes are 13. Research in this paper concludes that BPNN and RNN give the best accuracy and the model will have the potential to predict the scales of a greater range of wildfire occurrences.

The results of this study indicate that it is feasible to use meteorological data to forecast the extent of forest wildfires, which will be useful for forest fire prevention and rescue, especially for wildfires occurring in forests. Fire rescuers and firefighters will be able to make efficient and sufficient steps to minimize the damages incurred by forest wildfires, in line with the severity of the fire as expected at its initial stages.

REFERENCES

[1] Arnida L. Latifah, Ayu Shabrina, Intan N. Wahyuni, and Rifki Sadikin, "Evaluation of Random Forest model for forest fire prediction based on climatology over Borneo", IEEE, 2019.

[2] Jerry Gao, Kshama Shalini, Navit Gaur, Xuan Guan Sean Chen, Jesse Hong, and Medhat Mahmoud, "Data-driven Forest Fire analysis", IEEE, 2017.

[3] Gao Demin, Lin Haifeng, Jiang Anna, and Wu Guoxin, "A FOREST FIRE PREDICTION SYSTEM BASED ON RECHARGEABLE WIRELESS SENSOR NETWORKS", IEEE, 2014.

[4] Mochammad Anshori, Farhanna Mar'I, Mukhammad Wildan Alauddin and Wayan Firdaus Mahmudy, "Prediction of Forest Fire using Neural Network based on Extreme Learning Machines (ELM)", IEEE, 2019.

[5] Osama Elsarrar, Marjorie Darrah and Richard Devine, "Analysis of Forest Fire Data using Neural Network Rule Extraction with Human Understandable Rules", IEEE, 2019.

[6] Mrs. Vidya Zope, Tarun Dadlani, Ashutosh Matai, Pranjali Tembhurnikar, and Richa Kalani, "IoT Sensor and Deep Neural Network-based Wildfire Prediction System", IEEE, 2020.

[7] Mikhail D. Molovtsev and Irina S. Sineva, "Classification Algorithms Analysis in the Forest Fire Detection Problem", IEEE, 2019.

[8] Martina PETKOVIĆ, Ivan GARVANOV, Dragan KNEŽEVIC, and Slavoljub ALEKSIĆ, "Optimization of Geographic Information Systems for Forest Fire Risk Assessment", IEEE, 2020.

[9] R. Rishickesh, A. Shahina and A. Nayeemulla Khan, "Predicting Forest Fires using Supervised and Ensemble Machine Learning Algorithms", International Journal of Recent Technology and Engineering (IJRTE), Volume-8 Issue-2, July 2019.

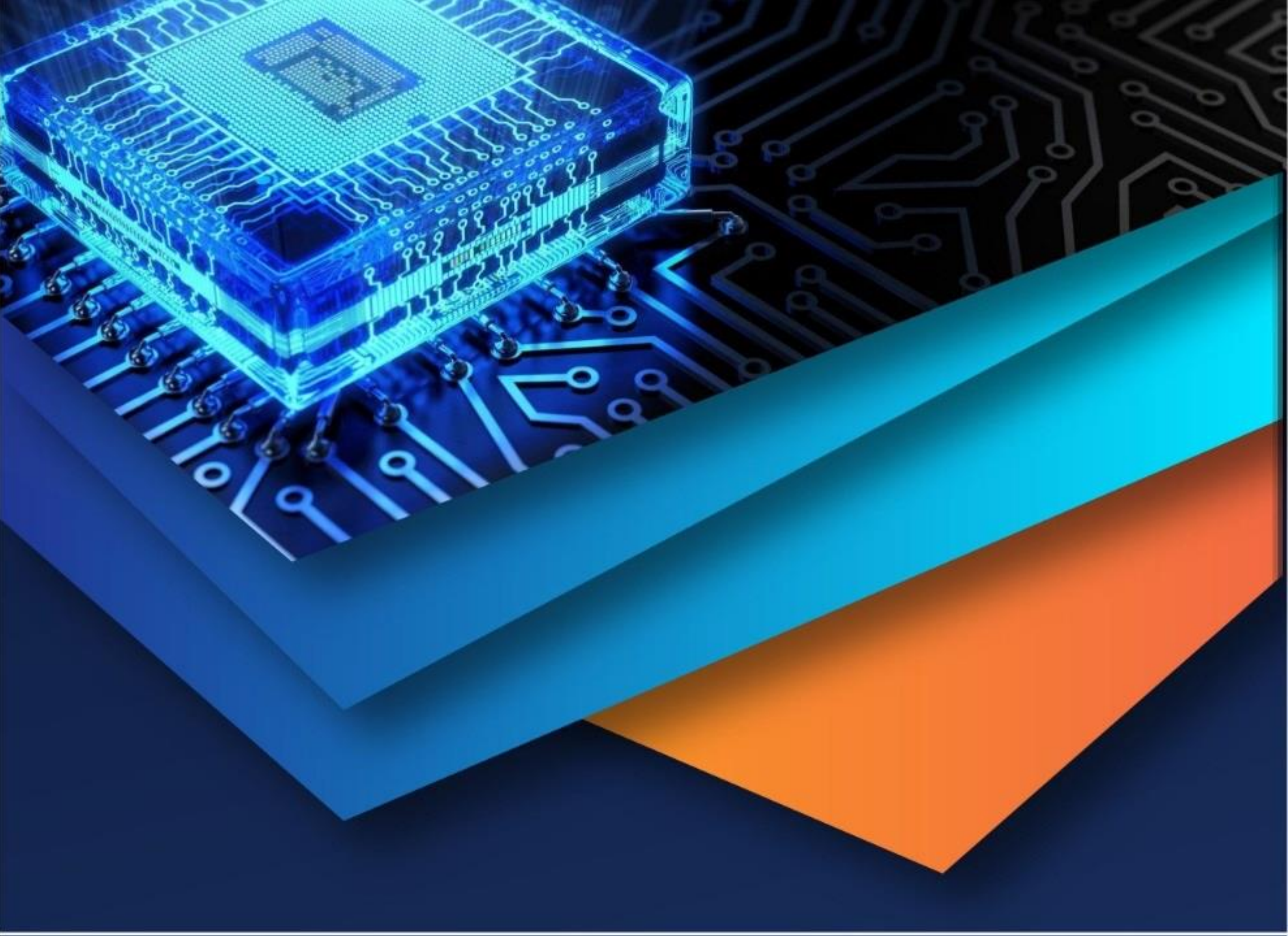
[10] J. H. Friedman, "Greedy function approximation: A gradient boosting machine", Annals of Statistics, 2000.

[11] Jing Li, Ji-hang Cheng, Jing-yuan Shi, and Fei Huang, "Brief Introduction of Back Propagation (BP) Neural Network Algorithm and Its Improvement", Advances in CSIE, Vol. 2, AISC 169, pp. 553–558.

[12] https://www.ibm.com/cloud/learn/recurrent-neural-networks

[13] J. H. Friedman, "Greedy function approximation: A gradient boosting machine", Annals of Statistics, 2000.

[14] https://archive.ics.uci.edu/ml/datasets/forest+fires



10.22214/IJRASET

INDEX COPERNICUS
45.98

IMPACT FACTOR:
7.129

TOGETHER WE REACH THE GOAL
IMPACT FACTOR:
7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY