International Journal of Engineering Sciences & Emerging Technologies, Oct. 2023.
ISSN: 22316604 Volume 11, Issue 2, pp: 431-437 ©IJESET

DESIGNING AND IMPLEMENTATION OF FOREST FIRE PREDICTION MODEL USING DEEP LEARNING

Shobit Agrawal[1], Aditya Pratap Singh[2],
Abhay Pratap Singh Bhadauria[3*] Amit Kumar[4]

[1]Department of Computer Science Engineering, GL Bajaj Group of Institutions, Mathura
shobit.agrawal@glbajajgroup.org

[2]Department of Computer Science Application, GLA University, Mathura
aditya.singh2_cs19@gla.ac.in

[3*]Department of Computer Science, Gurukula Kangri (Deemed to be) University, Haridwar
rs.abhaypratapsingh@gkv.ac.in

[4]Department of Computer science Engineering, Quantum University, Roorkee

amit.cse@quantumeducation.in

ABSTRACT

Every year, forest fires ravage a huge region of forest cover, having a devastating impact on the surrounding flora and fauna. Each year, forest fires play a significant part in the extinction of hundreds of species of wildlife. Machine learningcan be used to accurately predict forest fires and save species since it helps us predict the future. The FFMC, DMC, DC, ISI, Temp, RH, Wind, Area, Rain are the main determinants of any fire. This study seeks to forecast the likelihood of a forest fire occurring given the levels of factors at a certain location. In this suggested effort investigates the use of machine learning techniques to forecast forest fires and to find out which algorithms are best for prediction. The dataset has 518 records with 13attributes. Therefore, various techniques of machine learning like Logistic Regression, Linear Regression, Artificial Neural Network (ANN), and Support Vector Machine (SVM) are investigated to make a comparative analysis of these with respect to their accuracies. ANN's model outperforms other algorithms and achieved the highest accuracy. Additionally, a single page website that would use user input to forecast the likelihood of forest fires in real time is displayed.

KEYWORDS: *Forest Fire, Linear Regression, SVM, Random Forest, Deep Learning.*

1. INTRODUCTION

Forest and wildlife fires are uncontrolled blazes in flammable vegetation. Based on the size of the fire, it may be classed as a bush fire, forest fire, etc. A forest fire is a natural or human- caused fire that occurs in a forest or wooded area, often characterized by flames and smoke. It can start from lightning strikes, high temperatures, dry weather conditions, or human activities like campfires, cigarettes, or intentional arson. These fires can spread rapidly, causing damage to the environment, wildlife, and human life. It can also have long-lasting effects on the ecosystem, changing the soil composition, and disrupting natural processes. Effective measures such as prevention, early detection, and efficient firefighting techniques can help in minimizing the damage caused by forest fires.

Forest fires are highly dangerous and harmful to the environment, wildlife, and human life. It can spread rapidly, destroying large areas of forests and vegetation, leading to the loss of habitats and biodiversity. Forest fires also release harmful smoke and gases, affecting air quality and causing

respiratory problems for humans and animals alike. The heat generated by forest fires can also cause soil erosion, affecting the soil's ability to support new vegetation growth. Inaddition, the loss of trees and vegetation can increase the risk of landslides, mudflows, and flooding.

It is imperative that a strategy should be devise to combat them,as It can pose a significant threat to animals. Now theprimary problem is to detect or predict a wildfire before itactually occurs, as it becomes extremely difficult to extinguish forest fires once this have caused extensive irrevocable damage.

Machine learning is the process of learning from data to predict the future [3, 4]. Consequently, then model some criteria essential for any forest fire to occur in order to estimate the likelihood of a forest fire occurring.

This study's next section delves into the literature review, providing an overview of several methodologies, including their limitations and advantages whencompared to previous studies. The subsequent section presents the proposed approach. The fourth section presents the results and provides a critical analysis. The fifth and final section of this paper highlights potential areas of future research and concludes the study. Lastly, all references are listed.

2. RELATED WORK

In this research, a variety of approaches have been investigated, including but not limited to:

D. Rosadi et al., [1] considered the application of numerous machine learning approaches for the prediction of forest fire occurrence in the peatland areas. authors have considered some classical classification methods, such as SVM, K-NN, Logistic Regression, Decision Tree and Naïve Bayes. To compare the algorithms, authors have also considered the AdaBoost (DT Based) approach. Researchers have also considered topographical and meteorological data, and data contains 202 cases. authors have preprocessed the data and then classify it by splitting data into training and testing data into various ratios. Then applied machine learning approaches to both training and testing data.

T. Preeti et al., [2] presented the system processes employing climatic variables like temperature, rain, wind, and humidity. Authors have suggested Regression in random forest and Hyperparameter tuning by Randomized-Search-CV and suggested different subsamples of dataset that arrange different decisions tree and employs averages to reduce over-fitting and increase the forecast accuracy. Then compared the different algorithms like, Random Forest, Decision Tree, Artificial Neural Network, Support Vector Machine. After all the calculations, authors concluded that high number of fires in forests compared to other surface areas. Mean Absolute Error (MAE) is also calculated for all of these algorithms.

M. Anshori et al., [5] Temperature, wind, humidity, and rainfall were utilized in their research to forecast which geographic regions may be impacted by forest fires. Researchers used neural network with Extreme Learning Machines (ELM) training model. Based on the research carried out, the ELM method was able to provide predictions of forest fire datasets taken from National Park in Portugal with several 517 forest fire incidents. The ELM method can produce an RMSE value of 63.09 with computation time of 25.9 ms. The number of hidden neurons is 20. The best accuracy value is obtained from error calculations using RMSE.

B. Arteaga et al., [6] examined the effectiveness of several Convolutional Neural Network (CNN) models that were already trained to classify photographs of forest fires and may be used with economic development cards like the Raspberry Pi. As a result of their study, With the use of transfer learning and data augmentation, CNN are highly helpful for resolving picture categorization issues. Authors used a database of 1800 photos of woods and forest fires for their research, and got great results.

3. PROPOSED WORK

This section provides a brief explanation of the proposed work. The dataset used is initially introduced, followed by a step-by-step overview of the research methodology. All the visual representation featured in this study were generated using the Python libraries on PyCharm. For the proposed study, the dataset was sourced from Kaggle [7] in CSV format, comprising of 518 records with 13 numerical or predictive attributes. The class attributes are defined as follows:

The detection of a wildfire is primarily dependent upon several factors -

- FFMC: It is a statistical assessment of the moisture content of cured litter and other fine fuels.
- DMC: It is a quantitative assessment of the average moisture content of moderately dense, loosely compacted organic layers.
- DC: It is a numerical measurement of the average moisture content of dense, deep organic layers.
- Temp: It is a physical measure that quantifies the sensations of warmth and coldness.
- Relative Humidity: It expresses the relationship between the relative humidity of air at a certain temperature and the relative humidity of saturated air at the same temperature.
- Wind: Wind is air in motion and is caused by variations in atmospheric air pressure. High-pressure air goes toward low-pressure regions.
- Area: The region bound by the shape of an object is the area. In a plane, the area of the figure or any two- dimensional geometric shape is the area of the figure.
- Rain: Rain is precipitation consisting of liquid water falling from the sky. When clouds get saturated with or filled with water droplets, precipitation occurs.

Table 1: DATASET SAMPLE VIEW

1	month	day	FFMC	DMC	DC	ISI	temp	RH	wind	rain	area
2	mar	fri	86.2	26.2	94.3	5.1	8.2	51	6.7	0	0
3	oct	tue	90.6	35.4	669.1	6.7	18	33	0.9	0	0
4	oct	sat	90.6	43.7	686.9	6.7	14.6	33	1.3	0	0
5	mar	fri	91.7	33.3	77.5	9	8.3	97	4	0.2	0
6	mar	sun	89.3	51.3	102.2	9.6	11.4	99	1.8	0	0
7	aug	sun	92.3	85.3	488	14.7	22.2	29	5.4	0	0
8	aug	mon	92.3	88.9	495.6	8.5	24.1	27	3.1	0	0
9	aug	mon	91.5	145.4	608.2	10.7	8	86	2.2	0	0
10	sep	tue	91	129.5	692.6	7	13.1	63	5.4	0	0
11	sep	sat	92.5	88	698.6	7.1	22.8	40	4	0	0
12	sep	sat	92.5	88	698.6	7.1	17.8	51	7.2	0	0
13	sep	sat	92.8	73.2	713	22.6	19.3	38	4	0	0
14	aug	fri	63.5	70.8	665.3	0.8	17	72	6.7	0	0
15	sep	mon	90.9	126.5	686.5	7	21.3	42	2.2	0	0
16	sep	wed	92.9	133.3	699.6	9.2	26.4	21	4.5	0	0
17	sep	fri	93.3	141.2	713.9	13.9	22.9	44	5.4	0	0

Table I is the collection of some data samples were used in this study given below. The suggested work's flow is shown in Figure 1, and a graphical representation of how it was executed is shown in the figure given below.

Following the dataset collection, machine learning models including Linear Regression, Logistic Regression, Support Vector Machine, and Random Forest are employed. These models are implemented to both training and testing- set. The suggested work's flow is shown in Figure 1, and a graphical representation of how it was executed is shown in the figure given below.

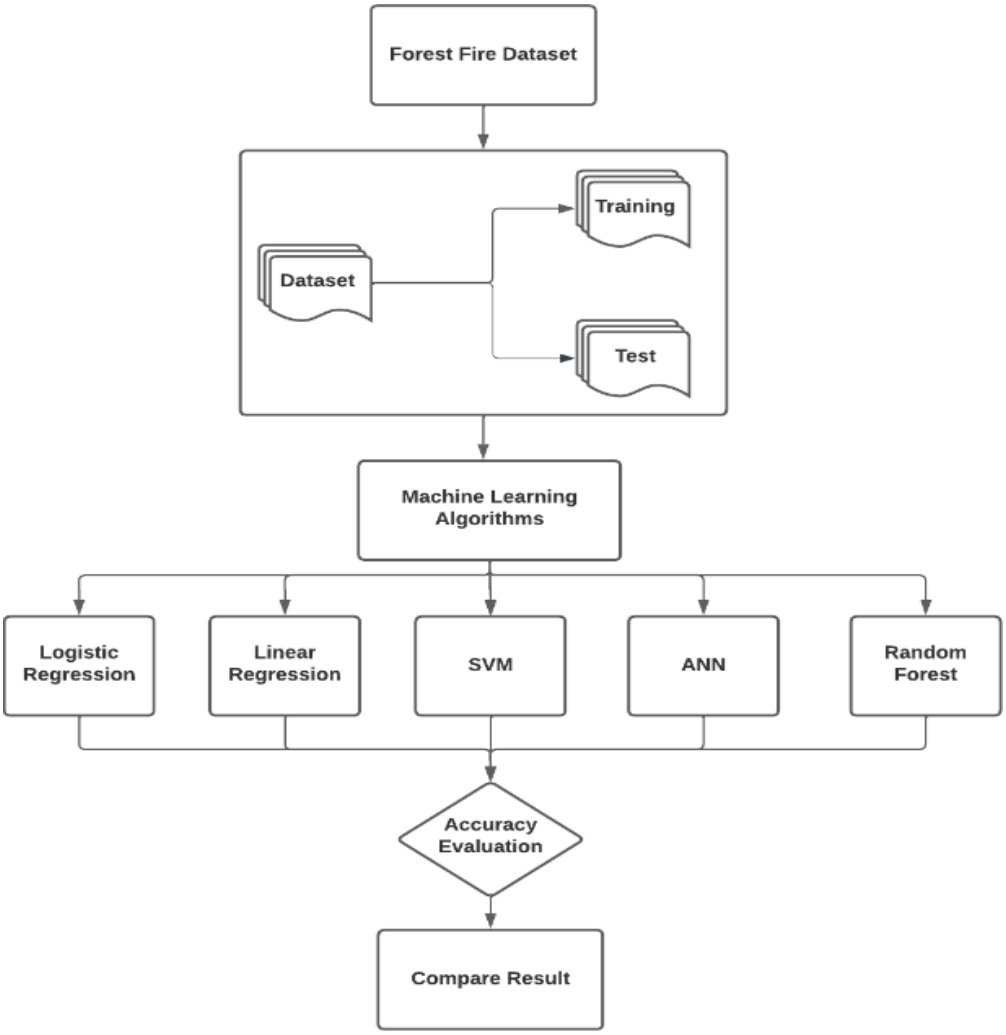

Figure 1: Proposed Architecture of Forest Fire Prediction

This work's execution is described in the following steps:

Step 1: The dataset has been obtained from Kaggle [7] in this first stage. after the collection of datasets, it has been preprocessed using various preprocessing techniques and then step 2nd taken into consideration.

Step 2: following the previous step, the dataset is divided in two parts, such as the train set and the test set, in various combinations, but the ratio of 80 to 20 yields the best results in terms of the confusion matrix.

Step 3: In this step, machine learning algorithms like Linear Regression, Logistic Regression, Random Forest, andANN are applied to the training dataset after the dataset has been classified. The proposed method's prototype states that each algorithm is independently applied to thedataset before the model is trained. The following stage is then.

Step 4: During this step, the accuracy of each algorithm is assessed by feeding the testing dataset to the

model and predicting values.

Step 5: Following the outcome of model, a comparison is made with the actual values, and a confusion matrix is created as depicted in Figure 2. Based on the accuracy, graphs are generated for each algorithm, and the results are then summarized.

In order to estimate the likelihood of a forest fire, multiple algorithms were applied. In figure 3, the

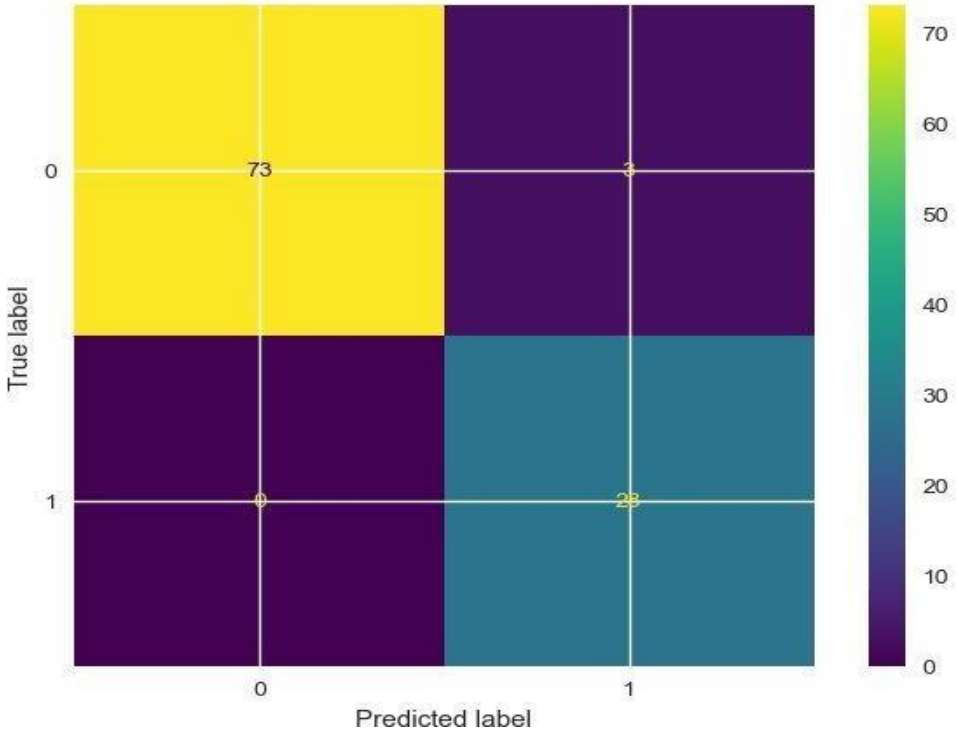

Figure 2: Heatmap of Confusion Matrix

calculated accuracy is displayed. Then applied ANN's Sequential model to train and test the data. Sequential model is a type of neural network architecture commonly used for predicting sequential data, such as time-series data. It is made up of a series of linearly stacked layers, with the output of one layer acting as the input for the subsequent layer. The Sequential model has 2 hidden layers, 3 dense layers and activation functions – 'ReLU' and 'sigmoid'. With this Sequential model, accuracy of 99.80% in training and 95.20% in testing and validation has achieved. The heatmap of the confusion matrix is shown in Fig. 2. There are four instances of the confusion matrix, i.e., ((00=true positive=73), (01=false negative=3), (10=false positive=0), (11=true negative=28)). The noticed class and anticipated class are drawn in a matrix with an overall suitable result.

4. RESULT AND DISCUSSION

Outcome demonstrates that machine learning techniques can be used to predict forest fires in an efficient manner. The accuracy of these five models and algorithms is calculated. And it is discovered that ANN's Sequential model beats the other four techniques [8, 9]., having the highest accuracy (95.20%), compared to Logistic Regression's (85.80%), Linear Regression's (83.85%), Random Forest's (83.41%), and SVM's (86.93%). After ANN's Sequential model, SVM algorithm has the highest accuracy with 86.93%. Agrawal et.al [10] suggested the comparison of multiple algorithms and concluded the use of deep learning a part of machine learning is the best suitable algorithm for classification of data. In [11] suggested the comparison of different model that also concluded deep learning is the best approach for classification. Fig. 3, which compares the accuracy of the suggested work, ANN algorithm performs better than the other described algorithms.

Figure 3: Comparison of used algorithms on the basis of accuracy

Table 2 also reflects the statistical accuracy of algorithms which attained in this research work.

Table 2: Experimental Accuracy Analysis

Algorithms	Accuracy
Logistic Regression	85.80%
Linear Regression	83.85%
Random Forest	83.41%
Support Vector Machine	86.93%
ANN Model	95.20%

5. CONCLUSION

In this section of the study, the optimal algorithm or model for predicting forest fires is reviewed along with an overview of forest fires, their negative impacts on both people and the environment, and why it is crucial to forecast them. Then identified the dataset sources used by machine learning models and algorithms. The workflow of the research is also outlined. Then the dataset sample, architecture, heatmap and accuracy graph are also discussed. Then compared the accuracy of used algorithms i.e., Logistic Regression, Linear Regression, Random Forest, Support Vector Machine, and ANN. As a result, it can be said that the ANN's Sequential model performs better in terms of accuracy than any other listed techniques. ANN holds 95.20% accuracy. In the end, it was found that deep learning has a high prediction accuracy and that it can be enhanced by doing better work.

REFERENCES

[I] D. Rosadi, W. Andriyani, D. Arisanty, “Prediction of Forest Fire Occurrence in Peatlands using Machine Learning Approaches”, στο 2020 3rd International Seminar on Research of Information Technology and Intelligent Systems (ISTITI), 2020.

[II] T. Preeti, S. Kanakaraddi, A. Beelagi, “Forest Fire Prediction using Machine Learning Techniques”, στο 2021 International Conference on Intelligent Technology (CONIT), 2021.

[III] A. P. Singh, and M. Singh. "A comparative review of malware analysis and detection in HTTPs traffic." International Journal of Computing and Digital Systems 10.1 (2021): 111-123.

[IV] A. Singh "Classification of Malware in HTTPs Traffic Using Machine Learning Approach." El-Cezeri 9.2 (2022): 644-655.

[V] M. Anshori, F. Mar’i, M. W. Alauddin, “Prediction of Forest Fire using Neural Network based on Extreme Learning Machines (ELM)”, στο 2019 International Conference on Sustainable Information

Engineering and Technology (SIET), 2019.

[VI] B. Arteaga, M. Diaz, M. Jojoa, "Deep Learning Applied to Forest FireDetection", στο 2020 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT), 2020.

[VII] https://www.kaggle.com/code/sahilbisht/forest-fire-analysis.

[VIII] R. Fan, M.Pei, "Lightweight Forest Fire Detection Based on Deep Learning", στο 2021 IEEE 31st International Workshop on Machine Learning for Signal Processing (MLSP), 2021.

[IX] R. Ghali, M. A.Akhloufi, M. Jmal, "Forest Fires Segmentation using Deep Convolutional Neural Networks", στο 2021 International Conference on Systems, Man, and Cybernetics (SMC), 2020.

[X] S. Agrawal and D. K. Sharma, "Handling High-Dimensional Data and Classification Using a Hybrid Feature Selection Approach," in 2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), 2022, pp. 2168–2172.

[XI] A. P. Singh and S. Agrawal, "Accuracy Prediction on Detection of Breast Cancer Using Machine Learning Classifiers," in 2022 14th International Conference on Computational Intelligence and Communication Networks (CICN), 2022, pp. 401–405.