Predicting Wildfires Using Neural Network

Trupti Bhuta
Information Technology
St. Francis Institute of Technology
Mumbai, India

Srishti Shirgavi
Information Technology
St. Francis Institute of Technology
Mumbai, India

Amrita Mathur
Information Technology
St. Francis Institute of Technology
Mumbai, India

Manali Vichare
Information Technology
St. Francis Institute of Technology
Mumbai, India

Abstract—**Wildfire could be a fire in a part of flammable vegetation occurring in rural areas. Forest fires are chiefly caused because of lighting and sun, or by ignition of dry fuel like wood or dry leaves. Destruction caused because of wildfires are various and wide-ranging. They will have vital impacts on the economy, surroundings, heritage and social cloth of rural areas. For this purpose, wildfire prediction before an unimaginable destruction is caused has become progressively vital. Forest Fire Prediction, employing Artificial Neural Networks to forge a benevolent system that could to an astonishing extent, serve a noble cause in an exemplary fashion.Data used for prediction will include the vegetation and environmental parameters. The data has been taken from UCI repository. The model will learn using this data and predict the possibility of a probable fire if similar environmental conditions occur in future.**

Keywords*— *Forest Fire Prediction, Ecosystem, Artificial neural network

I. INTRODUCTION

Each year, wildfires damage billions of dollars' wellworth of infrastructure and claim the lives of humans caught in their path. Studies indicate climate change will cause more severe drought in vulnerable areas, leading to more dry and dead vegetation, increasing the duration and frequency of wildfires. According to the data, the recent Australia's fires (year 2019) are 46% bigger than last year's Brazilian Amazon blazes. Australian bushfires have completely wiped out anestimated 18.6 million hectares .Extreme temperatures and drought conditions contributed to the fires' extraordinary scale and intensity. The carbon dioxide the blazes send into the atmosphere contributes to climate change, raising the risk of larger fires in the future.Casting light on the recent surges in the number of incidences involving catastrophic forest fires, it is mandatory to have a system capable of portending wildfires that might occur, which would help in getting acquainted with the required resources on time to prevent it. Artificial Neural Networks makes it possible to develop a system that could predict Forest fires with the help of available meteorological and vegetation data. Artificial Neural Networks, a computational algorithm for machine learning is being used to fore-see forest fires and itsrate of spread. It considers different parameters which includes humidity, moisture, temperature, ignition points, drought factor, fuel index, etc.

II. LITERATURE REVIEW

1. Integrating wildfires propagation prediction into early warning of electrical transmission line outage:

The paper uses a cell geography cellular automata model to predict the rate of spread of fires and the threat it proposes to the electrical transmission line. This is done by taking into account the starting point of the fire and an assumed buffer point along with various other parameters [1].

2. Fire-risk evaluation of Austrian pine stands in Hungary − effects of drought conditions and slope aspect on fire spread and fire behaviour:

This paper focuses on effects of drought conditions and terrain inclination on fire behaviour.These two parameters were changed within specific intervals while other input parameters like Air temperature, relative humidity, wind speed, amount of fuel,etc.were kept constant [2].

3. CFFDRS: An R Package for the Canadian forest-fire danger rating system:

Here,fire weather index system and the fire behavior prediction system, have been studied to predict wildfire. It focuses on prediction of forest fire using a Canadian Forest Fire Danger Rating System (CFFDRS) system which provides a broad set of R functions for the calculation of the various CFFDRS subsystems [4].

4. Data-driven Forest Fire analysis:

In this paper, the author discusses thevarious real time forest fire detection as well as prediction approaches. The algorithms used for fire detection takes into consideration not only the geometry, but also the characteristics of the images generated.Three kinds of approach were seen i.e. Video based, sensor based, and satellite based.Satellite based approach is considered to be abetterapproach because the quality of video camera and sensor may make lot of unnecessary differences [5].

5. A Data Mining Approach to predict forest fires using meteorological data:

The paper proposes a Data Mining (DM) approach that uses meteorological data, to predict the area burnt due to forest fires. Many different DM techniques were used from which SVM gave the best results with the help of four inputs temperature, relative humidity, rain and wind [7].

III. EXPLORATORY DATA ANALYSIS

The dataset is of the Northeastern region of Portugal. The following Table I. mentions all the attributes that are present in this dataset.

TABLE I. DATASET DESCRIPTION

Feature	Explanation
X	Spatial coordinate (x-axis) within the Montesinho park
Y	Spatial coordinate (y-axis) within the Montesinho park
Month	'January' to 'December'
Day	'Monday' to 'Sunday'
FFMC	Fine Fuel Moisture Code – It contains the moisture content of litter and other cured fine fuels numeric rating. It indicates the relative ease of ignition & flammability of fine fuel.
DMC	Duff Moisture Code- consists of the numeric rating of the average moisture content of loosely compacted organic layers of moderate depth. DMC indicates of fuel consumption in moderate duff layers and medium-size woody material.
DC	Drought Code- Contains the numeric rating of the average moisture content of deep, compact organic layers. Seasonal drought effects on fuels in the forest and the amount of smoldering in deep layers and large logs is indicated by DC.
ISI	Initial Spread Index- It is the numeric rating of rate of fire spread. ISI combines the effects of wind and the FFMC on rate of spread without any influence of variable quantities of fuel
Temp	Temperature measured in degree Celsius
RH	Relative Humidity in percentage
Wind	Wind speed in km/hr
Rain	Outside rain in mm/m2

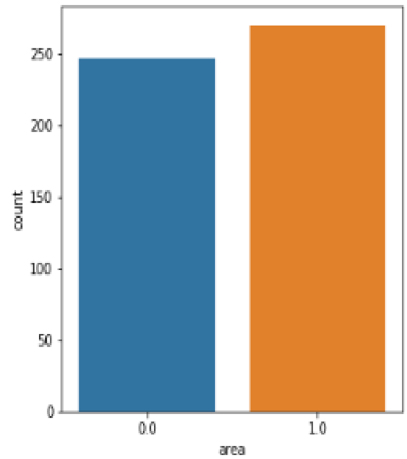

Fig.1. Occurrence of forest fire shown using Bar graph

IV. METHODOLOGY

Explanatory Analysis:

Different algorithms tested:

1) **Logistic Regression**

This is a machine learning algorithm based on the concept of probability which is used to classify problems when the dependent variable is binary. Initially this algorithm was used to first train and then test the dataset, which resulted in a very low testing accuracy as depicted in Table II. The sigmoid function is as follows-

$$\varphi(z) = \frac{1}{1+e^{-z}} \qquad (1)$$

2) **Decision Tree**

It is a type of supervised machine learning technique that has a tree-like structure of decisions that consists of only conditional control statements. According to certain parameters, the data is split continuously based on the decision nodes, to finally produce leaf nodes as the output. Fig. 2. is a pictorial representation of the same.

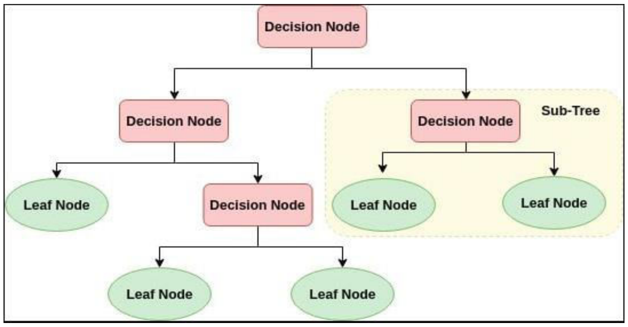

Fig.2. Decision Tree algorithm

3) **Support Vector Machine (SVM)**

This is a supervised learning algorithm which helps to categorize data into two parts. Since it is supervised learning, it needs to be trained with a dataset which already consists of two categories. Here, the data is plotted into a N-Dimensional space and there is a hyperplane (a distinguishing line) which separates these points into two classes which are on either side of this hyperplane. This illustration is shown in Fig. 3.

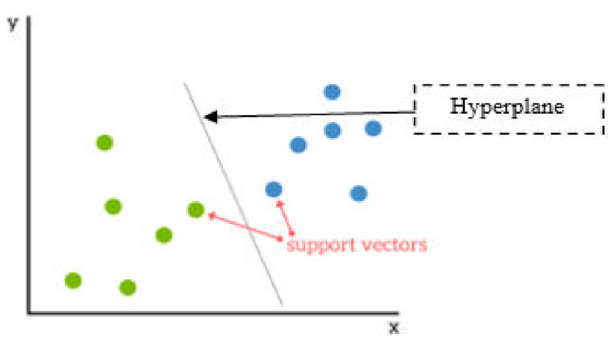

Fig.3. Support Vector Machine

The results of the analysis carried using the above three algorithms are depicted in Table II.

TABLE II. RESULTS WITH RESPECT TO THE ABOVE-MENTIONED ALGORITHMS

Algorithm	Results
Logistic regression	0.52
Decision Tree	0.49
Support Vector Machine	0.57

As seen, the accuracies achieved from the algorithms is quite low. Hence the following is a proposed system that uses neural networks.

ArtificialNeural Network:

The type of ANN model used is a feed-forward network. In this network, neurons are grouped into two layers which are as follows-

Layer 1: First layer

It consists of 256 neurons, with weights uniformly initialized, to which the input data is passed. The activation function used here is the Relu activation function.

Layer 2: Output layer

It consists of a single neuron which gives the output of whether a fire was caused or not.The activation function used here is the sigmoid activation function.

V. IMPLEMENTATION

Fig. 4 shows the implementation of the proposed model. Python programming is the language chosen for its implementation.

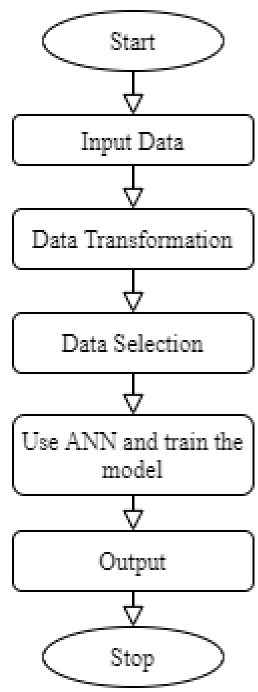

Fig.4. Flowchart

Step 1: Input data

The data is taken from the UCI machine learning repository, the detailed explanation of which is shown in TableI.

Step 2: Data Transformation

First, all the string values are converted into numeric values. Next, since there is a vast difference in the values of the dataset, normalization is done using the Min-Max Scaler, to scale the values in the range of [0,1] or [-1,1].

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (2)$$

Where, x denotes the value that needs to be normalized

Min(x) is the minimum value of 'x' in the column.

Max(x) is the maximum value of 'x' in the column.

Step 3: Data Selection

The most relevant attributes, which contribute to the occurrence of the fires are selected, and the remaining is dropped.

Step 4: Data Modeling

The input data to the model now consists of only those attributes that are contributing to the wildfires. This data is split into two parts- training and testing set. It is split in the ratio 80:20. A model is created using the training data. The training is done using 1200 epochs and to optimize the model, Adaptive learning rate optimization (Adam) optimizer is used.

Step 5: Output

The output was generated using the trained model with the help of test data. This output is shown in the next section.

VI. RESULT ANALYSIS

The model is trained with 1200 epochs. TableIII. shows the output of 6 epochs and the corresponding accuracy.

TABLE III. TRAINING ACCURACY

Epoch Number	Accuracy
1	0.4552
255	0.6901
507	0.7458
780	0.8015
1199	0.8475
1200	0.8547

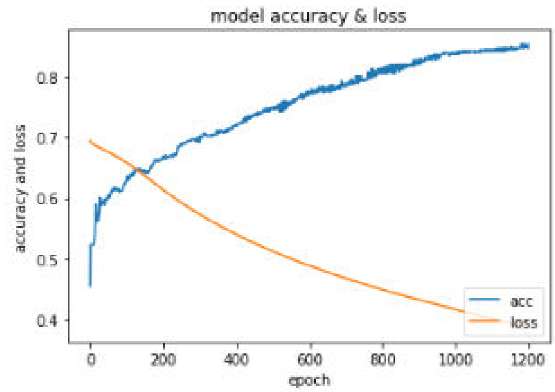

Fig.5. Training path loss exponent and accuracy

The number of epochs is chosen such that overfitting does not occur. As seen from above, as the number of epochs increases, the loss reduces and hence the accuracy increases. This model gives the test accuracy of 62.5%.

VII. CONCLUSION AND FUTURE SCOPE

In this paper, the forest fire dataset which consists ofmeteorological and other data was used to train different algorithms such as Logistic regression, Decision tree and SVM. From this it was incurred that the accuracy was around 50%. The same data was also used to build a neural network which showed better performance than the above-mentioned algorithms by giving a higher accuracy of . The work canfurther be improved by predictingits rate of spread, if a fire does occur.

REFERENCES

[1] S. Dian et al. "Integrating Wildfires Propagation Prediction into Early Warning of Electrical Transmission Line Outages," in IEEE Access, vol. 7, pp. 27586-27603, 2019.

[2] P. Csontos and I. Cseresnyes, "Fire-Risk Evaluation of Austrian Pine Stands in Hungary effects of drought conditions and slope aspect on Fire spread and Fire behavior," Carpathian Journal of Earth and Environmental Sciences, August 2015, Vol. 10, No 3, pp. 247-254, 2015.

[3] M. Salehi, L. Rusu, T. Lynar and A. Phan, kdd.org, "Dynamic and Robust Wildfire Risk Prediction System: An Unsupervised Approach" Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining pp. 245-254, 2016

[4] X.Wang et al. "CFFDRS: An R package for the Canadian Forest Fire Danger Rating System" Wang et al. Ecological Processes (2017) Vol. 6, 6:5 ,2017

[5] J. Gao et al., "Data-driven forest fire analysis," 2017 IEEE Smart World Intelligence and Computing, Advanced and Trusted Computed,Scalable Computing,Cloud,Big Data-Computing, San Francisco CA, 2017, pp.1-7

[6] S. Lall and B. Mathibela, "The application of artificial neural networks for wildfire risk pre-diction," 2016 International Conference on Robotics and Automation for Humanitarian Ap-plications (RAHA), Kollam, 2016, pp. 1-6

[7] Cortez, Paulo and Morais, A., "A Data Mining Approach to Predict Forest Fires using Meteorological Data, Department of Information Systems/R and Algoritmi Centre, University of Minho,", ResearchGate, 2007

[8] O. Satir, S. Berberoglu, and C. Donmez, "Mapping regional forest fire probability using artificial neural network model in a Mediterranean forest ecosystem," Geomatics, Natural Hazards and Risk, vol.7, no.5, pp. 1645-1658, 2015