Real-Time Wildfire Monitoring and Alert System

Dr. B. Mathivanan¹, S. Keerthana², B. R. Naveen Kumar³, K. Nishal⁴

¹Associate Professor In Sri Ramakrishna Engineering College In The Department of Computer Science And Engineering, Tamilnadu, India
²,³,⁴Under Graduate Student In Sri Ramakrishna Engineering College In The Department of Computer Science And Engineering, Tamilnadu, India

ABSTRACT

Forest fire leads to imbalances in nature causing difficulties to the living organism. This gives rise to the urgent need to predict forest fires as accurate as possible. This system uses a powerful machine learning algorithm as a potential solution to the challenge of early prediction of forest fires. The device presented makes use of various sensors attached, solar recharging mechanism and wireless data transmission. These collected data are analysed through the K-Nearest Neighbour (KNN) and Random forest (RF) algorithm and when forest fire is predicted, intimation is given to forest range office where they can prevent the disaster. The system is efficient and green fulfilling the need for its creation.

Keywords: Forest Fire Prediction, KNN, Machine Learning, RF.

I. INTRODUCTION

Forest fire is a major environmental issue that creates economical and ecological damage and endangers human lives. Early detection of forest fire is the efficient way of controlling these causes. To achieve this, building a system using sensors for monitoring and building a strong communication system in the forest area is one of the solution. Meteorological conditions such as CO₂, temperature influence forest fires and fire indexes namely the forest Fire Weather Index (FWI), use these data. Many technologies can be implemented to monitor the forest fire but the efficiency and the accuracy of the results vary. The system is built using the machine learning algorithms and it is strengthened with the ensemble learning to improve the accuracy of the results.

II. METHODOLOGY

A. Machine Learning

Machine Learning is one of the branch in science that deals with the way of developing a machine that automatically learns and improves its performance with its experience. Machine Learning is the process of recognizing and understanding the input data and results in wise decision based on conditions. Machine Learning can be implemented by two techniques namely supervised and unsupervised learning. Supervised learning deals with developing a function with the available training or labelled data whereas unsupervised learning deals with developing a function with unlabelled data without having predefined dataset. Supervised learning includes regression, decision tree, random forest and classification. Classification can be done by using algorithms namely KNN, Tress, Logistic Regression, Naïve Bayes and SVM. Unsupervised Learning involves clustering, association analysis and hidden Markov model. Clustering can be done by using SVD, PCA and K means. Apart from supervised and unsupervised learning it involves another type of learning called reinforcement learning. Reinforcement learning is a reward-based learning. In this type of learning machine learns how to act in a certain environment to maximize the rewards.

B. Random Forest Algorithm:

Random Forest is a classification algorithm which approaches in an ensemble learning method. Random forest is an algorithm which is tree based and involves in building several trees (decision trees), to combine the output to enhance the model’s ability to popularize. The Random Forest process repeatedly develops the decision trees which are utilized in the segregation or regression issues. The reason for choosing the random forest is as follows: Random Forest supports multiclass classification whereas SVM needs multiple models for the same. SVM reduces its efficiency when dealing with categorical values. It does not need feature scaling like neural network. Naïve Bayes works well for small training data and it cannot be applicable in real life scenarios but random forest works well for real time scenarios.

Working of algorithm:

1) Construction of decision tree and prediction of result for every sample.
2) Voting is performed for every predicted result.
3) Select the most Voted prediction result as the final prediction result.

C. K-Nearest Neighbour Algorithm:

The KNN algorithm is held by the family of instance-based, competitive learning and lazy learning algorithms. KNN is well-built because the data assumption is never done, except that the distance measure is calculated constantly between any two instances. Therefore, it is referred as non-parametric or non-linear as the functional form is not assumed. Reasons for the selection of KNN algorithm is as follows: If the training data is much larger than number of features then KNN is better than SVM. Naive Bayes is much faster than KNN due to KNN’s real time execution but Naive Bayes is not suitable for real time. Neural Networks needs large training data compared to KNN to achieve sufficient accuracy.

Working of Algorithm:

1) The value of K is chosen.
2) For each point in test data
   Calculate the distance between test data and each row of training data using Euclidean method.
   Based on distance, sort data in ascending order.
   Choose top K rows from the sorted array.
3) Assign a class to test point based on frequent class of these rows.

III. LITERATURE SURVEY


Exploratory data analysis such as boxplot and bar graph is conducted on all the attributes of the dataset to discover patterns, to spot anomalies and to check assumptions with the help of summary statistics and geographical representations. Normalization techniques such as principle component analysis (PCA) which is used to convert higher dimensional correlated features into lower dimensional uncorrelated features. Feature scaling such as min-max scaling is done to dataset to standardize the range of the variable of a data. Lable encoding is done to change the categories yes and no to 1 and 0 to support algorithms that do not intake categorical values .Once the dataset is cleaned. Machine learning algorithms such as Logistic Regression, SVM, Random forest, KNN bagging classifier [decision tree]. Boosting classifier is used to predict the forest fire. Results with respect to each machine learning algorithm after performing PCA and before performing PCA was discussed. Gradient boosting classifier gave the highest accuracy about 68.38 in both the methods.


Proposed application based on WSN for monitoring forest fires and preventing further spread and disaster.3 factors-fuel, oxygen and heat source .This paper provides a proof of concept for an application that will monitor the forest and prevent the potential fire by using IOT technologies system involves the deployment of several sensors in a forest area. The gathered data from the sensors is then transmitted to a web database application which will process the data and report the results to the end users. The main functionalities are monitoring process, production process. Clustering method along with LEACH (Low-energy adaptive clustering hierarchy) is used in the WSN network. The application is structured in 3 levels: data acquisition level, server level, client level. Data acquisition level is responsible to implement the network with auto configuration, low energy consumption, fault tolerance, scalability. These are attained using the LEACH algorithm. Server level is configured with data processing scripts and inserting new values in the database. Client level is added to typical client-server application. Browser-client is added to the server. The browser shows the graphs representing the risk index of forest fire, alarms for rainfall, temperature, risk index of prediction. This application will help users to react faster when forest fire appears using the friendly user interface.


Data mining approach is used to predict the forest fire. SVM (Support Vector Machine) can be utilized in anticipating forest fire with more accuracy. Set of four sensors is used to collect thermal data, relative dampness, rain and wind velocity. Fog computing technology is used to increase computation speed. Cumulative data is transferred through ZigBee tool via fog nodes. Since satellites and scanners are costly to implement the application is developed using data mining approach. Grouping and regression data analysis is done. Temperature, mist and humidity is measured continuously. FWI concept method is used to conclude the parameters involved. Five data mining approach is
implemented namely: Multiple regression, SVM, Neural Network, Random Forest, Decision tree. The project is initiated by collecting the data from the sensors, if the values reaches the threshold value then high priority is given and values is transferred through fog nodes and backed up in cloud. Prediction model is built using SVM. It predicts the forest prone area and output in given as input in CSV (Comma Separated Values). Data that is backed up in the cloud can be used for future uses. When the fire is predicted it notifies the forest department.

IV PROPOSED SYSTEM

The proposed solution recommends stand-alone boxes which are to be deployed throughout a forest and the boxes are connected through wireless sensor networks. Each box contains different types of sensors and ZigBee module configured either as router, end device or coordinator. These units communicate wirelessly through the ZigBee mesh network. Every wireless device is equipped with smoke sensor, moisture sensor, temperature humidity sensor, Arduino and ZigBee. MQ5 grove-gas sensor is used to detect the gas emission that is occurred during the forest fire. This gas sensor detects CO, methane and also CO2 gases. Carbon monoxide and methane is the major released gases during the forest fire. During the forest fire there will be huge increase in the temperature and the humidity of the air will be decreased in a high range. The temperature and humidity will be sensed using DHT11 sensor. Sometimes the dry weather conditions would be the cause of the fire. So, such conditions of the forest weather are monitored using the moisture sensor. These sensors are always in active mode. It is not practical to power up these wireless sensors using electricity or not rechargeable batteries. Hence, it is preferred for these devices to have a renewable form of energy that charges the battery such as solar energy system. In order to achieve the best fire detection results with the fastest time, a number of detection devices should be deployed in different spots in the forest. These spots in the forest are known as the hotspots. This number of hotspots is not randomly chosen, but varies from one forest to another. Accordingly, a mesh network links all these devices together.

The sensors at the transmitter end senses the weather conditions continuously and the data is collected, shared, and transmitted to a sync node, located outside the forest and it is sent to the server (here pc) at the receiver end through the stand-alone units. The prediction algorithm is developed to predict the forest fire percentage and to alert the authority. The dataset consisting of the historical data about the weather conditions before the occurrence of fire is given as input for data analysis. This dataset is pre-processed in order to remove the missing values and unnecessary attribute. The dataset is cleaned by using the fillna function. This function will remove the NaN values and fill it with the values. Finally, the cleaned and pre-processed data is used for the prediction. The current sensed weather conditions are analysed against the inbuilt dataset using the ensembled KNN and random forest algorithm and the forest fire prediction percentage is found. The machine learning algorithm KNN and random forest is used. Ensemble learning technique is deployed to increase the accuracy percentage. KNN algorithm uses the Euclidean distance concept and predicts the percentage. Once the four parameters enter into the server the prediction process gets started. For each parametric value the Euclidean distance is calculated between the current value and the values in the dataset and the column is updated with the new distance value.

The minimum distance value is taken for the prediction and the percentage is calculated. Random Forest algorithm is developed by building four trees. Each tree is used to predict each parametric value namely temperature, moisture, humidity and gas. When the sensor values enter into the prediction process through the RS232 data logging each parametric value enters into each tree and the class is predicted. Finally, the voting of the 4 trees is done and majority voting is considered for predicting the forest fire. In order to avoid the false prediction, the aggregated result of four trees in taken into consideration. Finally, the aggregated result of the two algorithms (KNN and Random forest) is used in the final prediction process. For the process of aggregation, the ensemble learning method is used. KNN and Random forest algorithm is ensembled based on the prediction percentage and aggregated result is found. When the fire prediction percentage crosses the threshold percentage then the alert is made to the forest fire department using the Twilio process. The Twilio account is created in the Twilio website and the numbers are verified. At the server end the web application is built. The login page consists of validation process to verify that the user is an admin. It consists of admin name and the password. Once the validation process is completed the user gets the admin privilege.

The home page consists of links to the following pages: range dataset, date wise, parameter wise, prediction, chart, change password and logout pages. The range dataset page displays the dataset that is fed into the application. The Euclidean distance gets updated in the last column when the prediction process gets started. The date wise page shows the sensed data from the three sensors with respect to the dates. Once the date is selected the parametric values gets displayed based on the date and time. This helps to know whether there is a gradual change in the sensed values. Parameter wise page will display the sensed data continuously that is read from the RS232 data logger. This page can be downloaded as an excel sheet which can be viewed later. The clear data option in this page will clear the data which is displayed. Sometimes the noise may interrupt during the transmission so in order to remove that noise clear data option is used. Chart page displays the sensed values that are plotted in a graph. These charts will help to interrupt the sensor details in an easier manner. Change password page will helps to change the password of the admin privilege.
This page is used for security purposes. Logout button is used to logoff of the application. The prediction page consists of the option to update the mail id to which the mail has to be sent when forest fire is predicted. The result of the prediction process is displayed in this page and alert is sent to the concerned department through mail and call. Mail is sent through the SMTP process and call is made through the Twilio process.

Fig (a): Overview of Wildfire Prediction System

IV RESULT AND CONCLUSION

The benefit of using this system is that it reduces the occurrence of forest fire in deep dense forest. If any initial occurrence of forest fire is predicted, an immediate alert is sent to the nearby forest range office in order to prevent from major loss of natural resource and living organism. The screenshot of the application’s home page containing the options described in the proposed system is added. The result contains the KNN algorithm percentage, Random forest algorithm percentage and the aggregated result percentage.
Fig (b): Screenshot of the home page

Fig (c): Screenshot of the predicted result

REFERENCES


