An Efficient And Optimal Machine Learning Algorithm For Real-Time Forest Fire Prediction Using Weatherstack API

Submitted by
Reddy jyothi sri D
(RA2011028010108)

AIM AND SCOPE:

To create an efficient ML model that predicts the chances of forest fire using the weather information retrieved from the third party API's with minimal budget.

ABSTRACT:

Forest fires destroy natural life and demolish lands that provide a livelihood. Accurate prediction of fire occurrence is required in order to reduce its adverse effects.This paper presents an effective and reliable solution to predict the forest fires so as to minimize their effects .The current solution to this problem is to gather info through sensors which are placed at several locations inside of the forest and predict the chances of occurrence of fire. But this method is quite expensive and requires maintenance .And also the harsh conditions of forests might disable or displace the sensors which might no longer be useful . So , an effective solution is to use the weather info provided by the satellite to predict the chances of forest fire.The proposed system uses a third party API to retrieve the weather info and feed it to the machine learning model to predict the occurrence of forest fire.

https://www.youtube.com/watch?v=F6kmIpWWEdU&list=PLeo1K3hjS3uuASpe-1LjfG5f14Bnozjwy&index=2

OBJECTIVES :

1) Create an efficient Machine Learning model that predicts forest fires with utmost accuracy using the real time weather info from a 3rd party API.
2) To reduce the budget and maintenance as much as possible.

Literature Survey:

1) A Comprehensive Study on Weather Forecasting using Machine Learning

   by-Deepti Mishrav and Pratibha Joshi

   Abstract - Weather forecasting which is a key player in everyday life is a remarkable advantage of science and technology. Weather prediction can support people as protecting the assets and lives of them. The persons involved in outdoor occupations can be benefited by the weather prediction as they needed to know the weather previously. In the paper, the concept of supervised learning is used, which is one of the learning techniques in machine learning. The paper presents the algorithm, which applies the concept of linear regression and artificial neural network, to predict the variation of temperatures such as high and low values as features in a linear combination. The main features focus on maximum or minimum temperature, mean values of humidity, and atmospheric pressures considering the previous two days. The paper focuses on prediction and analysis of weather forecasting which further entirely applies the concept of machine learning

2) An Efficient and Optimal Clustering Algorithm For Real-Time Forest Fire Prediction with Sensor Networks and Data Mining

   By -Divya.T.L, Manjuprasad.B, VijayaJakshmLM.N, Andhe Dharani

   Abstract - Grouping of forest fire images into meaningful categories to reveal useful information is a challenging task. In order to overcome this challenge, data mining techniques can be used with wireless sensor network which can detect and forecast forest fire more promptly than the satellite-based detection approach. This paper proposes an efficient image clustering algorithm using real time data for predicting of the occurrence of forest fire, with a new mechanism for secure information transmission in wireless sensor networks by minimizing the threat attacks caused by malicious nodes in wireless sensor networks.

3) Leveraging Machine Learning to Predict Wild Fires

by-K Venkata Murali Mohan,Mallikharjuna Rao K,Aravapalli Rama Satish

Abstract - A raging wildfire is a catastrophic event which damages forests, which has a serious effect on people, fauna and flora that are dependent on the forest ecosystem. A study of the size of wildfires in a Canadian Province in USA i.e. Alberta is seen in this article. A variation of the duration of the fire and the area it burns defines the scale of a fire. Our predictive algorithm helps wildfire rescue workers to use their foreseen level in the initial phases in order to mitigate destruction inflicted by a forest fire. Modeling information has been gathered from Natural Resources Canada's realtime dataset, including forest fire and weather information for Alberta, Canada. To evaluate the severity of flames, the dimensions of the region affected with fire and the timeframe of the flames have been used. The information was split into training and evaluation environments after multi-linearity validation and function normalization. In addition, the climatic variables were used to create predictive model by using inputs, a Neural Network for Back Propagation (BPNN), a type of artificial neural network i.e. Recurrent Neural Network (RNN) and a type of RNN i.e. Long Short-Term memory (LSTM). LSTM showed the greatest precision, 95.9 percent, of these classification models. The findings suggest that the scope of a wildfire can be forecast using climatic knowledge at the outset of the event

4) A Data Mining Approach to Predict Forest Fires using Meteorological Data

by-Paulo Cortez and Anibal Morais1

Abstract-Forest fires are a major environmental issue, creating economical and ecological damage while endangering human lives. Fast detection is a key element for controlling such phenomenon. To achieve this, one alternative is to use automatic tools based on local sensors, such as provided by meteorological stations. In effect, meteorological conditions (e.g. temperature, wind) are known to influence forest fires and several fire indexes, such as the forest Fire Weather Index (FWI), use such data. In this work, we explore a Data Mining (DM) approach to predict the burned area

of forest fires. Five different DM techniques, e.g. Support Vector Machines (SVM) and Random Forests, and four distinct feature selection setups (using spatial, temporal, FWI components and weather attributes), were tested on recent real-world data collected from the northeast region of Portugal. The best configuration uses a SVM and four meteorological inputs (i.e. temperature, relative humidity, rain and wind) and it is capable of predicting the burned area of small fires, which are more frequent. Such knowledge is particularly useful for improving firefighting resource management (e.g. prioritizing targets for air tankers and ground crews).

5) Artificial Neural Network for Weather Forecasting: A Review

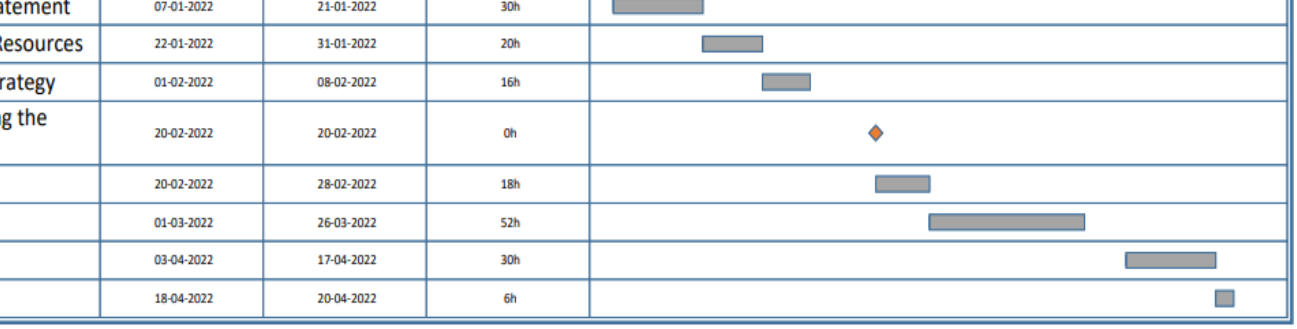
By- Ushakiran Huiningsumbam,Anjali Jain, Neelam Verma

Abstract - Forecasting is considered one of the most imperative and perplex proceedings in the present-day. Weather forecasting is a natural process which entails the prediction of the change in the atmospheric conditions with the passage of time. It has become a core area of studies and analysis for the scientist owing to the brusque forecasts and approaches of weather. Weather forecasting has always been a challenging task. In several cases, scientists and researchers had attempted in establishing a linear association between the input and targeted data of the weather. Since, weather is non-linear and dynamic in nature, the target has shifted to prediction of non-linear weather data. Though weather prediction is relatively a statistical measure and is automated, with the traditional tools, its result is rather uncertain and not always accurate. Due to its non-linearity and complex process, the best approach for resolving such problems is with the use of Artificial Neural Network (ANN). ANN simplify the weather predictions with its better efficiency, reliability and accuracy. The features of ANN are not just to analyse the past data but also to acquire future predictions rendering to be much ideal for weather forecasting. NN is rather a complicated network which are pliable and flexible in nature. It is autodidactic with its existing

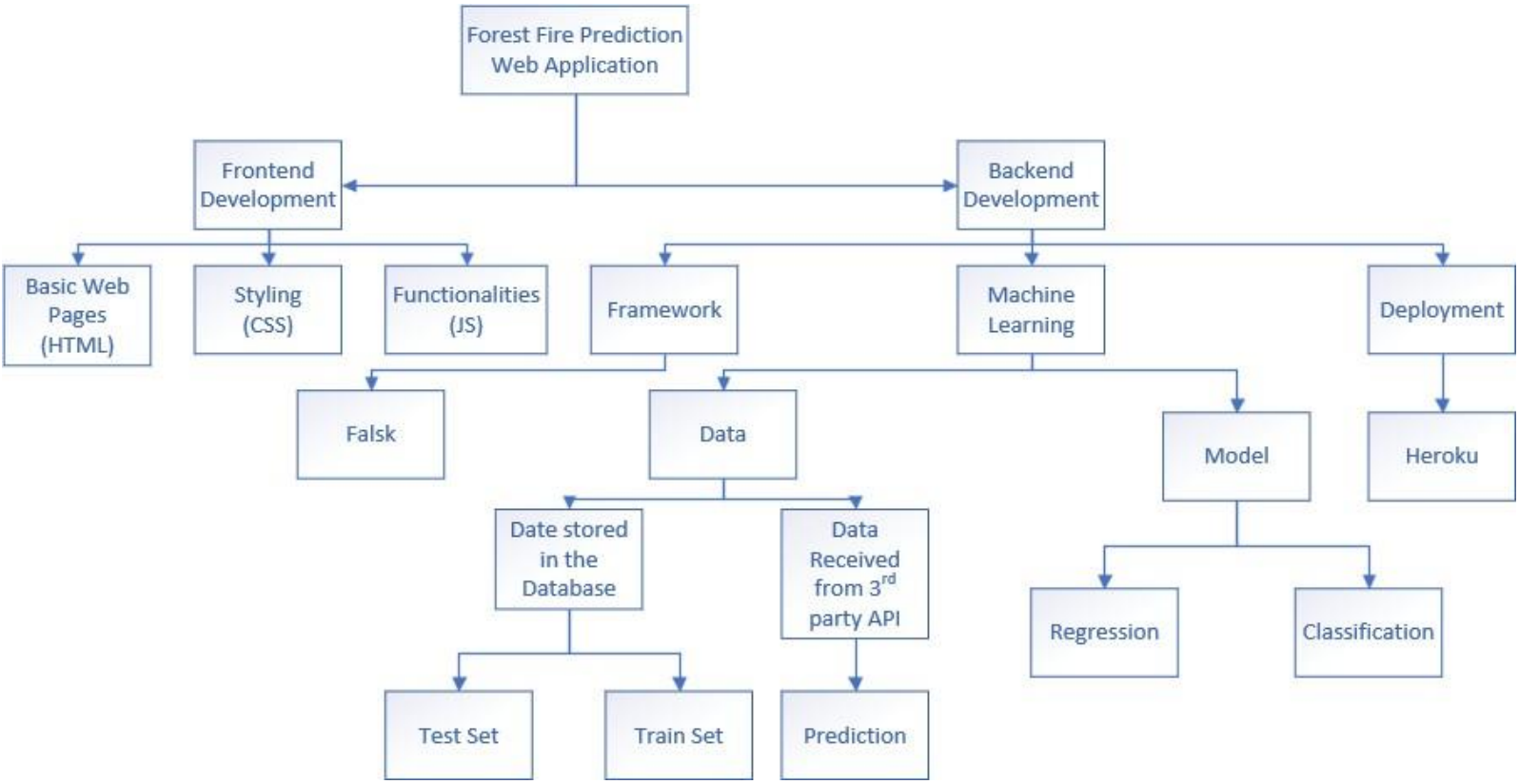
training data consequently forging a new smart pattern useful for predicting the weather. Survey of various techniques of NN for weather prediction is provided and it is observed that by simply increasing the number of hidden layers the trained NN can predict and classify the weather variables with minimal error. In this paper, predictive analysis algorithm is incorporated with back propagation network(BPN) to predict future weather by training the network. The technical milestone, where various researchers have acquired on this discipline has been reviewed and presented in these surveys.

SCHEDULE DIAGRAM OF THE PROJECT:

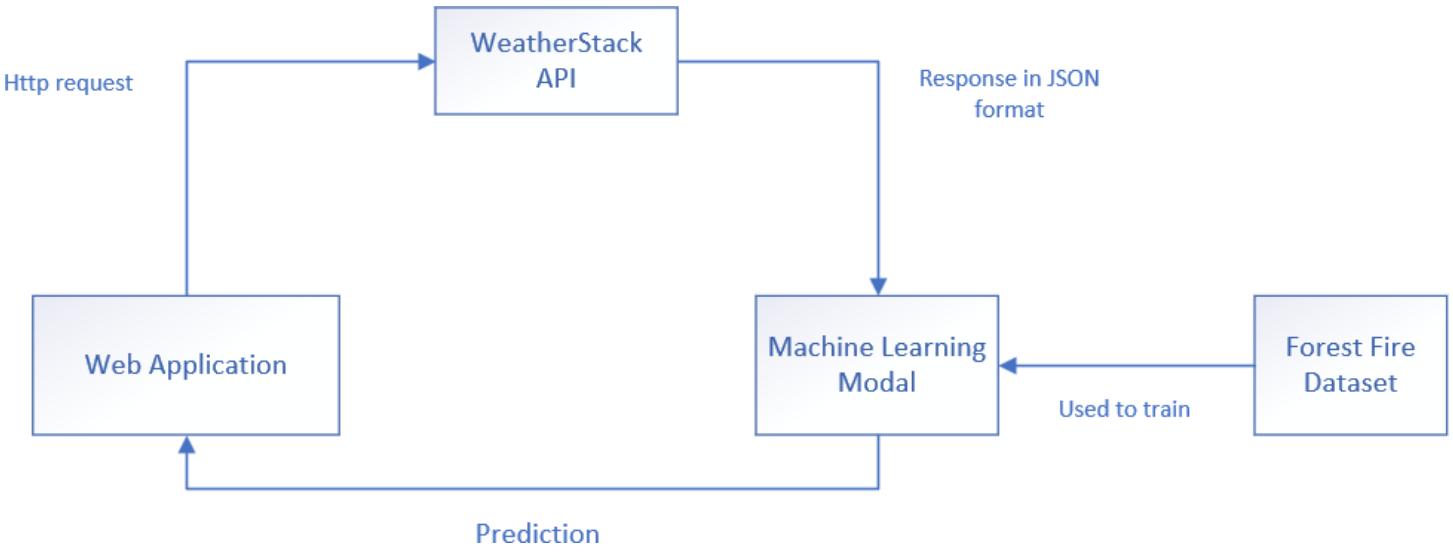
ID	Task Name	Start	Finish	Duration
1	Team Formation	03-01-2022	06-01-2022	8h
2	Defining the problem Statement	07-01-2022	21-01-2022	30h
3	Acquiring the Data and Resources	22-01-2022	31-01-2022	20h
4	Developing a Solution Strategy	01-02-2022	08-02-2022	16h
5	Designing and Developing the algorithm	20-02-2022	20-02-2022	0h
6	Pre Processing the Data	20-02-2022	28-02-2022	18h
7	Creating an efficient ML model	01-03-2022	26-03-2022	52h
8	Deploying the Model	03-04-2022	17-04-2022	30h
9	Testing and Debugging	18-04-2022	20-04-2022	6h



WBS OF THE PROPOSED METHOD:



BLOCK DIAGRAM / SYSTEM ARCHITECTURE:



METHODOLOGY:

- **Pre Processing the Data :**
    - The Dataset that would be used by the ML model would contain a lot of unnecessary information . So , preprocessing is necessary.
    - Only the required attributes are selected to form the data frame.
    - The rows with NaN values would either be deleted or would be filled with the median value of that column.
    - Outliers must be removed before it is fed into the model.

- **Creating the Machine Learning Model :**
    - Various Machine Learning Models including Regression and Classification would be developed that would predict the chances of forest fires .
    - The Model would be trained using the training set obtained after pre processing the dataset.
    - Efficiency of each model will be checked using the test dataset and accordingly the best model will be selected.

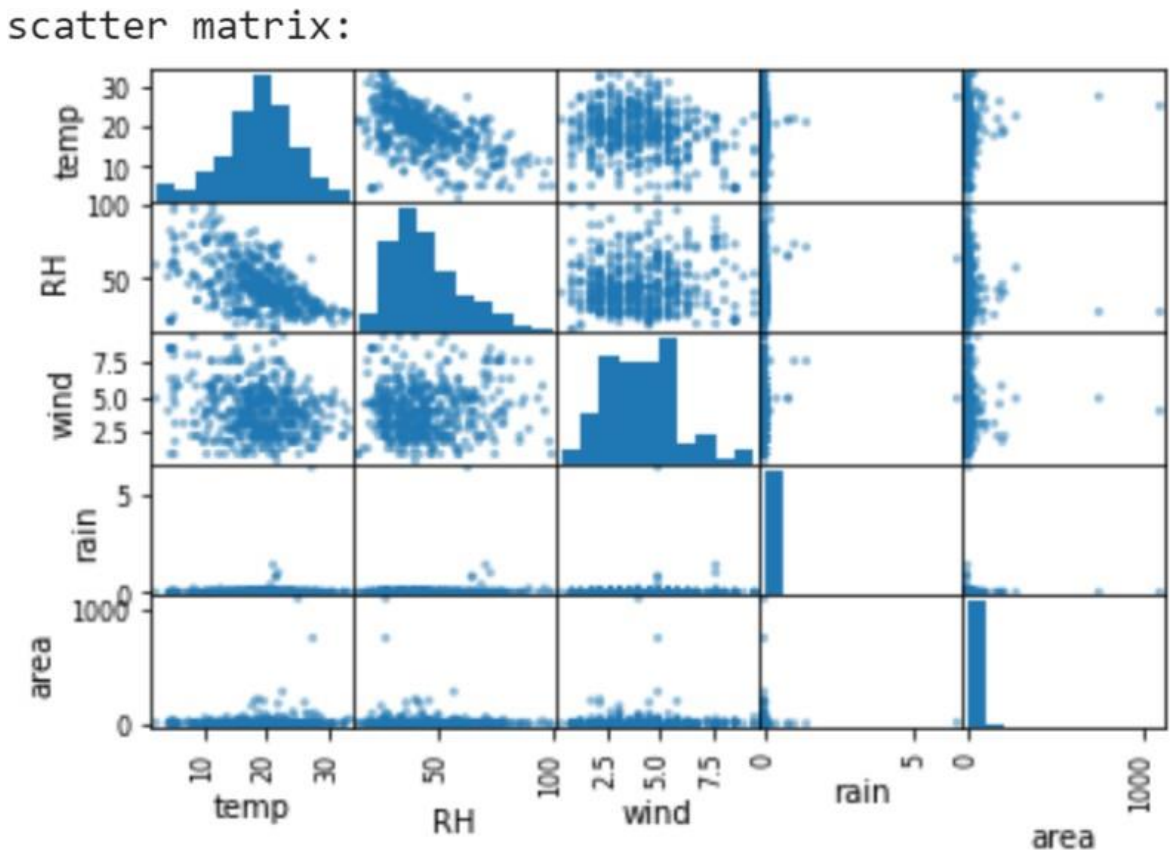
- **Creating the web application to deploy our model :**
    - The web application would send a http request to the Weather Stack API , that would send us the weather data in JSON format.
    - We will retrieve the necessary information and feed it to the ML model at the backend .
    - The model will predict the chances of the forest fire and send it back to the web application which would then be displayed on the web page.
    - Every time the web page is visited or refreshed , the same process would be continued.

DELIVERABLES:

1) The main objective is to create a web application , which takes the real time weather info as an input from a third party API and predicts the chances of forest fire.
2) The budget required for this project is almost nil.

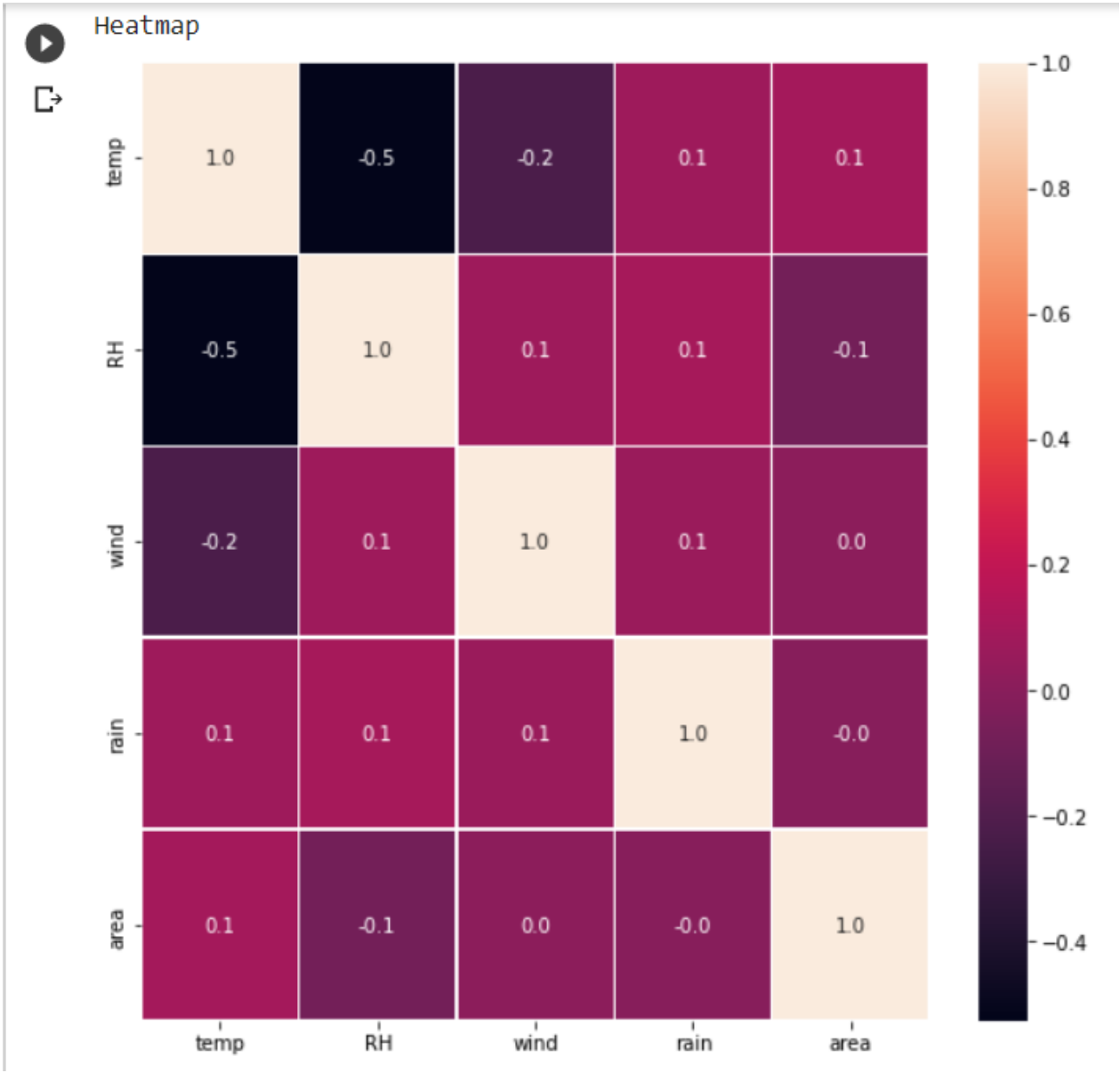
METRICS:

Scatter matrix shows the relation between each parameter



HEATMAP:

Heat maps in Python is a type of a graph which represents different shades of a colour to distinguish the values in the graph. The higher values are represented in the darker shades and the lesser values are represented in lighter shades.



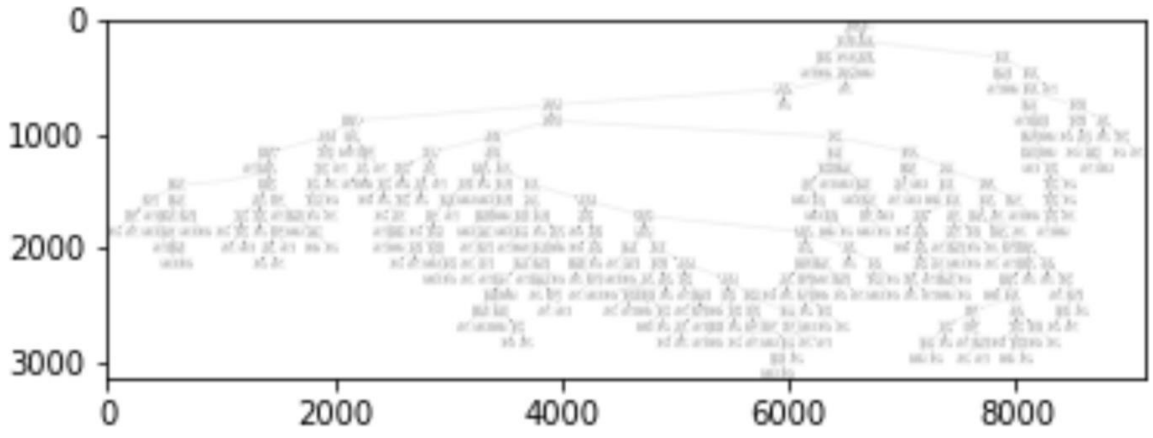
ACCURACY:

```
from sklearn.ensemble import ExtraTreesRegressor

extratree = ExtraTreesRegressor()
extratree = extratree.fit(X,y)
pred = extratree.predict(X)
score = explained_variance_score(y,pred)
meanabserr = mean_absolute_error(pred,y)
print("variance IN extratreeregressor :",score)
print("mean absolute error IN extra tree regressor :",meanabserr)
```

```
variance IN extratreeregressor : 0.9228459701734151
mean absolute error IN extratreeregressor : 0.03803997421018697
```

DECISION TREE:



1) Testing

The data is splitted into training data and testing data

```
df = pandas.read_csv("/content/forestfirex.csv")
df1 = pandas.read_csv("/content/forestfirex1.csv")

print(df)
features = ['temp', 'RH', 'wind', 'rain']

X = df[features]
y = df1['area']
x_train,x_test,y_train,y_test = train_test_split(X,y,train_size =0.9)
```

x_train = this is the dataframe that consists of temperature ,relative humidity,rain,windspeed used in training the data model.
x_test =this is the dataframe that consists of temperature ,relative humidity,rain,windspeed used in testing the data model. **y_train=**this is the dataframe that consists of area burnt used in training the data model.
y_test=this is the dataframe that consists of area burnt used intesting the data model.

The training dataset is used to train in the extratreesregressor model while the testing data set is used to predict the outcome of the trained model

```
from sklearn.ensemble import ExtraTreesRegressor

extratree = ExtraTreesRegressor()
extratree = extratree.fit(X,y)
pred = extratree.predict(X)
score = explained_variance_score(y,pred)
meanabserr = mean_absolute_error(pred,y)
print("Accuracy IN extratreeregressor :",score)
print("mean absolute error IN extra tree regressor :",meanabserr)
```

```
Accuracy IN extratreeregressor : 0.9228459701734151
mean absolute error IN extra tree regressor : 0.03803997421018697
```

We got an Accuracy of 0.92 (due to less available data we got high accuracy)

Tools and techniques

- **Machine learning**
  - We have used scikit learn and matplotlib for the implementation of machine learning
  - **Algorithms**
    We have used decision tree algorithms which gives more accuracyfor less data
    1) **extratreeregressor**
    2) **decision tree classifier**
    3) **random forest regressor**
    4) **bagging regressor**

    Among these machine learning models extra tree regression give best accuracy so we have implemented it in ourapplication

  - **Flask**
    we have used Flask in back-end
  - **Sendgrid**
    We used sendgrid to send mail when ever a forest fire impact is high
  - **Weatherstack API**
    We used weather stack api to retrieve the required parameters using http request in json format

2) **Storage**

- MySql Database is used for storing the data.
- The database is updated every time a http request is made to the weather stack api for the weather details.
- The dataset is updated from time to time and the model is tested with the new data ,to make it more reliable.

- Even in case of system failure or any other disaster , the data remains safe as it is stored on the MySql database.

```
mysql> use forestfire;
Database changed
mysql> select * from forestinfo01;
+------+------+------+------+---------+
| temp | RH   | wind | rain | area    |
+------+------+------+------+---------+
|  8.2 |   51 |  6.7 |    0 |       0 |
|   18 |   33 |  0.9 |    0 |       0 |
| 14.6 |   33 |  1.3 |    0 |       0 |
|  8.3 |   97 |    4 |  0.2 |       0 |
| 11.4 |   99 |  1.8 |    0 |       0 |
| 22.2 |   29 |  5.4 |    0 |       0 |
| 24.1 |   27 |  3.1 |    0 |       0 |
|    8 |   86 |  2.2 |    0 |       0 |
| 13.1 |   63 |  5.4 |    0 |       0 |
| 22.8 |   40 |    4 |    0 |       0 |
| 17.8 |   51 |  7.2 |    0 |       0 |
| 19.3 |   38 |    4 |    0 |       0 |
|   17 |   72 |  6.7 |    0 |       0 |
| 21.3 |   42 |  2.2 |    0 |       0 |
| 26.4 |   21 |  4.5 |    0 |       0 |
| 22.9 |   44 |  5.4 |    0 |       0 |
| 15.1 |   27 |  5.4 |    0 |       0 |
| 16.7 |   47 |  4.9 |    0 |       0 |
| 15.9 |   35 |    4 |    0 |       0 |
|  9.3 |   44 |  4.5 |    0 |       0 |
| 18.3 |   40 |  2.7 |    0 |       0 |
| 19.1 |   38 |  2.7 |    0 |       0 |
|   21 |   44 |  4.5 |    0 |       0 |
| 19.5 |   43 |  5.8 |    0 |       0 |
| 23.7 |   32 |  5.8 |    0 |       0 |
| 16.3 |   60 |  5.4 |    0 |       0 |
|   19 |   34 |  5.8 |    0 |       0 |
| 19.4 |   48 |  1.3 |    0 |       0 |
| 30.2 |   24 |  2.7 |    0 |       0 |
| 22.8 |   39 |  3.6 |    0 |       0 |
| 25.4 |   24 |  3.6 |    0 |       0 |
| 11.2 |   78 |  7.6 |    0 |       0 |
| 20.6 |   37 |  1.8 |    0 |       0 |
| 17.7 |   39 |  3.6 |    0 |       0 |
| 21.2 |   32 |  2.7 |    0 |       0 |
| 18.2 |   62 |  4.5 |    0 |       0 |
| 21.7 |   24 |  4.5 |    0 |       0 |
| 11.3 |   60 |  5.4 |    0 |       0 |
| 17.8 |   27 |    4 |    0 |       0 |
| 14.1 |   43 |  2.7 |    0 |       0 |
| 23.3 |   37 |  3.1 |    0 |       0 |
| 18.4 |   42 |  6.7 |    0 |       0 |
| 16.6 |   54 |  5.4 |    0 |       0 |
| 19.6 |   48 |  2.7 |    0 |       0 |
| 12.9 |   74 |  4.9 |    0 |       0 |
| 25.9 |   24 |    4 |    0 |       0 |
| 14.7 |   70 |  3.6 |    0 |       0 |
|   23 |   36 |  3.1 |    0 |       0 |
```

Cost-Benefit Analysis

- We have many papers regarding forest fires which used various sensors to find various parameters .so it is quite expensive to maintain those type of plans.
- Our application uses the data that is retrieved from the weatherstack api
  Which reduces the cost of maintenance

Cost analysis:

We invest money during the starting stage for servers and acquiring data for the machine learning model then after that there will be low maintenance cost because we don't use sensors

3) Implementation

Step 1 : when we visit the webpage , a http request is sent to the weather stack API , requesting for the current weather details.

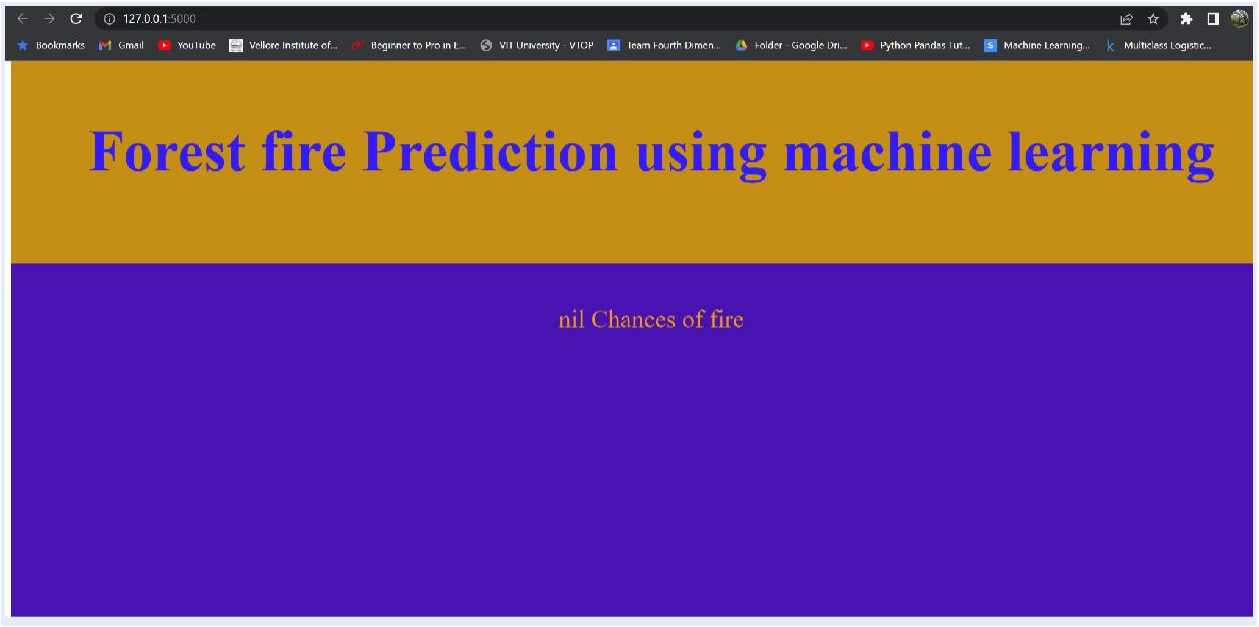
Step 2 : As a response , the API sends us the weather info of that particular place , in the following JSON format


```
{
    "request": {
        "type": "City",
        "query": "San Francisco, United States of America",
        "language": "en",
        "unit": "m"
    },
    "location": {
        "name": "San Francisco",
        "country": "United States of America",
        "region": "California",
        "lat": "37.775",
        "lon": "-122.418",
        "timezone_id": "America/Los_Angeles",
        "localtime": "2019-09-03 05:35",
        "localtime_epoch": 1567488900,
        "utc_offset": "-7.0"
    },
    "current": {
        "observation_time": "12:35 PM",
        "temparature": 16,
        "weather_code": 122,
        "weather_icons": [
            "https://assets.weatherstack.com/images/symbol.png"
        ],
        "weather_descriptions": [
            "Overcast"
        ],
    "wind_speed": 17,
    "wind_degree": 260,
    "wind_dir": "W",
    "pressure": 1016,
    "precip": 0,
    "humidity": 87,
    "cloudcover": 100,
    "feelslike": 16,
    "uv_index": 0,
    "visibility": 16
    }
}
```

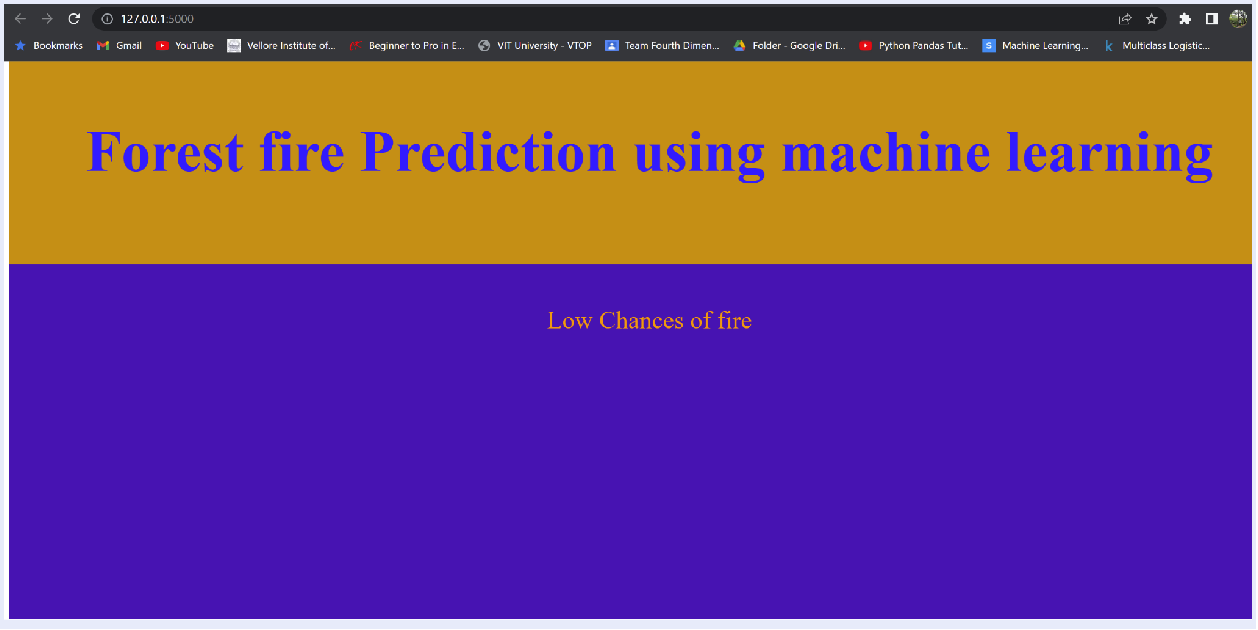
Step 3 : From all the data received , we retrieve the necessary parameters , which are- temperature, wind speed , relative Humidity and precipitation. This retrieved data is given as an input to the Machine learning model (Extra tree regressor) and as the result , the model predicts the area that might be affected.

Step 4 : based on the predicted area , the following messages will be displayed onthe website.

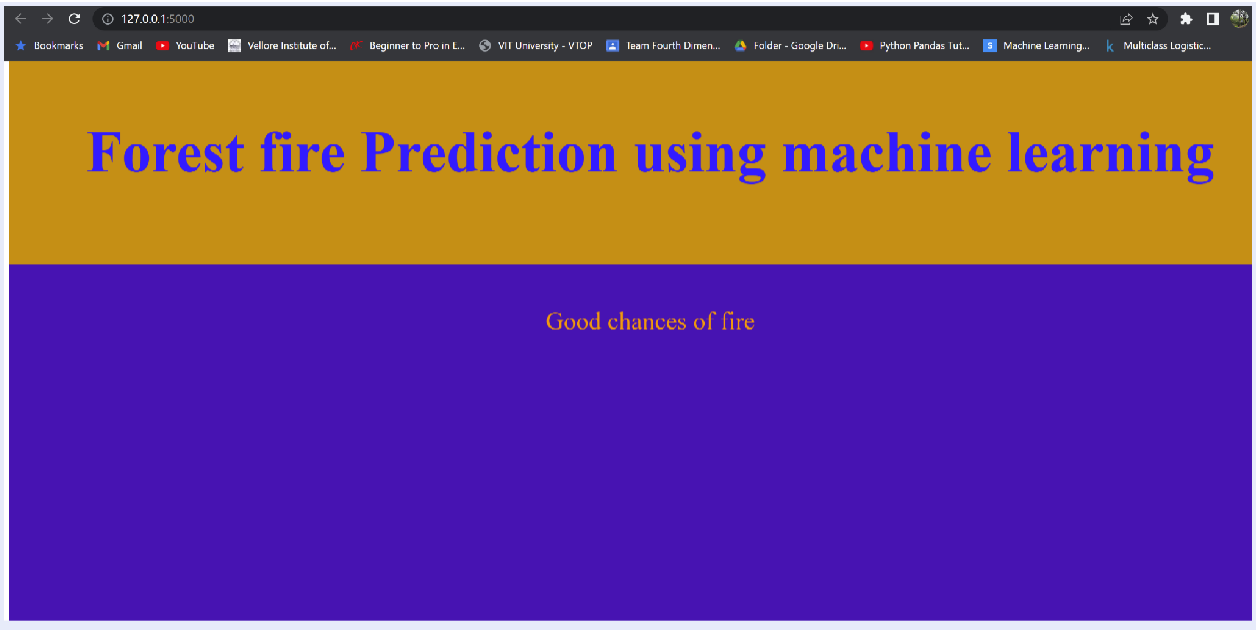
- If the affected area is less than 10 acers , then ‘nil chances of fire’.



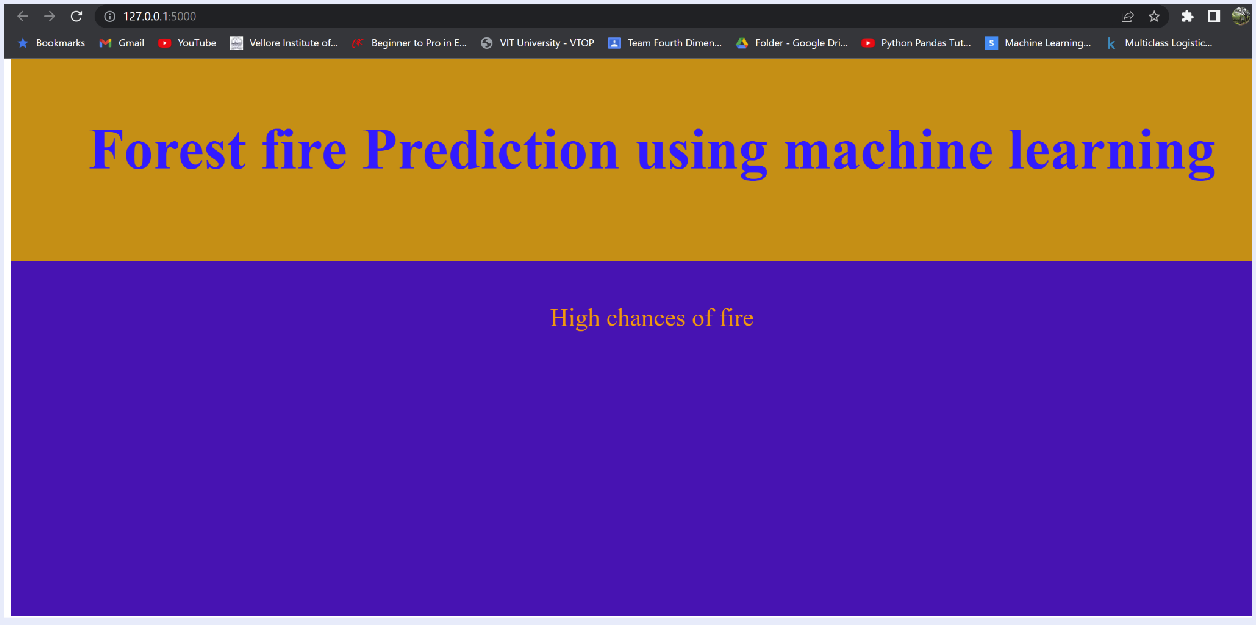
- If the affected area is greater than 10 and less than less than 60 , then ‘low chances of fire’.



- If the affected area is greater than 60 and less than 120 , then ‘Good chances of fire’.

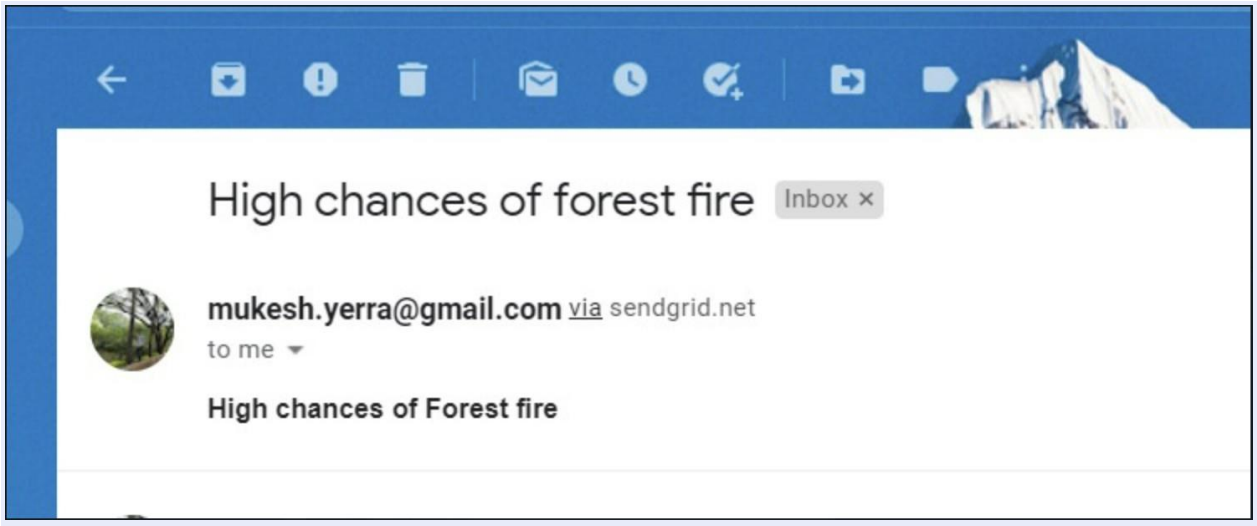


- If the affected area is greater than 120 ,the ‘High chances of fire’.



NOTIFICATIONS

Step 5 : If there are High chances for the forest fire Occurrence , then an alert mailwill be sent to the concerned authorities.



REFERENCES:

1) K Venkata Murali Mohan,Aravapalli Rama Satish,Mallikharjuna Rao K,Rokesh Kumar Yarava,G Charles Babu Leveraging Machine Learning to Predict WildFires.978-1-6654-3368-6/21/$31.00 ©2021 IEEE

2) Divya.T.L, Manjuprasad.B, VijayaLakshmI.N,, Andhe Dharan .An Efficient and Optimal Clustering Algorithm For Real-Time Forest Fire Prediction with Sensor Networks and Data Mining i.978-1-4799-3358-7114/$31.00 ©2014 IEEE

3) Dedi Rosadii ,Widyastuti,Andriyani ,Dina Agustina ,Deasy Arisanty .Prediction of Forest Fire Occurrence in Peatlands using Machine Learning Approaches .978-1-7281-8406-7/20/$31.00 ©2020 IEEE

4) N. Saranya, S. Sahana, B. Suganthi, R. K. Vijaynigilesh, T. Vivin .IOT Based Forest Fire Prediction and Detection,,Retrieval Number: F3914049620/2020©BEIESP .DOI: 10.35940/ijitee.F3914.049620

5) George E. Sakr, Imad H. Elhajj , George Mitri .Efficient forest fire occurrence prediction for developing countries using two weather parameters. b,0952-1976/$ - see front matter & 2011 Elsevier Ltd. All rights reserved. doi:10.1016/j.engappai.2011.02.017

6) A. L. Latifah, A. Shabrina, I. N. Wahyuni and R. Sadikin, "Evaluation of Random Forest model for forest fire prediction based on climatology over Borneo," 2019 International Conference on Computer, Control, Informatics and its Applications (IC3INA), 2019, pp. 4-8, doi: 10.1109/IC3INA48034.2019.8949588.

7) B. K. Singh, N. Kumar and P. Tiwari, "Extreme Learning Machine Approach for Prediction of Forest Fires using Topographical and Metrological Data of Vietnam," 2019 Women Institute of Technology Conference on Electrical and Computer Engineering (WITCON ECE), 2019, pp. 104-112, doi: 10.1109/WITCONECE48374.2019.9092926.