MAJOR PROJECT
MID SEM REPORT
ON
FOREST FIRE PREDICTION SYSTEM BASED ON GEOSPATIAL IMAGE ANALYTICS

BACHELOR OF TECHNOLOGY

in

COMPUTER SCIENCE

Specialization in

Oil And Gas Informatics

Submitted By :

Name	Roll No	Branch
Amit Singh	R970216010	CSE OGI
Ankit Jha	R970216013	CSE OGI
Ankit Sinha	R970216014	CSE OGI
Rahul Raj	R970216050	CSE OGI

Under the guidance of

Dr Ashish Kailashchandra Sharma

Assistant Professor(SS)

School of Computer Science

UNIVERSITY OF PETROLEUM AND ENERGY STUDIES

Dehradun-248007

2019-20

Approved By

(Dr Ashish Kailashchandra Sharma)
Project Guide

(Dr. T P Singh)
Department Head

Project Title

FOREST FIRE PREDICTION SYSTEM BASED ON GEOSPATIAL IMAGE ANALYTICS

Abstract

The forest fires are common events over the Central Himalayan region during the pre-monsoon season (March - June) of every year. Forest fire plays a crucial role in governing the vegetation structure, ecosystem, climate change as well as in atmospheric chemistry. In regional and global scales, the combustion of forest and grassland vegetation releases large volumes of smoke, aerosols, and other chemically active species that significantly influence Earth's radiative budget and atmospheric chemistry, impacting air quality and risks to human health. During the year 2016, massive forest fires have been recorded over the Central Himalayan region of Uttarakhand which continues for several weeks. To study this event we used the multi-satellite observations of aerosols and pollutants during pre-fire, fire and post-fire period over the central Himalayan region. So there are may further consequences for example, many physical, chemical, mineralogical, and biological soil properties can be affected by forest fire. So we are making the forest fire prediction model that will predict the hotspot for fire in the forest with the help of some parameter like Instant Air Temperature(o C), Maximum Air temperature, Soil temperature, Humidity, Wind speed, Atmospheric Pressure, Accumulated Solar Pressure with the help of spatial image analytics.

- **KEYWORD** :- Spatial image analytics, Forest fire, foci, hotspot, risk, Forest fire prediction.

TABLE OF CONTENTS

Contents

LIST OF FIGURES

List of Figures

1 Introduction

The factors that influence fire occurrence in the boreal forest include the properties of the forest vegetation, weather, and ignition agents . Forest fire mostly occurs in Central Himalyan Region that has very large area. Forest fire plays a crucial role in governing the vegetation structure, ecosystem, climate change as well as in atmospheric chemistry. In regional and global scales, the combustion of forest and grassland vegetation releases large volumes of smoke, aerosols, and other chemically active species that significantly influence Earth's radiative budget and atmospheric chemistry, impacting air quality and risks to human health. During the year 2016, massive forest fires have been recorded over the Central Himalayan region of Uttarakhand which continues for several weeks. To study this event we used the multi-satellite observations of aerosols and pollutants during pre-fire, fire and post-fire period over the central Himalayan region. Since there are various post-fire consequences a risk to life and extant infrastructure. These fires can cause extensive economical damage, they also threaten human life. Furthermore, the aftermath of forest fires can have other far-reaching consequences. For example, many physical, chemical, mineralogical, and biological soil properties can be affected by forest fire. Negative effects resulting from high levels burn severity include significant removal of organic matter, deterioration of both soil structure and porosity, considerable loss of nutrients through volatilization, and erosion also the release of hazardous chemicals significantly impacts human health and increases the risk of future diseases. Wildfire smoke is accompanied by high concentrations of carbon dioxide, which can result in consequences such as headache, mental confusion, nausea, disorientation, come, and even death. So we are creating the model that will help us to predict the place where fire may occurred with the help of spatial image analysis. Image analysis is the extraction of meaningful information from images; mainly from digital images by means of digital image processing techniques. Image analysis tasks can be as simple as reading bar coded tags or as sophisticated as identifying a person from their face. So from the image we will perform the image analysis and thus predict the suitable hotspot where fire may occur.

2 Background Study

- In this paper [1] by George E. Sakr it describe a significant component of forest fire management. It plays a major role in resource allocation, mitigation and recovery efforts. This paper presents a description and analysis of forest fire prediction methods based on artificial intelligence. A novel forest fire risk prediction algorithm, based on support vector machines, is presented. The algorithm depends on previous weather conditions in order to predict the fire hazard level of a day. A forest fire risk zone map was constructed using a four-category risk scale. The risk range is from very high to low. The resulting map was found to have a strong correlation with the highly affected fire sites. This method is also dependent on the studied area and the model cannot be generalized to other lands. It is also notable that the use of satellite images and GIS is imperative in the construction of the models.

- In this paper [2] by Prof. K.Angayarkkani Fire plays a vital role in a majority of the forest ecosystems. Forest fires are serious ecological threats that result in deterioration of economy and environment apart from jeopardizing human lives. Thus forest fires need to be detected as early as possible in order to inhibit from being spread. This paper intends to detect forest fires from the forest spatial data. The approach makes use of spatial data mining, image processing and artificial intelligence techniques for the detection of fires. A fuzzy rule base is formed for the detection of fires, from the spatial data with the presence of fires. The digital images from the spatial data are converted to YCbCr color space and then segmented by employing anisotropic diffusion to identify fire regions. Subsequently, a fuzzy set is created with the color space values of the fire regions. Further, fuzzy rules are derived on basis of fuzzy logic reasoning. Extensive experimental assessment on publicly available spatial data illustrated that the proposed approach efficiently detects forest fires.

- In this paper [3] by Daniela Stojanova this study was to learn to predict forest fires in Slovenia using different data mining techniques. We used predictive models based on data from a GIS (geographical information system), the weather prediction model - Aladin and MODIS satellite data.Fire prevention is the first step in reducing the damage caused by fire and estimation of fire movement is very important for successful fire prevention, organization of prevention measures and optimal storage of firefighting resources.An important tool for fire movement estimation is modeling of the relations between the fire threat and the influence factors. Because these factors are more or less geograpicaly determined, these types of models are usually developed within GIS (Geographical Information System). GIS is a computer system capable of capturing, storing,analyzing, and displaying geographically referenced information; that is, data identified according to their location.
- In this paper [4] by Vasanth Iyer the technical basis of Machine learning with Data mining is studied with the evidence collected uniformly over many years and which allow using users perspective in collected evidence. This model helps in probabilistically forecasting fires and help forest department in planing day to day schedules. Using a model to predict future events reliably one needs to collect samples from sensors and select a feature, which does have any particular bias. Another is the unbalanced nature of the problem of the many forest fire events many are of the burnt area is very small and gives skewed distribution. Most of the examples naturally group into batches, which are collected from evidence satellite photography and collaborative reports from national parks departments. The second set of database was collected from the meteorological weather station about several weather observations, which are located very close to the reported fires. Finally, the compiling task is to serve as a filter and provide the user to vary the false alarm rate. We show by regression analysis of the compiled dataset that the forest fire classifier has a minimum false alarm rate when includingtemporal features.

3 Problem Statement

Since Wildfire risk models are useful for wildfire suppression planning activities such as the allocation of initial attack resources among geographic regions. So Fire occurrence models are important for estimating both expected loss and the likelihood of an extreme fire event and we are thus creating the model to reduce the consequences of forest fire problem by predicting the hotspot so that we can predict where the fire will take place with the help of some parameter like like Instant Air Temperature, Maximum Air temperature, Soil temperature, Humidity, Wind speed, Atmospheric Pressure, Accumulated Solar Pressure. This will not only prevent the major loss due to the forest fire but also we can take some initial step so that the fire will minimized after it catches the fire. This will also help the government to monitor through this model and take some crucial action that will manage the after consequence of the fire. By setting some Veterinary Hospitals near the forest and some fire station near the forest.

4 Objectives

We will prepare the respective model which will successfully predict the hotspot of the forest fire using various Machine learning classification algorithm.

5 Methodology

The various methodology used in this project are.

- **STUDY AREA**

Uttarakhand state is situated in the northern part of India and shares an international boundary with China in the north and Nepal in the east. Uttarakhand has an area of 53,483 Km sq. and lies between 2843 N to 3127 N latitude and 7734 E to 8102 E longitude and the recorded forest area in the state is 34,651 Km2,which constitute 64.79 of of its total geographical area covered by the state. The state lies in the western part of the great Himalaya range. Climate and vegetation vary greatly with elevation of the region, from the glaciers at the highest elevations to tropical forests at the lower elevations. Physiographically the state can be divided into three zones, the Himalayas, the Shiwaliks and the Terrain region and has a temperate climate in hilly regions, whereas in plain areas, climate is tropical with temperatures ranging from sub-zero in the higher regions to 43 degrees in the plains [SFR, 2009]. The average rainfall in the state is 1550 mm [SFR, 2009]. The state consists of major occurring forest types such as Tropical Moist Deciduous, Tropical Dry Deciduous, Sub Tropical Pine, Himalayan Moist temperate, Himalayan Dry Temperate, Sub Alpine and Alpine forests.

- **DATA COLLECTION**

We shall collect the data from the naksha website. Then we will view the Toposheet of Uttarakhand and then download the Topographical map of the state.
The collected data is analyzed in ArcGIS tool.

6 Implementation

6.1 Pseudocode

Import necessary libraries

1. matplotlib.pyplot is a collection of command style functions that make matplotlib work like MATLAB. Each pyplot function makes some change to a figure: e.g., creates a figure, creates a plotting area in a figure, plots some lines in a plotting area, decorates the plot with labels, etc. sklearn libraries are implemented for datasets.

Load the data set: Attributes included in the data set are as follows

1. X - x-axis spatial coordinate within the Montesinho park map: 1 to 9.
2. Y - y-axis spatial coordinate within the Montesinho park map: 2 to 9
3. month - month of the year: "jan" to "dec"
4. day - day of the week: "mon" to "sun"
5. FFMC - FFMC index from the FWI system: 18.7 to 96.20
6. DMC - DMC index from the FWI system: 1.1 to 291.3
7. DC - DC index from the FWI system: 7.9 to 860.6
8. ISI - ISI index from the FWI system: 0.0 to 56.10
9. temp - temperature in Celsius degrees: 2.2 to 33.30
10. RH - relative humidity in : 15.0 to 100
11. wind - wind speed in km/h: 0.40 to 9.40
12. rain - outside rain in mm/m2 : 0.0 to 6.4
13. area - the burned area of the forest (in ha): 0.00 to 1090.84

Statistical & Correlation Analysis of dataset is conducted and extract features from the dataset

1. All the values selected as useful attributes in the dataset are extracted.

Plot the Histogram and Scatterplot for the locations and plot the distribution of values for the dataset

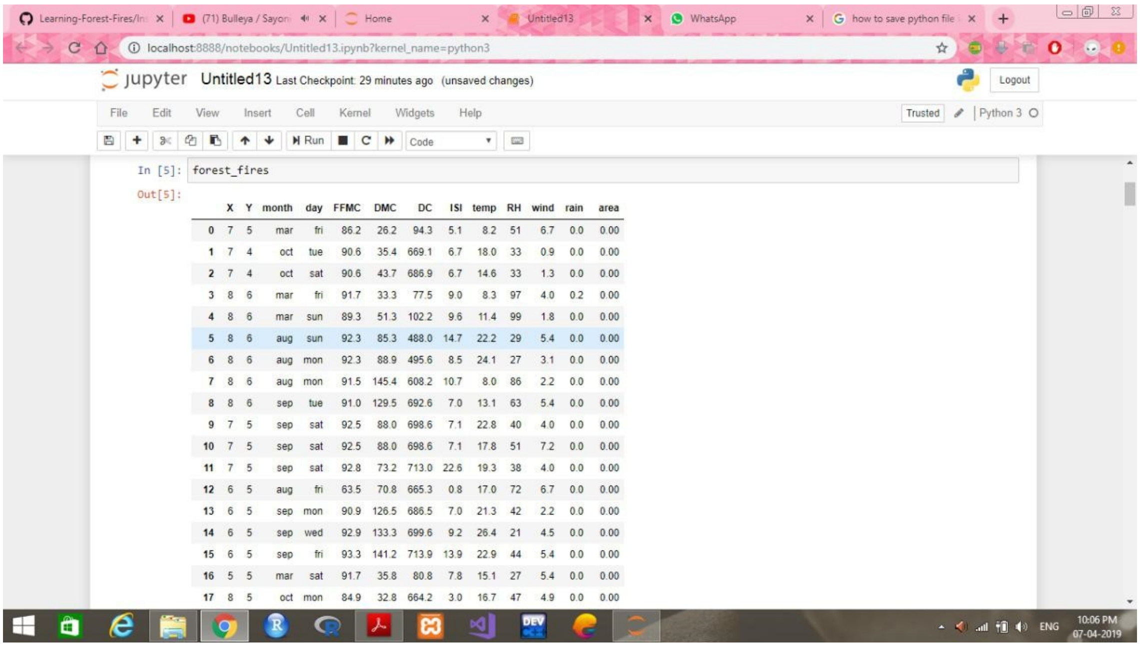
1. Histogram plots showing the distribution of attributes like Month, Day, FFMC, DMC, DC, ISI, Temperature, RH, Wind, Rain, Burned Area are created for insight.

Learning Model Creation

1. The dataset is split into two parts test and train. The actual and predicted outputs are printed on screen. Linear Regression Model is implemented.

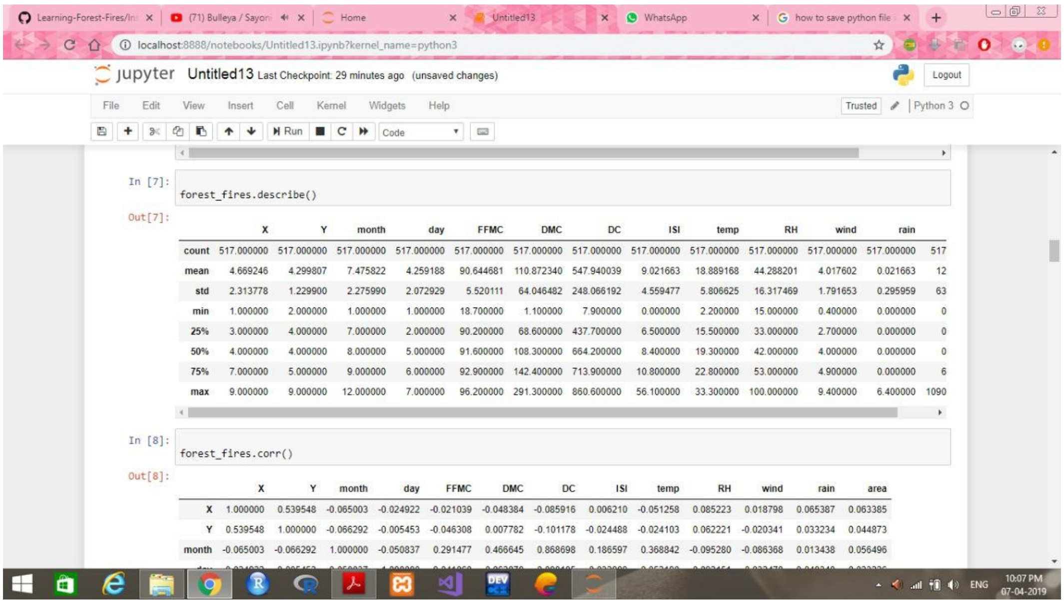
For future, more insight shall be gained by creating more models to find the different Mean squared error and variance score in models such as MLP, SVM, Bayesian Ridge Model, etc. The learning models will be then implemented on the processed data. Normalisation odf the data will be applied and the results will be then visualised in order to make the forest fire prediction model.

6.2 Output Screen



	X	Y	month	day	FFMC	DMC	DC	ISI	temp	RH	wind	rain	area
0	7	5	mar	fri	86.2	26.2	94.3	5.1	8.2	51	6.7	0.0	0.00
1	7	4	oct	tue	90.6	35.4	669.1	6.7	18.0	33	0.9	0.0	0.00
2	7	4	oct	sat	90.6	43.7	686.9	6.7	14.6	33	1.3	0.0	0.00
3	8	6	mar	fri	91.7	33.3	77.5	9.0	8.3	97	4.0	0.2	0.00
4	8	6	mar	sun	89.3	51.3	102.2	9.6	11.4	99	1.8	0.0	0.00
5	8	6	aug	sun	92.3	85.3	488.0	14.7	22.2	29	5.4	0.0	0.00
6	8	6	aug	mon	92.3	88.9	495.6	8.5	24.1	27	3.1	0.0	0.00
7	8	6	aug	mon	91.5	145.4	608.2	10.7	8.0	86	2.2	0.0	0.00
8	8	6	sep	tue	91.0	129.5	692.6	7.0	13.1	63	5.4	0.0	0.00
9	7	5	sep	sat	92.5	88.0	698.6	7.1	22.8	40	4.0	0.0	0.00
10	7	5	sep	sat	92.5	88.0	698.6	7.1	17.8	51	7.2	0.0	0.00
11	7	5	sep	sat	92.8	73.2	713.0	22.6	19.3	38	4.0	0.0	0.00
12	6	5	aug	fri	63.5	70.8	665.3	0.8	17.0	72	6.7	0.0	0.00
13	6	5	sep	mon	90.9	126.5	686.5	7.0	21.3	42	2.2	0.0	0.00
14	6	5	sep	wed	92.9	133.3	699.6	9.2	26.4	21	4.5	0.0	0.00
15	6	5	sep	fri	93.3	141.2	713.9	13.9	22.9	44	5.4	0.0	0.00
16	5	5	mar	sat	91.7	35.8	80.8	7.8	15.1	27	5.4	0.0	0.00
17	8	5	oct	mon	84.9	32.8	664.2	3.0	16.7	47	4.9	0.0	0.00

Figure 1: The Forest Fire dataset



```
In [7]: forest_fires.describe()
```

Out[7]:

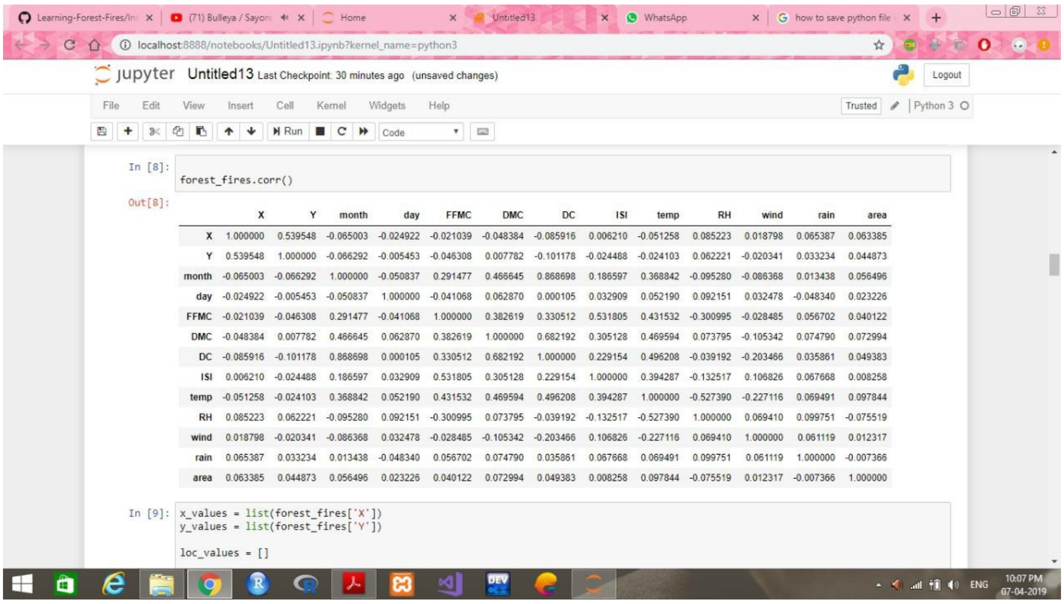
	X	Y	month	day	FFMC	DMC	DC	ISI	temp	RH	wind	rain	
count	517.000000	517.000000	517.000000	517.000000	517.000000	517.000000	517.000000	517.000000	517.000000	517.000000	517.000000	517.000000	517
mean	4.669246	4.299807	7.475822	4.259188	90.644681	110.872340	547.940039	9.021663	18.889168	44.288201	4.017602	0.021663	12
std	2.313778	1.229900	2.275990	2.072929	5.520111	64.046482	248.066192	4.559477	5.806625	16.317469	1.791653	0.295959	63
min	1.000000	2.000000	1.000000	1.000000	18.700000	1.100000	7.900000	0.000000	2.200000	15.000000	0.400000	0.000000	0
25%	3.000000	4.000000	7.000000	2.000000	90.200000	68.600000	437.700000	6.500000	15.500000	33.000000	2.700000	0.000000	0
50%	4.000000	4.000000	8.000000	5.000000	91.600000	108.300000	664.200000	8.400000	19.300000	42.000000	4.000000	0.000000	0
75%	7.000000	5.000000	9.000000	6.000000	92.900000	142.400000	713.900000	10.800000	22.800000	53.000000	4.900000	0.000000	6
max	9.000000	9.000000	12.000000	7.000000	96.200000	291.300000	860.600000	56.100000	33.300000	100.000000	9.400000	6.400000	1090

```
In [8]: forest_fires.corr()
```

Out[8]:

	X	Y	month	day	FFMC	DMC	DC	ISI	temp	RH	wind	rain	area
X	1.000000	0.539548	-0.065003	-0.024922	-0.021039	-0.048384	-0.085916	0.006210	-0.051258	0.085223	0.018798	0.065387	0.063385
Y	0.539548	1.000000	-0.066292	-0.005453	-0.046308	0.007782	-0.101178	-0.024488	-0.024103	0.062221	-0.020341	0.033234	0.044873
month	-0.065003	-0.066292	1.000000	-0.050837	0.291477	0.466645	0.868698	0.186597	0.368842	-0.095280	-0.086368	0.013438	0.056496

Figure 2: Statistical Analysis



```
In [8]: forest_fires.corr()
```

Out[8]:

	X	Y	month	day	FFMC	DMC	DC	ISI	temp	RH	wind	rain	area
X	1.000000	0.539548	-0.065003	-0.024922	-0.021039	-0.048384	-0.085916	0.006210	-0.051258	0.085223	0.018798	0.065387	0.063385
Y	0.539548	1.000000	-0.066292	-0.005453	-0.046308	0.007782	-0.101178	-0.024488	-0.024103	0.062221	-0.020341	0.033234	0.044873
month	-0.065003	-0.066292	1.000000	-0.050837	0.291477	0.466645	0.868698	0.186597	0.368842	-0.095280	-0.086368	0.013438	0.056496
day	-0.024922	-0.005453	-0.050837	1.000000	-0.041068	0.062870	0.000105	0.032909	0.052190	0.092151	0.032478	-0.048340	0.023226
FFMC	-0.021039	-0.046308	0.291477	-0.041068	1.000000	0.382619	0.330512	0.531805	0.431532	-0.300995	-0.028485	0.056702	0.040122
DMC	-0.048384	0.007782	0.466645	0.062870	0.382619	1.000000	0.682192	0.305128	0.469594	0.073795	-0.105342	0.074790	0.072994
DC	-0.085916	-0.101178	0.868698	0.000105	0.330512	0.682192	1.000000	0.229154	0.496208	-0.039192	-0.203466	0.035861	0.049383
ISI	0.006210	-0.024488	0.186597	0.032909	0.531805	0.305128	0.229154	1.000000	0.394287	-0.132517	0.106826	0.067668	0.008258
temp	-0.051258	-0.024103	0.368842	0.052190	0.431532	0.469594	0.496208	0.394287	1.000000	-0.527390	-0.227116	0.069491	0.097844
RH	0.085223	0.062221	-0.095280	0.092151	-0.300995	0.073795	-0.039192	-0.132517	-0.527390	1.000000	0.069410	0.099751	-0.075519
wind	0.018798	-0.020341	-0.086368	0.032478	-0.028485	-0.105342	-0.203466	0.106826	-0.227116	0.069410	1.000000	0.061119	0.012317
rain	0.065387	0.033234	0.013438	-0.048340	0.056702	0.074790	0.035861	0.067668	0.069491	0.099751	0.061119	1.000000	-0.007366
area	0.063385	0.044873	0.056496	0.023226	0.040122	0.072994	0.049383	0.008258	0.097844	-0.075519	0.012317	-0.007366	1.000000

```
In [9]: x_values = list(forest_fires['X'])
        y_values = list(forest_fires['Y'])

        loc_values = []
```

Figure 3: Correlation Analysis

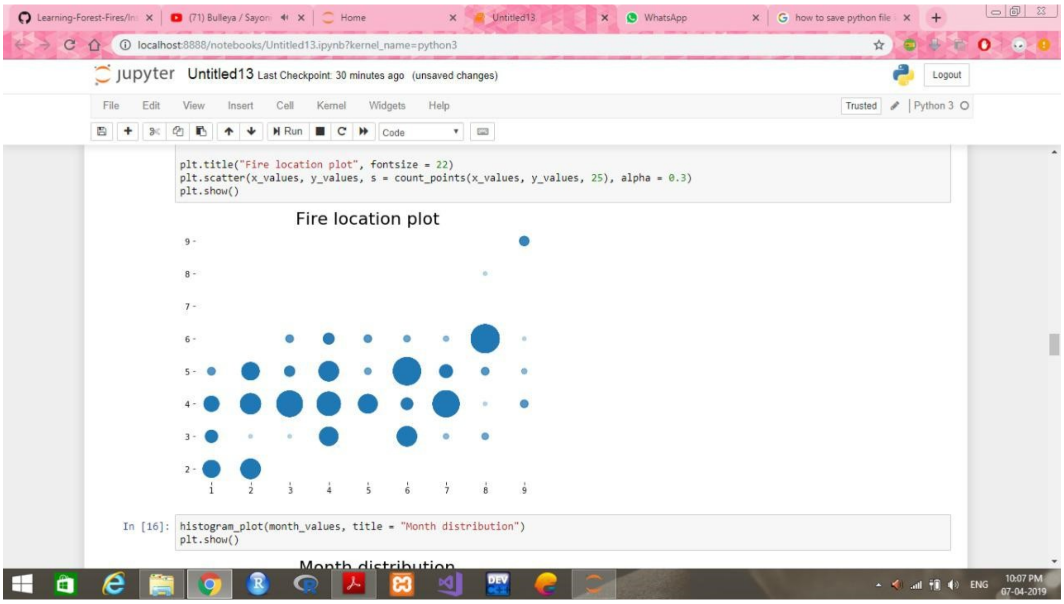


Figure 4: Scatterplot

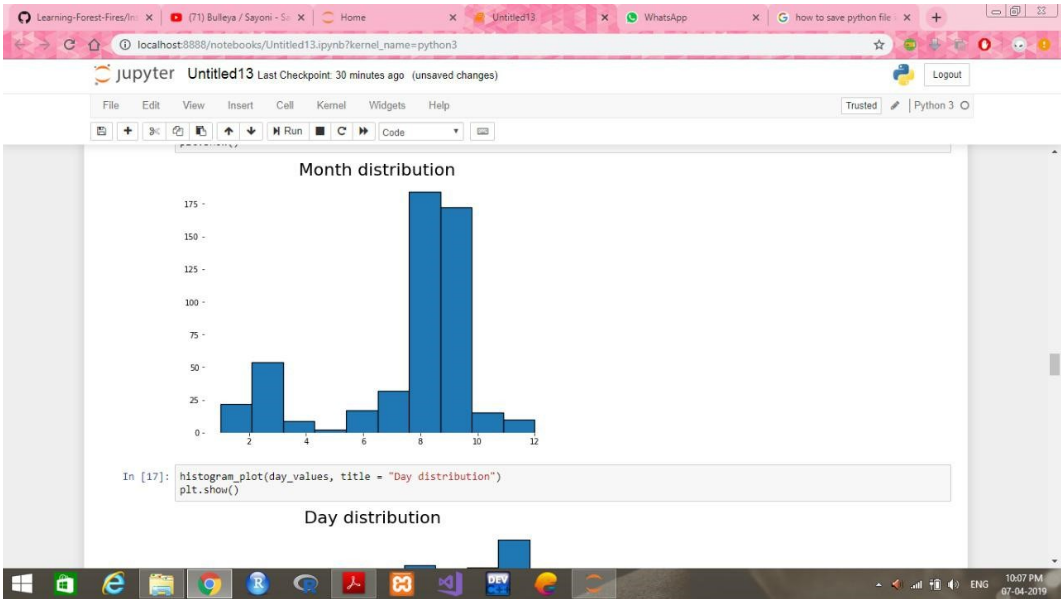


Figure 5: Monthly Distribution Histogram

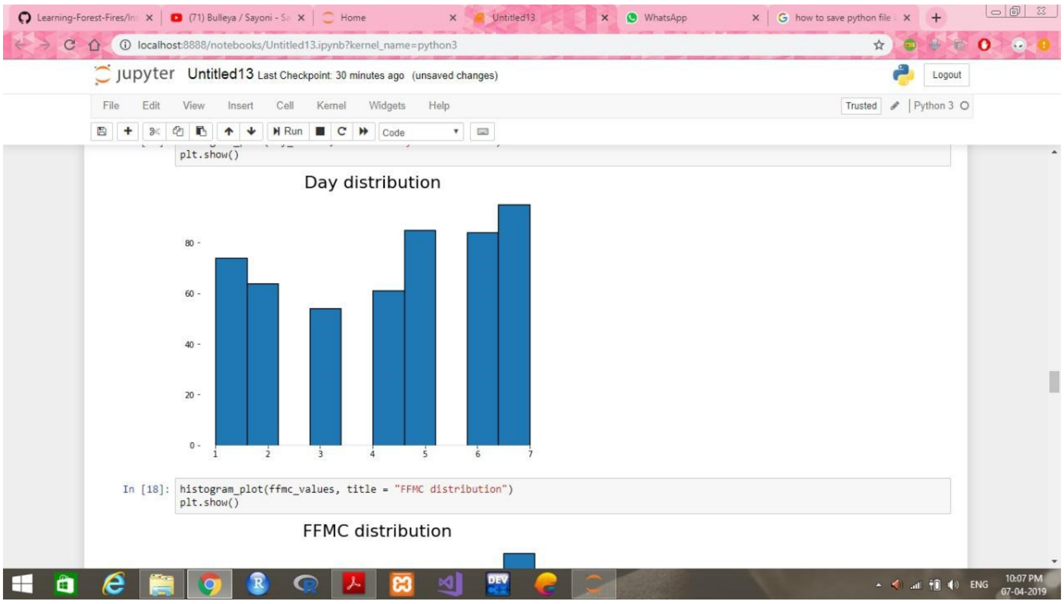


Figure 6: Day Distribution Histogram

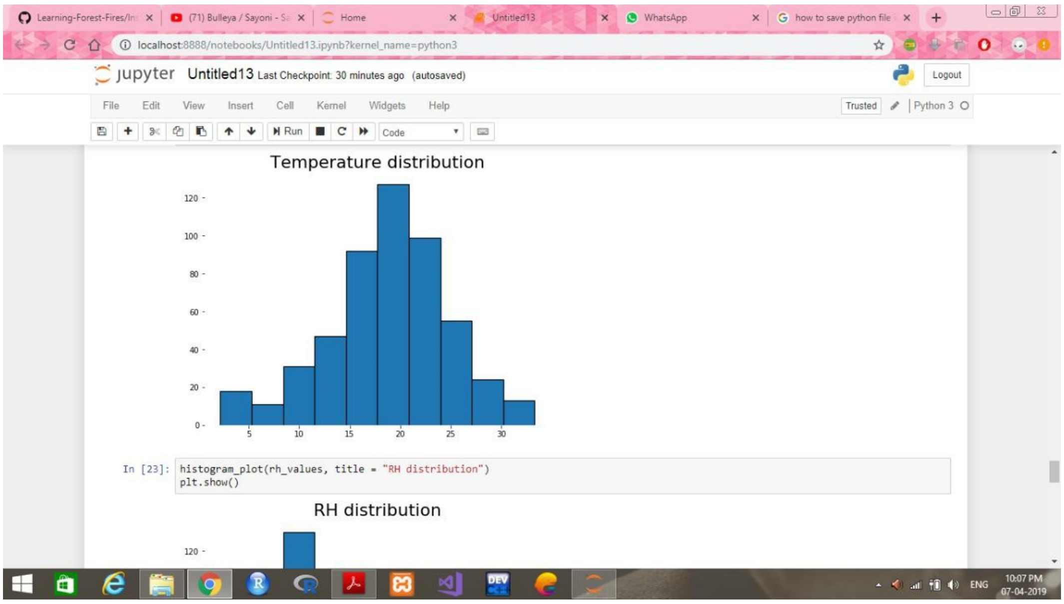


Figure 7: Temperature Distribution Histogram

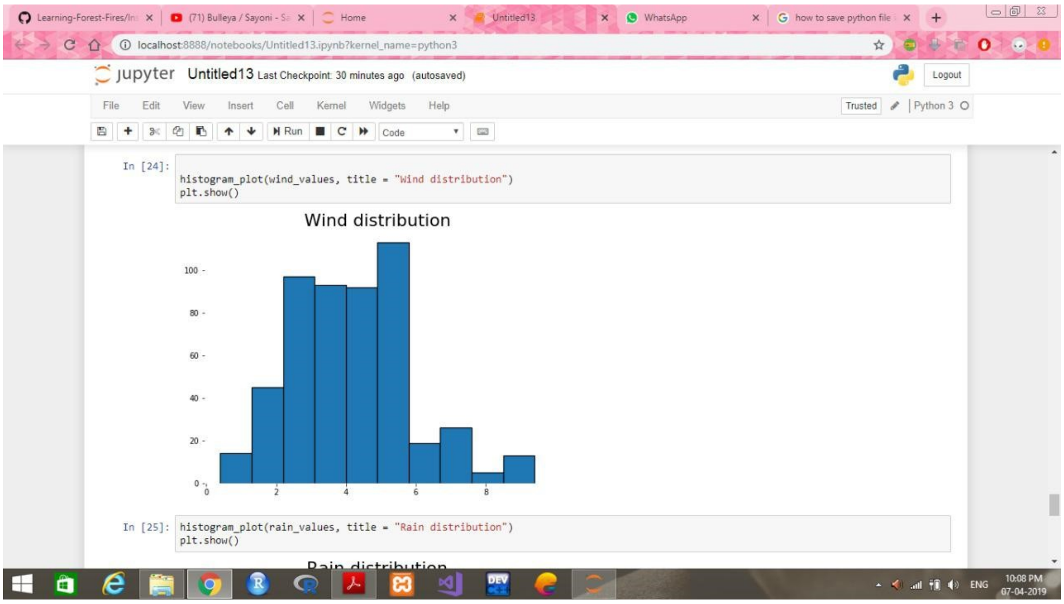


Figure 8: Wind Distribution Histogram

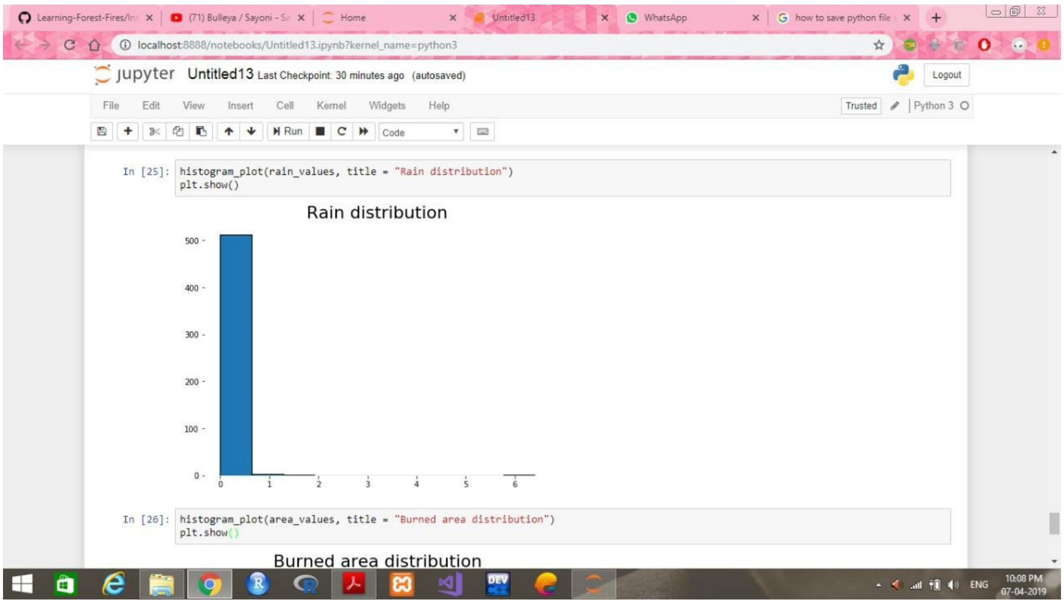


Figure 9: Rain Distribution Histogram

References

[1] G. E. Sakr, I. H. Elhajj, G. Mitri, and U. C. Wejinya, "Artificial intelligence for forest fire prediction," *IEEE/ASME International Conference on Advanced Intelligent Mechatronics, AIM*, no. September 2015, pp. 1311–1316, 2010.

[2] P. K. Angayarkkani, "Efficient Forest Fire Detection System : A Spatial Data Mining and Image Processing Based Approach," *Journal of Computer Science*, vol. 9, no. 3, pp. 100–107, 2009.

[3] V. Iyer, S. S. Iyengar, N. Paramesh, G. R. Murthy, and M. B. Srinivas, "Machine learning and datamining algorithms for predicting accidental small forest fires," *SENSORCOMM 2011 - 5th International Conference on Sensor Technologies and Applications and WSNSCM 2011, 1st International Workshop on Sensor Networks for Supply Chain Management*, no. c, pp. 116–121, 2011.

[4] D. Stojanova, A. Kobler, S. Džeroski, and K. Taškova, "Learning To Predict Forest Fires," *In 9th International multiconference Information Society*, no. March 2016, pp. 3–6, 2006.