**Mini-Project – 2B Web based on ML**

**(ITM 601)**

**FOREST FIRE PREDICTION SYSTEM**

**T. E. Information Technology**

By

**HIMANSHU CHUARASIYA 45**

**VIKAS CHAURASIYA 46**

**MUKESH GUPTA 47**

**ASHLY JOHN 48**

Under Guidance of

**Dr VAISHALI JADHAV**

Associate Professor



Department of Information Technology

St. Francis Institute of Technology

(Engineering College)

University of Mumbai

2021-2022

**CERTIFICATE**

This is to certify that the project entitled **“Forest Fire Prediction System”** is a bonafide work of “ **Himanshu Chaurasiya, Vikas Chaurasiya, Mukesh Gupta, Ashly John Roll no 45,46,47,48”** submitted to the University of Mumbai towards completion of mini project work for the subject of **Mini Project – 2B Web Based on ML (ITM 601).**

**Dr.Vaishali Jadhav**

**Guide**

**Dr. Joanne Gomes**

**HOD-IT**

Examiners

1.---------------------------------------------

2.---------------------------------------------

Date:

**DECLARATION**

We declare that this written submission represents our ideas in our own words and where others' ideas or words have been included, we have adequately cited and referenced the original sources. We also declare that we have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in our submission. We understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

--------------------------------------- ---------------------------------------

(Name of student and Signature) (Name of student and Signature)

--------------------------------------- ---------------------------------------

(Name of student and Signature) (Name of student and Signature)

**ABSTRACT**

Forest wildfires are becoming more common all around the world. In the previous five decades, the frequency of large wildfires has increased by more than four and six times, respectively. For resource allocation, mitigation, and recovery, forest fire forecasting is critical. Our report presents a description and analysis of forest fire prediction methods based on Machine Learning Techniques. By using machine learning techniques and algorithms such as Linear Regression, Random Forest ,Logistic Regression, and Multi Layer Perceptron, we propose a system to predict if our forest is safe or not based on various climatics parameters entered by the user in the front end.

**INDEX**

| **Chapter**  **No.** | **Contents** | **Page No.** |
| --- | --- | --- |
| **1** | **Introduction** | 9 |
|  | **1.1 Background** | 9 |
|  | **1.2 Scope of the project** | 9 |
|  | **1.3 Objectives and Problem Statement** | 9 |
| **2** | **Literature Review** | 10-11 |
| **3** | **Proposed Work** | 12-13 |
|  | **3.1 Architectural Details** | 12 |
|  | **3.1.1 Data Collection** | 13 |
|  | **3.1.2 Dataset Preprocessing** | 13 |
|  | **3.1.3 Prediction Algorithm** | 13 |
|  | **3.1.4 Testing and Evaluation** | 13 |
| **4** | **Implementation** | 14-21 |
|  | **4.1 Dataset Details** | 14-15 |
|  | **4.2 Algorithm Details** | 16-18 |
|  | **4.3 Web based Project details** | 19-20 |
|  | **4.4 Screenshots of GUI with Explanation** | 20 |
|  | **4.5 Performance Metrics Details** | 21 |
| **5** | **Results and Discussions** | 22-23 |
| **6** | **Conclusion and Future Scope** | 24 |
| **7** | **References** | 25 |
| **8** | **Acknowledgement** | 26 |

**List of Abbreviations**

| **Sr. No.** | **Abbreviation** | **Full Form** |
| --- | --- | --- |
| 1 | RH | Relative Humidity |
| 2 | FFMC | Fine Fuel Moisture Code |
| 3 | DMC | Duff Moisture Code |
| 4 | DC | Drought code |
| 5 | ISI | Initial Spread Index |
| 6 | BUI | Build Up index |
| 7 | FWI | Fire Weather Index |

**List of Figures**

| **Fig. No.** | **Figure Name** | **Page No.** |
| --- | --- | --- |
| 1.1 | Architectural Details | 12 |
| 1.2 | Dataset Details | 15 |
| 1.3 | Linear regression | 16 |
| 1.4 | Multi layer Perceptron | 17 |
| 1.5 | Graphical representation of sigmoid function | 17 |
| 1.6 | Overview of the two decision trees that constitute a Random forest | 18 |
| 1.7 | GUI Interface | 20 |
| 1.8 | GUI Interface | 20 |
| 1.9 | Accuracy of random forest regressor | 22 |
| 2.0 | Accuracy of Logistic regression | 23 |
| 2.1 | Accuracy of MLP classifier | 23 |

**List of Tables**

| **Table. No.** | **Table Name** | **Page No.** |
| --- | --- | --- |
| 1 | Literature Review | 10 |
| **2** | Performance of regression task | 21 |
| **3** | Performance of classification task | 21 |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

**Chapter 1**

Introduction

* 1. Background

Forest fires have recently become one of the most regular natural disasters, resulting in the devastation of thousands of hectares of forest. They endanger not only forest resources, but also the entire ecosystem, including fauna and plants, disrupting biodiversity, ecosystems, and the ecology of a region. It is packed with dried senescent leaven and dwelling that can burst into flames triggered by a single spark in the summer when there has been no rain for months. Forecasting forest fires is an essential part of forest fire management. control. It has a big influence on how resources are allocated.

1.2 Scope of the project

A unique forest fire prediction approach is created using linear regression, random forest regression, logistic regression, and MLP Classifier. Weather parameters are used by the algorithm to predict the severity of the fire hazard. The technique is used to painstakingly evaluate the threat of a forest fire incident using data from a dataset.

# 1.3 Objectives and Problem Statement

# Our Project aims to build a forest fire prediction system which will predict whether your forest is safe or not by taking in consideration various climatic values .

● To create a machine learning model.

● To predict forest fires.

● To create a user interface for predicting forest fires and display an alert

message .

**Chapter 2**

Literature Review

| **Sr No** | **Title** | **Methodology** | **Future/Scope** |
| --- | --- | --- | --- |
| 1 | Forest Fire Risk Zone Modeling Using Logistic Regression and GIS | In this research paper demonstrates that logistic regression modeling and geographic information system are suitable for determining the forest fire risk zone, and therefore management of it.  The analysis has revealed that the elevation, slope, annual precipitation and distance to farms have high significant correlation with fires | The logistic regression method combined with GIS and inventory data is useful for fire risk mapping, because each factor influencing forest fire risk is analyzed in this method. This modeling and mapping provides valuable information about areas most likely to be affected by fire. Forest fire risk zone mapping is a useful tool in forest fire prevention and management in order to minimize wildfire risk and damage, allowing forest fire managers to identify high fire risk locations easily and manage these areas effectively |
| 2. | Forest fire prediction using ML and AI. | In this research paper different prediction algorithms are used to predict the output like Logistic Regression,Random Forest etc.the system here  predict the percentage of fire occurrence based on various  parameters such as temperature, humidity, and oxygen data  entered by the user in the front end | Quick fire detection and reaction are efficient methods for decreasing fire damage. Various studies are conducted in try to improve early fire prediction and detection systems, which aid within the development of  fireside response methods. It signifies that early caution detection is provided. The accurate prediction of results supported defined parameters is expounded to early caution detection. |
| 3 | Prediction of forest fires using Artificial neural networks. | The research paper builds a prediction system which uses artificial neural networks. Multilayer perceptron is used for prediction.The system is object oriented software coded in C++ language under a Linux environment. | The adaptation and use of the same approach to other kind of prediction problems, encountered in other fields such as economics and social sciences |
| 4 | Predicting Forest Fires With Different Data Mining Techniques. | This paper suggests four different types of data mining techniques which can also be implemented for the prediction of forest fires like Logistic Regression, Support Vector Machine, Decision Trees, Random Forest. | The FMS system is widely used around the world, further research is need to confirm if direct weather conditions are preferable than  accumulated values, as suggested . |

**Chapter 3**

Proposed Work

3.1 Architectural Details

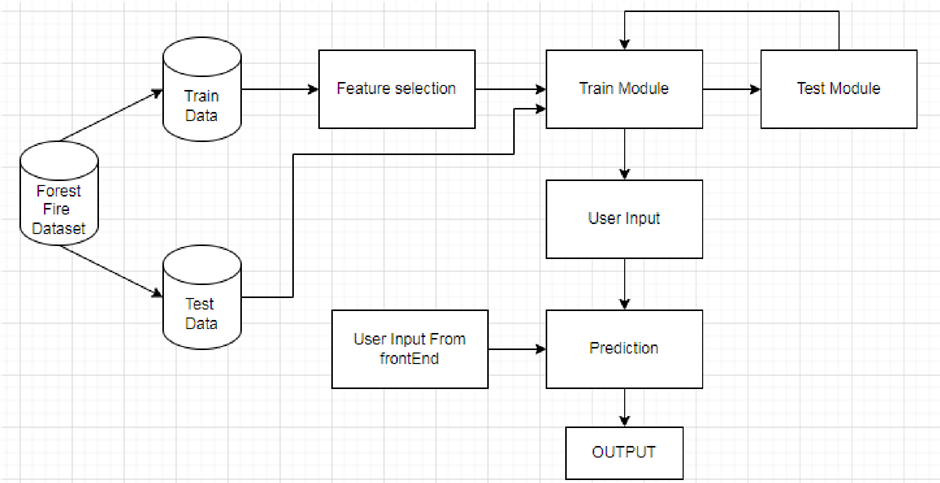


figure 1

As shown in the fig 1.1 the forest fire dataset is split into two dataset train dataset and test data . Feature Selection is the process where you automatically or manually select those features which contribute most to your prediction variable or output in which you are interested in. Feature selection is done on the train data in which we have selected the variables which we are going to use. After which we trained the module and tested the module.Then we applied the appropriate prediction algorithm to predict our output which is the probability of fire occurrence. In the GUI the user would have to input various variables from the frontend on which prediction algorithm would be applied and we would get the appropriate output.

### 

### 3.1.1 Data Collection

Kaggle Forest Fire Dataset

<https://www.kaggle.com/datasets/elikplim/forest-fires-data-set>

The dataset used is a public dataset which is available for research purposes. This dataset was created by Paulo Cortez and Aníbal Morais (Univ. Minho) @2007. The forest fire dataset is a multivariate dataset which is a data set consisting of two or more than two variables is referred to as a multivariate dataset. It has 13 attributes having 517 instances. The total file size f the kaggle Forest Fire dataset is 7KB

3.1.2 Dataset Preprocessing

Data preprocessing is a process of preparing the raw data and making it suitable for a machine learning model. It is the first and crucial step while creating a machine learning model. In the Forest Fire dataset there are a total of 517 instances with 13 attributes. All the 517 instances are not null instances. Out of the 13 attributes 9 attributes are float data type, 2 are of integer data type and 2 Object data type.

3.1.3 Prediction Algorithm

In our model we have tried to use mainly two types of predictive algorithm regression and classification. Algorithms which were applied are linear regression, random Forest Regressor, LogisticRegression and MLP classifier. Depending on the performance of the algorithm we have chosen the appropriate algorithm which in our case was MLP classifier. MLPClassifier stands for Multi-layer Perceptron classifier.

3.1.4 Testing and Evaluation

We have split the train and test model into a ratio of 60:40 and then we have applied the algorithm. When we applied the algorithm the r2 score for the regression task was very low. The classification task also showed a minimum accuracy but the accuracy was better with MLP classifier. In the web based implementation we have used the MLP classifier as the main algorithm.

**Chapter 4**

Implementation

## 4.1 Dataset Details

The forest fire dataset is a multivariate dataset which is a data set consisting of two or more than two variables is referred to as a multivariate dataset. It has 13 attributes having 517 instances. The attributes are explained as follows:

1. X - x-axis spatial coordinate within the Montesinho park map: 1 to 9 where the fire occurred.

2. Y - y-axis spatial coordinate within the Montesinho park map: 2 to 9 where the fire occurred.

3. Month - month: 'Jan' to 'Dec' of the year

4. Day - day: 'mon' to 'sun' of the week

5. FFMC - FFMC index from the FWI system: 18.7 to 96.20. It is the Fine Fuel Moisture Code which denotes the moisture content of surface litter and influences ignition and fire spread.

6. DMC - DMC index from the FWI system: 1.1 to 291.3. It is the Duff Moisture Code which represents the moisture content of shallow and deep organic layers, which affect fire intensity.

7. DC-DC index from the FWI system: 7.9 to 860.6. It is the Drought Code that represents the moisture content of shallow and deep organic layers, which affect fire intensity.

8. ISI - ISI index from the FWI system: 0.0 to 56.10. It is the Initial Spread Index that determines and correlates it with fire velocity spread.

9. Temp - Temperature in Celsius degrees: 2.2 to 33.30 of the area.

10. RH - relative humidity in %: 15.0 to 100 in the air.

11. Wind - wind speed in km/h: 0.40 to 9.40 at the time of fire.

12. Rain - outside rain in mm/m2: 0.0 to 6.4.

13. Area - The forest area (in ha): 0.00 to 1090.84 that burned during the forest fire.

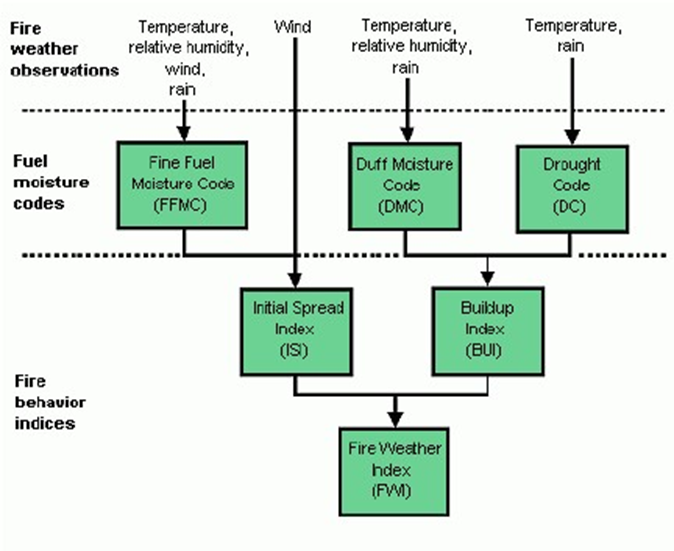
****

figure 2 : Dataset Details

### 

### 4.2 Algorithm Details

**Linear Regression:** Linear Regression is a machine learning algorithm based on supervised learning. It performs a regression task. Regression models a target prediction value based on independent variables. It is mostly used for finding out the relationship between variables and forecasting. Different regression models differ based on – the kind of relationship between dependent and independent variables they are considering, and the number of independent variables getting used.

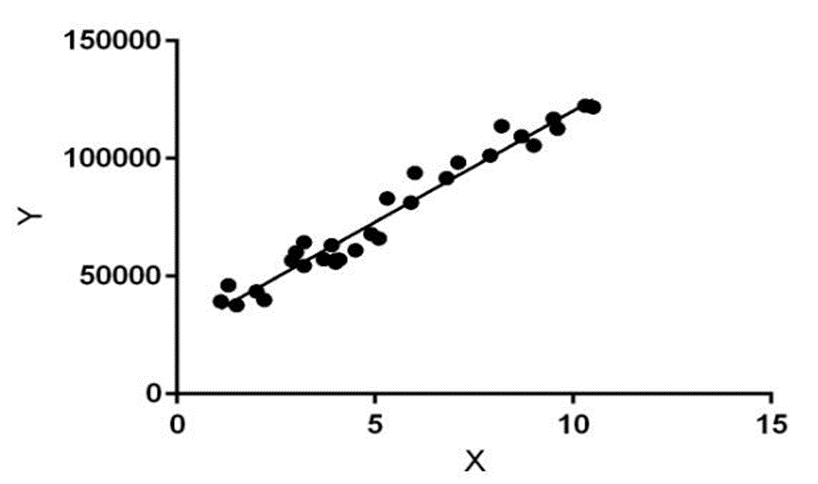


figure 3 : linear regression

Linear regression performs the task to predict a dependent variable value (y) based on a given independent variable (x). So, this regression technique finds out a linear relationship between x (input) and y(output). Hence, the name is Linear Regression.

**Multi-layer Perceptron:** Multi-layer perceptron is also known as MLP. It is fully connected dense layers, which transform any input dimension to the desired dimension. A multi-layer perception is a neural network that has multiple layers. To create a neural network we combine neurons together so that the outputs of some neurons are inputs of other neurons.

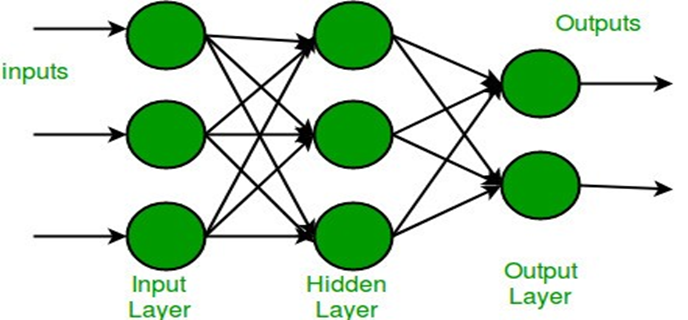
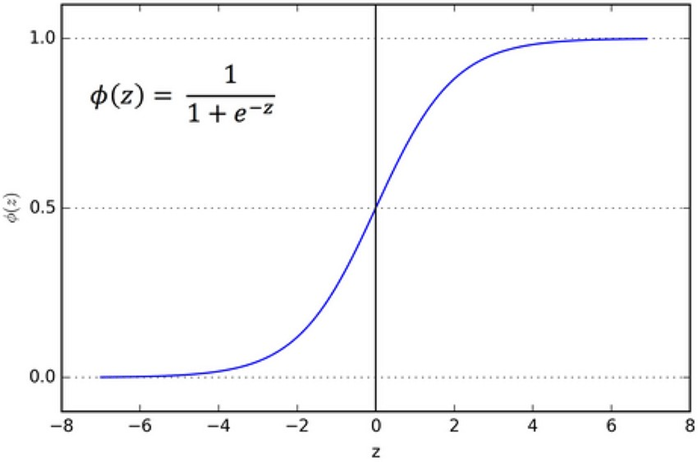


figure 4 : Multi layer Perceptron

**Logistic regression:** Logistic regression is a statistical method for predicting binary classes. The outcome or target variable is binary in nature. For example, it can be used for cancer detection problems. It computes the probability of an event occurrence. It is a form of binomial regression. The relationship between the dependent variable and the independent variable helps it to predict the target variable. To determine their probability and map them to some discrete values, the logistic regression uses the sigmoid function. The sigmoid function is as follows:-



### figure 5 : Graphical representation of sigmoid function

**Random forest:** Random forest, like its name implies, consists of a large number of individual decision trees that operate as an ensemble. Each individual tree in the random forest spits out a class prediction and the class with the most votes becomes our model’s prediction. The advantage of random forest over decision tree is that they correct the over-fitting nature of the decision trees.

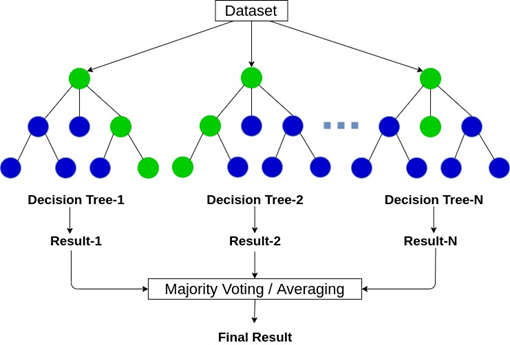


figure 6 : Overview of the two decision trees that constitute a Random forest

### 

### 4.3 Web based Project details

* For the Implementation of this Web based project we have created a form.Where the form tag starts inside action We’ve written ‘/predict’ which says that whenever we click the submit button the data send to the given URL and the form data send to the POST request.
* There are 8 input boxes such as FFMC, DMC, DC, ISI, Temperature,Relative humidity , Wind and Rain. All of these are input text types and I've created a submit button. Whenever I click the button the entire form data is going to be submitted to the URL of ‘/predict’.
* Inside the prediction program we have to import numpy,pandas,MLPClassifier, and joblib.Then we are reading our CSV file and dropping some columns which are not required like x-axis, y-axis, month, area and day.
* After this we have split our dataset train and test data and the size of train data 60% and test is 40% and calling the MLC classifier and giving the training set as the input.We don't want this entire python program run again and again whenever using the website so we will train our model once and importing that module somewhere else from where we can directly fetch it so that pkl\_file comes in.
* When we come to our main program which is app.py in which we have imported flask , numpy and the joblib library. After that we have created a machine learning model and store into the forestfiremodel.pkl in the variable called model.When the default function gets triggered which is hello world and open the HTML file called ‘index.html’.
* Next we have defined what happens when we get the form data? The form data is sent to the URL called ‘/predict’. Inside the predict we have defined how to get data from the form so we have created a list for features after taking the input then converting that list into a numpy array and storing the variable called final.
* After that we have created a variable called output which formats the prediction value and its print up to two decimal values so now if the output is greater than 0.5 then it will send the alert that your forest is in danger if less than your forest is safe.

### 

### 4.4 Screenshots of GUI with Explanation

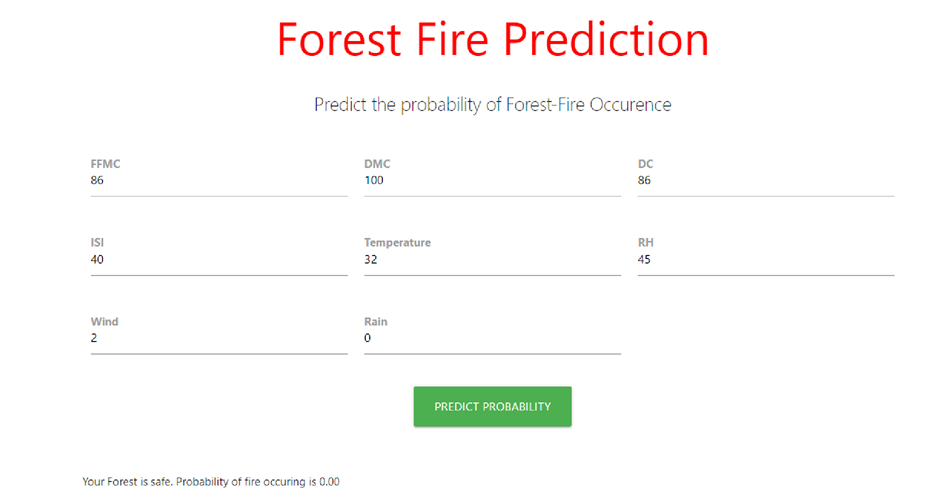


figure 7 : GUI Interface

In the above fig there are 8 input boxes which are FFMC, DMC,DC,ISI, Temperature , RH, Wind and Rain and we have entered their value in the given range. There is also a button called Predict Probability.When we click on the button then in the background MLP algorithm runs and it predicts the probability of fire occurring.In the above image our forest is safe. Probability of a fire occurring is 0.00.



figure 8 : GUI Interface

This is a footer of our ML Project in which we have introduced us and group member names.

4.5 Performance Metrics Details

### Performance of regression task

| Sr no | Algorithm | R2 Score |
| --- | --- | --- |
| 1 | Linear regression | 0.0137 |
| 2 | Random Forest | -61.62 |

Performance of Classification task

| Sr no | Algorithm | Accuracy |
| --- | --- | --- |
| 1 | Logistic regression | 53.21 % |
| 2 | MLP classifier | 62.38 % |

**Chapter 5**

Results and Discussions

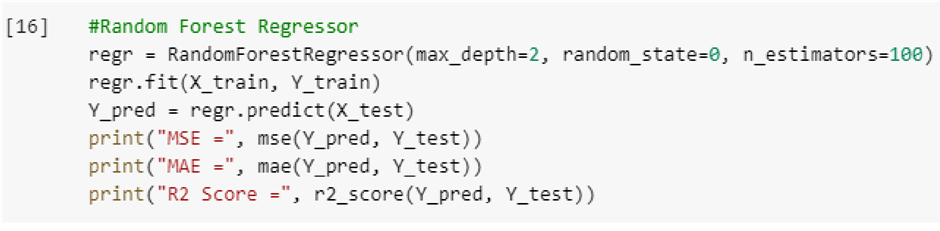


figure 9 : Accuracy of random forest regressor

In the above image we can see that we have trained and tested the model with a random forest regressor algorithm. To check its accuracy we have to use mean absolute error, mean square error and r2 score. we can see that mse is very high ,mae and is lower and the r2 score is negative. r2 score should be between 0 to 1 which is preferred and the mse and mae score should be closer to zero. So random forest cant be performed for prediction.

**Formulas:**

****

Mean Absolute Error = (1/n) \* ∑|y i – x i |

where,

● Σ: Greek symbol for summation

● y i: Actual value for the ith observation

● x i: Calculated value for the ith observation

● n: Total number of observations

MSE = (1/n) \* Σ(actual – forecast)2

where:

●Σ – a symbol that means “sum”

●n – sample size

●actual – the actual data value

●forecast – the predicted data value

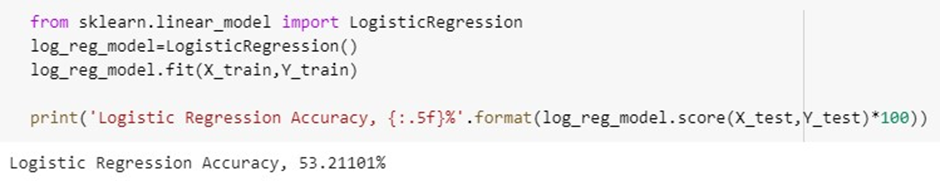


figure 10 : Accuracy of logistic regression

The output got for for logistic regression model which is a classification task is 53 %

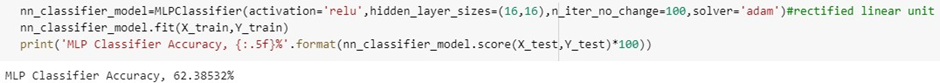


figure 11 : Accuracy of MLP classifier

The output of the mlp classifier, which is a classification algorithm, may be seen in the image above. It produced a result of 62%, which is considered better than 53 %.

**Chapter 6**

Conclusion and Future Scope

Although the particular cause or spark for a wildfire occurrence cannot always be predicted, wildfire risk can be anticipated to some extent by understanding how climate, geography, weather, and land cover influence fire behavior and the likelihood of a fire spreading.

Forest fires cause significant environmental damage while threatening human lives. A substantial effort was made in the last two decades to build automatic detection tools that could assist Fire Management Systems (FMS). The three major trends are the use of satellite data, infrared/smoke scanners and local sensors (e.g.meteorological).

However, this work opens room for the development of automatic tools for fire management support. Since the FMS system is widely used around the world, further research is need to confirm if direct weather conditions are preferable than accumulated values as suggested by this study.

**Chapter 7**

References

[1] Elshewey, Ahmed & Elsonbaty, Amira. (2020). Forest Fires Detection Using Machine Learning Techniques. Xi'an Jianzhu Keji Daxue Xuebao/Journal of Xi'an University of Architecture & Technology. XII. 2020.

[2] Adithi M. Shrouthy, Syed Matheen Pasha, Yamini S. R. E., Navya Shree S., Lisha U.. Forest fire prediction using ML and AI, International Journal of Advance Research, Ideas and Innovations in Technology, [www.IJARIIT.com](https://www.ijariit.com/)

[3] Safi, Youssef & Bouroumi, A.. (2013). Prediction of forest fires using Artificial neural networks. Applied Mathematical Sciences. 7. 271-286.10.12988/ams.2013.13025.

[4] Madhurima De,Linika Labdhi,Bindu Garg..PREDICTING FOREST FIRES WITH DIFFERENT DATA MINING TECHNIQUES,International Journal of Scientific Development and Research (IJSDR) www.ijsdr.org

[5][https://www.sciencedirect.com/science/article/pii/S2666222121000046#tbl01](https://www.sciencedirect.com/science/article/pii/S2666222121000046#tbl0001)

[6] <https://ieeexplore.ieee.org/document/9498448>

[7] [https://archive.ics.uci.edu/ml/datasets/forest+fires](https://ieeexplore.ieee.org/document/9498448)

[8][https://www.analyticsvidhya.com/blog/2021/10/forest-fire-prediction-using-achine-learning](https://ieeexplore.ieee.org/document/9498448) [/](https://ieeexplore.ieee.org/document/9498448)

[9][https://www.researchgate.net/publication/261272818\_Artificial\_intelligen ce\_for\_forest\_fire\_prediction](https://ieeexplore.ieee.org/document/9498448)

[10] [https://fireecology.springeropen.com/articles/10.4996/fireecology.11011](https://ieeexplore.ieee.org/document/9498448) [06#:~:text=The%20main%20objective%20was%20the,the%20amount%20 of%20burned%20area.](https://ieeexplore.ieee.org/document/9498448)

**Chapter 8**

Acknowledgement

We are highly indebted to our institute ,**St Francis Institute of Technology**, for providing us with this learning opportunity and also to our **Director Bro. Jose Thuruthiyil**, **Principal Dr. Sincy George** and **H.O.D. DR. Joanne Gomes**.

Our effort would be incomplete without the mention of our guide, **Dr. Vaishali Jadhav** , who guided and constantly supervised us and also provided us with necessary information concerning the betterment of the report.