

## REGRESSION:

### Title: Predict the Burned Area of Forest Fires

**Description:** Forest fire causes serious damage to the Flora and fauna of a country. This is one of major environmental concerns. Early prediction of fires saves large numbers of Flora and fauna and prevents the ecosystem. The meteorological conditions of the forest are the key factors of the forest fire. These climatic data can be obtained using the local sensors which are incorporated in the nearest meteorological stations.

**Problem Type:** This is a Regression problem, where the aim is to predict the burned area of forest fires, in the northeast region of Portugal, by using meteorological and other data.

**Where / for what purpose could your solutions be used:** The key purpose for this prediction is to help the fire management team in proper resource allocation and to help the firefighters in the best possible way.

### Overall approach to the problem and techniques intended to be used:

First it is important to know which variables can potentially affect the decision. So, from the obtained dataset it is observed that the variables (Independent variables) that can potentially affect the decision are FPMC, DMC, DC, ISI. Then cleaning the dataset for applying appropriate technique. After that various classification techniques can be used to build the model and then selecting the best model by evaluating its performance using certain metrics like accuracy, precision and recall and then make appropriate decisions.

### Declaration to be given that data is available and accessible:

The dataset is available and accessible. The file is available in csv format. We are adding 10% of null values in the dataset.

### Description of the dataset:

Number of Rows: 517

Number of Attributes: 13

1. X - x-axis spatial coordinate within the Montesano park map: 1 to 9
2. Y - y-axis spatial coordinate within the Montesano park map: 2 to 9
3. month - month of the year: 'Jan' to 'Dec'
4. day - day of the week: 'mon' to 'sun'
5. FPMC - FPMC 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.
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

## **CLASSIFICATION:**

**Title: Classification of Movie Genre.**

**Description:** Genre categorizes movies. Categorizing movies makes it easier for the viewer to discover what he or she likes and will want to see. Putting a movie into a particular genre or category does not diminish the quality of the movie. Movies often have genres that overlap, such as adventure in a spy movie, or crime in a science fiction movie. But one genre is predominant.

**Problem Type:** Classification Problem

**Where / for what purpose could your solutions be used:** This model can be used to classify the genre based on the movie plot which helps the audience to make choices accordingly.

### **Description of the Dataset:**

Features : Title, Date, Duration, Language, Genre, Cast, Rating, Synopsis

Number of rows : 5500+

Number of Attributes : 8

**Overall approach to the problem and techniques intended to be used:** We are planning to use multi-label classification approach. On the basis of that we are going to build a model that can predict the Genre of a movie using Synopsis. Firstly, we will clean the Synopsis like removing commas and converting all tags into lowercase etc. For which we will use 'nltk' library. Once, we will get the cleaned data, we will divide our dataset into training and testing part and then apply various classification algorithms.

### **Declaration to be given that data is available and accessible:**

The data for the experiment is to be scraped from website:

<https://timesofindia.indiatimes.com/entertainment/top-rated-movies/hindi/bestmovies/2019/2742919>

<https://www.imdb.com/list/ls074451163/>

## **TEAM MEMBERS:**

SRI LALANA – lalisri01@gmail.com

PARSIS PRESSWALA--- parsispresswala@gmail.com

TRAPTI KHANDELWAL— officialtrapti@gmail.com