MINOR PROJECT REPORT

Submitted by:

Savidh Tandon,20103013
Nehal,20103029
Deepanshu Gond,20103112
Nikhilesh Vardhan Gupta,20103129

Mentors:

Dr.Sudesh
Dr.Trilok

Index

Problem statement

Detect forest fires, in the mediterranean region, by using meteorological and other data.

Introduction

- Why the Mediterranean region?
    - The Mediterranean has been identified by WWF as one of the most important regions in the world for its outstanding biodiversity features. Mediterranean forests, situated in a transitional zone between the European, African and Asian continents, are one of the planet's centres of plant diversity, with 25,000 floral species representing 10% of the world's flowering plants on just over 1.6% of the Earth's surface. They also play host to an amazing faunal diversity.
- Forest fires in the Mediterranean have increased in recent years…
    - But recently,mediterranean forests are under serious threat, with forest fires, in most cases deliberately set, playing a major role in their degradation and bringing about huge social, economic and environmental effects. There is a strong need to put in place an effective policy of prevention to address the root causes of this phenomenon.
- What can we do?
    - The development of accurate strategies to prevent potential impacts and minimize the occurrence of disastrous events as much as possible requires modeling and forecasting severe conditions.

- The outcome of this report is to develop a prediction map of forest fire risk areas using input features which can provide crucial support for the management of Mediterranean forest ecosystems.
- The worst affected are the countries of Portugal where wildfires have destroyed 75000 acres of forest so far.
- A recent news from The Hindu related to forest fire in mediterranean region in 2022:
- https://www.thehindu.com/news/international/watch-whats-causing-wildfires-across-the-mediterranean-region/article65663393.ece
- This points out the major cause of forest fire being the climate change.So it’s natural we take weather conditions like temperature,wind,rain as attributes in our dataset.

Methodology

- Got a dataset combined from 3 sources.
    - Dataset source:
        - Paulo Cortez, pcortez '@' dsi.uminho.pt, Department of Information Systems, University of Minho, Portugal.
        - Aníbal Morais, araimorais '@' gmail.com, Department of Information Systems, University of Minho, Portugal.
        - Data on forest fires in Turkey,taken from NASA portal.Satellite:N Suomi National Polar-orbiting Partnership (Suomi NPP)
    - Other information about the dataset used:
        - Number of Instances: 517
        - Number of Attributes: 12 + output attribute
        - Attribute information:

X	x-axis spatial coordinate within the Montesinho park map: 1 to 9
Y	y-axis spatial coordinate within the Montesinho park map: 2 to 9
month	month of the year: "jan" to "dec"

day	day of the week: "mon" to "sun"
FFMC	FFMC index from the FWI system: 18.7 to 96.20(FWI:Fire weather index system)
DMC	DMC index from the FWI system: 1.1 to 291.3
DC	DC index from the FWI system: 7.9 to 860.6
ISI	ISI index from the FWI system: 0.0 to 56.10
temp	temperature in Celsius degrees: 2.2 to 33.30
RH	relative humidity in %: 15.0 to 100

wind	wind speed in km/h: 0.40 to 9.40
rain	outside rain in mm/m2 : 0.0 to 6.4
area	the burned area of the forest (in ha): 0.00 to 1090.84

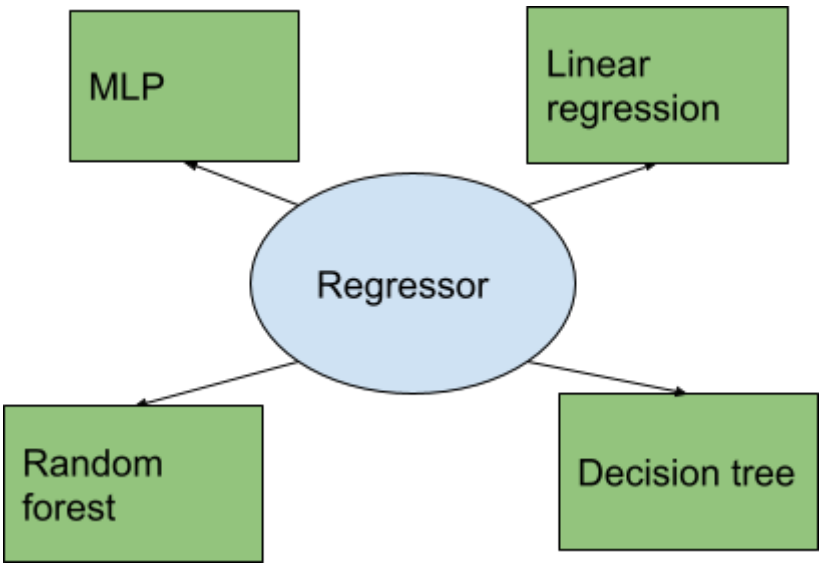
(the area is skewed towards 0.0,so it was first transformed with a ln(x+1) function)

- Missing Attribute Values: None
- Note: Several of the attributes may be correlated, thus it makes sense to apply some sort of feature selection.

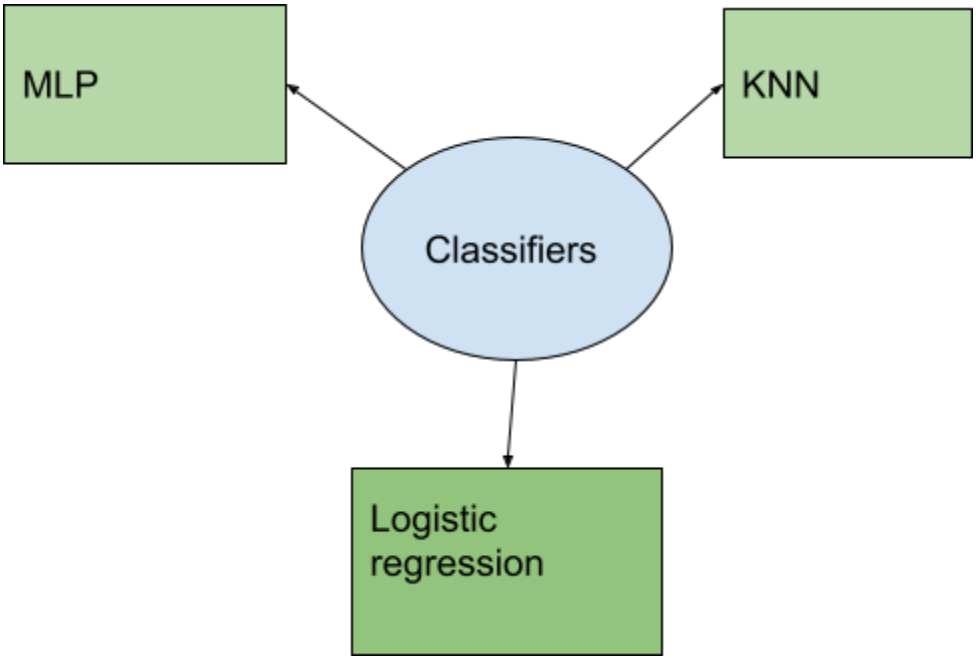
- Data preprocessing
  - Removed irrelevant columns from the dataset like X,Y,month,day.
  - Merged correlated columns into a single relevant column.However in our dataset,since we didn't have enough features we skipped this step.Also we had area as our predicted continuous label.It was skewed towards 0

since small fires were large in number as compared to larger ones.

- Compared different regression models for prediction of burnt area using pycaret.
    - **(PyCaret** is an open-source, low-code machine learning library in Python that automates machine learning workflows.Fast + Explainable + Scalable)
    - Some regression models used:
        - Linear regression
        - MLP regressor
        - Decision tree regressor
        - Random forest regressor

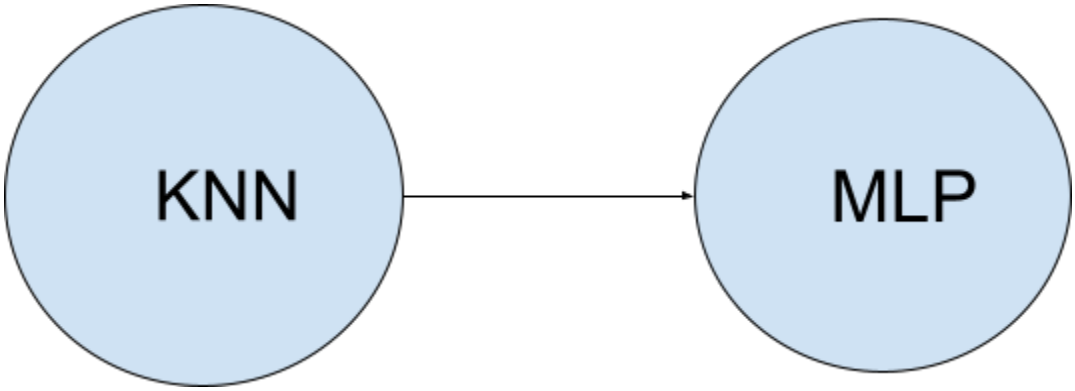


- Converted area feature which was continuous to a discrete.
    - Wherever the burnt area was greater than 0,it was labeled 1 and others labeled as 0.
    - Better approach:Define more labels as 0,1,2,3,4 etc since area ranged from 0 to 1000 units
- Compared different binary classification models for prediction of occurrence of forest fire or not using pycaret.
    - Some binary classification models used:
        - Logistic regression
        - MLP classifier
        - KNeigbours classifier



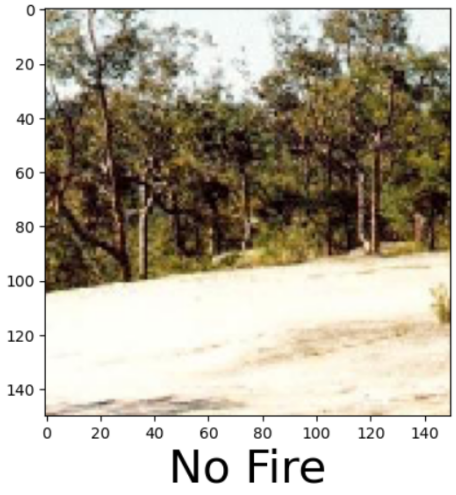
- Accuracy:KNeighbours classifier followed by MLP classifier achieved highest accuracy.

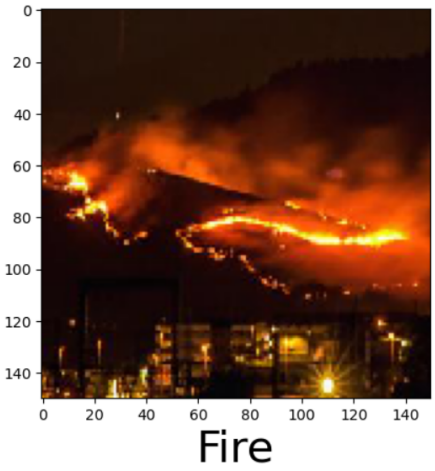
- Final Classifier ensemble model using Stacking



    - There are many classifier models such as SVM(Support Vector Machine), KNN(K-Nearest Neighbours), Decision Tree, Random Forest, etc. Each model will be performing in its unique way delivering different scores. Hence, an ensemble classification model will be developed that will consider the performance of all machine learning algorithms. It is called Stacking Generalization also known as the Stacking Ensemble Machine Learning algorithm.
    - Stacking ensemble classification uses a meta classifier to understand the combination of predictions from multiple machine learning algorithms to deliver the best accurate output.
    - This helps us harness the unique capabilities of well-performing models on a classification or regression task and make predictions that have better performance than any individual model in the ensemble.
- Last but not the least,we also input fire and non-fire images and the resultant model using deep learning to form multi-relationship layers of the parameter that shall detect the fire images from the mix of fire, no-fire images.
    - Under artificial intelligence, the concepts of computer vision will be executed. Computer vision helps us to reconstruct, interpret and understand a visual image, in terms of the properties of the structure present in the scene and draw meaningful descriptions of physical objects from their

image. The tools will be used to process the images and perform additional resizing and conversions. Tensorflow creates a deep learning network that solves image recognition problems and in this project, a sequential model is used that is going to create a list of layers that are passed to the Sequential constructor. It builds a stack of layers where each layer has exactly one input and output tensor.



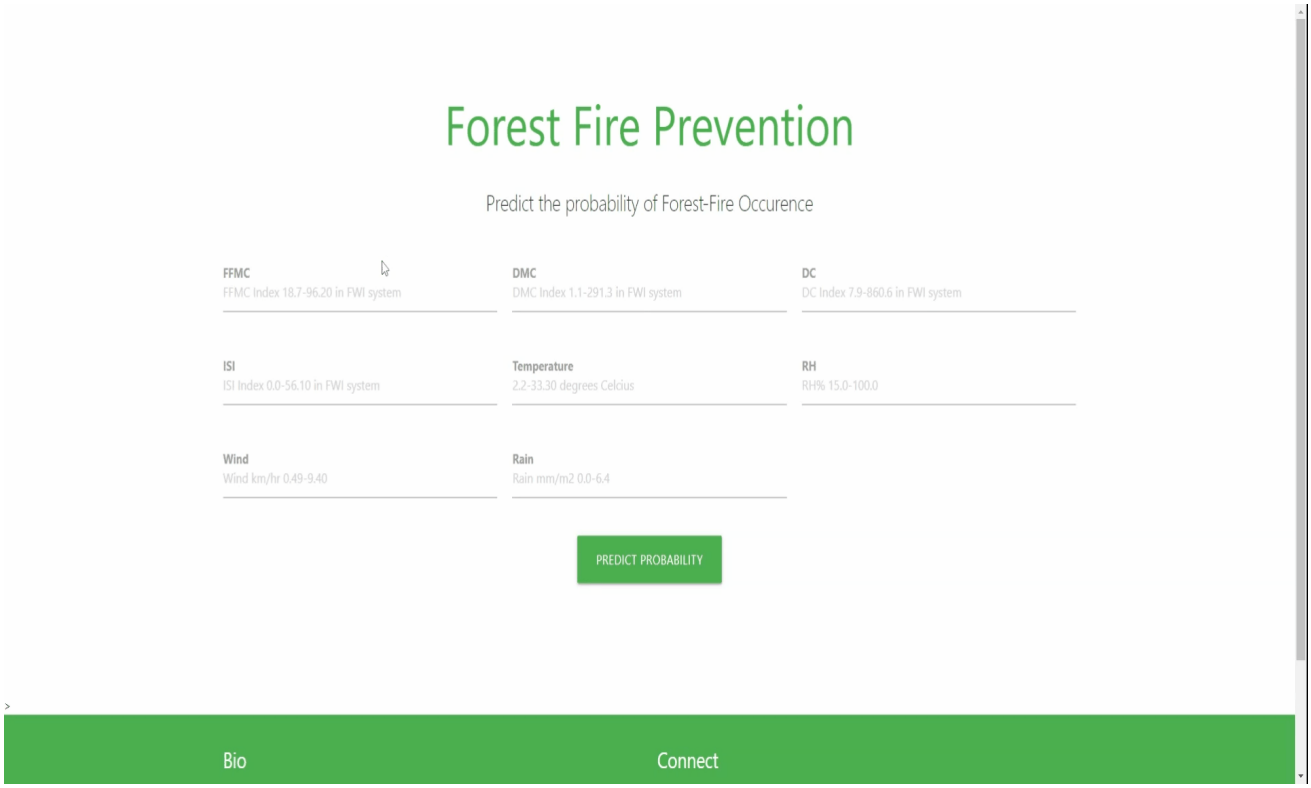


Implementation

- Python libraries used:
  - Numpy
  - Pandas
  - Matplotlib
  - Seaborn
  - PyCaret

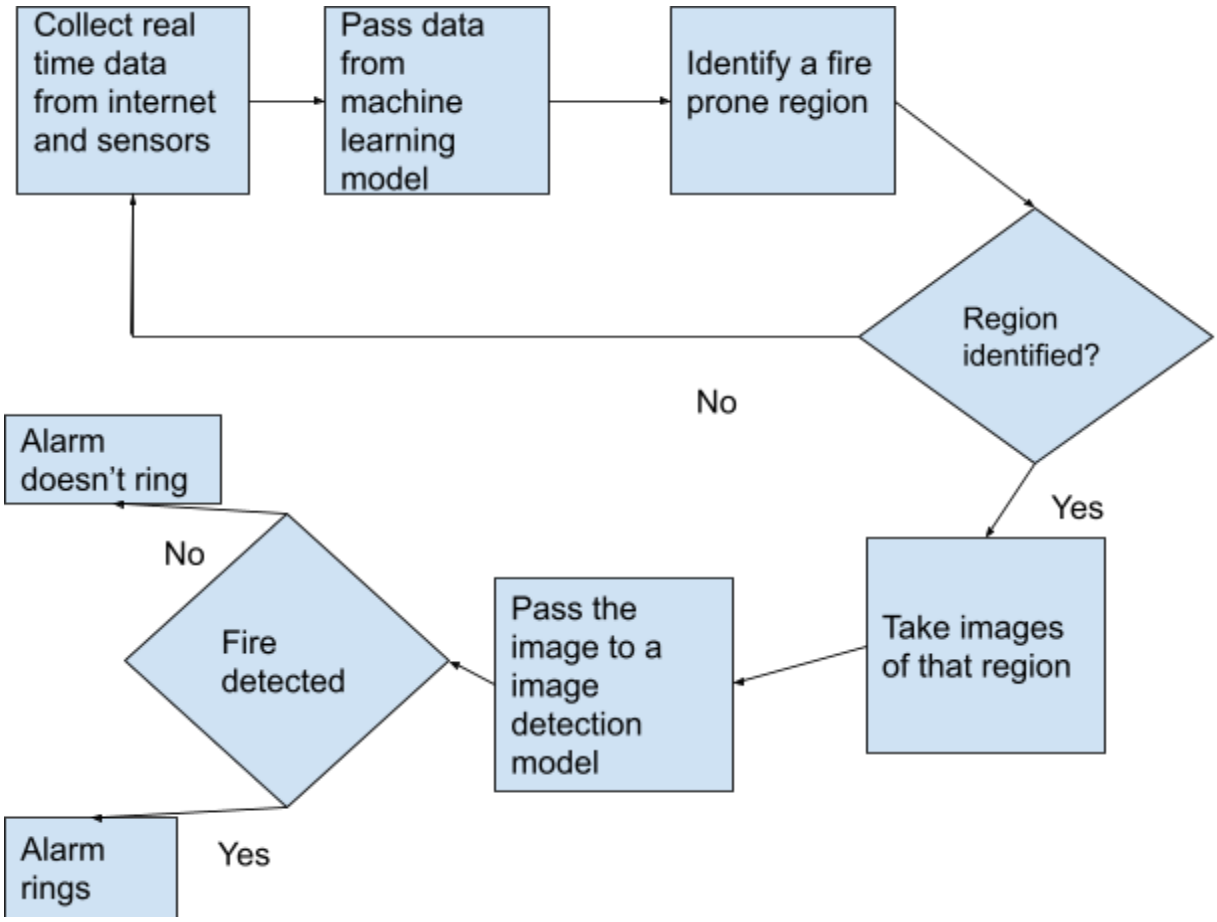
  - sklearn
- Joblib framework(Speeds up python pipelines)

- Tensorflow
- Keras
- Flask


Forest Fire Prevention
Predict the probability of Forest-Fire Occurence
FFMC
FFMC Index 18.7-96.20 in FWI system
DMC
DMC Index 1.1-291.3 in FWI system
DC
DC Index 7.9-860.6 in FWI system
ISI
ISI Index 0.0-56.10 in FWI system
Temperature
2.2-33.30 degrees Celcius
RH
RH% 15.0-100.0
Wind
Wind km/hr 0.49-9.40
Rain
Rain mm/m2 0.0-6.4
PREDICT PROBABILITY
Bio
Connect

Future scope

- Develop real time sensing drone for surveillance in mediterranean region.
- It senses weather conditions in forest prone regions and then takes a picture of that area where there are high chances of fire and alarms if fire is detected.
- Incorporating wireless sensing techniques into our model for making it real time and implementable.



Results

MLP classifier model achieved an accuracy of about 61% on training data and 52% on test data.

CNN model achieved an accuracy of about 90%.

References

https://www.sciencedirect.com/science/article/pii/S1470160X21005343

https://ec.europa.eu/environment/forests/pdf/meeting140504_wwffirstdocument.pdf

https://www.mdpi.com/1999-490
7/13/7/1129

.....**END OF REPORT**.....