EasyChair Preprint

№ 9901

# Prediction of Forest Fires Using Logistic Regression

Adhikari Durga Venkata Madhav, Addanki Gargeya, Amanchi Sravan and Chimata Suman



March 31, 2023

# Prediction of forest fires using logistic regression

Abstract: An important aspect of managing forest fires is forest fire prediction. It is crucial to efforts for resource allocation, mitigation, and recovery. In this essay, machine learning-based strategies for predicting forest fires are described and analysed. The research innovative logistic regression-based forest fire risk prediction system is presented. The system uses historical weather data to forecast the likelihood of a fire on a given day. Through the use of historical data, a correlation between forest fire producing elements and forest fire occurrence is being established in this study project. We can find out when there is a high risk of forest fires by using the system, and forest guards can pay particular attention to preventing forest fires at those times.

Key words: forest fire, logistic regression, prediction,weather, machine learning, historical data

## Introduction:

Several terrestrial ecosystems, including savannas, grasslands, Mediterranean ecosystems, temperate forests, and boreal forests, among others, depend heavily on forest fires. A large portion of all fires that happen worldwide are in the Mediterranean region. Prediction, prevention, and control strategies for forest fires are becoming more and more crucial. Systems for predicting forest fire danger are a crucial tool for predicting forest fire hazards, supporting the monitoring and extinction phase of forest fires, and helping with resource allocation and fire management planning. Forest fire databases are being used by various fire risk models to build and evaluate probabilistic models.Based on each location's fire history, elevation above sea level, and the matching dates of fire days and non-fire days. In addition to measures to avoid, detect, and anticipate forest fires, strategies to deal with them can also be used. Warming brought on by an increase in earth's average temperature and human irresponsibility are the main factors that cause forest fires to break out.

High ambient temperatures and dryness (low humidity) provide ideal conditions for the onset of a fire. Man-made causes - When an ignition source, such as a naked flame, a cigarette or bidi, an electric spark, or any other source of ignition, comes into contact with combustible material, a fire results.

The economy will suffer a loss of nearly $23 trillion over the next 80 years, according to the Dynamic Integrated Model of Climate and the Economic (DICE). Forest fires are caused by human activities such as agriculture and animal husbandry in Southeast Asia, New Zealand, South America, and Africa. Many tools, including physical models and mathematical models, are available today to simulate flames and predict their spread. To describe and forecast the growth of fire in many places, these models rely on data gathered during forest fires, modelling, and lab trials. Simulator tools recently. Despite the widespread application of these methods, many regions of the world have not yet been delineated in terms of fire susceptibility. Further, no single model/method has been yet identified to capture fire behavior in all regions due to the variation of training data from different regions . To fill this significant gap in fire prediction efforts, we aimed this study to develop a suite of predictive models based on the four machine learning methods, namely Bayes Network, Naïve Bayes, Decision Tree, and Multivariate Logistic Regression for the prediction of fire susceptibility in the Pu Mat National Park of Vietnam. Although these methods have been broadly investigated in environmental studies, particularly for the prediction of landslides and floods , their joint application and comparison have not yet been reported for forest fire prediction. The outcomes from this study allow researchers to determine if a particular predictive model derived from machine learning methods aligns with their objectives for modeling and mapping of forest fire susceptibility.

## Literature survey:

Initially, we've found terms relating to our article, such as "forest fire," "logistic regression," "prediction," "weather," and "historical data."

Using the selected keywords, we looked for past studies on forest fire prediction in Google Scholar, IEEE, and other databases.

First, we developed our prediction model using logistic regression, which we later finished with excellent accuracy. We then incorporated the prediction model into an HTML website, and this page would calculate the likelihood of a forest fire.

So, for this study endeavour, we have collected data from several websites and developed our own dataset. When incorporating a dataset into a model, several steps must be taken, such as data cleansing, data type conversion, formatting cleanup, handling of missing values, and data reduction.

There are various steps you may take to improve the model's accuracy.Very often, the dataset is crucial to the accuracy aspect. Increasing the amount of historical data or historical information will improve the prediction model's accuracy. We produced a unique dataset. As our model is not dependent on cities, we have not included any of them. Temperature, oxygen, fire occurrence, humidity, and fire occurrence is reliant on the remaining three elements are among the aspects we have taken into consideration.

Several study articles have been written in the past on this subject, but owing to a lack of data and the lack of alternative methodologies, their accuracy was poor. Our study's accuracy is a little bit higher than that of other studies, and we do this by employing the well-known and highly accurate machine learning technique Logistic Regression.

Moreover, this logistic regression mostly uses probability and is employed for making predictions.

Pickle is a tool we've used in Python, and it helps. Pickling is the process of turning a Python object hierarchy into a byte stream, while unpickling is the process of turning a byte stream (from a binary file or object that looks like bytes) back into an object hierarchy.

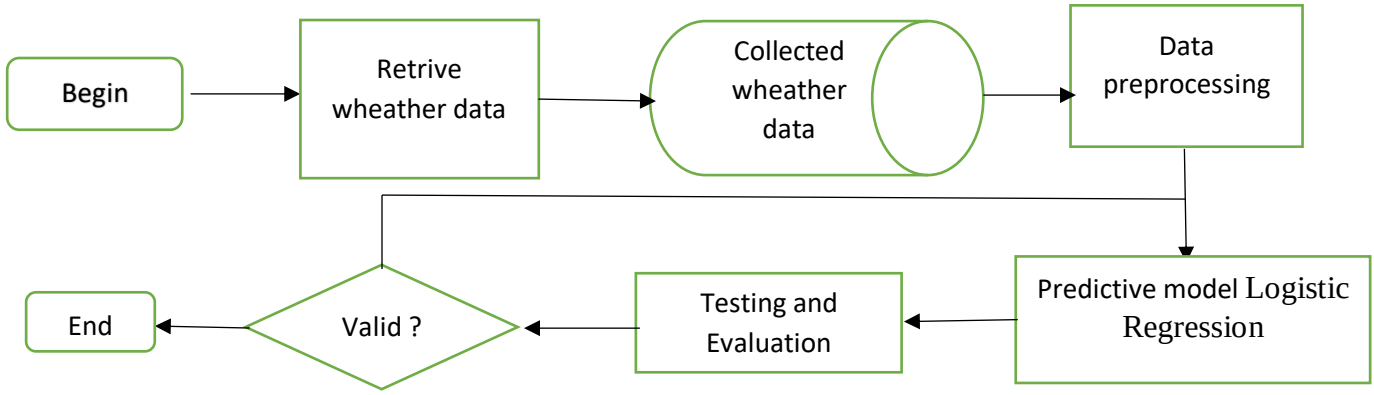
Flasks have been utilised, and this Python-based Flask, which is implemented on Werkzeug and Jinja2, is used to create web applications. Using the Flask framework has the following benefits: There is an integrated development server and a quick debugger available.

Since the webpage code was written in HTML and the programme code was written in logistic regression, we imported Flask, created a folder structure, and then used app.py to embed Flask into your application from HTML webpages. To produce an HTML file, we must develop a template, and these templates are shown on websites.

Next, we had to run the programme before submitting the application to a public server.

## RESEARCH METHODOLOGY:

The methodology of machine learning technique




Dataset: For the current Research work We are building our own dataset. Factors in the dataset include oxygen, temperature, humidity, and fire occurrence.we summarized the attributes and corresponding values.

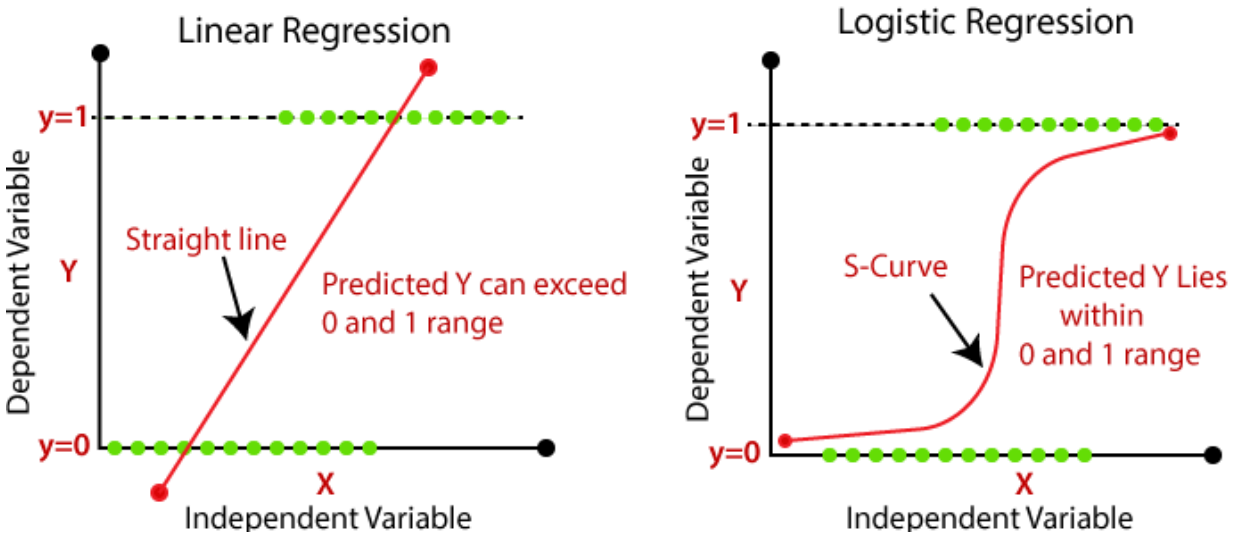
Apply Machine Learning Techqniues : once the data is ready for modelling, we apply the most popular prediction algorithm Logistic Regression to predict the forest fire

## About Logistic Regression:

- One of the most often used Machine Learning algorithms, within the category of Supervised Learning, is logistic regression. With a predetermined set of independent factors, it is used to predict the categorical dependent variable.
- The output of a categorical dependent variable is predicted via logistic regression. The result must thus be a discrete or categorical value. It can be either True or False, Yes or No, 0 or 1, etc., but rather than providing the precise values of 0 and 1, it provides the probability values that fall between 0 and 1.
- The main difference between linear regression and logistic regression is how they are used. In order to solve regression issues, one uses linear regression, whereas classification difficulties are solved using logistic regression.

- In logistic regression, we fit a "S" shaped logistic function, which predicts two maximum values, rather than a regression line (0 or 1).

Although it goes by the name of logistic regression and is used to classify samples, classification algorithms are what it comes under since it employs the notion of predictive modelling as regression.




## Type of Logistic Regression:

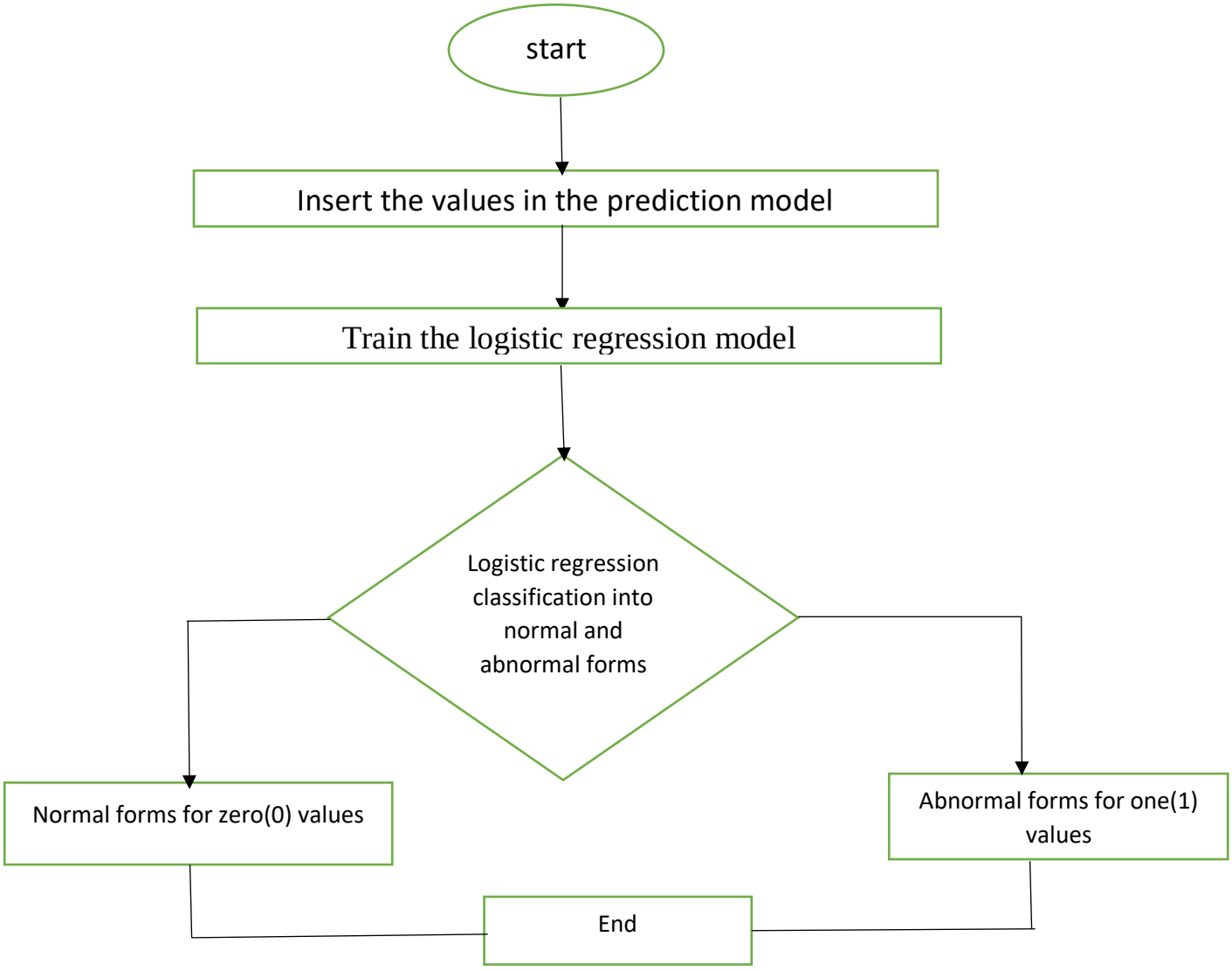
On the basis of the categories, Logistic Regression can be classified into three types:

Binomial: In a binomial logistic regression, the dependent variables can only be one of two potential kinds, such as 0 or 1, Pass or Fail, etc.

Multinomial: In multinomial logistic regression, the dependent variable may be one of three or more potential unordered kinds.

Ordinal: In ordinal Logistic regression, there can be 3 or more possible ordered types of dependent variables, such as "low", "Medium", or "High".

## Alogorithm for Logistic Regression:




## Performance Metrics :

We need some measures in order to evaluate a machine learning model's performance. These metrics are statistical standards that may be used to gauge and track a model's effectiveness.

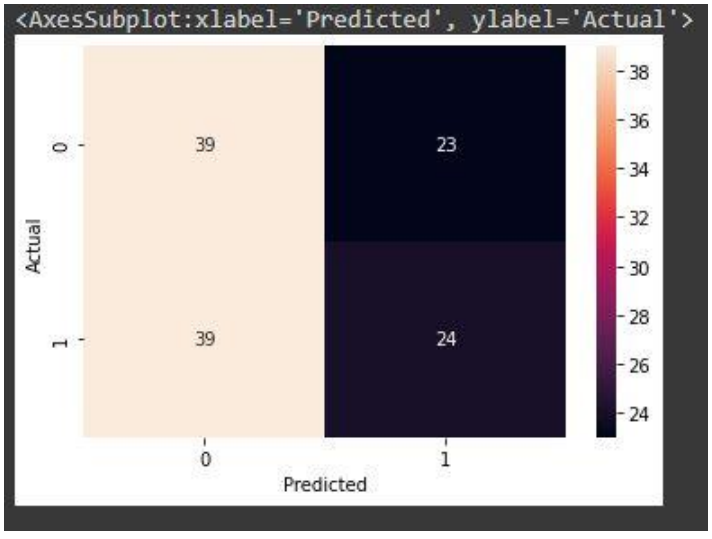
The following is a list of the performance evaluation measures utilised in this study.

- If TP and FP are true positive and false positive rates, respectively, then precision is defined as precision = (TP) / (TP + FP) in a formal sense.
- The definition of recall is as follows, where FN stands for false negative rate.
- Recall is equal to (TP)/(TP + FN)
- Accuracy = (TP + FN) / (TP + TN + FP + FN)

```
0.504
              precision    recall  f1-score   support

           0       0.50      0.63      0.56        62
           1       0.51      0.38      0.44        63

    accuracy                           0.50       125
   macro avg       0.51      0.50      0.50       125
weighted avg       0.51      0.50      0.50       125

[[39 23]
 [39 24]]
```

Confusion Matrix:

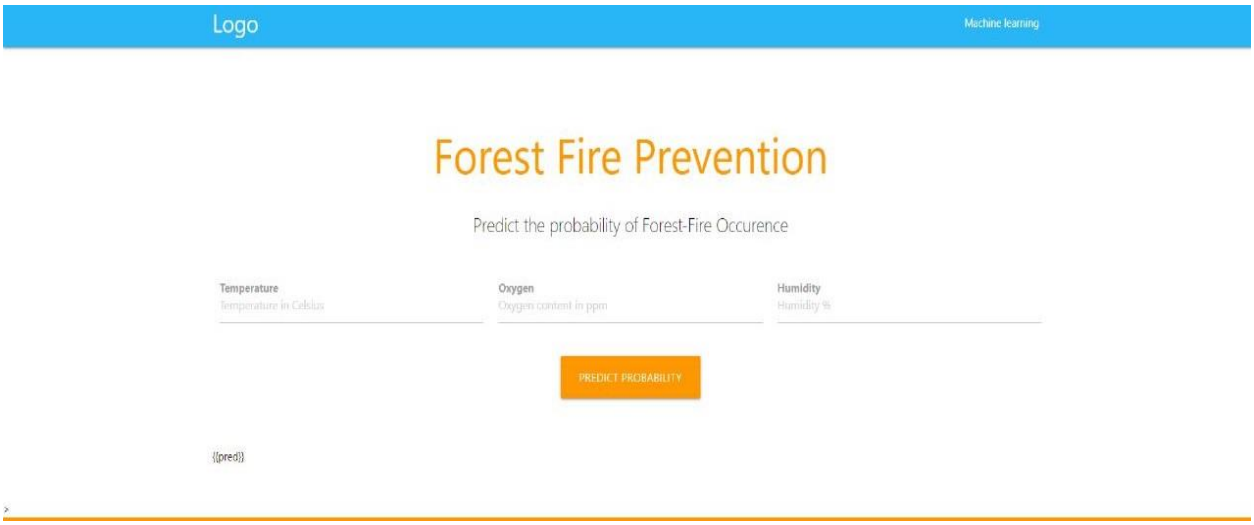
An evaluation of a classification algorithm's performance is done using a table called a confusion matrix. A confusion matrix depicts and summarises a classification algorithm's performance.



## Result Analysis:

The dataset contains factors like oxygen, temperature, and humidity, which are the main factors that contribute to forest fires. By providing these values to the webpage (or) algorithm, the machine predicts the forest fire and outputs values between 0 and 1. This raises the question: Why does the machine predict values between 0 and 1 when the majority of probability is between 0 and 1? Because the algorithm we used is logistic regression, which predicts values between 0 and 1, the machine will respond

The user must input the values into the model in order for it to predict the values and inform you whether the climate is nice or not. If the prediction is below 0.5, it is low; if it is above 0.7, there is a greater likelihood of a wildfire; and between 0.5 and 0.7, it is moderate. In this research, the machine only simply predicts the values; it does not intimate you whenever you want. A wildfire is more likely to occur in the summer and less likely to occur in the winter.




## Conclusion:

In this research project, we utilised logistic regression to forecast forest fires. Since logistic regression primarily relies on probability, the numbers that should be entered into the webpage we constructed in order for the algorithm to function properly are temperature, humidity, and oxygen.Because of temporal variations, these values do not remain constant.

Because of the dynamic environment, unpredictability, and change of factors like humidity and temperature, predicting forest fires is a difficult endeavour.

Accurate fire probability forecasting helps forest managers create more effective firefighting plans and restructure existing regulations for the sustainable management of forest resources.

## Reference:


1.Syed Mohammed Khalid, Raini Hassan

Prediction of the Level of Air Pollution During Wildfires Using Machine Learning Classification Methods

International Journal on Perceptive and Cognitive Computing (IJPCC)

Vol 6, Issue 2 (2020)

2.B S Negara,R Kurniawan, M Z A Nazri, S N H S Abdullah, R W Saputra and A Ismanto

Riau Forest Fire Prediction using Supervised Machine Learning,ICCAI 2019

3.Ahmed M. Elshewey,Amira. A. Elsonbaty

Forest Fires Detection Using Machine Learning Techniques,ISSN No : 1006-7930

https://www.researchgate.net/publication/344462171

4.Yongqi Pang , Yudong Li, Zhongke Feng ,Zemin Feng,Ziyu Zhao,Shilin Chen and Hanyue Zhang

Forest Fire Occurrence Prediction in China Based on Machine Learning Methods

https://www.mdpi.com/journal/remotesensing

Remote Sens. 2022, 14, 5546

5.Bahareh Kalantar,Naonori Ueda,Mohammed O. Idrees,Saeid Janizadeh,Kourosh Ahmadi and Farzin Shabani

Forest Fire Susceptibility Prediction Based on Machine Learning Models with Resampling Algorithms on Remote Sensing Data

https://www.mdpi.com/journal/remotesensing

Remote Sens. 2020, 12, 3682

6.Suwei Yang,Massimo Lupascu,Kuldeep S.Meel

Predicting Forest Fire Using Remote Sensing Data And Machine Learning

7.Binh Thai Pham ,Abolfazl Jaafari, Mohammadtaghi Avand, Nadhir Al-Ansari,Tran Dinh Du, Hoang Phan Hai Yen , Tran Van Phong , Duy Huu Nguyen , Hiep Van Le ,

Davood Mafi-Gholami, Indra Prakash, Hoang Thi Thuy and Tran Thi Tuyen

Performance Evaluation of Machine Learning Methods for Forest Fire Modeling and Prediction

https://www.mdpi.com/journal/symmetry

Symmetry 2020, 12, 1022

8.George E. Sakr, Imad H. Elhajj, George Mitri and Uchechukwu C. Wejinya

Artificial Intelligence for Forest Fire Prediction

2010 IEEE/ASME International Conference on Advanced Intelligent Mechatronics Montréal, Canada, July 6-9, 2010

9.Mochammad Anshori,Farhanna Mar'i,Mukhammad Wildan Alauddin,Wayan Firdaus Mahmudy

Prediction of Forest Fire using Neural Network based on Extreme Learning Machines (ELM)

Authorized licensed use limited to: University of Exeter.

10.Aditi Kansal, Yashwant Singh, Nagesh Kumar, Vandana Mohindru

Detection of Forest Fires using Machine Learning Technique: A Perspective

11.Gao Demin, Lin Haifeng, Jiang Anna, Wu Guoxin

A FOREST FIRE PREDICTION SYSTEM BASED ON RECHARGEABLE WIRLESS SENSOR NETWORKS

12.B.K. Singh,Nikhilesh Kumar,Pratima 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)

13.George E. Sakr, Imad H. Elhajj, George Mitri and Uchechukwu C. Wejinya

Artificial Intelligence for Forest Fire Prediction

2010 IEEE/ASME International Conference on Advanced Intelligent Mechatronics Montréal, Canada, July 6-9, 2010

14. - Kansal A, Singh Y, Kumar N, Mohindru V, "Detection of forest fires using machine learning technique: A perspective", In Third International Conference on Image Information Processing (ICIIP), IEEE, 2015

15. Stojanova D, Panov P, Kobler A, Džeroski S, Taškova K, "Learning to predict forest fires with different data mining techniques", In Conference on Data Mining and Data Warehouses, SiKDD, 2006