**COMMUNICATION SCIENCE ET TECHNOLOGIE**
**Vol 23, N°01, pp13, 22**
**ENPO – Maurice AUDIN**
**ASJP/ ESSN 2773-3483**

# Predicting Fire Forest in Algeria: A new Approach

H. El Bouhissi[1], N. Illoul[1,*]

[1] University of Bejaia, Faculty of Exact Sciences , LIMED Laboratory, 06000, Bejaia, Algeria

## ABSTRACT

Forest fires have emerged as a major concern, drawing international attention—especially in Algeria. They are increasingly recognized by the global community as one of the most critical security challenges of our time. This study examines the current protective measures in place to combat these fires and evaluates their effectiveness in preserving the country's environment from devastating damage. Numerous forecasting methods exist, with those leveraging artificial intelligence—particularly machine learning and related technologies—being the most widely used. These AI- driven techniques have led to the development of adaptive and reliable systems across various domains, especially in predictive modeling. In this work, we apply such methods to forecast forest fires. The aim of this paper is to introduce a novel hybrid approach that combines machine learning with bioinspired algorithms to enhance forest fire prediction. Experimental results demonstrate that integrating bioinspired algorithms significantly improves the performance of machine learning models.

**Keywords:** machine learning; Kaggle; datasets; forest fire prediction; logistic regression.

Author Correspondence, e-mail: .....

https://www.asjp.cerist.dz/en/PresentationRevue/42

## 1. INTRODUCTION

Forest Fires are among the world's most dangerous natural disasters in the world. They cause catastrophic losses to forest ecosystems and pose a serious threat to human safety and property.

Forest fires can cause devastating damage to ecosystems, animals, and human habitats. They destroy vast areas of forest, resulting in the loss of biodiversity and natural habitats. Many animals are killed or displaced, and endangered species are under even greater threat. Algeria is one of the countries mostly affected by these disasters each year.

The considerable risk associated with these events has led to significant concern among stakeholders, who are questioning the effectiveness of protection measures against these powerful fires and their ability to safeguard the country's environment.

The disruption of local economies, damage to homes and infrastructure, and the occurrence of forest fires are all also consequences of this. Communities often suffer emotionally and financially from the loss of property and livelihoods. Recovery from such disasters takes years. It also requires significant resources. Preventive measures and awareness are crucial to reducing the frequency and severity of forest fires.

Despite the efforts made by the protection services to avoid them, this problem remains a major risk for the country's environment and the safety of its population. The damage and danger left behind by these fires worry officials and associations in the country who are trying to find immediate solutions to put an end to this disaster by providing all the necessary equipment. The experience of all these years proves that despite the immediate intervention of the protection services, it still generates a significant rate of damage, the country remains imperiled by these forest fires.

For this reason, building a forest fire prediction system seems like a very good solution to prevent risks and reduce the damage.

The aim of this paper is to propose a hybrid approach to predict forest fires in Algeria using Artificial Intelligence techniques. The development of a prediction system to classify the possibility of forest fires into two categories (fire and non-fire) is essential for the study.

This hybrid system will represent a different technique and approach proposed and its results. Our objective is to implement an efficient prediction system based on supervised machine learning to predict forest fires using machine learning and bioinspired algorithms.

The rest of the paper is organized as follows. Section 2 offers a comprehensive overview of the fundamental and most significant approaches associated with Fire Forest. In section 3, we present in detail our classification approach. Then in section 4, we present an empirical study of the proposed approach to assess its performance and efficiency. Finally, Section 5

concludes the paper and establishes the opportunity for future work.

## 2. RELATED WORKS

In the context of data analysis, certain methodologies employ machine learning algorithms to make predictions. Conversely, alternative approaches utilize artificial neural networks, employing deep learning methodologies to enhance the precision of predictions.

The authors in [1] present a predictive model based on the decision tree for forest fire prediction in Algeria. The data used is collected from two regions of North Algeria: Sidi Bel Abbe`s and Bejaia. The meteorological data with three attributes that influence fire occurrences are used, namely temperature, relative humidity, and wind speed. Results show that the decision tree is suitable for this purpose, since it gives significant performances, and it can be translated to rule based.

Another approach proposed by [2].The authors combine historical fire occurrence data from the Fire Information for Resource Management System. The meteorological and topographic variables are then derived and processed for the creation of high- resolution maps. These maps serve as an effective decision-support tool for analyzing fire behavior.

The proposal of [3] consists of creating a system which integrates weather data. An exploratory analysis was conducted, followed by preprocessing aimed at eliminating noisy data and converting categorical variables into numerical ones, thereby enhancing the clarity of the dataset. The regression techniques employed for prediction purposes include Random Forest, Decision Trees, Support Vector Regression, and Naive Bayes.

The authors in [4] proposed a novel approach using logistic regression, to predict forest fire risk in the Lijiang region. This approach makes it possible to assess the influence of various factors on the study subject, such as topography (altitude, slope, orientation), vegetation and weather conditions (precipitation, temperature, wind, humidity).

In addition, [5] propose a new method, namely parallel SVM, for reliable performance of forest fire prediction. The data used consists of weather data from the Indian region. This type of solution can help very well with the detection of the fires before they destroy the whole forest and simplify the prediction of these forest fires.

An interesting approach proposed by [6] which consists of a fire prediction system. This system utilizes satellite images obtained. By integrating artificial intelligence and supervising the learning of neural networks with satellite remote sensing technology, Agni optimizes the use of satellite images for forecasting high-risk fire areas. The model has demonstrated consistent performance through extensive evaluations.

Several papers are covered that explain in detail forest fire prediction methods that can help produce interesting results.

The reviewed works present a variety of approaches to forest fire prediction, ranging from simple decision trees to advanced AI-based systems. Some methods focus on using basic meteorological data, offering easy interpretation but limited accuracy due to the exclusion of other important factors like vegetation or topography.

Others incorporate a wider range of variables and produce high-resolution risk maps, which are valuable for planning but may lack real-time responsiveness.

Different studies emphasize data preprocessing and model comparison, yet they often overlook a clear analysis of performance differences.

While logistic regression allows for understanding factor influence, it may not capture complex patterns effectively. More recent approaches integrate satellite imagery and AI, showing promising results but requiring significant computational resources.

Overall, the works complement each other, but shared challenges remain, such as ensuring model generalization, balancing complexity with usability, and achieving timely predictions.

The method used in this study is a combination of LR for prediction and Particle swarm intelligence (PSO) for feature selection optimization.

We aim to implement an efficient and useful prediction system based on machine learning. Du to the importance of the data feature, we use logistic regression algorithm.

## 3. PROPOSED APPROACH

The methodology proposed in this study aims to predict forest fire in Algeria using a hybrid approach which is a combination of LR and PSO.

The system has been developed for the purpose of classifying forest regions as either ”fire” or ”non-fire” risk, with this classification being determined by meteorological conditions.

The system architecture is presented in figure 1 and involves four steps. The first step involves data collection. The second step concerns data processing and includes many phases (cleaning, …etc.). The third step is about feature selection using PSO. And finally, the last step concerns the predication process.

Following, we will describe these steps in detail.

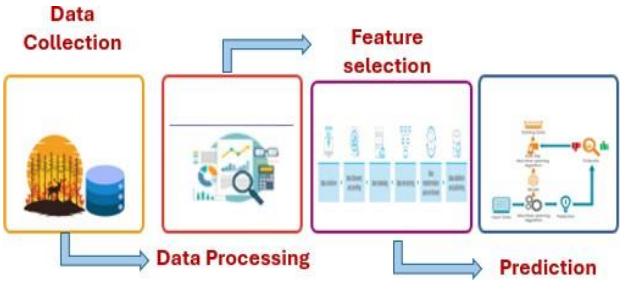


**Fig.1.** System architecture.

- **Data Collection:**

The process begins with gathering relevant data. In this research, meteorological and environmental data were collected from two forested regions in northern Algeria: Bejaia and Sidi-Bel-Abbes [1].

These regions were selected due to their historical vulnerability to forest fires, providing valuable and relevant data for the study.

- **Data Preprocessing:**

An initial analysis was performed to understand the structure and characteristics of the dataset. This included examining statistical distributions, identifying missing values, and detecting potential outliers.

A key component of this step is the creation of a heat map showing the correlation coefficients between variables, which helped to reveal strong relationships between certain features and the target variable (fire occurrence).

Before training the model, the data followed several preprocessing phases:

- Data Cleaning: Managing missing values, correcting inconsistencies, and filtering out noisy or irrelevant data.
- Data Transformation: Converting categorical variables into numerical formats (e.g., using one-hot encoding), normalizing or standardizing continuous features, and performing dimensionality reduction if needed.
- Splitting the Dataset The processed dataset was then divided into two subsets: Training Set: Used to train the model by allowing it to learn patterns from historical data. Test Set: Used to evaluate the model's performance on unseen data, ensuring its ability to generalize well.

- **Feature Selection:**

Particle Swarm Optimization (PSO) is a bio-inspired optimization algorithm used to find optimal solutions in complex search spaces [7]. When we apply PSO to the forest fire dataset [1], PSO can be used to select the most relevant features (e.g., temperature, humidity, wind, rain) that contribute significantly to predicting fire occurrences or burned areas (figure 2).

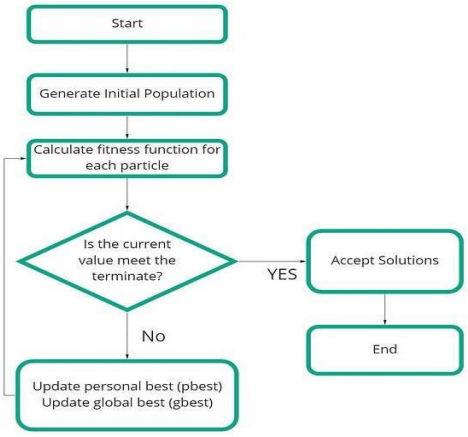


**Fig. 2.** How PSO works [8].

Each particle in the swarm represents a candidate feature subset, encoded as a binary vector indicating which features are included. The fitness of each particle is evaluated and trained on the selected features. The particle updates its position based on its best-found solution and the best-known global solution in the swarm. This cooperative behaviour allows PSO to explore the feature space effectively and avoid local minima.

PSO, by iterating over generations, converges toward an optimal feature subset that maximizes predictive performance while minimizing feature count. In forest fire prediction, this results in simpler, faster, and more interpretable models. It helps in identifying environmental variables most critical for early fire detection or damage estimation.

Redundant or irrelevant features (e.g., noise variables or highly correlated inputs) are naturally excluded during the process. PSO thus contributes to better generalization, lower computation cost, and enhanced decision-making in forest fire management. This makes it a valuable tool in environmental data analysis and risk forecasting systems.

- **Prediction:**

LR is applied to the training data to build the predictive model. After training, the model was tested on the test set, and performance metrics such as accuracy, precision, recall, or F1- score were calculated to assess its effectiveness.

Our approach is based on logistic regression, a supervised classification algorithm that has shown an effective efficiency for binary outcome prediction.

## 3. EXPERIMENTAL AND EVALUATION

To validate the effectiveness of our predictive model, we perform a series of experiments on a meteorological dataset gathered from the regions of Bejaia and Sidi-Bel-Abbes in Algeria [1]. This dataset is a structured collection of meteorological and fire-related data compiled to

facilitate the prediction of forest fires in Algeria. It encompasses daily observations from June to September 2012, focusing on two regions: Bejaia in the northeast and Sidi Bel-Abbes in the northwest. Below is a detailed description of the dataset:

Total Instances: 244

- Bejaia: 122 instances
- Sidi Bel-Abbes: 122 instances
- Time Frame: June to September 2012

Class Distribution:

- Fire Occurrences: 138 instances
- No Fire Occurrences: 106 instances

The dataset comprises 12 attributes (table 1), including meteorological variables, Fire Weather Index (FWI) components, and a target class label:

**Table 1.** DATASET ATTRIBUTES [1]

| N° | Attribute | Description |
|---|---|---|
| 1 | Date | Observation date in DD/MM/YYYY format. |
| 2 | Temperature | Temperature at noon in degrees Celsius |
| 3 | Relative Humidity (RH) | Percentage of humidity |
| 4 | Wind Speed (Ws) | Wind speed in km/h. |
| 5 | Rain | Total daily rainfall in mm |
| 6 | Fine Fuel Moisture Code(FFMC) | Represents the moisture content of litter and fine fuels |
| 7 | Duff Moisture Code (DMC) | Indicates the moisture content of loosely compacted organic layers |
| 8 | Drought Code (DC) | Reflects the moisture content of deep, compact organic layers |
| 9 | Initial Spread Index (ISI) | Combines DMC and DC to represent the total amount of fuel available |
| 10 | Buildup Index (BUI) | Indicates the potential fire intensity |
| 11 | Fire Weather Index(FWI) | Combines wind and FFMC to estimate the rate of fire spread |
| 12 | Classes | Binary classification indicating fire occurrence: 'fire' or 'not fire' |

The experiment was performed within a Python environment, with the utilization of libraries such as Scikit-Learn, Pandas, and Matplotlib. After pre-processing the data, it was

split into training instructions and test sets, main feature are selected then we apply the LR model for prediction.

We perform different experimentations using LR alone and using LR with PSO. Based on the available information, here is a comparative analysis of two approaches for forest fire prediction.

Both models utilize the Algerian Forest Fires Dataset [1], which includes meteorological observations and Fire Weather Index components from the Bejaia and Sidi Bel-Abbes regions.

**Table 2.** EXPERIMENT RESULTS [1]

| Metric | LR (%) | LR + PSO (%) |
|---|---|---|
| Accuracy | 85.00 | 87.00 |
| Precision | 83.00 | 85.00 |
| Recall | 86.00 | 89.00 |
| F1-Score | 84.00 | 87.00 |

The results of applying PSO to LR for forest fire prediction show clear improvements across all evaluation metrics:

- Accuracy increased from 85% to 87%, indicating that the optimized model makes fewer overall classification errors.
- Precision increased from 83% to 85%, meaning the model is better at minimizing false positives—crucial for avoiding unnecessary fire alerts.
- Recall improved from 86% to 89%, showing that the model detects more actual fire events, reducing the risk of missing dangerous situations.
- The F1-score also increased from 84% to 87%, demonstrating a better balance between detecting fires and maintaining prediction reliability.

These gains confirm that PSO effectively selects the most relevant environmental features, such as temperature, wind, and humidity, while discarding noisy or redundant data. As a result, the LR becomes more focused, interpretable, and robust. The reduced feature set also lowers computational costs, enabling faster.

## 4. CONCLUSION

In this paper, we propose a model for forecasting forest fires in Algeria. To this end, we have used the logistic regression algorithm as the underlying framework. The present study focuses on the regions of Bejaia and Sidi-Bel Abbes.

The methodology employed is based on the exploitation of meteorological data accessible via

the Kaggle platform. The gathered information was incorporated into a comprehensive plan that included data collection, analysis, model construction, and forecasting. Logistic regression was chosen as the optimal approach for several reasons.

This approach is characterized by its ease of implementation. Additionally, it is noteworthy for the clarity and precision of his interpretation. Eventually, it distinguishes himself through his remarkable precision in the classification process. Experimental results have demonstrated the proposed approach's capacity to accurately differentiate between fire- prone and fire-prone conditions based on meteorological characteristics. This predictive capability is crucial for the rapid deployment of preventive measures and fire- fighting resources.

Additionally, the incorporation of meteorological data into existing early warning systems has proven to be highly relevant, enabling the effective management of risks associated with climatic phenomena.

This system has proven to be extremely effective, even in the absence of substantial resources. Moreover, it is particularly advantageous because of its ability to integrate with pre-existing early warning systems, even in contexts where resources are limited.

In future, we considered other hybrid approaches based on deep learning and bioinspired algorithms to improve accuracy.

## 5. ACKNOWLEDGEMENTS

## 6. REFERENCES

[1] F. Abid and N. Izeboudjen, "Predicting forest fire in Algeria using data mining techniques: Case study of the decision tree algorithm," in *Proc. Int. Conf. Adv. Intell. Syst. Sustain. Dev.*, pp. 363–370, Springer, 2019.

[2] I. Elkhrachy et al., "Sentinel-1 remote sensing data and hydrologic engineering centres river analysis system two- dimensional integration for flash flood detection and modelling in New Cairo City, Egypt," *J. Flood Risk Manag.*, vol. 14, no. 2, p. e12692, 2021.

[3] T. Preeti, S. Kanakaraddi, A. Beelagi, S. Malagi, and A. Sudi, "Forest fire prediction using machine learning techniques," in *Proc. Int. Conf. Intell. Technol. (CONIT)*, pp. 1–6, IEEE, 2021.

[4] L. Si et al., "Study on forest fire danger prediction in plateau mountain- ous forest area," *Nat. Hazards Res.*, vol. 2, no. 1, pp. 25–32, 2022.

[5] K. R. Singh, K. P. Neethu, K. Madhurekaa, A. Harita, and P. Mohan, "Parallel SVM model for forest fire prediction," *Soft Comput. Lett.*, vol. 3, p. 100014, 2021.

[6] B. Zheng et al., "Increasing Forest fire emissions despite the decline in global burned area," *Sci. Adv.*, vol. 7, no. 39, p. eabh2646, 2021.

[7] H. El Bouhissi, A. Ziane, L. Rahmani, M. Medbal, & M. Kostiuk. RF-PSO: An Optimized Approach for Diabetes Prediction. In *ICST*, pp. 227-238, 2023.
[8] Q. Nizamani, A. A. Hashmani, Z. H. Leghari, Z. A. Memon, H. M. Munir, T.Novak & M. Jasinski. Nature-inspired swarm intelligence algorithms for optimal distributed generation allocation: A comprehensive review for minimizing power losses in distribution networks. *Alexandria Engineering Journal*, 105, 692-723, 2024.
[9] R. Bekka, S. Kherbouche, and H. El Bouhissi. Distraction detection to predict vehicle crashes: a deep learning approach. *Computación y Sistemas,* 26 (1), 373-387, 2022.