Article

Predict the duration of forest fires using machine learning methods

Constantina Kopitsa[1], Ioannis G. Tsoulos[2]*, Vasileios Charilogis[3], Athanassios Stavrakoudis[4]

[1] Department of Informatics and Telecommunications, University of Ioannina, Greece;k.kopitsa@uoi.gr
[2] Department of Informatics and Telecommunications, University of Ioannina, Greece; itsoulos@uoi.gr
[3] Department of Informatics and Telecommunications, University of Ioannina, Greece; v.charilog@uoi.gr
[4] Department of Economics, University of Ioannina, Ioannina, Greece; astavrak@uoi.gr
* Correspondence: itsoulos@uoi.gr

Abstract: For thousands of years forest fires played the role of regulator in the ecosystem. Forest fires contributed to the ecological balance by destroying old and diseased plant material, but in the modern era fires are a major problem that tests the endurance not only of governments agencies, around the world, but also the effect on climate change. Forestfires have become more intense, more destructive, and deadlier, known as mega – fires. They can cause major economic and ecological problems, especially in the summer months (dry season). Although, in the dystopian future, humanity has developed a weapon that can: predict the fire event, to detect in time, but also to predict the duration of it. The weapon is: Artificial Intelligence, specifically, through Machine Learning, which is one part of (AI). Consequently, this paper is briefly mentioned in several methods of machine learning: in predicting forest fires, and in early detections, submitting an overall review of present models. Our main overall objective is to venture into the novelty field: predict the duration of ongoing forest fires. Our contribution is offering a new way to manage forest fires, using accessible open data, available from the Hellenic Fire Service. In particular, we imported over 72.000 data, from a 10-year period (2014 – 2023) using machine learning techniques. The experimental and validation results were more than encouraging, with the Random Forest achieving the lower value on error range (8 -13%), that means it was (87 – 92%) accurate on the prediction of forest fire duration. In the end, there are some future directions in order to extend this research.

Keywords: Forest fires; Machine learning; Neural networks; Decision trees

Citation: Kopitsa, C.;Tsoulos, I.G.; Charilogis V., Stavrakoudis A. Predict the duration of forest fires using machine learning methods. *Journal Not Specified* **2024**, *1*, 0. https://doi.org/

Received:
Revised:
Accepted:
Published:

1. Introduction

Forests play an important role in the ecological balance [1] of our planet as well as in our everyday life [2]. However, these ecosystems are threatened by various risks, the most important of which are fires [3–5]. Forest fires destroy the forest ecosystem [6–8] and can have devastating effects on local economies [9,10], with a significant impact also on tourism development [11–13] as well as in human health [14–16].

Since the risks of fires are great, governments must take measures and review them in the direction of fire prevention by analyzing data collected from fires that have broken out in recent history [17–19]. Also, local authorities have used techniques for forest fire monitoring, such as small UAVs [20], usage of a monitoring system based on GPRS and ZigBee wireless network [21], the iForestFire system [22] etc. Merino et al. suggested an Unmanned Aircraft System (UAS) [23] for forest fire monitoring. Also, Aslan et al. proposed a system [24] of wireless sensor networks for forest fire detection and monitoring. Recently, Serna et al. suggested a distributed system for fire monitoring using wireless sensor networks [25].

During recent years, machine learning techniques have started to play an important role in the prevention and treatment of forest fires. For example, Dwiasnati and Devianto proposed the usage of various machine learning methods for the classification of forest fire areas [26]. Also, Pang et al. suggested the usage of a series of machine learning models to forest fire occurrence prediction in China [27]. Dampage et al. suggested a system of wireless sensor networks with data handled by machine learning models for the detection of forest fires [28]. Shao et al. proposed a mapping of China's forest fire risks using a series of machine learning models [31]. A parallel SVM model is also suggested by Singh et al. [29] for forest fire prediction on data collected from India and Portugal. A survey on machine learning models used for forest fire prediction can be found in the work of Abid [30].

In addition, image processing has been established as a fire detection method. In this direction, a multitude of techniques have been presented that also take advantage of machine learning methods, such as the work of Vicente and Guillemant that presented a method for early smoke source detection [32]. Also, Yan et al. proposed a method [33] that combined image processing techniques and neural networks for forest fire recognition. Mubarak et al. suggested a rule - based image processing algorithm [34] for forest fire detection. Convolution neural networks were utilized in the work of Wang et al. [35] for forest fire image recognition. Also, wavelet analysis was used in the work of Jiao et al. [36] for forest fire detection. Also, Jain et al. underlines that the field of ML has undergone an explosion of new algorithmic advances in recent years and is deeply connected to the broader field of artificial intelligence (AI) [37]. From this research paper the diagram of Figure 1 was obtained to outline the incorporation of machine learning methods in wildfire management. ML models, a subset of artificial intelligence, harness the power of data and algorithms to learn from past experiences [38]. In the context of wildfire management, these models are invaluable by analyzing historical data and leveraging it to create predictive models capable of forecasting the spread of future fires. Although more complex than their statistical and physical counterparts, these models stand out for their ability to incorporate a broad array of variables. Understanding the requirements of ML models is crucial, particularly in terms of the variables they depend on for accurate predictions [39].

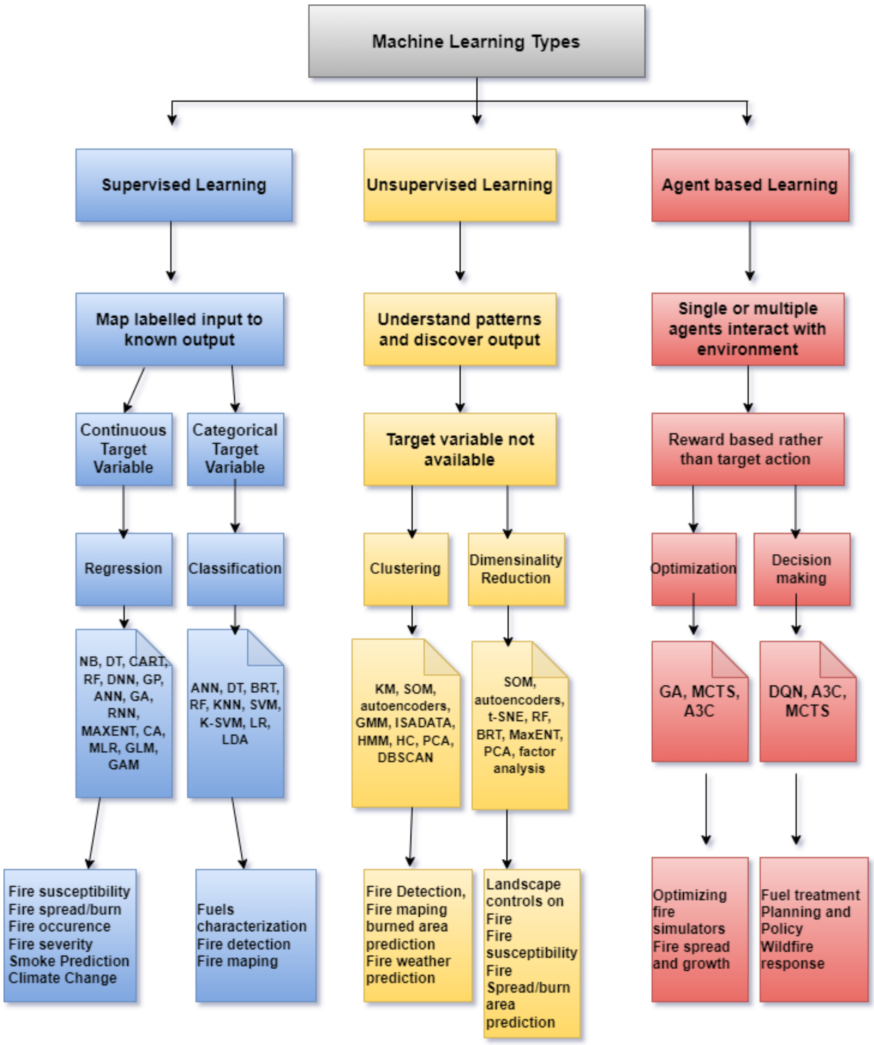

Figure 1. Machine learning techniques used in wildfire management.

This research work focuses on the use of machine learning techniques to predict the duration of forest fires, which occurred in Greece from 2014 until 2023. The data was collected by the Hellenic Fire Service and then, after clearing missing records, the data was digitized and one of three categories was assigned to every pattern: fires of short duration, fires of medium duration and fires of long duration. Figure 2 shows the burned areas in Greece per year for the last 10 years, according to the Global Wildfire Information System (GWIS).

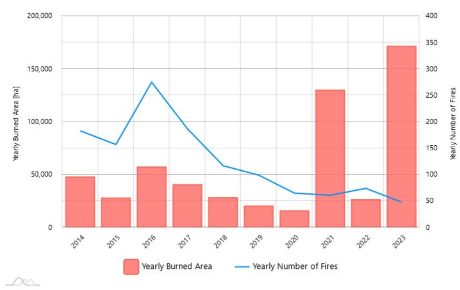

Figure 2. Burned area in Greece for the last ten years.

Farid et al. pointed out the vulnerability of the Greek ecosystem (pine forest), to extent forest fires, along with other Mediterranean countries, such as Spain, Portugal, Italy and France [40]. The fires in the Mediterranean region have become very intense, and dangerous, with scientists reporting that we have: a sixth generation mega fire clearly linked to global change. This new type of fire broke out, for the first time in Portugal, and Spain in 2017, with over 120 deaths, and the next year in Greece, with 104 deaths. The common characteristics of the new type of fire are: extreme, uncontrollable, lethal as indicated in the WWF report of 2019, available from http://awsassets.panda.org/downloads/wwf__the_mediterranean_burns_2019_eng_final.pdf. In order to be better prepared from fire hazards, this paper will estimate better forest fire management, through the use of machine learning methods. The prediction of the duration of a fire is important as in this way, on the one hand, an estimate can be made of the expected damage that will be caused in the area, and on the other, the human resources required to extinguish the fire can be calculated. Similar works in this area include the work of Liang et al. that used the duration of a wildfire and the burnt area to determine the scale of wildfires using neural networks [41]. Also, KC et al.. proposed a Surrogate model [42] to model the size of a wildfire over time, using data collected from wildfires in Tasmania. Furthermore, Xi et al. proposed [43] the application of joint mixture models to model the duration and the size of wildfires. In this work, a number of machine learning models were used, which have been successfully tested on a wide range of problems in the modern literature. The purpose of these models is the satisfactory separation of the categories of the problem through stochastic techniques that adjust the parameters of the above models.

In this paper we review the ML applications in forest fires management. Our main overall objective is to improve brand awareness of ML methods among fire researchers and managers, and illustrate the open data that Hellenic Fire Services provide. Besides that, the current state - of - the - art method in forest fire duration, added benefit in the decision making regarding the fire fighting resources, for the firefighters. The US Forest management, underlines the importance of predictive services, that gives information to the fire managers in order to anticipate and determine the need in resources, such as: firefighters, engines, airplane, e.t.c. [44] Also, the European State Forest Association, addressing the challenge of forest fires requires a concerted effort that combines scientific research, practical management strategies, and strong community engagement [45]. Therefore, technology has become a valuable ally for the environmental sciences. On that, recent reviews demonstrate the increase, in the last ten years, in the application of ML models in the environmental sciences [46] and forest ecology [47]. The Canadian specialist in fire science, Jain,. also points out that for wildfire management, it is crucial a better predicting way. Consequently, there has been a growing interest in the use of machine learning (ML) methodologies in wildfire science and management in recent years [48] Forest fires & the management is a unique science field, with six domains problem, according to Jain et al: Fire detection, fuels characterization, and mapping. Climate change, fire weather. Fire susceptibility, occurrence, and risk. Fire behavior prediction. Fire effects. Fire management. In our paper, we are focusing on fire management, thus there appears to be few studies in this domain problem, according to Jain et al. Finley showed that Fire management is a type of risk management that aims to maximize fire benefits while minimizing costs and losses [49]. Fire management decisions are crucial on a variety of scales, including long-term strategic decisions about resource procurement and location control in large regions, medium-term tactical decisions about resource acquisition, relocation, or release during the fire season, and short-term real-time operational decisions regarding resource deployment and usage on specific events [48]. Xiao indicates that fire Management groups struggle to effectively respond in a limited amount of time. It would be wise to keep an eye out for potential large fires [50].

This paper presents a fast-decision model for predicting the duration of ongoing fires. As a result, this review will help practitioners and researchers in the wildfire community who are interested in using machine learning techniques by offering guidance and information. It will also give ML researchers the chance to find potential uses in the field of wildfire science and management.

The field of wildfire duration prediction is impoverished, as researchers tend to focus more on fire occurrence and early detection. The domain we are focusing on appears to have great potential in two ways: one for wildfire management and another for machine learning researchers. The objective of this paper is to introduce an innovative approach by incorporating the number of firefighters, vehicles, and aerial forces used in each of the more than 72,000 fire incidents as key data points. Our research stands out by focusing not only on the occurrence of fires but also on the critical role of human and material resources in managing them.The application of machine learning (ML), and especially the Random Forest algorithm, in our project has proven highly valuable, enabling us to accurately analyze and predict critical parameters such as fire duration, while also considering human and material resources. Our findings are reliable, offering substantial support for optimizing wildfire management strategies.Our contribution can enhance the understanding of both material needs and human resources for fighting a wildfire, at both the local level and the European level, through the European Civil Protection Mechanism.

The rest of this article is divided as follows: in section 2 the used dataset is described as well as the incorporated machine learning methods, in section 3 the experimental results are fully described and finally in section 4 some conclusions are discussed accompanied by some guidelines for future research.

Also, we have run an additional experiment where a random noise

2. Materials and Methods

This section presents the datasets that will be used in the experiments as well as the machine learning techniques that will be applied to these datasets.

2.1. The used datasets

In this research work, open data was used which is available from the Hellenic Fire Service at the relevant link https://www.fireservice.gr/en_US/synola-dedomenon(accessed on 2 October 2024). The data was obtained for the years 2014-2023 and data preprocessing techniques were applied before inputting the data into machine learning models. The data used for this paper, concerns the strengthening of the European transparency legislation 2013/37/EE. Therefore, the data are neither type nor location biased, and concern all fires in the Greek (Hellenic) area. The information provided from the Hellenic Fire Service, are easily accessible, allow for analysis, updated, accurate, and includes all participating parties.

The initial datasets contained both numerical and alphanumeric information. For example, they included data on the area where the forest fire occurred, as well as information about the fire station that participated in the suppression efforts. Therefore, the first step in data preprocessing was the digitization of the columns containing alphanumeric data, specifically those with numerical information. This involved replacing categorical data with discrete integer values, enabling their use by machine learning algorithms that require numerical input. The next crucial step in data preprocessing involved handling missing values. Specifically, records that contained missing values in important features, such as climatic data or other relevant variables, were removed from the dataset. This typically occurred when a value was unavailable at the time of recording, which could lead to biased or unreliable results. Additionally, records with a fire duration of zero were excluded, as they were considered unrealistic and inappropriate for analysis. To define the output category, the duration of the forest fire was converted from hours or other time units into minutes, providing greater precision in classification. Subsequently, three distinct categories were created based on the logarithmic value of the fire duration in minutes. This logarithmic transformation allowed for better management of the large variations in fire duration, ensuring that both shorter and longer fires were appropriately considered without overemphasizing extremely large values. The three resulting categories were used as target values for the execution of experiments, enabling the classification of forest fires based on their duration. In the present work and for the Greek data on forest fires, the following fire classification has been used

1. Up to 6 hours is considered to be a fire of short duration
2. Up to 2 days is a forest fire of medium duration
3. From 2 days or more it is considered a long duration fire.

In this way, the data were adequately prepared to be analyzed using machine learning methods, ultimately achieving greater accuracy in predicting the duration of forest fires. The preprocessing steps are graphically illustrated in Figure 3.

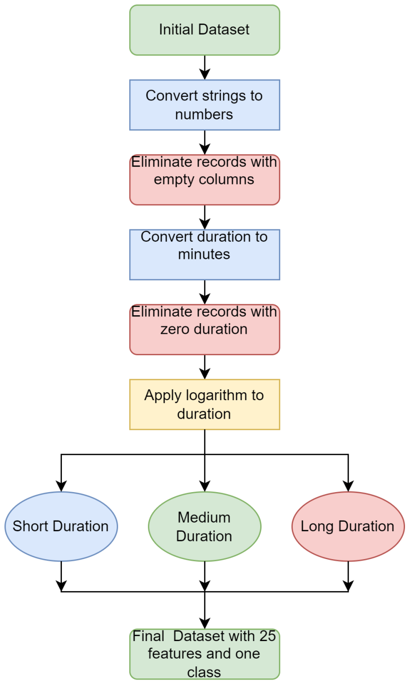

Figure 3. The steps of the preprocessing that were applied on the original datasets.

Having performed the previously mentioned preprocessing steps, the final datasets contain 25 features and the following information about the forest fires:

1. Fire department.
2. Province.
3. Season.
4. Burnt area: forest area, grove, grasslands, reeds/swamps, agricultural lands, cover crop, garbage dumps.
5. Personnel: Firefighters, volunteers, army, etc.
6. Vehicles: firefighting, tanks, etc.
7. Aerial means: helicopters and other aircrafts.

A schematic representation of the used dataset is outline in Figure 4.

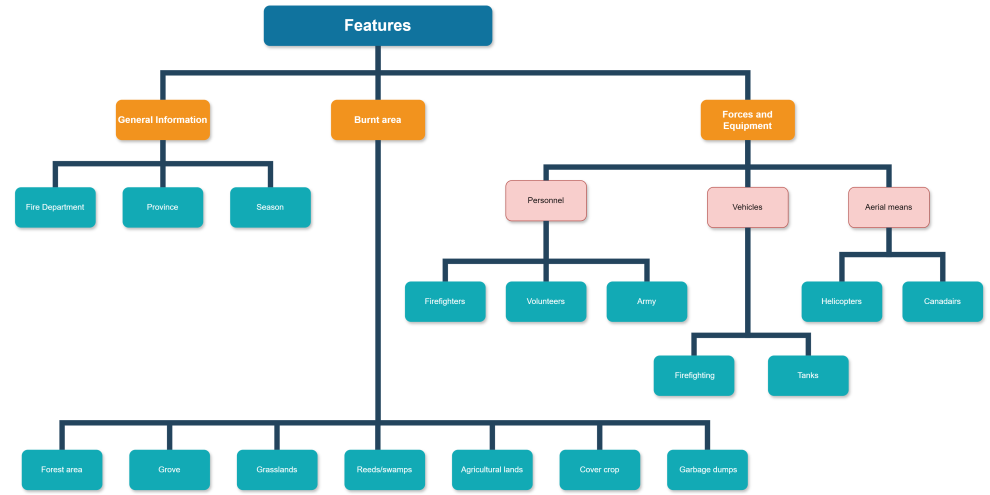

Figure 4. The used dataset after the preprocessing steps.

2.2. The used machine learning methods

A number of machine learning techniques were used to efficiently find classes in the datasets of the previous subsection. These techniques cover a wide range of techniques available in the field of machine learning and are presented in more detail below.

2.2.1. Bayesian Networks

Bayesian networks are probabilistic models based on direct acyclic graphs [52,53] and they have been applied with success in various cases. For example, Friedman et al. used Bayesian Networks to analyze expression data [54]. Also, Cai et al. used Bayesian Networks in fault diagnosis [55] and Barton et al. proposed the use of Bayesian Networks to environmental problems [56]. In the case of forest fires, Bayesian Networks have been used in many cases, such as to predict and analyze possible fire causes [57]. The study was conducted in Mugla of Turkey. Also, Bayesian networks were used to model the cascading impacts of drought and forest fire in a recent study [58]. Also, Bayesian Networks were combined with deep learning for detection of fires from video frames [59].

2.2.2. Naïve Bayes

The Naïve Bayes is a supervised machine learning algorithm, used for classification tasks. This classifier, uses principles of probability in order to perform classification tasks [60,61]. This algorithm has been incorporated in many research areas, such as document classification [62], traffic risk management [63], network intrusion detection [64] etc. Also, the Naive Bayes has been used in forest fire issues in a series of papers. For example, Nugroho et al. proposed a system for forest fire prevention using a combination of a wireless sensor network and a Naïve Bayes classifier [65]. A classification of hotspots causing forest fires using the Naive Bayes algorithm is proposed in the work of Zainul et al. [66]. Karo et al. proposed a methodology to classify wildfires using feature selection and the Naive Bayes among other machine learning methods [67]. Also, a variant of the Naïve Bayes Algorith was suggested by Shu et al. for forest fire prediction [68].

2.2.3. Logistic Regression

Like the previously mentioned algorithms, Logistic Regression, works also with machine learning classification and it can be considered as a data analysis technique used to predict probabilities [69]. Cabrera proposed the Logistic Regression for higher school decisions [70]. Also, Lawson et al. proposed the usage of Logistic Regression method to analyze customer satisfaction data [71]. Hu and Lo used the Logistic Regression technique to model urban growth in their paper [72]. This method has been used also in a series of issues involving forest fires, such as human - caused wildfire risk estimation [73], prediction of wildfire vulnerability [74], probabilistic modeling of wildfire occurrence [75], analysis of wildfire danger [76]etc.

2.2.4. Artificial neural networks

Artificial neural networks (ANNs) are parametric models [77,78], where a set of parameters, commonly called weights, must be calculated to be adapted to classification or regression data. This machine learning model has been utilized in a variety of scientific and real - world problems, such as physics problems [79–81], solving differential equations [82,83], solar radiation prediction [84], agriculture problems [85,86], problems appeared in chemistry [87–89], wind speed forecasting [90], economics problems [91–93], problems related to medicine [94,95] etc.

In the area of forest fire prediction and observation, a number of works using artificial neural networks have been published. Hossain et al. used ANNs to detect flames and smoke from static image features [96]. Lall and Mathibela utilized neural networks to predict the risk of wildfires in the city of Cape Town [97]. Also, Sayad et al. used neural networks among other machine learning techniques for predictive modeling of wildfires from data collected from NASA's Land Processes Distributed Active Archive Center (LP DAAC) [98]. Artificial neural networks and meteorological data were used in the work of Liang et al. to predict the scale of wildfires [99]. Also, a case study for predicting wildfires for a Chinese province using neural networks was published recently by Gao et al. [100].

2.2.5. The J48 algorithm

The J48 algorithm [101] is one of the most used supervised machine learning algorithms, used to construct decision trees for classification data. This method was tested on a series of classification problems, such as prediction of diabetes [102], network intrusion detection [103], classification of criminal data [104], fingerprint gender classification [105], fake news classification [106] etc. Also, the J48 algorithm was used to predict forest fires using data from Slovenia in a recent work [108]. A similar study was performed in Algeria using the J48 algorithm among other machine learning models [108].

2.2.6. Random Forests

Random Forest [109,110] is a popular supervised machine learning algorithm, used to construct decision trees for classification problems. The method of Random Forests has proven its adaptability and effectiveness in a number of difficult problems, such as remote sensing classification [111], ecology issues [112], bionformatics [113], text categorization [114], network intrusion detection [115] etc. Moreover, random forest was incorporated for forest fire prediction, such as in the work of Latifah et al., where random forests were applied to predict forest fires in Borneo [116]. Also, Malik et al. proposed the usage of Random Forests for wildfire risk prediction in Northern California [117]. Also, Gao et al. performed a forest fire risk prediction [118] in China using a combination of Random Forests and a neural network trained with the Back Propagation method [119].

3. Results

The experiments were conducted using the freely available programming tool of WEKA [120]. The software, which is written in the JAVA programming language to be portable, can be downloaded freely from https://ml.cms.waikato.ac.nz/weka/(accessed on 14 September 2024) or it can be found in the repositories of most Linux systems. The WEKA software is a collection of machine learning and data analysis tools and it contains also some visualization tools for modeling. The WEKA has been used with success in many cases, such as educational problems [121,122], medical problems [123,124] etc. The validation of the conducted experiments was performed using the ten - fold cross validation technique. The experiments were carried out on an AMD Ryzen 5950X with 128GB of RAM, running the Debian Linux operating system. The experimental results using the methods mentioned in the previous section and the 10 modified datasets from the Hellenic Fire Service are listed in Table 1. The following applies to the tables of experimental results:

1. The numbers in cells denote average classification error as calculated on the test set.
2. The column YEAR denotes the year where the machine learning methods were applied.
3. The column BAYESNET stands for the application of the Bayesian Network method.
4. The column NAIVEBAYES denotes the application of the Naïve Bayes algorithm.
5. The column LOGISTIC represents the application of the Logistic Regression algorithm.
6. The column MLP denotes the application of a neural network to the dataset.
7. The column J48 denotes the application of the J48 method to the forest fire data.
8. The column RANDOMFOREST denotes the usage of the Random Forest method to the data.
9. The row AVERAGE denotes the average classification error for all datasets.

Table 1. Experimental results using various machine learning models for 10 years of observations. The numbers in cells denote average classification error as measured on the test set.

YEAR	BAYESNET	NAIVEBAYES	LOGISTIC	MLP	J48	RANDOMFOREST
2014	11.44%	12.89%	9.81%	11.37%	10.04%	9.42%
2015	11.08%	11.26%	9.53%	10.65%	9.51%	8.95%
2016	25.71%	13.00%	3.41%	3.90%	3.65%	3.00%
2017	11.04%	11.51%	9.48%	10.08%	10.30%	9.29%
2018	11.20%	10.46%	9.09%	9.48%	9.27%	8.58%
2019	9.61%	9.25%	8.29%	8.53%	9.08%	8.01%
2020	18.00%	6.72%	5.54%	5.97%	6.09%	5.50%
2021	12.35%	14.15%	12.04%	13.59%	13.59%	11.92%
2022	10.25%	9.62%	9.01%	9.47%	9.04%	8.93%
2023	9.74%	9.19%	8.26%	8.77%	8.39%	7.66%
AVERAGE	**13.04%**	**10.81%**	**8.45%**	**9.18%**	**8.90%**	**8.13%**

Judging from the experimental results it is evident that the Random Forest technique excellently outweighs the others along with the Logistic Regression technique. This observations is reinforced from the box plot of Figure 5.

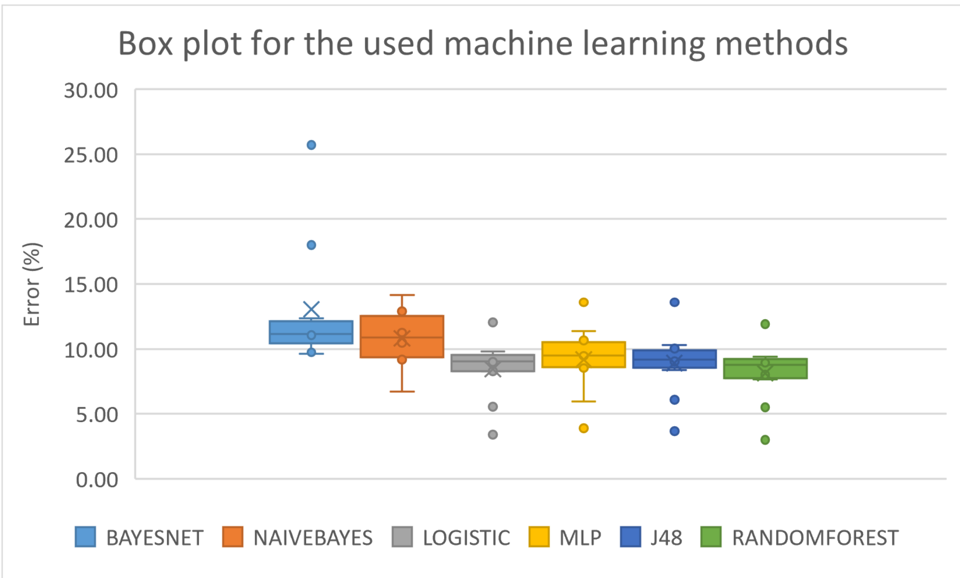

Figure 5. Box plot for the used machine learning techniques.

Also the precision and recall measures for every dataset and for each method are presented in Table 2.

	BAYES NET		NAIVEBAYES		LOGISTIC		MLP		J48		FOREST	
YEAR	PRECISION	RECALL	PRECISION	RECALL	PRECISION	RECALL	PRECISION	RECALL	PRECISION	RECALL	PRECISION	RECALL
2014	0.889	0.886	0.851	0.871	0.89	0.902	0.87	0.886	0.888	0.90	0.898	0.906
2015	0.892	0.889	0.87	0.887	0.891	0.905	0.876	0.893	0.892	0.905	0.901	0.91
2016	0.959	0.743	0.959	0.87	0.96	0.966	0.955	0.961	0.959	0.963	0.968	0.97
2017	0.897	0.89	0.869	0.885	0.893	0.905	0.886	0.889	0.886	0.897	0.899	0.907
2018	0.897	0.888	0.879	0.895	0.894	0.909	0.89	0.905	0.894	0.907	0.905	0.914
2019	0.914	0.904	0.894	0.907	0.903	0.917	0.903	0.915	0.898	0.909	0.912	0.92
2020	0.929	0.82	0.923	0.933	0.937	0.945	0.933	0.94	0.931	0.939	0.94	0.945
2021	0.879	0.876	0.835	0.858	0.865	0.88	0.846	0.864	0.849	0.864	0.871	0.881
2022	0.91	0.897	0.889	0.904	0.893	0.91	0.891	0.905	0.896	0.91	0.9	0.911
2023	0.912	0.903	0.894	0.908	0.905	0.917	0.899	0.912	0.906	0.916	0.916	0.923

Table 2. Precision and recall for every machine learning method.

The statistical comparison of the Random Forest with the other machine learning methods is depicted in Figure 6.

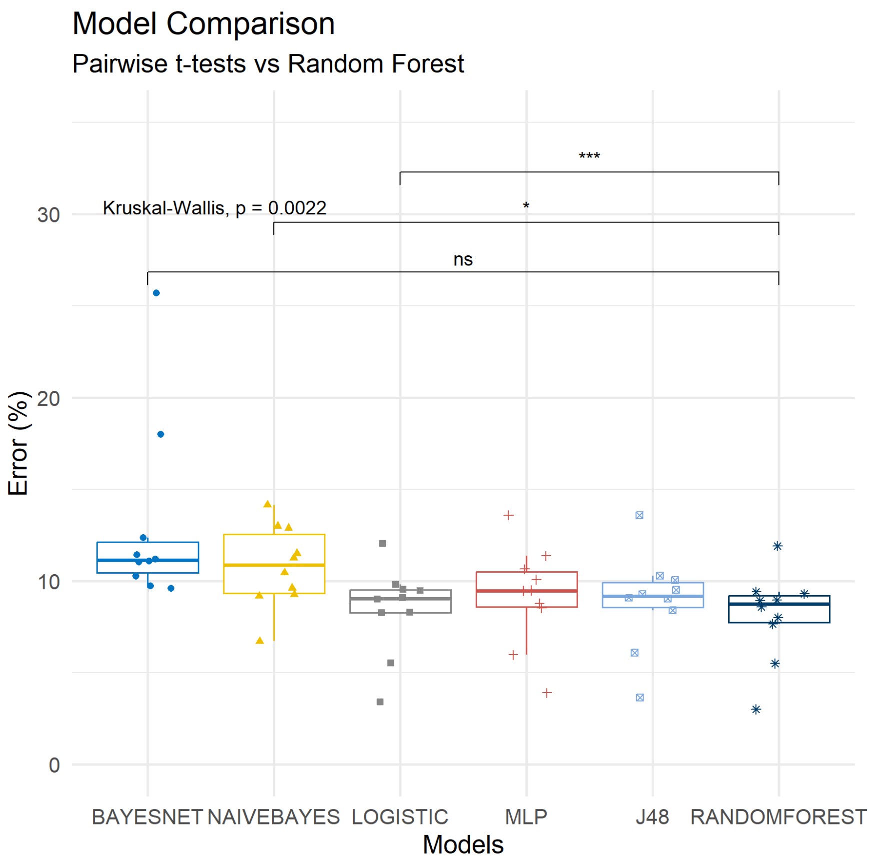

Figure 6. Statistical comparison between the Random Forest method and the other machine learning methods.

The Kruskal Wallis test was used because we had many groups to compare, specifically six different tree counts, making it necessary to use a test that could compare more than two groups simultaneously. In the statistical visualization of Figure 6, the BAYESNET model shows higher error rates in most years, with an exceptionally high error in 2016, reaching 25.71%. However, in other years, such as 2019 and 2023, its performance was quite close to RANDOMFOREST, although the latter remains slightly better. Although NAIVEBAYES performs relatively well in several years, it still shows higher errors compared to RANDOMFOREST, especially during the 2020-2021 period, where NAIVEBAYES had significantly increased errors. LOGISTIC records exceptionally low error rates in 2016, but remains consistently below RANDOMFOREST in most years. An exception is in 2020, where a divergence is observed, with better performance for RANDOMFOREST. MLP shows some fluctuations in errors, with significant improvement after 2016. Despite better performance in certain years, such as 2021, it remains generally inferior compared to Random Forest. J48 has fairly comparable error rates with RANDOMFOREST, especially after 2018, but RANDOMFOREST consistently proves to be the most efficient model in most years. RANDOMFOREST consistently emerges as the best model based on error rates, recording lower errors in most years compared to other machine learning models. Although other models, such as LOGISTIC and MLP, perform well in certain years, RANDOMFOREST maintains a more stable and reliable performance with fewer fluctuations.

Furthermore, in order to evaluate the performance and the effectiveness of the Random Forest technique, an additional test was carried out where the number of trees for this method increased from 10 to 200. The experimental results are outlined in Table 3.

YEAR	**10 trees**	**20 trees**	**50 trees**	**100 trees**	**150 trees**	**200 trees**
2014	10.05%	9.66%	9.59%	9.43%	9.24%	9.25%
2015	9.90%	9.44%	9.16%	8.95%	9.07%	9.00%
2016	3.34%	3.15%	2.97%	3.00%	3.01%	2.99%
2017	10.07%	9.54%	9.43%	9.30%	9.36%	9.25%
2018	9.27%	8.76%	8.52%	8.58%	8.64%	8.67%
2019	8.81%	8.18%	7.95%	8.02%	7.98%	7.89%
2020	6.11%	5.78%	5.53%	5.51%	5.58%	5.57%
2021	12.26%	11.67%	11.74%	11.92%	11.77%	11.77%
2022	9.43%	9.44%	9.14%	8.94%	8.86%	8.86%
2023	8.47%	7.85%	7.61%	7.67%	7.43%	7.43%

Table 3. Experimental results using different number of trees for the Random Forest technique.

Also, a statistical comparison for the previously mentioned results are shown in Figure 7.

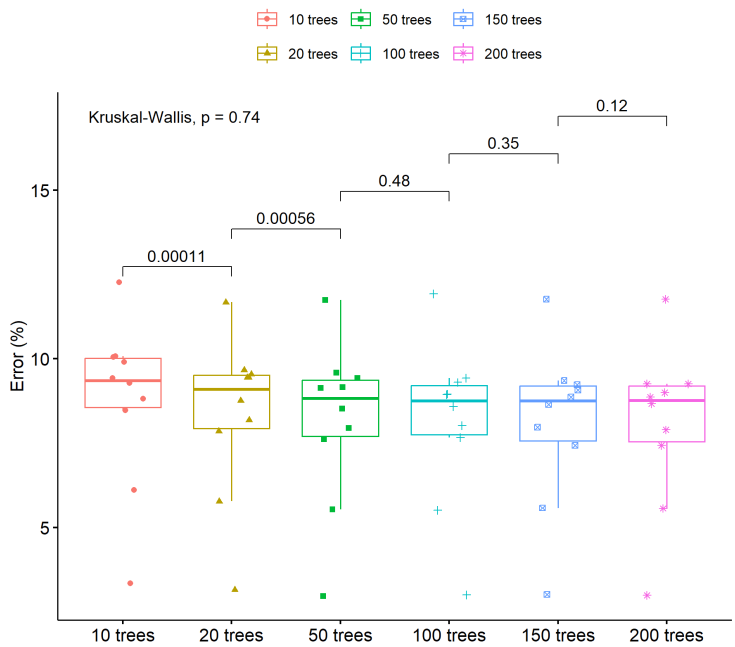

Figure 7. Statistical test for the experiment with different number of trees and the Random Forest technique.

In Table 3, it is observed that as the number of trees increases, the error decreases. This is expected in Random Forest models, as more trees typically improve the model's accuracy. For example, in 2014, the error decreases from 10.0477% for 10 trees to 9.253% for 200 trees, while in 2023, the error starts at 8.4684% for 10 trees and decreases to 7.4303% for 200 trees. In Figure 7, the paired comparisons using the t-test showed that for certain pairs, such as between 10 and 20 trees, the differences were statistically significant with p-values less than 0.05. This indicates that increasing from 10 to 20 trees results in a significant reduction in error. On the other hand, for larger numbers of trees, such as 150 and 200, the p-values were greater than 0.05, suggesting that the differences in errors are not statistically significant. This means that increasing the number of trees beyond a certain point (e.g., from 150 to 200) does not have a substantial impact on reducing the error. In conclusion, the analysis shows that increasing the number of trees in the Random Forest model leads to a reduction in error, especially for smaller tree counts, where the differences are statistically significant ($p < 0.05$). For larger tree counts, the differences in errors become smaller, and the p-values indicate that these differences are not statistically significant ($p > 0.05$).

In addition, to validate the experimental results, another comparison was made in which the Random Forest method participated as well as machine learning models, for the training of which the OPTIMUS optimization software was used. This software is freely available from https://github.com/itsoulos/GlobalOptimus/(accessed on 3 October 2024). The experimental results for this comparison are outlined in Table 4.

Table 4. Comparison of Random Forests against other machine learning models.

YEAR	RANDOM FOREST	SVM	MLP10_BFGS	MLP10_LBFGS	RBF10
2014	9.43%	14.19%	11.61%	10.28%	13.76%
2015	8.95%	13.08%	10.86%	9.71%	12.34%
2016	3.00%	5.03%	4.36%	3.61%	7.25%
2017	9.30%	13.20%	10.60%	9.77%	12.87%
2018	8.58%	11.50%	10.16%	9.15%	11.12%
2019	8.02%	10.67%	9.27%	8.26%	9.83%
2020	5.51%	8.18%	6.69%	5.62%	10.39%
2021	11.92%	17.98%	14.12%	12.71%	18.17%
2022	8.94%	11.16%	9.86%	8.98%	11.05%
2023	7.67%	11.37%	9.55%	8.63%	11.53%
AVERAGE	**8.13%**	**11.64%**	**9.71%**	**8.67%**	**11.83%**

The following notation is used this notation:

1. The column RANDOM FOREST denotes the results with the method Random Forest implemented by the WEKA package.
2. The column SVM denotes the application of the support vector machines (SVM) method [125] using the libsvm software package [126].
3. The column MLP10 BFGS stands for the results obtained by a artificial neural network trained by a BFGS optimization method as modified by Powell [127]. This neural network is equipped with 10 processing nodes.
4. The column MLP10 LBFGS represents the results obtained by a neural network with 10 processing nodes that was trained using the Limited Memory BFGS optimization method [128].
5. The column RBF10 denotes the results produced by the training of an Radial Basis Function (RBF) [129] network with 10 processing nodes.

Also, a statistical comparison for the previously presented results is outlined in Figure 8.

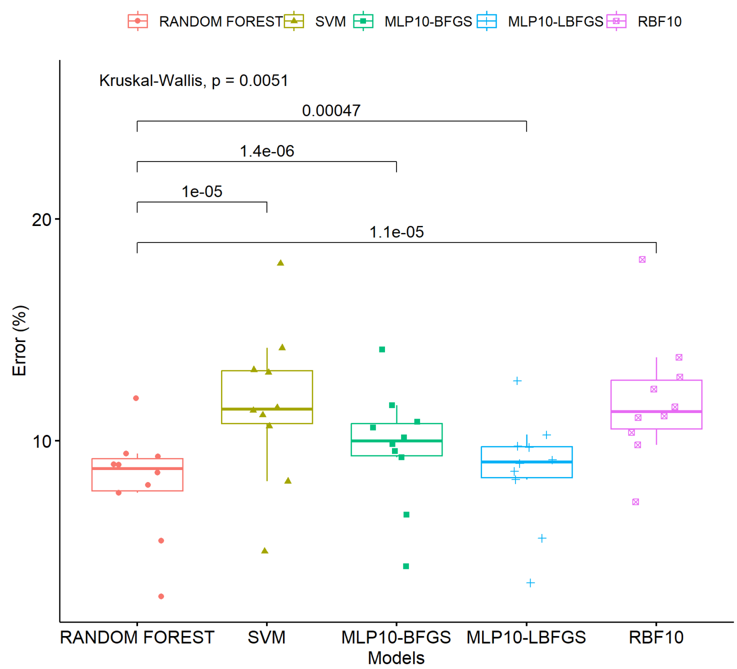

Figure 8. Statistical comparison for the methods: MLP, RBF and Random Forest used to predict the duration of fires.

In Table 4, we observe that the RBF10 model generally exhibits higher error rates compared to the other models, with greater variability from year to year. Its highest error occurs in 2021 at 18.17%, while its lowest is in 2016 at 7.25%. The MLP10-BFGS model, although showing some fluctuations, is more stable than RBF10 and has lower error rates in most years. Its lowest error is recorded in 2016 at 4.36%, while the highest is in 2021 at 14.12%. The MLP10-LBFGS model generally performs better than MLP10-BFGS, with slightly lower error rates in most years. Its lowest error is 3.61% in 2016, while the highest is in 2021 at 12.71%. Finally, the Random Forest model shows the best overall performance, with the lowest error recorded in 2016 (3%) and the highest in 2021 (11.92%). Overall, Random Forest consistently maintains lower error rates compared to the other models in nearly all years. Comparing the models using the t-test (Figure 8), the statistical analysis shows that the differences between the models are statistically significant, especially when comparing RBF10, MLP10-BFGS, and MLP10-LBFGS to Random Forest. The p-value is lower than the specified threshold ($p < 0.05$), confirming that Random Forest statistically outperforms the other models.

Furthermore, to measure the effectiveness of the artificial neural network on the proposed datasets one more experiment was conducted where the BFGS method was used to train a neural network, where the number of processing nodes was in the range $[2, 20]$. The results from this experiment are outlined in Table 5.

Table 5. Experiments with different number of processing nodes for the artificial neural network case. The BFGS optimization method was used to train the neural network.

YEAR	MLP2 BFGS	MLP5 BFGS	MLP10_BFGS	MLP15_BFGS	MLP20 BFGS
2014	12.88%	12.40%	11.61%	11.58%	11.43%
2015	11.78%	10.98%	10.86%	12.71%	12.66%
2016	5.08%	4.40%	4.36%	4.13%	4.14%
2017	12.14%	10.71%	10.60%	10.80%	10.46%
2018	10.46%	10.28%	10.16%	10.04%	9.90%
2019	9.61%	9.21%	9.27%	9.12%	9.01%
2020	7.49%	6.67%	6.69%	6.36%	6.40%
2021	16.29%	14.82%	14.12%	14.05%	13.95%
2022	10.39%	10.06%	9.86%	9.67%	9.63%
2023	10.51%	9.96%	9.55%	9.45%	9.47%
AVERAGE	**10.66%**	**9.95%**	**9.71%**	**9.79%**	**9.71%**

From the generated results it is evident that the artificial neural network model shows improved performance when the nodes increase from 2 to 5 or to 10 but from this point onwards there is no noticeable difference.

4. Conclusions

In the present research work, a study was made of the duration of forest fires using open data for the Greek area. This data contained information such as the area of the fire, the time it erupted, the destruction it caused as well as the human resources involved in extinguishing it. Machine learning models were then used to estimate the duration of a fire. The successful prediction of the duration contributes to the proper management of such catastrophic events by the state mechanisms. The machine learning models used included models such as artificial neural networks, decision trees, etc. Most of the machine learning techniques used achieved significantly low error values for each year of experimental data. On average, these classification error values are in the range of 8-13% with the Random Forest technique achieving the lowest value. Furthermore, additional tests were executed to measure the effectiveness of random forests and neural networks. In addition, the stability of the used data was assessed in the presence of random noise with encouraging results.

The present work could be extended in the future in various research directions, such as:

1. Incorporation of more machine learning methods from the relevant literature.
2. Use feature selection or construction techniques from recent literature to identify the most important factors influencing the classification process.
3. Usage of methods that create classification rules, in order to discover any hidden relationships between the data and the classes of the datasets.
4. Usage of data that also includes meteorological data, in order to identify a possible correlation of the categories with the meteorological conditions that prevailed at the time of the fire.
5. Parallel programming techniques may be incorporated to speed up the optimization process, such as MPI [130] or the OpenMP library [131].

Author Contributions: C.K., V.C. and I.G.T. conceived of the idea and the methodology, and C.K. and V.C. implemented the corresponding software. C.K. conducted the experiments, employing objective functions as test cases, and provided the comparative experiments. A.S. performed the necessary statistical tests. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: This research has been financed by the European Union: Next Generation EU through the Program Greece 2.0 National Recovery and Resilience Plan, under the call RESEARCH–CREATE–INNOVATE, project name "iCREW: Intelligent small craft simulator for advanced crew training using Virtual Reality techniques" (project code: TAEDK-06195).

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Z. Qiang, E.Z. Meka, R.C. Anderson, Y. Kakabadse, Forests nature at your service. UNEP report. The magazine of the United Nations Environment Program, 2011.
2. A.S. Mori, K.P. Lertzman, L. Gustafsson, Biodiversity and ecosystem services in forest ecosystems: a research agenda for applied forest ecology, Journal of Applied Ecology **54**, pp. 12-27, 2017.
3. B.J. Stocks, Mason, J. A., Todd, J. B., Bosch, E. M., Wotton, B. M., Amiro, B. D., ... & Skinner, W. R. (2002). Large forest fires in Canada, 1959–1997. Journal of Geophysical Research: Atmospheres, 107(D1), FFR-5.
4. M.D. Flannigan, B.D. Amiro, K.A. Logan, B.J. Stocks, B.M. Wotton, Forest fires and climate change in the 21 st century. Mitigation and adaptation strategies for global change **11**, pp. 847-859, 2006.
5. O. Sahar, Wildfires in Algeria: problems and challenges. IFOREST **8**, pp. 818–826, 2015.
6. G. Certini, Effects of fire on properties of forest soils: a review, Oecologia **143**, pp. 1–10, 2005.
7. G.R. Van Der Werf, J.T. Randerson, G.J. Collatz, L. GIGLIO, Carbon emissions from fires in tropical and subtropical ecosystems. Global Change Biology, 9: 547-562, 2003.
8. A.A. Agbeshie, S. Abugre, T. Atta-Darkwa et al, A review of the effects of forest fire on soil properties, J. For. Res. **33**, pp. 1419–1441, 2022.
9. P. Aleksić, M. Krstić , G. Jančić, Forest fi res - ecological and economic problem in Serbia, Botanica Serbia **32**, pp. 169-176, 2009.
10. D. Wang, D. Guan, S. Zhu. et al, Economic footprint of California wildfires in 2018, Nat Sustain **4**, pp. 252–260, 2021.
11. P. W. Hystad, P.C. Keller, Towards a destination tourism disaster management framework: Long-term lessons from a forest fire disaster, Tourism Management **29**, pp. 151-162, 2008.
12. G. Boustras, N. Boukas, Forest fires' impact on tourism development: a comparative study of Greece and Cyprus, Management of Environmental Quality **24**, pp. 498-511, 2013.
13. V. Otrachshenko, L.C. Nunes, Fire takes no vacation: impact of fires on tourism, Environment and Development Economics **27**, pp. 86-101, 2022.
14. N. Sastry, Forest fires, air pollution, and mortality in Southeast Asia, Demography **39**, pp. 1–23, 2002.
15. E. Frankenberg, D. McKee, D. Thomas, Health consequences of forest fires in Indonesia, Demography **42**, pp. 109–129, 2005.
16. D. Bowman, G. Williamson, J. Abatzoglou. et al, Human exposure and sensitivity to globally extreme wildfire events, Nat Ecol Evol **1**, 0058, 2017.
17. M. Zhong, W. Fan, T. Liu, P. Li, Statistical analysis on current status of China forest fire safety, Fire Safety Journal **38**, pp. 257-269, 2003.

18. D. Avila-Flores, M. Pompa-Garcia, X. Antonio-Nemiga. et al., Driving factors for forest fire occurrence in Durango State of Mexico: A geospatial perspective, Chin. Geogr. Sci. **20**, pp. 491–497, 2010.
19. R. Lovreglio, V. Leone, P. Giaquinto, A. Notarnicola, Wildfire cause analysis: four case-studies in southern Italy. iForest **3**, pp. 8-15, 2010.
20. D. W. Casbeer, R. W. Beard, T. W. McLain, Sai-Ming Li and R. K. Mehra, "Forest fire monitoring with multiple small UAVs," Proceedings of the 2005, American Control Conference, 2005., Portland, OR, USA, 2005, pp. 3530-3535 vol. 5, doi: 10.1109/ACC.2005.1470520.
21. Guozhu Wang, J. Zhang, Wenbin Li, Dongxu Cui and Ye Jing, "A forest fire monitoring system based on GPRS and ZigBee wireless sensor network," 2010 5th IEEE Conference on Industrial Electronics and Applications, Taichung, 2010, pp. 1859-1862, doi: 10.1109/ICIEA.2010.5515417.
22. M. Stula, D. Krstinic, L. Seric, Intelligent forest fire monitoring system, Inf Syst Front **14**, pp. 725–739, 2012.
23. L. Merino, E. Caballero, J.R. Martínez-de-Dios al., An Unmanned Aircraft System for Automatic Forest Fire Monitoring and Measurement, J Intell Robot Syst **65**, pp. 533–548, 2012.
24. Y.E. Aslan, I. Korpeoglu, Ö. Ulusoy, A framework for use of wireless sensor networks in forest fire detection and monitoring, Computers, Environment and Urban Systems **36**, pp. 614-625, 2012.
25. M.Á. Serna, R. Casado, A. Bermúdez, N. Pereira, S. Tennina, Distributed Forest Fire Monitoring Using Wireless Sensor Networks, International Journal of Distributed Sensor Networks **11**, 10, 2015.
26. S. Dwiasnati, Y. Devianto, Classification of forest fire areas using machine learning algorithm, World Journal of Advanced Engineering Technology and Sciences, **3**, pp. 008-015, 2021.
27. Y. Pang, Y. Li, Z. Feng, Z. Feng, Z. Zhao, S. Chen, H. Zhang, Forest Fire Occurrence Prediction in China Based on Machine Learning Methods, Remote Sensing **14**, 5546, 2022.
28. U. Dampage, L. Bandaranayake, R. Wanasinghe et al, Forest fire detection system using wireless sensor networks and machine learning, Sci Rep **12**, 46, 2022.
29. K.R. Singh, K.P. Neethu, K. Madhurekaa, A. Harita, P. Mohan, Parallel SVM model for forest fire prediction, Soft Computing Letters **3**, 100014, 2021.
30. F. Abid, A survey of machine learning algorithms based forest fires prediction and detection systems, Fire technology **57**, pp. 559-590, 2021.
31. Y. Shao, Z. Feng, L. Sun, X. Yang, Y. Li, B. Xu, Y. Chen, Mapping China's Forest Fire Risks with Machine Learning, Forests **13**, 856, 2022.
32. J. Vicente, P. Guillemant, An image processing technique for automatically detecting forest fire, International Journal of Thermal Sciences **41**, pp. 1113-1120, 2002.
33. Q. Yan, P. Bo, Z. Juanjuan, Forest Fire Image Intelligent Recognition based on the Neural Network. Journal of Multimedia **9**, pp. 469-475, 2014.
34. Mahmoud, Mubarak A. I., Ren, Honge, Forest Fire Detection Using a Rule-Based Image Processing Algorithm and Temporal Variation, Mathematical Problems in Engineering, 2018, 7612487, 8 pages, 2018.
35. Y. Wang, L. Dang, J. Ren, Forest fire image recognition based on convolutional neural network, Journal of Algorithms & Computational Technology **13**, 2019.
36. Z. Jiao, Y. Zhang, J. Xin, Y. Yi, D. Liu and H. Liu, "Forest Fire Detection with Color Features and Wavelet Analysis Based on Aerial Imagery," 2018 Chinese Automation Congress (CAC), Xi'an, China, 2018, pp. 2206-2211, doi: 10.1109/CAC.2018.8623473.
37. Jain, P., Coogan, S. C., Subramanian, S. G., Crowley, M., Taylor, S., & Flannigan, M. D. (2020). A review of machine learning applications in wildfire science and management. Environmental Reviews, 28(4), 478-505.
38. Dogan, A., & Birant, D. (2021). Machine learning and data mining in manufacturing. Expert Systems with Applications, 166, 114060.
39. Singh, H., Ang, LM., Lewis, T. et al. Trending and emerging prospects of physics-based and ML-based wildfire spread models: a comprehensive review. J. For. Res. 35, 135 (2024).
40. A. Farid, M.K. Alam, V.S.N.S. Goli, I.D. Akin, T. Akinleye, X. Chen, Q. Cheng, P. Cleall, S. Cuomo, V. Foresta et al, A Review of the Occurrence and Causes for Wildfires and Their Impacts on the Geoenvironment, Fire **7**, 295, 2024.
41. H. Liang, M. Zhang, H. Wang, A Neural Network Model for Wildfire Scale Prediction Using Meteorological Factors, IEEE Access **7**, pp. 176746-176755, 2019.
42. U. KC, J. Aryal, J. Hilton, S. Garg, A Surrogate Model for Rapidly Assessing the Size of a Wildfire over Time, Fire **4**, 20, 2021.
43. D.D. Xi, C.B. Dean, S.W. Taylor, Modeling the duration and size of wildfires using joint mixture models. Environmetrics **32**, e2685, 2021.
44. US Forest Service. Department of Agriculture. Science and Technology. Managing Fire, Fire Forecasting. Available from https://www.fs.usda.gov/science-technology/managing-fire
45. eustafor. European State Forest Association. Forest Fires in Europe: a Growing Challenge. Towards a Resilient Future. Available from https://eustafor.eu/forest-fires-in-europe-a-growing-challenge/
46. Karpatne, A., Ebert-Uphoff, I., Ravela, S., Babaie, H.A., and Kumar, V. 2017. Machine learning for the geosciences: challenges and opportunities. Available from http://arxiv.org/abs/1711.04708.

47. Liu Z., Peng C., Work T., Candau J.N., DesRochers A., and Kneeshaw D. 2018. Application of machine-learning methods in forest ecology: recent progress and future challenges. Environ. Rev. 26(4): 339–350.
48. Jain P., Coogan S., Subramanian S.G., Crowley M., Taylor S., Flannigan M.D. 2020. A review of machine learning applications in wildfire science and management. Environ. Rev. 28(4): 478 - 505. Available from https://cdnsciencepub.com/doi/10.1139/er-2020-0019
49. Finney M.A. 2005. The challenge of quantitative risk analysis for wildland fire. For. Ecol. Manage. 211: 97–108.
50. Xiao H., 2023. Estimating Wildfire Duration using regression - models. Available from: https://arxiv.org/pdf/2308.08936
51. H.M. Hammouri, R.T. Sabo, R. Alsaadawi, K.A. Kheirallah, Handling Skewed Data: A Comparison of Two Popular Methods. Applied Sciences **10**, 6247, 2020.
52. Ben-Gal, I. (2008). Bayesian Networks. In Encyclopedia of Statistics in Quality and Reliability (eds F. Ruggeri, R.S. Kenett and F.W. Faltin).
53. Koski, T., & Noble, J. (2011). Bayesian networks: an introduction. John Wiley & Sons.
54. N. Friedman, M. Linial, I. Nachman, D. Pe'er, Using Bayesian networks to analyze expression data, In: Proceedings of the fourth annual international conference on Computational molecular biology, pp. 127-135, 2000.
55. B. Cai, L. Huang, M. Xie, Bayesian Networks in Fault Diagnosis, IEEE Transactions on Industrial Informatics **13**, pp. 2227-2240, 2017.
56. D.N. Barton, S. Kuikka, O. Varis, L. Uusitalo, H.J. Henriksen, M. Borsuk, A. de la Hera, R. Farmani, S. Johnson, J.D. Linnell, Bayesian networks in environmental and resource management. Integr Environ Assess Manag, **8**, pp. 418-429, 2012.
57. V. Sevinc, O. Kucuk, M. Goltas, A Bayesian network model for prediction and analysis of possible forest fire causes. Forest Ecology and Management **457**, 117723, 2020.
58. F. Chen, H. Jia, E. Du, Y. Chen, L. Wang, Modeling of the cascading impacts of drought and forest fire based on a Bayesian network, International Journal of Disaster Risk Reduction **111**, 104716,2024.
59. B. Kim, J. Lee, A Bayesian network-based information fusion combined with DNNs for robust video fire detection. Applied Sciences **11**, 7624, 2021.
60. Bayes, T. (1968). Naive bayes classifier. Article Sources and Contributors, 1-9.
61. G.I. Webb, E. Keogh, R. Miikkulainen, Naïve Bayes, Encyclopedia of machine learning **15**, pp. 713-714, 2010.
62. S.L. Ting, W.H. Ip, A.H. Tsang, Is Naive Bayes a good classifier for document classification, International Journal of Software Engineering and Its Applications **5**, pp. 37-46, 2011.
63. H. Chen, S. Hu, R. Hua, et al, Improved naive Bayes classification algorithm for traffic risk management, EURASIP J. Adv. Signal Process. 2021, 30, 2021.
64. M. Panda, M.R. Patra, Network intrusion detection using naive bayes. International journal of computer science and network security **7**, pp. 258-263, 2007.
65. A. Andi Nugroho, I. Iwan, K. Iroh Nur Azizah, F. Hakim Raswa, Peatland Forest Fire Prevention Using Wireless Sensor Network Based on Naïve Bayes Classifier, KnE Social Sciences **3**, pp. 20–34, 2019.
66. M. Zainul, E. Minggu, Classification of Hotspots Causing Forest and Land Fires Using the Naive Bayes Algorithm, Interdisciplinary Social Studies **1**, pp. 555-567, 2022.
67. I.M.K. Karo, S.N. Amalia, D. Septiana, Wildfires Classification Using Feature Selection with K-NN, Naïve Bayes, and ID3 Algorithms, Journal of Software Engineering, Information and Communication Technology (SEICT) **3**, pp. 15-24, 2022.
68. L. Shu, H. Zhang, Y. You, Y. Cui, W. Chen, Towards fire prediction accuracy enhancements by leveraging an improved naïve bayes algorithm, Symmetry **13**, 530, 2021.
69. S. Sperandei, Understanding logistic regression analysis, Biochemia medica **24**, pp. 12-18, 2014.
70. A.F. Cabrera, Logistic regression analysis in higher education: An applied perspective, Higher education: Handbook of theory and research **10**, pp. 225-256, 1994.
71. C. Lawson, D.C. Montgomery, Logistic Regression Analysis of Customer Satisfaction Data, Qual. Reliab. Engng. Int. **22**, pp. 971-984, 2006.
72. Zhiyong Hu and C.P. Lo, Modeling urban growth in Atlanta using logistic regression, Computers, Environment and Urban Systems **31**, pp. 667-688, 2007.
73. L. Vilar del Hoyo, M.P. Martín Isabel, F.J. Martínez Vega, Logistic regression models for human-caused wildfire risk estimation: analysing the effect of the spatial accuracy in fire occurrence data, Eur J Forest Res **130**, pp. 983–996, 2011.
74. P.P. de Bem, O.A. de Carvalho Júnior, E.A.T. Matricardi, R.F. Guimarães, R.A.T. Gomes, Predicting wildfire vulnerability using logistic regression and artificial neural networks: a case study in Brazil's Federal District. International journal of wildland fire **28**, pp. 35-45, 2018.
75. Eufrásio João Sozinho Nhongo, Denise Cybis Fontana, Laurindo Antonio Guasselli and Carolina Bremm, Probabilistic modelling of wildfire occurrence based on logistic regression, Niassa Reserve, Mozambique, Geomatics, Natural Hazards and Risk **10**, pp. 1772--1792, 2019.
76. W. Peng, Y. Wei, G. Chen, G. Lu, Q. Ye, R. Ding, P. Hu, Z. Cheng, Analysis of Wildfire Danger Level Using Logistic Regression Model in Sichuan Province, China. Forests **14**, 2352, 2023.
77. C. Bishop, Neural Networks for Pattern Recognition, Oxford University Press, 1995.

78. G. Cybenko, Approximation by superpositions of a sigmoidal function, Mathematics of Control Signals and Systems **2**, pp. 303-314, 1989.
79. P. Baldi, K. Cranmer, T. Faucett et al, Parameterized neural networks for high-energy physics, Eur. Phys. J. C **76**, 2016.
80. J. J. Valdas and G. Bonham-Carter, Time dependent neural network models for detecting changes of state in complex processes: Applications in earth sciences and astronomy, Neural Networks **19**, pp. 196-207, 2006
81. G. Carleo,M. Troyer, Solving the quantum many-body problem with artificial neural networks, Science **355**, pp. 602-606, 2017.
82. Y. Shirvany, M. Hayati, R. Moradian, Multilayer perceptron neural networks with novel unsupervised training method for numerical solution of the partial differential equations, Applied Soft Computing **9**, pp. 20-29, 2009.
83. A. Malek, R. Shekari Beidokhti, Numerical solution for high order differential equations using a hybrid neural network—Optimization method, Applied Mathematics and Computation **183**, pp. 260-271, 2006.
84. A. Kumar Yadav, S.S. Chandel, Solar radiation prediction using Artificial Neural Network techniques: A review, Renewable and Sustainable Energy Reviews **33**, pp. 772-781, 2014.
85. A. Topuz, Predicting moisture content of agricultural products using artificial neural networks, Advances in Engineering Software **41**, pp. 464-470, 2010.
86. A. Escamilla-García, G.M. Soto-Zarazúa, M. Toledano-Ayala, E. Rivas-Araiza, A. Gastélum-Barrios, Abraham,Applications of Artificial Neural Networks in Greenhouse Technology and Overview for Smart Agriculture Development, Applied Sciences **10**, Article number 3835, 2020.
87. Lin Shen, Jingheng Wu, and Weitao Yang, Multiscale Quantum Mechanics/Molecular Mechanics Simulations with Neural Networks, Journal of Chemical Theory and Computation **12**, pp. 4934-4946, 2016.
88. Sergei Manzhos, Richard Dawes, Tucker Carrington, Neural network-based approaches for building high dimensional and quantum dynamics-friendly potential energy surfaces, Int. J. Quantum Chem. **115**, pp. 1012-1020, 2015.
89. Jennifer N. Wei, David Duvenaud, and Alán Aspuru-Guzik, Neural Networks for the Prediction of Organic Chemistry Reactions, ACS Central Science **2**, pp. 725-732, 2016.
90. G. Li, J. Shi, On comparing three artificial neural networks for wind speed forecasting, Applied Energy **87**, pp. 2313-2320, 2010.
91. Lukas Falat and Lucia Pancikova, Quantitative Modelling in Economics with Advanced Artificial Neural Networks, Procedia Economics and Finance **34**, pp. 194-201, 2015.
92. Mohammad Namazi, Ahmad Shokrolahi, Mohammad Sadeghzadeh Maharluie, Detecting and ranking cash flow risk factors via artificial neural networks technique, Journal of Business Research **69**, pp. 1801-1806, 2016.
93. G. Tkacz, Neural network forecasting of Canadian GDP growth, International Journal of Forecasting **17**, pp. 57-69, 2001.
94. Igor I. Baskin, David Winkler and Igor V. Tetko, A renaissance of neural networks in drug discovery, Expert Opinion on Drug Discovery **11**, pp. 785-795, 2016.
95. Ronadl Bartzatt, Prediction of Novel Anti-Ebola Virus Compounds Utilizing Artificial Neural Network (ANN), Chemistry Faculty Publications **49**, pp. 16-34, 2018.
96. F. M. A. Hossain, Y. Zhang, C. Yuan and C. -Y. Su, "Wildfire Flame and Smoke Detection Using Static Image Features and Artificial Neural Network," 2019 1st International Conference on Industrial Artificial Intelligence (IAI), Shenyang, China, 2019, pp. 1-6, doi: 10.1109/ICIAI.2019.8850811.
97. S. Lall and B. Mathibela, "The application of artificial neural networks for wildfire risk prediction," 2016 International Conference on Robotics and Automation for Humanitarian Applications (RAHA), Amritapuri, India, 2016, pp. 1-6, doi: 10.1109/RAHA.2016.7931880.
98. Y. O. Sayad, H. Mousannif, H. Al Moatassime, Predictive modeling of wildfires: A new dataset and machine learning approach, Fire Safety Journal **104**, pp. 130-146, 2019.
99. H. Liang, M. Zhang, H. Wang, A Neural Network Model for Wildfire Scale Prediction Using Meteorological Factors, IEEE Access **7**, pp. 176746-176755, 2019.
100. K. Gao, Z. Feng, S. Wang, Using multilayer perceptron to predict forest fires in jiangxi province, southeast china, Discrete Dynamics in Nature and Society **1**, 6930812, 2022.
101. N. Bhargava, G. Sharma, R. Bhargava, M. Mathuria, Decision tree analysis on j48 algorithm for data mining. Proceedings of international journal of advanced research in computer science and software engineering **3**, 2013.
102. G. Kaur, A. Chhabra, Improved J48 classification algorithm for the prediction of diabetes, International journal of computer applications **98**, 22, 2014.
103. S. Sahu, B.M. Mehtre, Network intrusion detection system using J48 Decision Tree, In: 2015 International Conference on Advances in Computing, Communications and Informatics (ICACCI), pp. 2023-2026, 2015.
104. N.N. Sakhare, S.A. Joshi, Classification of criminal data using J48-Decision Tree algorithm. Int. J. Data Warehous. Min **4**, pp. 167-171, 2015.
105. S.F. Abdullah, A.F.N.A. Rahman, Z.A. Abas, W.H.M. Saad, Fingerprint gender classification using univariate decision tree (j48), International Journal of Advanced Computer Science and Applications **7**, 2016.
106. Jehad, R., & Yousif, S. A. (2020). Fake news classification using random forest and decision tree (j48). Al-Nahrain Journal of Science, 23(4), 49-55.
107. D. Stojanova, P. Panov, A. Kobler, S. Džeroski, K. Taškova, Learning to predict forest fires with different data mining techniques. In Conference on data mining and data warehouses (SiKDD 2006), Ljubljana, Slovenia, pp. 255-258, 2006.

108. F. Abid, N. Izeboudjen, Predicting forest fire in algeria using data mining techniques: Case study of the decision tree algorithm. In Advanced Intelligent Systems for Sustainable Development (AI2SD'2019) Volume 4-Advanced Intelligent Systems for Applied Computing Sciences (pp. 363-370). Springer International Publishing, 2020.
109. L. Breiman, Random forests, Machine learning **45**, pp. 5-32, 2001.
110. S.J. Rigatti, Random forest, Journal of Insurance Medicine **47**, pp. 31-39, 2017.
111. M. Pal, Random forest classifier for remote sensing classification. International Journal of Remote Sensing **26**, pp. 217–222, 2005.
112. D.R. Cutler, T.C. Edwards, K.H. Beard, A. Cutler, K.T. Hess, J. Gibson, J.J. Lawler, random forests for classification in ecology. Ecology **88**, pp. 2783-2792, 2007.
113. Qi, Y. (2012). Random Forest for Bioinformatics. In: Zhang, C., Ma, Y. (eds) Ensemble Machine Learning. Springer, New York, NY. https://doi.org/10.1007/978-1-4419-9326-7_11
114. B. Xu, X. Guo, Y. Ye, J. Cheng, An improved random forest classifier for text categorization. J. Comput., **7**, pp. 2913-2920, 2012.
115. N. Farnaaz, M.A. Jabbar, Random Forest Modeling for Network Intrusion Detection System, Procedia Computer Science **89**, pp. 213-217, 2016.
116. A. L. Latifah, A. Shabrina, I. N. Wahyuni, R. Sadikin, Evaluation of Random Forest model for forest fire prediction based on climatology over Borneo, In: 2019 International Conference on Computer, Control, Informatics and its Applications (IC3INA), Tangerang, Indonesia, pp. 4-8, 2019.
117. A. Malik, M.R. Rao, N. Puppala, P. Koouri, V.A.K. Thota, Q. Liu, S. Chiao, J. Gao, Data-Driven Wildfire Risk Prediction in Northern California. Atmosphere, **12**, 109, 2021.
118. C. Gao, H. Lin, H. Hu, Forest fire risk prediction based on random forest and backpropagation neural network of Heihe area in Heilongjiang province, China, Forests **14**, 170, 2023.
119. D.E. Rumelhart, G.E. Hinton and R.J. Williams, Learning representations by back-propagating errors, Nature **323**, pp. 533 - 536 , 1986.
120. M. Hall, F. Frank, G. Holmes, B. Pfahringer, P. Reutemann, I.H. Witten, The WEKA data mining software: an update. ACM SIGKDD explorations newsletter **11**, pp. 10-18, 2009.
121. S.B. Aher, L.M.R.J. Lobo, Data mining in educational system using weka. In International conference on emerging technology trends, Foundation of Computer Science **3**, pp. 20-25, 2011.
122. S. Hussain, N.A. Dahan, F.M. Ba-Alwib, N. Ribata, Educational data mining and analysis of students' academic performance using WEKA. Indonesian Journal of Electrical Engineering and Computer Science **9**, pp. 447-459, 2018.
123. A.K. Sigurdardottir, H. Jonsdottir, R. Benediktsson, Outcomes of educational interventions in type 2 diabetes: WEKA data-mining analysis. Patient education and counseling **67**, pp. 21-31, 2007.
124. M.N. Amin, A. Habib, Comparison of different classification techniques using WEKA for hematological data. American Journal of Engineering Research **4**, pp. 55-61, 2015.
125. Suthaharan, S., & Suthaharan, S. (2016). Support vector machine. Machine learning models and algorithms for big data classification: thinking with examples for effective learning, 207-235.
126. Chang, C. C., & Lin, C. J. (2011). LIBSVM: a library for support vector machines. ACM transactions on intelligent systems and technology (TIST), 2(3), 1-27.
127. Powell, M.J.D. (1989). A Tolerant Algorithm for Linearly Constrained Optimization Calculations. Math. Program, 45, 547–566.
128. D.C. Liu, J. Nocedal, On the Limited Memory Method for Large Scale Optimization, Mathematical Programming B. **45**, pp. 503–528, 1989.
129. J. Park and I. W. Sandberg, Universal Approximation Using Radial-Basis-Function Networks, Neural Computation 3, pp. 246-257, 1991.
130. Gropp, W.; Lusk, E.; Doss, N.; Skjellum, A. A high-performance, portable implementation of the MPI message passing interface standard. Parallel Comput. 1996, 22, 789–828.
131. Chandra, R. Parallel Programming in OpenMP; Morgan Kaufmann: Cambridge, MA, USA, 2001.