Article

Predict the duration of forestfires 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: Forest and urban fires are a major problem in the modern era that tests the endurance of governments to extinguish them. Fires can cause economic and ecological problems especially in the summer months. In modern times, the rapid development of Artificial Intelligence can be a weapon for predicting the evolution of fires or even for their prevention. Specifically, through Machine Learning, which is one part of Artificial Intelligence several methods have been incorporated to detect the duration of fires using data which are freely available from the Fire Service of Greece for a period of 10 years. For this purpose, a wide range of machine learning techniques were used on this data and the experimental results were more than encouraging.

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

Citation: Kopitsa, C.;Tsoulos, I.G.; Charilogis V., Stavrakoudis A. Predict the duration of forestfires 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.

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. 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 [37]. Also, KC et al.. proposed a Surrogate model [38] to model the size of a wildfire over time, using data collected from wildfires in Tasmania. Furthermore, Xi et al. proposed [39] the application of joint mixture models to model the duration and the size of wildfires.

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.

2. Materials and Methods

2.1. Data pre - processing

For this study we have used data by the: **Hellenic Fire Service**, from the years 2014 – 2023. We have processed the preliminary raw data from years 2014 – 2023, to obtain usable data that will lead to a more efficient model. The first step involved: data cleaning. In the second step, we merged the columns, containing the fire start & time, as well as the columns containing the fire extinguishment & time, in order to perform a subtraction and calculate the fire duration. Then we identified three classes based on the fire duration (class 1, 3, 4). The data categorized was complemented with the classification, based on the Fire Department, the Region, and the Season. Our final dataset includes, 25 columns into an excel sheet**.** specifically: fire department, province, season, fuel (forest, forest area, grove, grasslands, reeds/swamps, agricultural lands, cover crop, garbage dumps) burnt area acres, manpower (Firefighters, volunteer, army, etc.), vehicles (firefighting, tanks, etc.), aerial means (helicopters, different aircrafts) duration, class and log. We converted all the Excel files since 2014, into an ARFF file, and using (nom) the class for fire duration, we evaluated the performance of all algorithms, in WEKA.

2.2. The proposed algorithms

2.2.1. Bayes Net

The Bayesian Networks, are Probabilistic Graphical Models, they can create: diagnostic models, causal models, decision making, prediction, e.t.c. [40]. Therefore, the Bayes Net, is considered as a useful tool, in prediction and detection area, of forest/ wildfires fires. Consequently, it follows a brief reference, of two papers.

- The specific algorithms, was used, in study: A Bayesian network model for prediction and analysis of possible forest fires causes, in 2020. The study was conducted in Mugla, of Turkey. The model showed, that the most effective factors, on forest fires ignition, were: the month, and the temperature [41].
- In 2021, it was published, on MDPI, the survey: A Bayesian Network – Based Information Fusion Combined with DNNS for Robust Video Fire Detection. The combination,

Regional – Convolutional Neural Network (R – CNN), Long – Short term memory (LSTM), and Bayesian Net, proved that the last one, not only improves the detection accuracy, of forest / wildfires**.** but also reduces the decision making [42].

2.2.2. Naive Bayes

The Naïve Bayes is a supervised machine learning algorithm, used for classification tasks. This classifier, use principles of probability in order to perform classification tasks [43]. One of the strong points, of the algorithm, is the amenable improvements, and modifications, as to achieve better results in research, such as the forest wildfires prediction.

- A proportional modification was found in the following research, which was published on MDPI: Towards Fire prediction Accuracy Enhancements by Leveraging an Improved Naïve Bayes Algorithm. In the aforementioned paper, there was an evolutionary algorithm, the Double Weighted Naïve Bayes with Compensation Coefficient (DWCNB)**.** which compared with Naïve Bayes, and Double Weighted Naïve Bayes. The results showed a prediction accuracy of 98.13%, higher than Naïve Bayes for 5.08%, and respectively 2.52% than Double weighted Naïve Bayes [44].
- In a recent research, 2024, in Turkey, it was used different algorithms so as to extract the highest accuracy, for forest / wildfires. The paper, was: Predicting forest fire vulnerability using machine learning approaches in the Mediterranean region: a case study in Turkiye. The study compared: Naïve Bayes, Decision Tree, Random Forest, Neural Networks, and Support Vector machines. The Random Forest algorithm, yielded the highest accuracy, while Naïve Bayes performed consistently, albeit lower than Random Forest, and Decision Tree [45].

2.2.3. Logistic Regression

Like the previous algorithm, Logistic Regression, works also with machine learning classification, and used to predict probabilities. This (ML) technique used in data sets with many features [46]. The Logistic Regression, is widely used in natural hazards, such as fire modeling by estimating, the probability of occurrences, according the following survey: A Survey of Machine Learning Algorithms based Forest Fires Prediction and Detection [47]

2.2.4. Mlp Network

The Multilayer Perceptron is a commonly used Neural Network. It composed of multiple layers, and contains a set of perception elements known as neurons. It is used in forecasting models, and image pattern recognition [48,49]. In a Chinese province, in 2022, was conducted the research: Using Multilayer Perceptron to Predict Forest Fires in Jiangxi Province, Southeast China. In this paper, several models were studied for the occurrence of forest fires. ROC plots were used to compare results from: (MLP), Logistic, and SVM. The (MLP) model scored the highest percentage, compared to the rest. Precisely, (MLP) scored 0.984, Logistic 0.933, and SVM 0.974 [50].

2.2.5. The J48 algorithm

J48, belong to the of Decision Tree algorithm, in supervised learning. Creates a decision tree, which breaks into subsets. It used in risk analysis, pattern recognition and makes predictions [51].

- J48, was selected among with Random Forest (RF), adaboostM1, and Bagging, in Algeria, 2020, for the project: Predicting Forest Fires in Algeria using Data Mining Techniques: Case study of the Decision Tree Algorithm. Although, the results showed best performance, with adaBoostM1 (84.21%), the researchers do not recommend it due it needs significant resources and effort to be translated to hardware implementation. Therefore, they recommend J48 with accuracy (82,89%). The (RF) came up to (72,36%), and Bagging to (78,94%) [52].
- In a study, to Slovenia, 2005, the J48, had the lowest result. Specifically, on the paper: Learning to Predict Forest Fires with Different data Mining Techniques, highlighted

the Bagging, as the most efficient in relation to: Logistic Regression, Random Forest, J48, Boosting [53].

2.2.6. Random Forests

Random Forest is a popular machine learning algorithm. Its ease of use and flexibility, in handles classification and regression problems, in more precise predictions, fueled its commonly adoption [54].

- On MDPI, 2023, it was published, the paper: Forest – Fire – Risk Prediction based on Random Forest and Backpropagation Neural Network of Heihe Area in Heilongjiang Province, China. There was a comparison, in the research, among the Random Forest, and the backpropagation Neural Network (BPNN). Both methods, were found suitable for predicting forest / wildfires. The (RF) prediction range between 87,91% - 88,98%, while the (BPNN) accuracy was 86,01% and 86,94% [55].
- Random Forest, presented satisfactory results also in this research: Using GIS and random Forest to identify fire drivers in a forest city, Yichun, China. The forecast, for a fire outbreak, meteorological data ranged from 71,2% to 76,5%. The prediction, from combined factors, had a higher percentage among 80,78% - 84,8%. The AUC value, based on meteorological factors, were 0,740 – 0,807 indicating a moderated model fit. On the other hand, the value with combined factors showed an excellent fit, with the model 0,886 – 0,906 [56].

3. Results

Table 1. Experimental results using various machine learning models for 10 years of observations.

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%**

4. Conclusions

1) Discover important features for predicting duration
2) create rules for determination of categories
3) using meteorological data

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. H. Liang, M. Zhang, H. Wang, A Neural Network Model for Wildfire Scale Prediction Using Meteorological Factors, IEEE Access **7**, pp. 176746-176755, 2019.
38. 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.
39. D.D. Xi, C.B. Dean, S.W. Taylor, Modeling the duration and size of wildfires using joint mixture models. Environmetrics **32**, e2685, 2021.
40. Koski, T., & Noble, J. (2011). Bayesian networks: an introduction. John Wiley & Sons.
41. Sevinc, V., Kucuk, O., & Goltas, M. (2020). A Bayesian network model for prediction and analysis of possible forest fire causes. Forest Ecology and Management, 457, 117723.
42. Kim, B., & Lee, J. (2021). A Bayesian network-based information fusion combined with DNNs for robust video fire detection. Applied Sciences, 11(16), 7624.
43. Bayes, T. (1968). Naive bayes classifier. Article Sources and Contributors, 1-9.
44. Shu, L., Zhang, H., You, Y., Cui, Y., & Chen, W. (2021). Towards fire prediction accuracy enhancements by leveraging an improved naïve bayes algorithm. Symmetry, 13(4), 530.
45. Purnama, M. I., Jaya, I. N. S., Syaufina, L., Çoban, H. O., & Raihan, M. (2024, March). Predicting forest fire vulnerability using machine learning approaches in The Mediterranean Region: a case study of Türkiye. In IOP Conference Series: Earth and Environmental Science (Vol. 1315, No. 1, p. 012056). IOP Publishing.
46. Rymarczyk, T., Kozłowski, E., Kłosowski, G., & Niderla, K. (2019). Logistic regression for machine learning in process tomography. Sensors, 19(15), 3400.
47. Abid, F. (2021). A survey of machine learning algorithms based forest fires prediction and detection systems. Fire technology, 57(2), 559-590.
48. C. Bishop, Neural Networks for Pattern Recognition, Oxford University Press, 1995.
49. G. Cybenko, Approximation by superpositions of a sigmoidal function, Mathematics of Control Signals and Systems **2**, pp. 303-314, 1989.
50. Gao, K., Feng, Z., & Wang, S. (2022). Using multilayer perceptron to predict forest fires in jiangxi province, southeast china. Discrete Dynamics in Nature and Society, 2022(1), 6930812.
51. J, Biju. Bannari Amman. Institute of Technology. Classification by tree – based Data Mining Algorithms. 2023. Retrieved from: https://www.bitsathy.ac.in/data-mining-algorithms/
52. Abid, F., & Izeboudjen, N. (2020). 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.
53. Stojanova, D., Panov, P., Kobler, A., Džeroski, S., & Taškova, K. (2006, October). 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).
54. Rigatti, S. J. (2017). Random forest. Journal of Insurance Medicine, 47(1), 31-39.
55. Gao, C., Lin, H., & Hu, H. (2023). Forest-fire-risk prediction based on random forest and backpropagation neural network of Heihe area in Heilongjiang province, China. Forests, 14(2), 170.
56. Su, Z., Hu, H., Wang, G., Ma, Y., Yang, X., & Guo, F. (2018). Using GIS and Random Forests to identify fire drivers in a forest city, Yichun, China. Geomatics, Natural Hazards and Risk, 9(1), 1207-1229.