MDPI

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

The Bayesian Networks, are Probabilistic Graphical Models, they can create: diagnostic models, causal models, decision making, prediction, e.t.c. [1]. 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 [2].
- 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 [3].

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 [4]. 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

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:

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 [5].

- 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 [6].

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 [7]. 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 [8]

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 [9,10]. 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 [11].

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 [12].

- 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%) [13].
- 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 [14].

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 [15].

- 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% [16].

- 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 [17].

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

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. Koski, T., & Noble, J. (2011). Bayesian networks: an introduction. John Wiley & Sons.
2. 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.
3. Kim, B., & Lee, J. (2021). A Bayesian network-based information fusion combined with DNNs for robust video fire detection. Applied Sciences, 11(16), 7624.
4. Bayes, T. (1968). Naive bayes classifier. Article Sources and Contributors, 1-9.
5. 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.

6. 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.
7. Rymarczyk, T., Kozłowski, E., Kłosowski, G., & Niderla, K. (2019). Logistic regression for machine learning in process tomography. Sensors, 19(15), 3400.
8. Abid, F. (2021). A survey of machine learning algorithms based forest fires prediction and detection systems. Fire technology, 57(2), 559-590.
9. C. Bishop, Neural Networks for Pattern Recognition, Oxford University Press, 1995.
10. G. Cybenko, Approximation by superpositions of a sigmoidal function, Mathematics of Control Signals and Systems **2**, pp. 303-314, 1989.
11. 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.
12. 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/
13. 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.
14. 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).
15. Rigatti, S. J. (2017). Random forest. Journal of Insurance Medicine, 47(1), 31-39.
16. 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.
17. 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.