Fire Dimension Prediction and Resource Allocation System for Portugal

Green Aid for Portugal

Diogo Risca
1220186@isep.ipp.pt

Diogo Teixeira
1180893@isep.ipp.pt

Gonçalo Lopes
1220207@isep.ipp.pt

Guilherme Mendes
1190641@isep.ipp.pt

João Azevedo
1220259@isep.ipp.pt

Miguel Panda
1190896@isep.ipp.pt

Department of Computer Engineering
Instituto Superior de Engenharia do Porto (ISEP)

***Abstract*—Forest Fires in Portugal are numerous and are an emergency that is attended to today. These are some of the most serious natural disasters in Portugal, not only because of the high frequency and scale they occur, but also because of the destructive effects they cause. In addition to the economic and environmental damage, creating a source of danger for people and property. In this report, an automatic algorithm will be presented to predict the fire severity which will help the firefighters to manage the necessary resources that will be needed to put out the fire. More specifically, the proposed method will be based on a forest fires dataset that will be divided into two, one for training and other for validation of the machine learning algorithm. Overall, the proposed solution shows potential in future applications for forest fires, being an asset for firefighters. The method at the end of predicting will give a proposed assignment of the firefighters to attend the fires that are current, giving the necessary resources to put out the fire, or in cases that the demand is higher than the available resources, a scaled attribution will be placed for the goal to equally divide resources between every fire.**

***Index Terms*—fire, dataset, machine learning, decision tree, prediction, allocation**

I. Introduction

Portugal has a forest coverage of about 39%, placing it at the third place on the European Greenest Countries [1]. Although, almost every year, mainly during the summer, Portugal suffers the impact of several forest fires. During the year of 2022, the data from EFFIS, places Portugal as the European country with the most burned area (measured in hectares [ha]) [2].

Forest fires in Portugal are numerous, especially in the summer. One of the most serious natural disasters, not only because of the high frequency and scale they occur, but for the destructive effects they cause. In addition to the economic and environmental damage, they can create a source of danger for people and property. Human intervention can play a decisive role in its origin and in limiting its development, a good resource management and the allocation of firefighting assets are the key to control the damage that the fire might cause. The study goal will be to predict the severity of the fire with the use of indexes with climate data. With this prediction, it will be possible to make a better management of resources and decision making, which will allow for the method to give a proper attribution of the resources to each fire with the goal to reduce the distance that the firefighters do to attend the fires, being an asset for firefighters to control the fire, since the resources will be available sooner.

The final purpose of this project is to implement a prototype system that provides a weighted distribution of available firefighting resources based upon the estimated severity of current active fires. As described before, this estimation is provided by an Extremely Randomized Trees Classifier model using a number of indexes that pertain to fire hazard previously collected. Enhancing the resources allocation when new fires happen has been one of the main concerns for the past years as global warming and the impact of human civilisation on the natural world has led to an overall increase in the number of forest fires as well as their severity. An implementation of this kind is not believed to be able to ever replace the need for a well trained and experienced firefighter command (or other civil protection service) that coordinates firefighting efforts and commands the firefighting means. It is believed by the group however that a system as the one proposed in this paper could aid in providing a baseline or framework that is capable of improving response times to fires and also making sure that the required and appropriate means are assigned to each fire. There have been many examples not only in Portugal but throughout Europe where fires that were initially deemed to be of a lower risk turned out to be some of the most devastating and deadly forest fires.

II. State of the Art

Nowadays there is no shortage of information on the weather, climate, or forest fires. Not only do these represent a national concern for countries around the world, but also international organisations like the EU. EFFIS, or the European Forest Fire Information System monitors forest fires in Europe, North Africa, and the Middle East. Since 2007, they have adopted and applied the Canadian FWI, or Fire Weather Index (2), to, using some known metrics, provide an estimated level of forest fire danger in a certain area of Europe. This is achieved through the use of several well-explored deterministic models using weather indexes and other information. The use of machine learning algorithms to assess fire risk and allow for prevention and preparation of firefighting means has been growing in recent years, as an alternative to these deterministic models.

A. *Croatia Fire Study Approach*

Many different approaches have been proposed and implemented, from using satellite images and computer vision systems [3] to infer information on the vegetation being observed and how prone it is to fire to using new metrics that have been proven to also play a role in the risk of fire, like distance to nearest road (a measure of fire engine accessibility).

B. *Greece Fire Study Approach*

Another approach [4] proposed was the use of new indexes that attempt to better represent the state of the forest (it's topography, vegetation density, and wetness/dryness, etc). This approach uses the long-term mean normalised difference vegetation index (NDVI) of the woody vegetation (NDVIW) and its trend (NDVIT). In this way, an attempt is made at capturing both information about the historical (long-term) state of the forest, but also its most recent seasons and years. Using the Greek Forest Fire season of 2007 as a data-set, several models were attempted and the conclusion was reached that the indexes did indeed improve the risk assessment of a given area of Greece, in particular using Gradient Boosting with the aid of the XGBoost open-source library.

C. *The weather indexes impact*

A forest fire is not an isolated event. Its start and evolution is related with different events. [5] In order to predict the severity of a fire it is important to understand which role those external factors can play. The different weather indexes are based in the combination of 4 factors (Temperature, Humidity, Wind and rainfall) [5]. A study that took place at the University of Bern, Switzerland explains the relations between the factors and the indexes as displayed in the table I.

Another important study result to consider is the seasonality of the indexes. The ISI, FFMC and FWI are the main indexes to consider when predicting the fire danger in spring. In the other hand, the BUI and DC perform better in autumn and winter.

D. *Optimising Resource Allocation*

In 2012, a study of Poggenpohl, F and Guttinger, D considering the Optimising Task Allocation on Fire Fighting, used an approach with a modified version of the simulated annealing algorithm to improve the resource allocation optimisation, comparing the algorithm performance with the human managers. With their approach the results lead to a smaller total processing time and to a more balanced workload of roles compared to action plan recommendations given by the six operational managers. [6].

Fire Indices	**Meteorological Input**			
	T	**H**	**U**	**P**
Fine fuel moisture code (FFMC)	●	●	●	●
Duff moisture code (DMC)	●	●		●
Drought code (DC)	●			●
Initial spread index (ISI)	●	●	●	●
Buildup index (BUI)	●	●		●
Fire weather index (FWI)	●	●	●	●
Keetch-Byram drought index (KBDI)	●	●		●
M68dwd	●	●		●
Me Arthur Mark 5 forest fire danger index (FFDI)	●	●	●	●
Sharples	●	●	●	
Fosberg fire weather index (FFWI)	●	●	●	
Angström index (Angström)	●	●		
Nesterov ingnition index (Nesterov)	●	●		
Baumgarthner index (Baumgartner)	●		●	●

Table I
FIRE INDICES [5]

E. *Available Technologies*

The IPMA (Instituto Português do Mar e da Atmosfera) [7] uses their measurements data on a daily basis to predict the fire danger based on temperature, humidity, wind and rainfall. On their website, the predicted fire danger is displayed on a Portugal coloured map.

III. DATASET

The dataset used in this study was obtained at the *Instituto de Conservação da Natureza e das Florestas* (ICNF) website and has the statistics of 185316 Portuguese fires that had happen between the year 2010 and 2021 [8]. The dataset has useful information about the fire, which will be used by the training part of the divided dataset to help predict the severity of the fire. A useful information given, was the meteorological data that has been presented by index, presented as follows: FFMC (Fine Fuel Moisture Code) that is an indicator of the sustained flaming ignition and the fire spread. DMC (Duff Moisture Code), that relates to the probability of a lightning ignition and fuel consumption. DC (Drought Code), this code can give important information about the effort needed to extinguish a fire and the consumption of heavier fuels. ISI (Initial Spread Index) that combines wind and FFMC representing rate of fire spread without considering the influence of fuel. BUI (Build-Up Index) is a combination of DMC and DC representing total fuel available to the spreading fire. FWI (Fire Weather Index) that is a combination of ISI and BUI measuring the intensity of the spreading fire as energy rate per unit length of fire front. DSR (Daily Severity Rating), a power function of FWI representing a measure of control difficulty for a fire [9].

A. *Dataset Treatment*

For the study purposes, the dataset had some irrelevant columns and it was missing crucial information in some of the parameters. A machine learning study and prediction model results are always linked with the quality of the data it is trained with. In terms of the dataset processing, some columns that were not relevant to this study or contained repetitive information were deleted that, such as: The SGIF and ANEPC code, since it could be replaced in this study by the automatic generated ID. IncSup24horas informs if the fire lasted at least 24 hours, which its not necessary since there already is a column with the duration of the fire, so this information was redundant. DTCCFR is a code which provides the location of the fire, not being a relevant data for this study, same goes for other information regarding the location of the fire. The Military X and the Military Y columns along with the X ETRS89, Y ETRS89, Latitude and Longitude were also removed. Since the goal of this project is to measure the severity of the fire, the cause was not relevant so the columns Cause Code, Cause Type, Cause Group, Cause Description, Alert Source were removed. Information about the firefighters contact with the fire has been removed along with the burned area, since the goal is to give the severity of the fire before the firefighers reach the location of the fire.

Apart from removing columns, a new one was created: The response time, a column that shows how long the firefighting forces took to arrive to the fire scene, because it is important to study the correlation between a long response time and the overall severity of the fire. The response time is obtained by the subtraction of the first response time and the alert time. All the other columns, specially the indexes were left at the dataset to be included in the study.

IV. Prediction Component

A. *Technologies used*

Sci-kit learn is a python library that is intended for learning purposes. This library, originally developed by David Cournapeu in 2010, made it possible to manipulate and treat the dataset, train several different types of models on said dataset as well as evaluate and study each of the models. It was an invaluable tool for the development of this project and report.

B. *Models tested and evaluations*

Several models were implemented and tested for the purpose of classifying the expected severity of new fires. Since the project presents a problem of classification, all of these were tested for the following metrics: Accuracy, Precision, Recall as well as the F1-score.

To balance the classes in the dataset it was used the Synthetic Minority Oversampling Technique (SMOTE), this algorithm is very used for imbalanced data. Instead of applying a simple replication of the minority class instances, the key idea of SMOTE is to introduce synthetic examples with some variance [10].

The following table II describes the results obtained for the best five models tested, a simple neural network with 3 layer, 8 neurons each and ReLu activation is also present on the table just for comparison.

Models	**Classification**		
	Accuracy	**Recall**	**f1-score**
Decision Tree Classifier	87	88	88
Random Forest Classifier	94	94	94
Extremely Randomized Trees Classifier	95	95	95
Bagging Classifier	94	94	94
XGBoost	88	88	88
Neural Network	47	47	45

Table II
Model Metrics

From the tested models, a clear superiority of Decision Tree Classifiers related models is apparent when applied the dataset. Decision Trees are well known and explored algorithm for the purpose of classification problems. They function as a conjunction of “nodes” and “branches”. In every node, a feature is evaluated and conditional upon its value one of the two outgoing branches leading to a different node is selected so that this process can then repeat itself until a prediction is reached. The two most successful models were, as stated before, ensemble models based upon decision trees, in particular the Random Forest and the Extremely randomized Trees classifiers. These models attempt to tackle the Decision Tree models’ tendency to over-fit the data. Random Forest models, as the name indicates, build a multitude of trees each being trained with only a subset of the original data, classification being decided as whatever result a majority of the trees have come to. Extremely Randomized Trees, the model type with which the best results were achieved, differ from Random Forest as, instead of every tree being trained upon only a subset of the data, all the dataset is used for the training of the multiple decision trees.
Although not described in table II, an important consideration for this project was also the very high levels of recall achieved by these decision tree models when classifying severity class 3 and class 4 fires, which was regarded as paramount for it guarantees that these high-severity fires were as correctly classified as possible.

V. The Decision component

Different approaches were studied during the course of the project, allocation techniques that emphasise different aspects, and different optimisation techniques to minimise the cost of resource allocation.

A. Allocation

Two main characteristics were valued on this implementation, including algorithms that provide a logical allocation combined with the minimal cost possible. Although, an alternative was also included, that on the case the resources are not enough to attend all the fires at their maximum necessity, it will, at least, provide the allocation of fire responding units to all of the fires, not leaving any fire unattended

1) Basic Allocation: The first approach was the development of a basic allocation algorithm that, with the the needed resources for each fire information through the existent fires, it provides the allocation resources, considering their distance from their zone and the fire. Logically, the nearest zones units will be the first option when responding to a fire in a nearby zone. This distance weights are giving by a distance matrix auto-generated, considering the zones location and distance from each other.

As an example, if there exists a fire with level 4 of severity in a forest zone there would be needed 80 jeeps, 50 trucks and 3 air support resources needed to respond optimally to the fire.

Using the basic allocation algorithm, this needed resources will be provided by the nearest zones.

If there is no air resources available at the 4 nearest zones from the fire, this algorithm converts the air support in the equality jeeps and trucks fire power and allocates more of this resources to the fire to help extinguish it the fastest way possible and reducing the impact on cost of bringing air support to a fire that is too many miles away.

2) Round Robin Fashioned Allocation: Considering the fact that the mean severity is too high or that the available resources are not enough to fulfil all the existent fire needs at his maximum, a second approach was developed that is preferred rather than the basic allocation if the above pre-requisites are verified.

This approach was inspired by the Round Robin algorithm, since by receiving the needed and available resources and the zones information, it distributes the resources alternately to all the existent fires.

As an example, if there exists 10 fires with level 4 of severity in a forest zone at the Porto district there may not have been the necessary fire resources available at the moment to deal with all the fires. Using the round robin fashioned allocation algorithm, all the 10 fires would get resources that will prevent the fire to spread uncontrollably and lead to catastrophically situations.

B. Optimisation

1) Simulated Annealing (SA): As third approach to resource allocation, the simulated annealing was developed, which will use the basic attribution method as an input and the required resources for each current fire as well. Following the basis of annealing algorithm, the main goal of this approach is to create an quantitative objective function containing the trade-offs that have to be made [11]. In this case, the goal is to have the lowest value possible, so the resources will travel the lowest amount possible. With this in mind, the objective function has the purpose to minimise the cost value to the minimum. The method will receive an initial allocation as input value with a certain cost, then it will start to allocate resources randomly through the zones to attend the proper fires. After that, a cost will be calculated. And finally, it will be compared to the last best value, and the one with the lowest value will be the current solution, the algorithm will continue to do this approach n-times, after that the final approach will be presented.

2) Modified Distribution (MODI): The MODI optimisation is an algorithm very used for allocation problems. The main objective is to minimise or maximise a cost, fulfilling every demand.

After the first allocation with the Round Robin Fashioned Allocation (see subsection V-A2), the demand obtained is used to feed a MODI algorithm, with the cost being the distance between the resource and the fire, and the supply being the free resources on each district. The first step of this module is to allocate again all the resources with a Vogel method. This method determines a starting basic feasible solution, then the solution is optimised with the MODI algorithm. Computing $ui + vj = cij$ (ui - lines, vj - columns), for all the cells with allocations will give an opportunity cost, if any of the unoccupied cells has a negative impact on the opportunity cost, then the resulting solution cannot be considered as the optimum solution, making further savings in transportation cost to be possible. Therefore, the cell having the least negative opportunity cost but not occupied, will be selected as the cell to be included in the solution to be computed next. An Optimal solution is found when no empty cell gives a negative opportunity cost [12].

For this project the combination of the Round Robin Fashioned Allocation and the MODI Optimisation, produced the best allocation.

VI. User Interaction

For this system to be useful in the real world, it was essential for it to be able to be used by anyone with the minimal amount of training. To achieve this, a graphical user-interface had to be developed and it should limit the number of possible interactions a user can have. After analysis, it was concluded that the best approach was to only allow the user to input the coordinates of a fire, the type of terrain the fire was on (urban or forest) and the district it was present. The need to introduce the type of the terrain and the district came to a problem in accessing the required API. In addition, the option to remove a fire was added. To ensure the application was compatible with the most amount of devices, it was decided that the user would interact through the system via an website. The website uses a REST API to handle communication with the rest of the system. The REST API was developed using the Django Framework, with Python facilitating communication with the other components of this system. The website was developed using JavaScript with the React Framework. This technology was used due to it having the necessary libraries and familiarity by the team members.

VII. Conclusion

This report demonstrated conceptually that an approach to AI aided decision making is possible in the context of fire fighting as it pertains to firefighting asset management and fire severity assessment. As stated, while never being able to replace the hard and thoughtful work that fire fighters and civil servants put into fire fighting operations on a daily basis, the use of AI to offset some of the load from some of the more managerial concerns of firefighting operations as well as the introductions of fail-safes from checking and guaranteeing a minimum amount of fire fighting assets to the curbing of underestimation of new fires, among others, it is believed that the further development of projects such as this could prove invaluable in not only Portugal's yearly struggle with fire but also all other nations, especially in the context of our worlds' worsening climate.

As pertaining to the implementation of the project itself, the team responsible, once again, confronted with fresh challenges, and, through the use of newly explored tools and processes learned throughout the last few months, believes it was able to achieve a good result that, if not proved the initial point, at least demonstrated the possible utility if this kind of implementation in the field. The team encountered some difficulties with the development of the annealing algorithm, while an effort was made to adapt it properly to the context of our problem, its' outputs were always inferior to those of the Round Robin. The website could provide a much bigger range of features taking advantage of all the capabilities of a system such as this, namely connecting to mapping APIs using proper distances instead of Euclidean distances. As another future improvement the system proposed would greatly benefit from having its database be updated and the model being retrained as a way to include more information on the evolving fight against forest and urban fires.

References

[1] "Floresta portuguesa ocupa mais de um terço do país -," 3 2022. [Online]. Available: https://florestas.pt/conhecer/floresta-portuguesa-ocupa-mais-de-um-terco-do-pais/

[2] "Effis - statistics portal." [Online]. Available: https://effis.jrc.ec.europa.eu/apps/effis.statistics/estimates/EU/2022/2006/202

[3] I. N. Kosović and D. Škurić Kuraži, "Machine learning approach in fire risk estimation," in *2021 International Conference on Software, Telecommunications and Computer Networks (SoftCOM)*, 2021, pp. 1–6.

[4] Y. Michael, D. Helman, O. Glickman, D. Gabay, S. Brenner, and I. M. Lensky, "Forecasting fire risk with machine learning and dynamic information derived from satellite vegetation index time-series," *Science of The Total Environment*, vol. 764, p. 142844, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0048969720363749

[5] O. M. S. B. Daniel Steinfeld, Adrian Peter, April 2022. [Online]. Available: https://egusphere.copernicus.org/preprints/2022/egusphere-2022-92/egusphere-2022-92.pdf

[6] F. G. Poggenpohl and D. Guttinger, "Optimizing task allocation on fire fighting," *Proceedings of the 2012 4th International Conference on Intelligent Networking and Collaborative Systems, INCoS 2012*, pp. 497–502, 2012.

[7] "Instituto Português do Mar e da Atmosfera." [Online]. Available: https://www.ipma.pt/pt/index.html

[8] "Icnf - instituto da conservação da natureza e das florestas." [Online]. Available: https://www.icnf.pt/florestas/gfr/gfrgestaoinformacao/estatisticas

[9] "Fwi (fire weather index-Índice de."

[10] A. Fernández, S. García, F. Herrera, and N. Chawla, "View of smote for learning from imbalanced data: Progress and challenges, marking the 15-year anniversary." [Online]. Available: https://www.jair.org/index.php/jair/article/view/11192/26406

[11] S. Kirkpatrick, . C. D. Gelatt, and . M. P. Vecchi, "Optimization by simulated annealing," pp. 671–680, 1983.

[12] M. Rekha and V. Joshi, "Optimization techniques for transportation problems of three variables," *IOSR Journal of Mathematics*, vol. 9, pp. 46–50. [Online]. Available: www.iosrjournals.org