**International Journal of Academic Engineering Research (IJAER)**
**ISSN: 2643-9085**
**Vol. 4 Issue 9, September – 2020, Pages: 68-72**

# Neural Network Approach to Predict Forest Fires using Meteorological Data

**Mutasim Mahmoud Al-Kahlout, Ahmed Mahmoud Abu Ghaly, Donia Zaher Mudawah, Samy S. Abu-Naser**

Department of Information Technology,
Faculty of Engineering and Information Technology,
Al-Azhar University, Gaza, Palestine

***Abstract.*** *Forest fires are a major environmental issue, creating economical and ecological damage while endangering human lives. Fast detection is a key element for controlling such phenomenon. To achieve this, one alternative is to use automatic tools based on local sensors, such as provided by meteorological stations. In effect, meteorological conditions (e.g. temperature, wind) are known to influence forest fires and several fire indexes, such as the forest Fire Weather Index (FWI), use such data. In this work, we explore a Just Neural Network (JNN) approach to predict the burned area of forest fires were tested on recent real-world data collected from the northeast region of Portugal. The best accuracy we achieved was 98.75 percent. Such knowledge is particularly useful for improving firefighting resource management (e.g. prioritizing targets for air tankers and ground crews).*

**Keywords:** JustNN, Fire Science, ANN, Neural Network

## 1. Introduction

One major environmental concern is the occurrence of forest fires (also called wildfires), which affect forest preservation, create economical and ecological damage and cause human suffering. Such phenomenon is due to multiple causes (e.g. human negligence and lightning) and despite an increasing of state expenses to control this disaster, each year millions of forest hectares (ha) are destroyed all around the world. In particular, Portugal is highly affected by forest fires [1]. From 1980 to 2005, over 2.7 million ha of forest area (equivalent to the Albania land area) have been destroyed. The 2003 and 2005 fire seasons were especially dramatic, affecting 4.6% and 3.1% of the territory, with 21 and 18 human deaths.

Fast detection is a key element for a successful firefighting. Since traditional human surveillance is expensive and affected by subjective factors, there has been an emphasis to develop automatic solutions. These can be grouped into three major categories: satellite-based, infrared/smoke scanners and local sensors (e.g. meteorological). Satellites have acquisition costs, localization delays and the resolution is not adequate for all cases. Moreover, scanners have a high equipment and maintenance costs. Weather conditions, such as temperature and air humidity, are known to affect fire occurrence [2]. Since automatic meteorological stations are often available (e.g. Portugal has 162 official stations), such data can be collected in real-time, with low costs.

In the past, meteorological data has been incorporated into numerical indices, which are used for prevention (e.g. warning the public of a fire danger) and to support fire management decisions (e.g. level of readiness, prioritizing targets or evaluating guidelines for safe firefighting). In particular, the Canadian forest Fire Weather Index (FWI) [3] system was designed in the 1970s when computers were scarce, thus it required only simple calculations using look-up tables with readings from four meteorological observations (i.e. temperature, relative humidity, rain and wind) that could be manually collected in weather stations. Nevertheless, nowadays this index highly used not only in Canada but also in several countries around the world (e.g. Argentina or New Zealand). Even though Mediterranean climate differs from those in Canada, the FWI system was correlated with fire activity in southern Europe countries, including Portugal [4].

## 2. Artificial Neural Network

Artificial neural networks (ANNs) are statistical models directly inspired by, and partially modeled on biological neural networks. They are capable of modeling and processing nonlinear relationships between inputs and outputs in parallel. The related algorithms are part of the broader field of machine learning, and can be used in many applications as discussed [5].
Artificial neural networks are characterized by containing adaptive weights along paths between neurons that can be tuned by a learning algorithm that learns from observed data in order to improve the model. In addition to the learning algorithm itself, one must choose an appropriate cost function [5].

The cost function is what’s used to learn the optimal solution to the problem being solved. This involves determining the best values for all of the tunable model parameters, with neuron path adaptive weights being the primary target, along with algorithm tuning parameters such as the learning rate. It’s usually done through optimization techniques such as gradient descent or stochastic gradient descent [6].

These optimization techniques basically try to make the ANN solution be as close as possible to the optimal solution,

which when successful means that the ANN is able to solve the intended problem with high performance.
Architecturally, an artificial neural network is modeled using layers of artificial neurons, or computational units able to receive input and apply an activation function along with a threshold to determine if messages are passed along.
In a simple model, the first layer is the input layer, followed by one hidden layer, and lastly by an output layer. Each layer can contain one or more neurons [6].

Models can become increasingly complex, and with increased abstraction and problem solving capabilities by increasing the number of hidden layers, the number of neurons in any given layer, and/or the number of paths between neurons. Note that an increased chance of overfitting can also occur with increased model complexity [6].

Model architecture and tuning are therefore major components of ANN techniques, in addition to the actual learning algorithms themselves. All of these characteristics of an ANN can have significant impact on the performance of the model.

Additionally, models are characterized and tunable by the activation function used to convert a neuron's weighted input to its output activation. There are many different types of transformations that can be used as the activation function.

The abstraction of the output as a result of the transformations of input data through neurons and layers is a form of distributed representation, as contrasted with local representation. The meaning represented by a single artificial neuron for example is a form of local representation. The meaning of the entire network however, is a form of distributed representation due to the many transformations across neurons and layers.

One thing worth noting is that while ANNs are extremely powerful, they can also be very complex and are considered black box algorithms, which means that their inner-workings are very difficult to understand and explain. Choosing whether to employ ANNs to solve problems should therefore be chosen with that in mind.

## 3. Literature Review

Artificial Neural Networks have been used many fields. In Education such as: Predicting Student Performance in the Faculty of Engineering and Information Technology using ANN, Prediction of the Academic Warning of Students in the Faculty of Engineering and Information Technology in Al-Azhar University-Gaza using ANN, Arabic Text Summarization Using AraBERT Model Using Extractive Text Summarization Approach[6].

In the field of Health such as: Parkinson's Disease Prediction, Classification Prediction of SBRCTs Cancers Using ANN, Predicting Medical Expenses Using ANN, Predicting Antibiotic Susceptibility Using Artificial Neural Network, Predicting Liver Patients using Artificial Neural Network, Blood Donation Prediction using Artificial Neural Network, Predicting DNA Lung Cancer using Artificial Neural Network, Diagnosis of Hepatitis Virus Using Artificial Neural Network, COVID-19 Detection using Artificial Intelligence[7].

In the field of Agriculture: Plant Seedlings Classification Using Deep Learning, Prediction of Whether Mushroom is Edible or Poisonous Using Back-propagation Neural Network, Analyzing Types of Cherry Using Deep Learning, Banana Classification Using Deep Learning, Mango Classification Using Deep Learning, Type of Grapefruit Classification Using Deep Learning, Grape Type Classification Using Deep Learning, Classifying Nuts Types Using Convolutional Neural Network, Potato Classification Using Deep Learning, Age and Gender Prediction and Validation Through Single User Images Using CNN[5].

In other fields such as: Predicting Software Analysis Process Risks Using Linear Stepwise Discriminant Analysis: Statistical Methods, Predicting Overall Car Performance Using Artificial Neural Network, Glass Classification Using Artificial Neural Network, Tic-Tac-Toe Learning Using Artificial Neural Networks, Energy Efficiency Predicting using Artificial Neural Network, Predicting Titanic Survivors using Artificial Neural Network, Classification of Software Risks with Discriminant Analysis Techniques in Software planning Development Process, Handwritten Signature Verification using Deep Learning, Email Classification Using Artificial Neural Network, Predicting Temperature and Humidity in the Surrounding Environment Using Artificial Neural Network, English Alphabet Prediction Using Artificial Neural Networks[6].

## 4. Methodology

By looking intensely through literature and soliciting the experience of human experts on pathological conditions, a number of factors have been recognized that have an impact on burned area of forest fires prediction. These factors were prudently studied and coordinated with an appropriate number for coding the computer within the modeling environment ANN. These factors were categorized as input variables and output variables that reflect some possible levels of burned area of forest fires status in terms of the assessment system. The data were entered into the JNN tool environment, determined the value of each of the variables using JNN (the most influential factor on burned area of forest fires), then the data were trained, validated, and tested.

## 5. Forest Fire Dataset

**International Journal of Academic Engineering Research (IJAER)**
**ISSN: 2643-9085**
**Vol. 4 Issue 9, September – 2020, Pages: 68-72**

The forest Fire Weather Index (FWI) is the Canadian system for rating fire danger and it includes six components [6]: Fine Fuel Moisture Code (FFMC), Duff Moisture Code (DMC), Drought Code (DC), Initial Spread Index (ISI), Buildup Index (BUI) and FWI. The first three are related to fuel codes: the FFMC denotes the moisture content surface litter and influences ignition and fire spread, while the DMC and DC represent the moisture content of shallow and deep organic layers, which affect fire intensity. The ISI is a score that correlates with fire velocity spread, while BUI represents the amount of available fuel. The FWI index is an indicator of fire intensity and it combines the two previous components. Although different scales are used for each of the FWI elements, high values suggest more severe burning conditions. Also, the fuel moisture codes require a memory (time lag) of past weather conditions: 16 hours for FFMC, 12 days for DMC and 52 days for DC.

This study will consider forest fire data from the Montesinho natural park, from the Tras-os-Montes northeast region of Portugal. This park contains a high flora and fauna diversity. Inserted within a supra-Mediterranean climate, the average annual temperature is within the range 8 to 12◦C. The data used in the experiments was collected from January 2000 to December 2003 and it was built using two sources. The first database was collected by the inspector that was responsible for the Montesinho fire occurrences. At a daily basis, every time a forest fire occurred, several features were registered, such as the time, date, spatial location within a $9\times9$ grid ($x$ and $y$ axis), the type of vegetation involved, the six components of the FWI system and the total burned area. The second database was collected by the Braganc¸a Polytechnic Institute, containing several weather observations (e.g. wind speed) that were recorded with a 30 minute period by a meteorological station located in the center of the Montesinho park. The two databases were stored in tens of individual spreadsheets, under distinct formats, and a substantial manual effort was performed to integrate them into a single dataset with a total of 517 entries.

Table 1 shows a description of the selected data features. The first four rows denote the spatial and temporal attributes. Only two geographic features were included, the **X** and **Y** axis values where the fire occurred, since the type of vegetation presented a low quality (i.e. more than 80% of the values were missing). After consulting the Montesinho fire inspector, we selected the **month** and **day** of the week temporal variables. Average monthly weather conditions are quite distinct; while the day of the week could also influence forest fires (e.g. work days vs. weekend) since most fires have a human cause. Next come the four FWI components that are affected directly by the weather conditions. The BUI and FWI were discarded since they are dependent of the previous values. From the meteorological station database, we selected the four weather attributes used by the FWI system. In contrast with the time lags used by FWI, in this case the values denote instant records, as given by the station sensors when the fire was detected. The exception is the **rain** variable, which denotes the accumulated precipitation within the previous 30 minutes.

The burned **area** denoting a positive skew, with the majority of the fires presenting a small size. It should be noted that this skewed trait is also present in other countries, such as Canada [7]. Regarding the present dataset, there are 247 samples with a zero value. As previously stated, all entries denote fire occurrences and zero value means that an area lower than $1ha/100 = 100m^2$ was burned. To reduce skewness and improve symmetry, the logarithm function $y = \ln(x + 1)$, which is a common transformation that tends to improve regression results for right-skewed targets [7], was applied to the **area** attribute. The final transformed variable will be the output target of this work.

**Table 1:** The preprocessed dataset attributes

| S.N. | Attribute | Description | Type |
|---|---|---|---|
| 1 | **X** | x-axis coordinate (from 1 to 9) | Input |
| 2 | **Y** | y-axis coordinate (from 1 to 9) | Input |
| 3 | **month** | Month of the year (January to December) | Input |
| 4 | **day** | Day of the week (Monday to Sunday) | Input |
| 5 | **FFMC** | FFMC index from the FWI system: 18.7 to 96.20 | Input |
| 6 | **DMC** | DMC index from the FWI system: 1.1 to 291.3 | Input |
| 7 | **DC** | DC index from the FWI system: 7.9 to 860.6 | Input |
| 8 | **ISI** | ISI index from the FWI system: 0.0 to 56.10 | Input |
| 9 | **temp** | temperature in Celsius degrees: 2.2 to 33.30 | Input |
| 10 | **RH** | relative humidity in %: 15.0 to 100 | Input |
| 11 | **wind** | wind speed in km/h: 0.40 to 9.40 | Input |
| 12 | **rain** | outside rain in mm/m2 : 0.0 to 6.4 | Input |
| 13 | **area** | the burned area of the forest (in ha): 0.00 to 1090.84 | Output |

## 6. Neural Network Evaluation

As mentioned above, the purpose of this experiment was to predict the burned area of forest fires. We used Backpropagation algorithm, which provides the ability to perform neural network learning and testing. Our neural

**International Journal of Academic Engineering Research (IJAER)**
**ISSN: 2643-9085**
**Vol. 4 Issue 9, September – 2020, Pages: 68-72**

network is the front feed network, with one input layer (12 inputs), one hidden layer and one output layer (1 output) as seen in Figure 2. The proposed model is implemented in Just Neural Network (JNN) environment[9]. The dataset for to predict the burned area of forest fires were gathered from the UCI machine Repository [8] which contains 517 samples with 13 attributes (as seen in Figure 1). This model was used to determine the value of each of the variables using JNN which they are the most influential factor on burned area of forest fires prediction as shown in figure 3. After training and validating, the network, it was tested using the test data and the following results were obtained. The accuracy of the burned area of forest fires predication was (98.87%). The average error was 0.010. The training cycles (number of epochs) were 116,115. The training examples were 378. The number of validating examples was 140 as seen in figure 4. The control of the parameter for burned area of forest fires prediction is shown in Figure 5 and the summary of the trained model is shown in Figure 6.

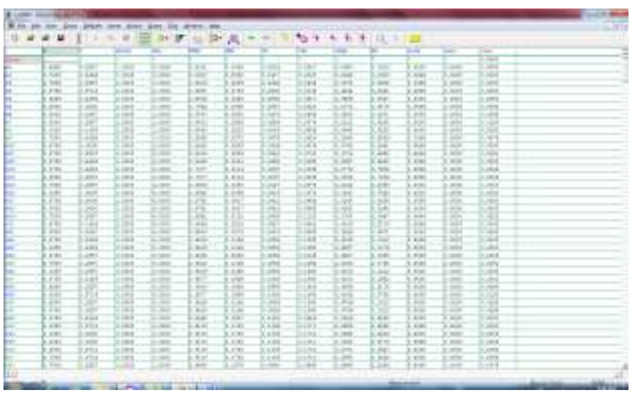

Figure 1: Attribute inputs and output

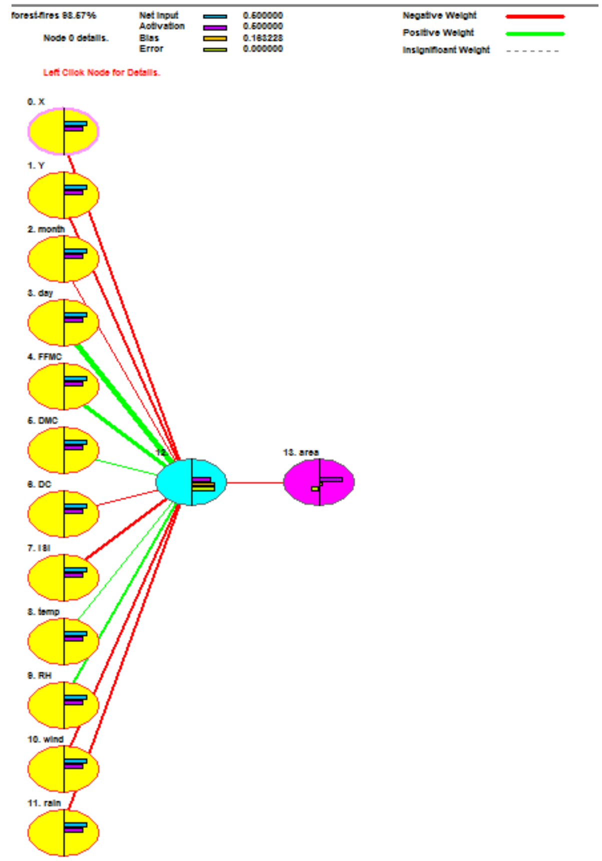


Figure 2: Artificial Neural Network Structure


Figure 3: Attributes importance (the most influential factor on burned area of forest fires)

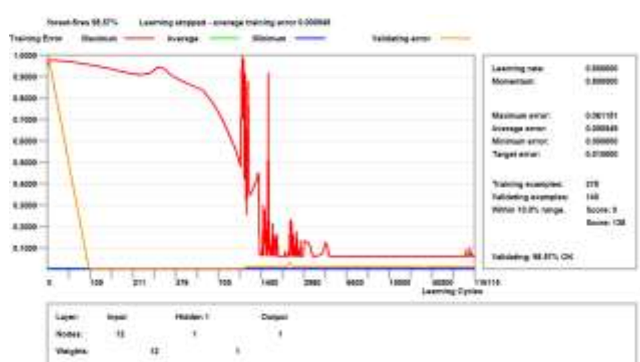

Figure 4: Learning progress

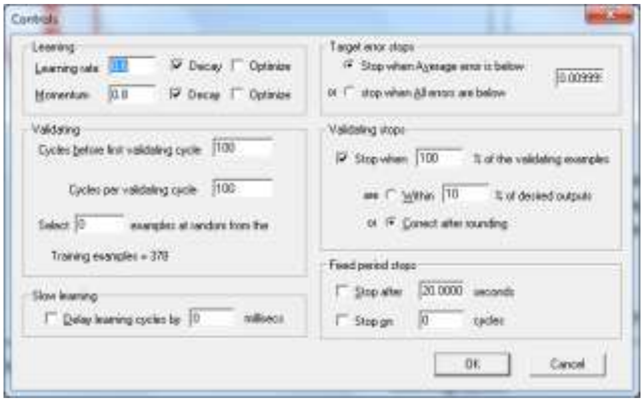

Figure 5: Control parameters of the ANN environment

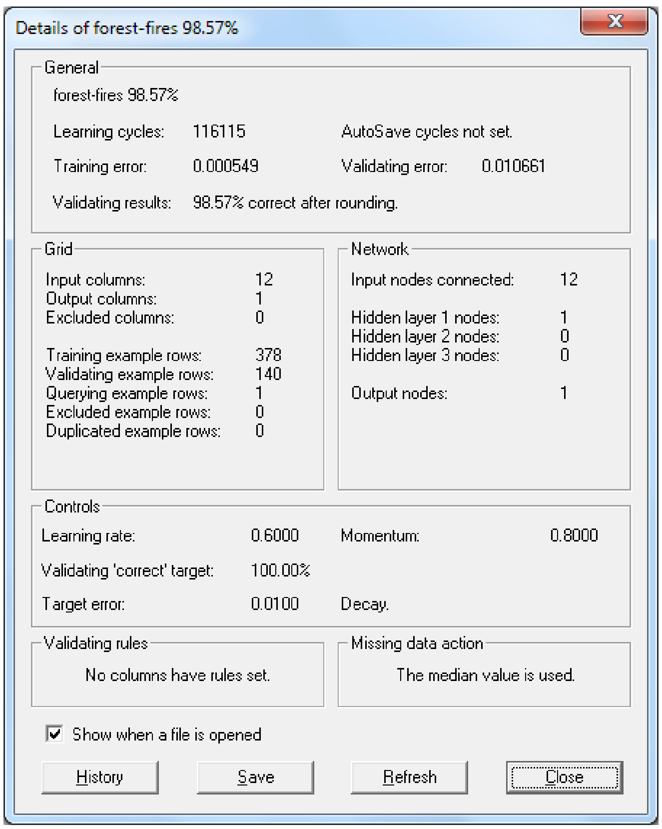


Figure 6: Summary of the trained model for predicting burned area of forest fires

## 7. Conclusions

Forest fires cause a significant environmental damage while threatening human lives. In the last two decades, a substantial effort was made to build automatic detection tools that could assist Fire Management Systems (FMS). The three major trends are the use of satellite data, infrared/smoke scanners and local sensors (e.g. meteorological). In this work, we propose an Artificial Neural Network model that uses meteorological data, as detected by local sensors in weather stations, and that is known to influence forest fires. The advantage is that such data can be collected in real-time and with very low costs, when compared with the satellite and scanner approaches. Recent real-world data, from the northeast region of Portugal, was used in the experiments. The database included spatial, temporal, components from the Canadian Fire Weather Index (FWI) and four weather conditions. This problem was modeled as a regression task, where the aim was the prediction of the burned area. The accuracy we got after training and testing the proposed model was 98.87%.

## References

1. Arrue, et al. (2000). “An Intelligent System for False Alarm Reduction in Infrared Forest-Fire Detection”. IEEE Intelligent Systems, 15(3):64–73, 2000.
2. Bi and K. Bennett. Regression Error Characteristic curves. In Proceedings of 20th International Conference on Machine Learning (ICML), pages 43–50, Washington DC, USA, 2003.
3. Breiman. Random Forests. Machine Learning, 45(1):5–32, 2001.
4. Breiman, et al. (1984). Classification and Regression Trees. Wadsworth, Monterey, CA, 1984.
5. Taylor and M. Alexander. Science, technology, and human factors in fire danger rating: the Canadian experience. International Journal of Wildland Fire, 15:121–135, 2006.
6. S. Menard. Applied Logistic Regression Analysis. SAGE, 2nd edition, 2001.
7. S. Taylor and M. Alexander. Science, technology, and human factors in fire danger rating: the Canadian experience. International Journal of Wildland Fire, 15:121–135, 2006.
8. UCI Machine Learning repository (https://archive.ics.uci.edu/ml/datasets.html)
9. JustNN Tool. http://www.justnn.com