Article

Predicting the damage of urban fires with Grammatical Evolution.

Constantina Kopitsa[1], Ioannis G. Tsoulos[2,*], Andreas Miltiadous[3],Vasileios Charilogis[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; a.miltiadous@uoi.gr
[4] Department of Informatics and Telecommunications, University of Ioannina, Greece; v.charilog@uoi.gr
* Correspondence: itsoulos@uoi.gr

Abstract: Fire, whether wild or urban, depends on the triad of oxygen, fuel, and heat. Urban fires, though smaller in scale, have devastating impacts, as evidenced by the 2018 wildfire in Mati, Attica (Greece), which claimed 104 lives. The elderly and children are the most vulnerable due to mobility and cognitive limitations. This study applies Grammatical Evolution (GE), a machine learning method that generates interpretable classification rules to predict the consequences of urban fires. Using historical data (casualties, containment time, meteorological/demographic parameters), GE produces classification rules in human readable form. The rules achieve over 85% accuracy, revealing critical correlations. For example, high temperatures (>35°C) combined with irregular building layouts exponentially increase fatality risks, while firefighter response time proves more critical than fire intensity itself. Applications include dynamic evacuation strategies (real-time adaptation), preventive urban planning (fire-resistant materials, green buffer zones), and targeted awareness campaigns for at-risk groups.Unlike "black-box" machine learning techniques, GE offers transparent, human-readable rules, enabling firefighters and authorities to make rapid, informed decisions. Future advancements could integrate real-time data (IoT sensors, satellites) and extend the methodology to other natural disasters. Protecting urban centers from fires is not only a technological challenge but also a moral imperative to safeguard human lives and societal cohesion.

Keywords: Urban fires; Machine learning; Neural networks; Genetic Programming; Grammatical Evolution.

Received:
Revised:
Accepted:
Published:

Citation: Kopitsa, C.; Tsoulos, I.G.; Miltiadous, A., Charilogis, V. . Predicting the damage of urban fires with Grammatical Evolution. *Journal Not Specified* **2025**, *1*, 0. https://doi.org/

1. Introduction

Whereas, fires can be triggered by various causes significant fires are frequently outcome from the following disasters: storms, transportation accidents, criminal activity / terrorism, droughts, hazardous materials spills [1] and forest fires. Also, Urban fires are predominantly attributed to negligent cooking practices, whereas rural fires often stem from faulty electrical installations, malfunctions in heating systems, or even natural causes such as lightning strikes [2]. Small scale urban fires often do not have a significant impact on an area, nevertheless on the other hand they still are equally hazardous to human life, in the socio-economic stability of a community, or even they contribute to increased insurance premiums [1].

Referring to human lives, let us examine the numerical data provided by the Hellenic Fire Service, which is available as open data in accordance with the European Union Directive (2013/37/EE) aimed at enhancing transparency [3]. Consequently, as presented in Figure 1, an increase in fatalities associated with urban fires, has been observed since 2020, with a slight decline in this trend, in 2023.

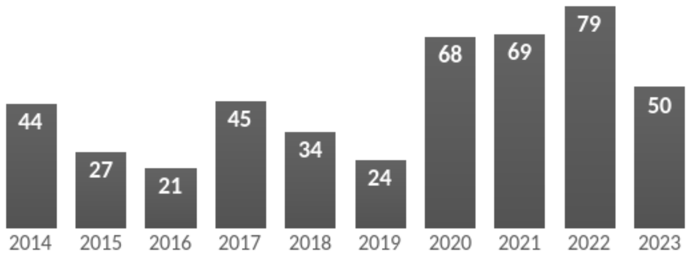

Figure 1. A graph presenting the deaths from fires in Greece for a period from 2014 to 2023.

Beyond this, it is observed that in 2018, the human casualties from the devastating fire in Mati, Attica, have not been accounted for, as previously mentioned in the Abstract [4,5] . This is likely due to the classification of this particular fire as a wildfire rather than an urban fire.

According to a study conducted, in America, by the National Fire Protection Association (NFPA), there was an approximate 4% rise in residential fires and a 13% increase in deliberately ignited structural fires in 2018 [2]. Thus, based on the official fire situation report, Fire Loss in the United States During 2017, published by the National Fire Protection Association, fire departments across the United States responded to an estimated 1,319,500 fire incidents in 2017. These incidents led to approximately 3,400 civilian fatalities, 14,670 civilian injuries, and an estimated $23 billion in direct property damage [2].

Additionally, the World Health Organization (WHO), an estimated three million fires occur globally each year, resulting in approximately 180,000 fatalities [6]. Moreover, the majority of these disasters take place in major urban centers, within low - income countries, where various economic, social, and environmental factors contribute to an increased risk of fire incidents.

Nevertheless, this conclusion does not exempt powerful nations from the impacts of climate change and natural disasters. Accordingly, the table 1 is introduced.

Table 1. How Climate change and Natural Disasters Effects even to powerful nations.

Country	Year	Causes of Forest Fires	Carbon Emissions	Hectares	Casualties
Australia (Black Summer) [7,8]	2019-2020	Dry winters, drought	900 million tons	19 million	33 people, 3.000 Houses & Buildings Billions of wild animals
Russia (Arctic fires)[9–11]	2019-2020	Dryer surface, higher temperature	31.1 megatones	24 million	No reports.
USA (L.A.) [13,14]	2025	High temperature	4.4 megatones	57000	30 people, 2.000.000 evacuated 16000 houses burnt
Japan (Ofunato) [15]	2025	High temperature	No reported.	2900	1 people, 4000 evacuated 210 buildings damaged

Climate change projections indicate that urban environments may face increasing fire hazards [12]. A stark example of the interconnectedness between climate change, and Forest fires are evident in the fires that occurred in Los Angeles in January 2025, and in Japan in February of the same year.

From these real events, it is inferred that an unpredictable natural disaster, such as a wildfire, does not discriminate between low and high-socioeconomic areas. Furthermore, it highlights the precarious balance (sword of Damocles) between climate change, and urban expansion into forested regions. This underscores the critical importance of our research in predicting the impacts damages of an urban fire.

Next, we will refer to relevant studies focused on the field of urban fire prediction, which employ statistical techniques, (GIS) spatial and temporal analyses, subjective evaluations through the Analytical Hierarchy Process (AHP), Multi Criteria Decision Making, and Probabilistic machine learning.

A team of researchers, used data from the Ankara region of Turkey, for analysis of the spatial and temporal patterns of residential fires, can enable decision-makers to strategically allocate resources for fire management, based on the intensity of fire clustering over time and across different locations [16]. Subsequently, the following study builds upon the methodology and effectiveness of a firefighter-led public education campaign on fire prevention, which successfully decreased both the frequency, and severity, of residential structure fires in high-risk areas of Surrey, British Columbia [17]. In the sequel, some studies have used GIS methods, to analyze the fire risk in urban areas. In Turkey, a study examines fires that have occurred in various locations across Turkey, including cold storage facilities, factories, and manufacturing plants. The case data was utilized to calculate risk scores using Geographic Information System (GIS), Analytical Hierarchy Process (AHP), and Inverse Distance Weight (IDW) methods [18]. In China, in order to select the best suitable fire brigade zone, they analyze: fire-risk areas, traffic congestion, land cover, and location. They employ various methods, including Geographic Information Systems (GIS), Multi-Criteria Decision Making (MCDM), and Location-Allocation (L-A) techniques, along with multi-source geospatial data such as land cover, points of interest, drive time, and statistical yearbooks. Additionally, they used Analytic Hierarchy Processes (AHP) to thoroughly assess undeveloped areas based on factors such as location, topography, and potential fire-risk zones [19]. Afterwards, a study seeks to assess fire risk in urban areas by analyzing 19 factors related to economic, social, and built environment aspects, as well as past fire incidents. It employs Multi-Criteria Decision Making (MCDM) techniques, specifically the Analytic Hierarchy Process (AHP) to determine the significance and weighting of each criterion. To illustrate the method's effectiveness, the research develops an urban vulnerability index map for Ardabil, Iran, using the Fuzzy-VIKOR approach within a Geographic Information System (GIS) framework [20]. Building upon, Analytic Hierarchy Process, the paper from Taiwan, evaluated the severity of building fires across 17 villages in Taishan District, New Taipei City. A comprehensive literature review was conducted to examine the influence of fire severity assessment criteria, which served as the foundation for identifying key factors, and developing evaluation items, within the (AHP) framework [21]. Turkish colleagues proposed a group decision-making (GDM) approach, integrating the recently developed Best–Worst Method (BWM) [22], a multi-criteria decision-making (MCDM) technique [23], with Geographic Information Systems (GIS) to identify optimal locations for new emergency facilities in Istanbul. Their analysis incorporated the input of two decision-makers [24]. The next study seeks to expand the limited empirical research on urban fires in the Global South by analyzing their causes and dynamics. Focusing on disaggregated fire incident data from Kathmandu Metropolitan City (KMC), Nepal, the research identifies key contributing factors and examines both the spatial and temporal distribution of urban fires [25]. In the sequel, we present studies that employ techniques belonging to the broader family of probabilistic machine learning. Studies from Japan integrate the earthquake factor as a cause of urban fires. The following article introduces the development of a stochastic model designed for time series forecasting of post-earthquake fire ignitions in buildings, aiming to enhance post-earthquake fire risk assessment [26]. The same researcher also developed a physics-based urban fire spread model that incorporates the stochastic occurrence of spot fires in the wooden residential areas of Itoigawa City. Utilizing the Monte Carlo method, they compared the simulated results with the actual fire damage recorded in 2016 [27]. Concluding with Japan, a probabilistic approach is introduced to evaluate the cascading risks associated with ground shaking and post-earthquake fires, on a regional scale [28].

We will proceed with related studies that make use of machine learning, and more specifically, supervised learning techniques.The first study examines existing research on the social, economic, and building stock characteristics associated with residential fire risk in urban neighborhoods [29]. From Australia in 2010, we have a paper utilizing the Bayesian approach, to produce detailed spatial forecasts of residential household fires, across metropolitan South-East Queensland [30]. Also, from Australia, a study employs a Markov chain approach to estimate the likelihood of residential fire occurrences based on historical fire data. Utilizing fire incident records collected over a decade in Melbourne, Australia, the spatially integrated fire risk model forecasts potential fire events by incorporating spatial and temporal variables as key predictive factors [31]. The next study was conducted in Greece, to develop a fire risk estimation model that integrates recent land cover changes alongside other critical risk factors. They implemented a Support Vector Machine (SVM) algorithm [32] combined with the Analytic Hierarchy Process (AHP), within a Geographic Information System (GIS) platform. This approach allowed for a more precise assessment of fire-prone areas. As a case study, they applied this methodology to the Dadia-Lefkimi - Soufli National Forest Park, ensuring a comprehensive evaluation of fire risk in the region [33]. Subsequently, American researchers, have introduced two machine learning models, utilizing Random Forest [34,35] and Extreme Gradient Boosting (XGBoost) [36], to forecast future service demand in urban areas based on spatial data analysis, in collaboration with Victoria Fire Department, U.S.A [37]. Moving forward, another fire risk model was implemented within the Pittsburgh Bureau of Fire (PBF), and an initial risk model was developed for predicting residential property fire risk [38]. Furthermore, the subsequent study, is a novel deep sequence learning model, referred to as the Fire Situation Forecasting Network (FSFN), is introduced to enhance the processing of information and the analysis of Spatio -temporal correlations within regional urban fire alarm datasets [39]. The following study, conducted in Iran, sought to apply machine learning algorithms to enhance the accuracy of predicting firefighting operation duration in urban areas, while also identifying the key factors that significantly impact this timeframe [40]. Moreover, the next study investigates urban fire incidents in Austin, Texas, by employing machine learning techniques, specifically Random Forest and time series modeling through the Autoregressive Integrated Moving Average (ARIMA) approach. The analysis reveals that ARIMA models generally perform better in forecasting most categories of fires, with the exception of vehicle-related fires. Furthermore, the findings underscore considerable variation in model accuracy across different urban districts, suggesting that localized factors significantly influence fire incidence prediction[41]. Building upon the need to address both spatial and temporal dimensions of urban fire risk, the following study introduces a deep neural network (DNN) framework designed to generate 30-day cumulative fire occurrence maps at a spatial resolution of 2.5 km × 2.5 km for the metropolitan area of Hangzhou, China. Drawing on a rich dataset spanning nine years (2015–2023), the proposed approach synthesizes diverse data sources—including meteorological variables, urban land use information, and historical daily fire incident records—to enhance predictive accuracy and provide a holistic view of urban fire dynamics [42].

In an effort to advance the understanding of fire risks in urban residential settings, the next paper introduces a predictive framework that integrates tree-based machine learning algorithms (Random Forest, AdaBoost, XGBoost, and CatBoost) with resampling strategies to estimate the likelihood of damage and casualties resulting from residential building fires. However, XGBoost was the most time-efficient [43]. The following study was conducted in Oregon and involved the analysis of over 48,000 reported structure fire incidents that occurred between January 2012 and August 2023. The dataset comprised 2,136 fires resulting in civilian casualties, including 317 confirmed fatalities. To assess the severity of injuries, bagged decision tree classifiers utilizing the random forest algorithm were employed. These models were used to evaluate the relative importance of various contributing factors, including socioeconomic conditions, population demographics, structural and behavioral incident characteristics, and the availability of local infrastructure [44]. Furthermore, colleagues from Seoul, Korea applied machine learning to predict fire-related property damage, and analyze contributing factors using three years of spatial fire data. Using k-fold cross-validation, the random forest algorithm achieved 83% accuracy in forecasting property damage [45].

Closing the introduction, a brief reference will also be made to machine learning, which employs unsupervised learning techniques. With this in mind, a weighted fire risk calculation method was developed, incorporating the frequency of fire occurrences, direct economic losses, and fire-related casualties. According to this approach, and with enhancements to the K-means clustering algorithm [46], this study introduced a fire risk K-means clustering model. This model offers an improved solution for the automated classification of fire risk levels [47]. The subsequent paper, employs an unsupervised deep learning (DL) approach to categorize hazard levels at fire sites and utilizes an autoregressive integrated moving average (ARIMA) model to predict temperature variations, leveraging the extrapolation capabilities of a random forest regressor [48]. In a related effort, the next article, introduces a methodology for forest fire detection utilizing unsupervised location-expert autoencoders in conjunction with Sentinel-1 SAR time series data. The models are trained on multitemporal SAR imagery from a designated reference period and are used to identify anomalous time series within the same region during a test period. Three variations of the autoencoder are presented, incorporating either temporal or spatiotemporal features, and their performance is compared against that of a state-of-the-art supervised autoencoder [49]. Aligned with similar methodologies, the following research may represent a pioneering effort in applying clustering analysis to explore the fire response of reinforced concrete (RC) columns. The results clearly demonstrate that unsupervised machine learning can yield valuable insights for fire engineering—insights that are often overlooked by conventional supervised learning approaches [50].

Following this approach, the subsequent study applies two unsupervised learning techniques, Principal Component Analysis (PCA) and K-means clustering utilizing Sentinel 2 satellite imagery, elevation data, and the Zagros Grass Index (ZGI) to detect areas at high risk, of wildfire in the increasingly vulnerable Kurdo Zagrosian forests. Among the two, PCA outperformed K-means by accurately identifying 80% of the areas burned between 2021 and 2023 as falling within moderate to high-risk fire zones [51].

This paper proposes the use of modern machine learning techniques based on Grammatical Evolution [52] to predict the potential damage caused by urban fires. This prediction was based on data that has been collected and subsequently digitized by the Greek Fire Service. After digitizing the original data, three categories were created depending on the size of the disaster that has been caused: small - scale disaster, medium - scale disaster and large - scale disaster. Therefore, the problem of predicting the magnitude of the disaster was transformed into a classification problem so that machine learning techniques could be applied to it. The techniques used in the conducted experiments include construction of neural networks [53,54], feature construction of artificial features from the original ones using Grammatical Evolution and production of classification rules. The obtained results are compared against the results from various traditional machine learning methods and a discussion is provided on the Results section of this manuscript. The methods utilized here include construction of artificial features from the original ones, construction of neural networks and production of classification rules. These techniques cover a wide range of machine learning techniques that have been presented in recent years and have shown high performance when applied to a variety of problems from various research areas. Furthermore, they can be used to effectively identify the most critical features of a problem, drastically reducing the number of inputs necessary for the efficient training of machine learning models. A key problem with classical machine learning techniques is the excessive number of features that a dataset can have in relation to the number of patterns that accompany it. The techniques used in this work can significantly reduce this number and select the most important ones for the effective training of machine learning techniques. Furthermore, in many cases some of the inputs to the problem may not contribute to the effective training of machine learning models and should be omitted.

The rest of this manuscript is divided as follows: in section 2 the proposed methods are presented in detail, in section 3 the experiments are illustrated and discussed and finally, in section 4 some conclusions are presented.

2. Materials and Methods

This section initiates with a detailed description of the used datasets and it continues with a brief presentation of the Grammatical Evolution technique and concludes with the full description of the used techniques.

2.1. The used datasets

The dataset utilized in this research was sourced from the Hellenic Fire Service in compliance with open data guidelines established by the European Union Directive (2013/37/EU), aiming to promote transparency and open access to governmental records. The dataset contains comprehensive records detailing urban fire incidents specifically for the calendar years 2014-2023. These datasets were downloaded from https://www.fireservice.gr/el_GR/synola-dedomenon(accessed on 29 April 2025).

For each urban fire event documented, detailed information was systematically collected, including the date and precise time of occurrence, allowing for temporal analysis and identification of patterns over various time intervals. The geographic location for each incident was also reported with a municipality code. In addition, specific characteristics relevant to each fire event were captured, including the probable cause or origin of the fire, which helps in identifying common fire risk factors within urban settings. Data regarding the structural properties involved, such as building type or property classification, was also documented, contributing to a comprehensive risk profile for urban infrastructure.

Furthermore, human casualty data detailing the number of fatalities and injuries associated with each incident were recorded. The dataset indicated an average of 0.002 fatalities per incident, with a maximum of 2 fatalities observed in a single event. Injuries averaged at approximately 0.0008 per incident, with a maximum of 1 injury recorded per event. Additionally, instances of burn injuries were relatively infrequent, averaging around 0.002 per incident, with a maximum count of 2 burn cases reported.

Information on the resources deployed was also comprehensively documented. On average, each incident involved approximately 1.65 firefighting vehicles, with a maximum of 24 vehicles responding to the most severe incidents. Personnel deployment averaged about 4.18 firefighters per incident, with an interquartile range from 2 to 5 firefighters, and up to 67 personnel attending a single event in extreme cases.

The dataset underwent thorough preprocessing procedures to ensure high-quality data for analysis. These procedures included validation checks for data accuracy, consistency, completeness, and the removal or correction of any identified errors or inconsistencies. Such rigorous preprocessing steps were critical for enhancing the reliability and validity of the analytical processes that followed. A review of data preprocessing techniques for neural networks is provided in the work of Nawi et al [55].

Based on data from the Hellenic Fire Service, we isolated and processed the pre- registered categories of Small, Medium, and Large Fires, excluding other categories that were not relevant to the scope of our research. The classification of a fire as Small, Medium, or Large may be determined in the field by firefighters, particularly when additional flammable materials are present in the surrounding area, when chemical substances that can accelerate the fire are involved, or when there is an exceptional risk to human lives. Furthermore, according to our data, urban fires have been recorded at 146 distinct locations, each assigned a unique identification number (code). For example, as an indicative reference, in 2023, the fire department responded to 21,606 fire incidents in apartment buildings (code 18), 9,106 on streets, 5,219 (code 22) in single-family homes (code 39), 4,971 in vacant lots (code 35), 2,389 in vehicles (code 6), 1,322 in duplex houses (code 37), and 1,270 in waste disposal areas (code 54). Beyond these, some unusual locations were also recorded, including 140 firefighting interventions in wells (code 143), 38 in cemetery (code 82), 106 in marine areas (code 69), 4 in public toilets (code 112), and so forth. Table 2 depicts the features used in this dataset.

Table 2. The used features.

Feature	Min Value	Max Value
Fire station	1	275
Region code	1	51
Month code	1	12
Season code	1	4
Area (Building type)	1	147
Persons involved	1	
Number of injuries	0	
Number of Burnt Victims	0	
Number of Fatalities	0	
Number of Vehicles involved	1	
Firefighters involved	1	

Hence, the problem of predict the extend of damage caused by urban fires can be considered as a classification problem, where any optimization method can be used to minimize the following training error:

$$E\left(M\left(\overrightarrow{x},\overrightarrow{p}\right)\right)=\sum_{i=1}^{K}\left(M\left(\overrightarrow{x_i},\overrightarrow{p}\right)-t_i\right)^2 \tag{1}$$

Where the function $M\left(\overrightarrow{x},\overrightarrow{p}\right)$ represents a machine learning model and the vector $\overrightarrow{p}$ represents the parameters of the model that should be estimated by any optimization method. The set $T=\left\{(x_1,t_1),(x_2,t_2),\ldots,(x_K,t_K)\right\}$ represents the training set of the input problem, where the vectors x_i stand for the input patterns and the values t_i are the expected outputs. The constant K represents the number of patterns in the train set.

2.2. Grammatical Evolution

The algorithm of Grammatical evolution can be considered as a genetic algorithm where the chromosomes, which are series of positive integer values, denote production rules of any given BNF (Backus–Naur form) grammar [56]. The method has been incorporated in various cases from real - world applications, such as data fitting [57,58], solution of trigonometric equations [59], composition of music [60], neural network construction [61,62], producing numeric constraints[63], video games [64,65], energy problems [66], combinatorial optimization [67], cryptography [68] etc. The BNF grammars are used to describe the syntax of programming languages and they can be defined as sets $G = (N, T, S, P)$ where

- The set N represents the non-terminal symbols of the grammar. Each non - terminal symbol can be replaced to a series of terminal symbols with the assistance of some associated production rules.
- The set T contains the terminal symbols.
- S is considers as the start symbol of the grammar with the assumption $S \in N$.
- The set P contains the production rules of the grammar, used to replace non - terminal symbols with series of terminal ones.

The production procedure of Grammatical Evolution starts from the symbol S and through a series of steps creates valid programs by replacing non - terminal symbols with series of terminal symbols following the selected rules. Every production rule is is selected usig the following steps:

- Obtain the next element V from the chromosome that is processed.
- Select the production rule as: Rule = V mod R, where R defines the total number of production rules for the non – terminal symbol that is under processing.

2.3. Neural Network Construction using Grammatical Evolution

The Neural Network Construction method was initial presented in the paper of Tsoulos et al [69] and it is used to determine the optimal architecture of artificial neural networks as well as the optimal set of parameters for the network. The Neural Network construction mechanism utilizes the Grammatical Evolution procedure in order to produce artificial neural networks in the form:

$$N(\overrightarrow{x}, \overrightarrow{w}) = \sum_{i=1}^{H} w_{(d+2)i-(d+1)} \sigma\left(\sum_{j=1}^{d} x_j w_{(d+2)i-(d+1)+j} + w_{(d+2)i}\right) \quad (2)$$

In this equation the term H represents the number of processing units (weights) of the network. The function $\sigma(x)$ stands for the sigmoid function. The grammar used by Grammatical Evolution to produce neural networks in the form of Equation 2 is outlined in Figure 2 which has been in the initial work for Neural Network construction with Grammatical Evolution [69] . This method has been incorporated in problems such as, chemistry problems [70], estimation of solutions of differential equations etc.

```
S:=<Sigval>
<Sigval>::=<Node>
           | <Node> + <Sigval>
<Node>::=<Number>*sig(<Sum>+<Number>)
<Sum>::= <Number>*<Xlist>
           |    <Sum>+<Sum>
<Xlist>::= x1        (0)
             |    x2  (1)
             ..............
             |    xd  (d-1)
<Number>::= (<Dlist>.<Dlist>)
             |    (-<Dlist>.<Dlist>)
<Dlist>::= <Digit>           (0)
             | <Digit><Dlist> (1)
<Digit>::= 0      (0)
             |  1 (1)
             ...........
             |  9 (9)
```

Figure 2. The proposed grammar neural network construction procedure.

The main steps of the algorithm used to produce neural networks are outlined below:

1. **Initialization Step**.
   (a) **Set** as N_c the number of chromosomes and as N_g the number of allowed generations.
   (b) **Set** as p_s the selection rate with $p_s \leq 1$ and as p_m the mutation rate with $p_m \leq 1$.
   (c) **Initialize** randomly each chromosome $c_i,\ i = 1, \ldots, N_c$ as a set of randomly selected integers.
   (d) **Set** $k = 0$, as the generation counter.
2. **Fitness Calculation Step**.
   (a) **For** $i = 1, \ldots, N_c$ **do**
      i. **Create** using the grammar of Figure 2 the corresponding neural network $N_i(x)$ for the chromosome c_i.
      ii. **Set** as the fitness f_i of the chromosome g_i the training error of neural network $N_i(x)$.
   (b) **End For**
3. **Application of Genetic Operations**.
   (a) **Application of selection**. The best $p_s \times N_c$ chromosomes are copied to the next generation. The remaining are substituted by chromosomes produced during crossover and mutation.
   (b) **Application of crossover**. During this procedure new chromosomes will be created from selected chromosomes from the current generation. For each pair (z, w) of produced chromosomes two chromosomes p_1 and p_2 will be selected from the current population using tournament. selection. The new chromosomes will be produced using one - point crossover [72], which is graphically illustrated in Figure 3.
   (c) **Application of mutation**. For every element of each chromosome a random number $r \leq 1$ is drawn. The corresponding element is altered randomly when $r \leq p_m$.
4. **Termination Check Step**.

   (a) **Set** $k = k + 1$
   (b) **If** $k < N_g$ **then go to** Fitness Calculation Step.

5. **Testing step**.
   (a) **Obtain** the chromosome c^* with the lowest fitness value in the population.
   (b) **Create** the corresponding neural network $N^*(x)$ and apply it to the test set and report the associated error.

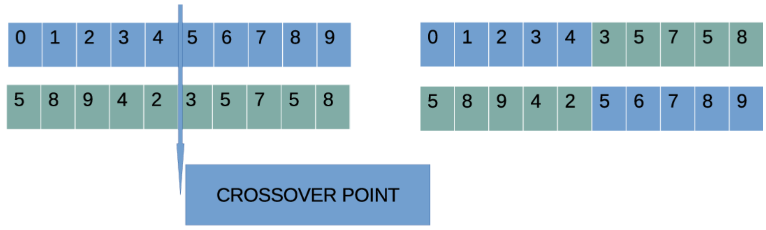

Figure 3. An example of the one - point crossover procedure.

2.4. Feature construction using Grammatical Evolution

The next method used in the performed experiments is the Feature Construction technique, initially presented in the work of Gavrilis et al [73]. The method utilizes the Grammatical Evolution procedure to create artificial features from the original ones and hence it can be used to enhance the effectiveness of any applied machine learning model to the artificial data. The artificial features are non - linear mappings of the original ones and the grammar used by the method to construct such features is shown in Figure 4, which also was initially presented in the original work of the feature construction method [73].

Figure 4. The grammar used in Feature Construction method.

```
S::=<expr>   (0)
<expr> ::=  (<expr> <op> <expr>)
           | <func> ( <expr> )
           |<terminal>
<op> ::=     +
           | -
           | *
           | /
<func> ::=   sin
           | cos
           |exp
           |log
<terminal>::=<xlist>
           |<digitlist>.<digitlist>
<xlist>::=x1
           | x2
           .........
           | xN
<digitlist>::=<digit>
           | <digit><digit>
           | <digit><digit><digit>
<digit>  ::= 0 | 1 | 2 | 3| 4| 5| 6| 7| 8| 9
```

The features produced by this procedure can be evaluated using any machine learning method, although the Radial Basis Function (RBF) networks [74,75] were used due to the speed of their corresponding training procedure. The main steps of this procedure are the following:

1. **Initialization step**.
   (a) **Define** as N_c the number of chromosomes and as N_g the number of allowed generations.
   (b) **Define** the selection rate p_s and the mutation rate p_m.
   (c) **Define** as N_f the number of constructed features.
   (d) **Initialize** the $c_i,\ i=1,\ldots,N_c$ chromosomes as vectors of randomly selected integers.
   (e) **Set** $k=0$, the generation counter.
2. **Fitness calculation step**.
   (a) **For** $i=1,\ldots,N_c$ **do**
      i. **Create** N_f artificial features $y_1,y_2,\ldots,y_{N_f}$ for the chromosome c_i. The production is performed using the grammar of Figure 4.
      ii. **Modify** the train set of the objective problem using the features $y_1,y_2,\ldots,y_{N_f}$.
      iii. **Apply** a machine learning model to the modified set and define as the fitness value f_i the corresponding training error.
   (b) **End For**
3. **Application of genetic operations**. Apply the same genetic operations as in the case of Neural Construction method of subsection 2.3.
4. **Termination check step**.
   (a) **Set** $k=k+1$
   (b) **If** $k<N_g$ go to Fitness calculation step.
5. **Testing step**.
   (a) **Obtain** the chromosome c^* with the lowest fitness value.
   (b) **Produce** the features $y_1^*,y_2^*,\ldots,y_{N_f}^*$ for this chromosome.
   (c) **Modify** the test set of the objective problem using the previously created features.
   (d) **Apply** any machine learning model to the test set and report the associated error.

2.5. Create classification rules using Grammatical Evolution

The third method used in the conducted experiments which is based on Grammatical Evolution is the method that produces classification rules [76]. This method has also been published as a software recently [77]. The BNF grammar used by this method is shown in Figure 5.

Figure 5. The grammar used by the method that produces classification rules using Grammatical Evolution.

```
<S>::=if(<BEXPR>) CLASS=0 else CLASS=1
<BEXPR>::=<XLIST><BOOLOP><EXPR>
         |!(<BEXPR>)
         |<XLIST><BOOLOP><EXPR>&<BEXPR>
         |<XLIST><BOOLOP><EXPR>|<BEXPR>
<BOOLOP>::= >
          | >=
          | <
          | <=
<EXPR>::= (<EXPR><BINARYOP><EXPR>)
        | <FUNCTION>(<EXPR>)
        | <TERMINAL>
<BINARYOP>::= +
            | -
            | *
            | /
<FUNCTION>::= sin | cos | exp | log
<TERMINAL>::= <XLIST>
            | <DIGITLIST>.<DIGITLIST>
            | (-<DIGITLIST>.<DIGITLIST>)
 <XLIST>::= x1 | x2 | ...|xD
<DIGITLIST>::=<DIGIT>
             | <DIGIT><DIGIT>
             | <DIGIT><DIGIT><DIGIT>
<DIGIT>::= 0 | 1 | 2 | 3 | 4 |5 |6 |7 |8 |9
```

The main steps of this method have as follows:

1. **Initialization step**.
   (a) **Define** as N_c the total number of chromosomes and with N_g the allowed number of generations.
   (b) **Define** the selection rate p_s and the mutation rate p_m.
   (c) **Initialize** as vectors of randomly selected integers the chromosomes $c_i,\ i = 1, \ldots, N_c$.
   (d) **Set** $k = 0$, the generation counter.
2. **Fitness calculation step**.
   (a) **For** $i = 1, \ldots, N_c$ **do**
      i. **Create** using the Grammatical evolution procedure and the grammar depicted in Figure 5 a classification program G_i for the corresponding chromosome c_i.
      ii. **Set** the fitness f_i as
$$f_i = \sum_{j=1}^{M} \left(G_i(x_j) - t_j\right)^2 \tag{3}$$
      for the corresponding training set $T = \{(x_1, t_1), (x_2, t_2), \ldots, (x_M, t_M)\}$. The values x_i denote the input patterns and the value t_i the expected outcome for pattern x_i.
   (b) **End For**
3. **Genetic operations step**. Apply the same genetic operations as in the case of Neural Construction method of subsection 2.3.
4. **Termination check step**.

   (a) **Set** $k = k + 1$
   (b) **If** $k < N_g$ then go to Fitness calculation step.
5. **Testing step**.
   (a) **Obtain** the best chromosome c^* and produce the associated classification program G^*.
   (b) **Apply** the classification program to the test set of the problem and report the result.

3. Results

The code used in the experiments was implemented in C++ programming language with the assistance of Optimus optimization environment, freely available from https://github.com/itsoulos/GlobalOptimus/ (accessed on 2 April 2025). Also, the freely available programming tool of WEKA [78] was used for some of the experiments, that can be downloaded also freely from https://ml.cms.waikato.ac.nz/weka/(accessed on 2 April 2025). For the validation of the experiments the method of ten - fold cross validation was incorporated. All the experiments were conducted on a machine running Debian Linux with 128GB of RAM. The values for the experimental settings are shown in Table 3.

Table 3. The values used for the experimental settings.

PARAMETER	MEANING	VALUE
N_c	Number of chromosomes	500
N_g	Maximum number of generations	200
p_s	Selection rate	0.10
p_m	Mutation rate	0.05
N_f	Number of produced features	2
H	Number or processing nodes	10

Also, the Table 4 contains the experimental results where the following notation is used:

1. The column YEAR denotes denotes the year for which the methods were applied.
2. The column PATTERNS denotes the number of patterns in the test set for every year.
3. The column BAYES NET denotes the application of the Bayesian Network method [79,80].
4. The column MLP represents the incorporation of an artificial neural network with $H = 10$ processing nodes, that was trained using the Back Propagation method [81,82].
5. The column RBF denotes the usage of a Radial Basis Function network with $H = 10$ processing nodes.
6. The column NNC represents the usage of the Neural Network construction method described in subsection 2.3.
7. The column FC stands for the usage of the Feature Construction method provided in subsection 2.4.
8. The column GENCLASS denotes the usage of the method that creates classification rules, described in subsection 2.5.
9. The final row AVERAGE represents the average classification error for all years between 2014 and 2023.

Table 4. Experimental results using various machine learning techniques. Numbers in cells represent average classification error as measured on the corresponding test set.

YEAR	PATTERNS	BAYES NET	MLP	RBF	NNC	FC	GENCLASS
2014	1287	9.01%	8.34%	11.97%	7.82%	7.35%	7.03%
2015	1686	8.75%	7.63%	10.79%	7.13%	6.78%	6.65%
2016	1735	8.99%	8.13%	10.73%	7.48%	7.10%	7.05%
2017	1736	8.43%	8.29%	10.78%	7.68%	7.24%	7.13%
2018	1637	8.38%	7.99%	9.33%	7.30%	7.11%	6.74%
2019	1971	7.31%	7.78%	24.81%	7.33%	7.15%	6.26%
2020	1990	8.70%	7.86%	10.03%	6.99%	6.66%	6.47%
2021	1883	8.58%	7.88%	10.37%	6.87%	6.55%	6.39%
2022	2036	8.80%	7.29%	8.45%	6.86%	6.52%	6.32%
2023	1978	8.46%	7.50%	10.56%	6.87%	6.58%	6.20%
AVERAGE		**8.54%**	**7.87%**	**11.78%**	**7.23%**	**6.90%**	**6.62%**

In Table 4 GENCLASS demonstrates the lowest average classification error (6.62%), making it the most reliable model for wildfire prediction. This method presents better results compared to other techniques as it can isolate the necessary features of the problem but also identify hidden correlations that could lead to lower classification errors through the automatic creation of classification rules. It is followed by FC with an average error of 6.90% and NNC with 7.23%. Traditional models such as BAYES NET (8.54%) and MLP(BP) (7.87%) exhibit higher error rates, while RBF performs the worst with an average error of 11.78%, including an exceptionally high value in 2019 (24.81%), likely due to overfitting or sensitivity to outliers. GENCLASS not only has the lowest error but also shows consistent improvement over time. Starting at 7.03% in 2014, it decreased to 6.20% in 2023, with minor fluctuations in between. FC and NNC also display a declining trend but with greater variability. In contrast, MLP(BP) and BAYES NET show no clear improvement, with BAYES NET even experiencing a slight increase in error in 2022–2023. RBF, despite improving after 2019, remains unstable and less reliable. When comparing the top three models (GENCLASS, FC, NNC), GENCLASS consistently outperforms the others in all years except 2019, where FC had a marginally better performance. NNC, while superior to traditional methods, lags behind GENCLASS and FC. The remaining models (BAYES NET, MLP(BP), RBF) appear less competitive in accuracy compared to newer techniques. The data strongly supports GENCLASS as the optimal model for wildfire prediction due to its consistently low and stable error rate, as well as its progressive improvement over time. FC and NNC remain viable alternatives with good performance, but GENCLASS maintains a clear advantage. The other models, particularly RBF, may require further optimization to enhance reliability. The steady performance of GENCLASS makes it the safest choice for practical applications.

Within the framework of statistical analysis, R language scripts were executed to extract significance levels (p-values) for performance differences between the classification models. The results, shown in Figure 6, reveal statistically significant differences between the compared models. Specifically, all pairwise model comparisons yielded p-values below the standard significance threshold (typically $p < 0.05$), indicating statistically significant performance differences [83]. The comparison between BAYES NET and MLP(BP) produced $p = 0.0098$, while BAYES NET's comparisons with RBF, NNC, FC and GENCLASS showed even smaller values ($p = 0.0039$ and $p = 0.002$), confirming that BAYES NET differs significantly from other models. Similarly, MLP(BP) comparisons with RBF, NNC, FC and GENCLASS all resulted in $p = 0.002$, demonstrating high statistical significance in their performance differences. The same holds true for RBF's comparisons with NNC, FC and GENCLASS ($p = 0.002$), as well as for comparisons between NNC, FC and GENCLASS. The

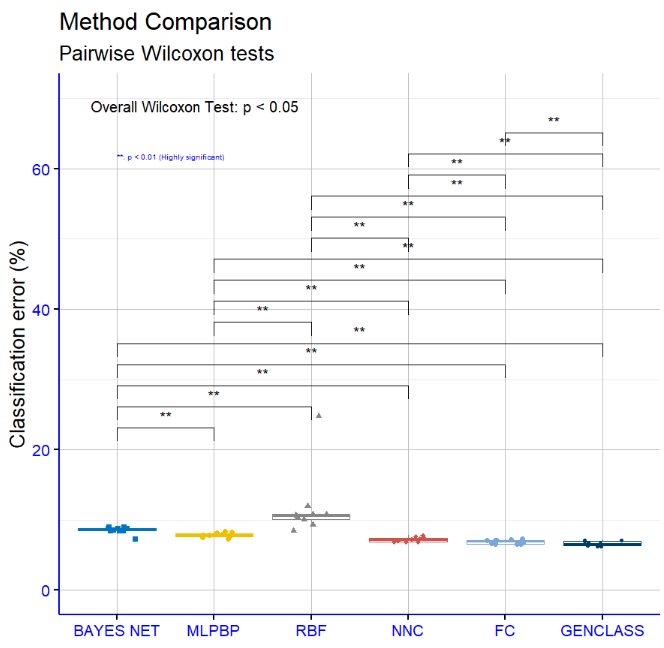

Figure 6. Statistical comparison of the experimental results obtained by various machine learning methods.

fact that all p-values are very small ($p \leq 0.0098$) confirms that the models do not perform equally and that statistically significant differences exist between them. This is particularly evident in comparisons involving GENCLASS, which - as previous analysis has shown - stands out for its high accuracy. These results reinforce the conclusion that certain models (such as GENCLASS and FC) are clearly superior to others (like BAYES NET and RBF), a finding that should be considered in practical machine learning applications.

Also, as an example consider the plot of Figure 7 where the train and test error for year 2016 are presented using the GENCLASS method.

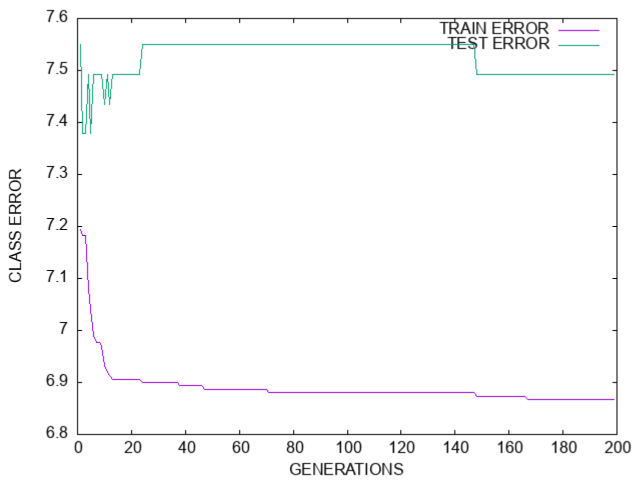

Figure 7. An example plot for the GENCLASS method.

As shown in this graph, the error in the training set gradually decreases as the generations increase and the error in the test set initially decreases to reach a constant value after some generations of the genetic algorithm.

4. Conclusions

This study implements the innovative technique of Grammatical Evolution to predict the consequences of urban fires, utilizing a decade of data from the Hellenic Fire Service. The results demonstrate the clear superiority of the GENCLASS method over other machine learning approaches, both in terms of accuracy and interpretability. The method's ability to generate human-readable classification rules constitutes a significant advantage over traditional "black-box" machine learning models. These rules reveal complex correlations between factors such as meteorological conditions, urban layout, and human activity, providing valuable insights for fire prevention and management. However, the research does face certain limitations that warrant discussion. The reliance on historical data reduces predictive capability in extreme or unprecedented scenarios, such as those caused by climate change. Additionally, the performance of some models, like RBF, shows significant instability in certain cases, likely due to overfitting or sensitivity to outliers. This underscores the need for further algorithm optimization and the integration of additional data to enhance reliability.

The findings of this research confirm that Grammatical Evolution, particularly the GENCLASS method, offers a robust solution for predicting the impacts of urban fires. This method not only achieves the lowest average classification error but also provides transparent, interpretable rules that can be directly utilized by fire departments and urban planners. The generated rules uncover critical dependencies, such as the influence of temperature, emergency response times, and demographic characteristics on the extent of damage. This paves the way for developing dynamic evacuation strategies, implementing preventive measures in vulnerable areas, and raising public awareness of fire risks. Moreover, the consistent improvement in GENCLASS's performance over time suggests its adaptability to changing conditions and its potential for enhancement with new data. However, it is important to recognize that the effectiveness of any predictive method depends heavily on the quality and completeness of available data, as well as the ability to account for new and unforeseen factors, such as climate change.

To further develop the findings of this research and strengthen their practical application, several directions for future exploration are proposed. First, the integration of real-time data from sensors and satellite systems could significantly improve the accuracy and timeliness of predictions, enabling dynamic model adjustments and rapid responses to emerging threats. Second, extending the method to other geographic regions with different urban and climatic characteristics could explore the generalizability of the findings and identify new factors influencing fire risk. Third, combining Grammatical Evolution with advanced deep learning techniques, such as neural networks for spatiotemporal analysis, could enhance predictive capabilities in complex scenarios involving multiple simultaneous hazards. Finally, developing simulations that account for the long-term impacts of climate change and urban expansion could provide valuable insights for designing more resilient and safer urban environments. These directions have the potential to transform the research findings into practical tools for safeguarding human lives and urban infrastructure.

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.M. 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.

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. Fires – Wildfires and Urban Fires. Juniata County Appendix CMulti-Jurisdictional Hazard Mitigation Plan Hazard Profiles. 2008. Pp. 15-19. Available from: https://juniataco.org/docs/hmp/Appendix%20C%20-%2004-Fire-Urban%20and%20Rural.pdf (Accessed on 07 March 2025).
2. M.R. Hossain, O. Smirnov, Analyzing the risk factors of residential fires in urban and rural census tracts of Ohio using panel data analysis, Applied geography **151**, 102863, 2023.
3. Hellenic Fire Service. Open Data. Incident record. Available from: https://www.fireservice.gr/el_GR/synola-dedomenon (Accessed on 07 March 2025).
4. Greek Wikipedia. The trial regarding MATI's wildfire. Available from: https://el.wikipedia.org/wiki/%CE%94%CE%AF%CE%BA%CE%B7_%CE%B3%CE%B9%CE%B1_%CF%84%CE%BF_%CE%9C%CE%AC%CF%84%CE%B9 (Accessed on 27 February 2025).
5. G. Xanthopoulos, M. Athanasiou, Uniting Our Global Wildfire Community. Wildfire, International Association of Wildland Fire. April 2019, volume 28.2.
6. World Health Organization (WHO). Burns. 13 October 2023. Available from: https://www.who.int/news-room/fact-sheets/detail/burns (Accessed on 13 March 2025).
7. Natural Hazards Research Australia. Understanding the Black Summer bushfires through research: a summary of key finding from the Bushfire and Natural Hazards CRC. January 2023. Available from: https://www.naturalhazards.com.au/sites/default/files/2023-01/Understanding%20the%20Black%20Summer%20bushfires%20through%20research_final_web_NHRA.pdf (Accessed on 13 March 2025).
8. Australian Government. Australian Public Service Commission. Black Summer. State of the Service Report 2019 – 20. Available from: https://www.apsc.gov.au/state-service/state-service-report-2019-20/chapter-1-commitment-service/black-summer (Accessed on 13 March 2025).
9. Fires char the Siberian Arctic. NASA, Earth observatory. 10 July 2024. Available from:https://earthobservatory.nasa.gov/images/153087/fires-char-the-siberian-arctic (Accessed on 17 March 2025).
10. Landsat Image Gallery. NASA. Available from: https://landsat.visibleearth.nasa.gov/view.php?id=153087 (Accessed on 17 March 2025) .
11. Raging Wildfires Devastate Russia's Far East Sakha Republic. The Moscow Times. By Leyla Latypova. 23 July 2024. Available from: https://www.themoscowtimes.com/2024/07/23/raging-wildfires-devastate-russias-far-east-sakha-republic-a85802 (Accessed on 17 March 2025) .
12. Cities face rising fire risks from climate change without emission cuts. ChosunBiz. By Park Keun-tae. 03 March 2025. Available from: https://biz.chosun.com/en/en-science/2025/03/05/FXRLKFRXJJB5LK4YXKPLVETVJM/ (Accessed on 17 March 2025) .
13. Here's how climate change fueled the Los Angeles fires. Npr org. By Lauren Sommer. 29 January 2025. Available from: https://www.npr.org/2025/01/29/nx-s1-5273676/la-fires-climate-change-rainfall-extreme-weather (Accessed on 17 March 2025).
14. Graphics Explain Los Angeles ' Rare and Devastating January Fires. World Resources Institute. Wri org. By James McCarthy & Jessica Richter. 05 February 2025. Available from: https://www.wri.org/insights/los-angeles-fires-january-2025-explained (Accessed on 17 March 2025).
15. Fire Grows Unusually Large in Japan. NASA. Earth observatory. Available from: https://earthobservatory.nasa.gov/images/154008/fire-grows-unusually-large-in-japan (Accessed on 17 March 2025).
16. E. Ceyhan, K. Ertugay, S. Duzgun, Exploratory and inferential methods for spatio-temporal analysis of residential fire clustering in urban areas, Fire safety Journal **58**, pp. 226-239, 2013.
17. J. Clare, L. Garis, D. Plecas, C. Jennings, Reduced frequency and severity of residential fires following delivery of fire prevention education by on-duty fire fighters: Cluster randomized controlled study, Journal of Safety Research **43**, pp. 123-128, 2012.
18. S. Alkis, E. Aksoy, K. Akpinar, Risk Assessment of Industrial Fires for Surrounding Vulnerable Facilities Using a Multi- Criteria Decision Support Approach and GIS, Fire **4**, 53, 2021.
19. Y. Jiang, A. Lv, Z. Yan, Z. Yang, A GIS-Based Multi-Criterion Decision-Making Method to Select City Fire Brigade: A Case Study of Wuhan, China, International Journal of Geo-Information. ISPRS **10**, 777, 2021.

20. S. Noori, A. Mohammadi, T. Ferreira, G. Miguel., A. Gilandeh, M. Ghaffari, S. Ardabili, J. Seyed, Modelling and Mapping Urban Vulnerability Index against Potential Structural Fire-related Risks: An Integrated GIS-MCDM Approach, Fire **6**, 107, 2023.
21. C-A. Lee, Y.-C. Sung, Y-S. Lin, G.K-K. Hsiao, Evaluating the severity of building fires with the analytical hierarchy process, big data analysis, and remote sensing, Natural Hazards **103**, pp. 1843-1856, 2020.
22. Pamučar, D., Ecer, F., Cirovic, G., & Arlasheedi, M. A. (2020). Application of improved best worst method (BWM) in real-world problems. Mathematics, 8(8), 1342.
23. Taherdoost, H., & Madanchian, M. (2023). Multi-criteria decision making (MCDM) methods and concepts. Encyclopedia, 3(1), 77-87.
24. Erden, Turan., Hopkins Nyimbili, Penjani. Comparative evaluation of GIS-based best – worst method (BMW) for emergency facility planning: perspectives from two decision-maker groups. Natural Hazards, volume 105, issue 1. January 2021. Pp. 1031-1067.
25. K. KC, R. Ardianto, P. Chhetri, J. Corcoran, Geographic patterns of urban fires in the global south: the case of Kathmandu, Nepal. GeoJournal **89**, 137, 2024.
26. T. Nishino, A. Hokugo, A stochastic model for time series prediction of the number of post – earthquake fire ignition in buildings based on the ignition record for the 2011 Tohoku Earthquake, Earthquake Spectra **36**, 2019.
27. T. Nishino, Physics – based urban fires spread simulation coupled with stochastic occurrence of spot fires, Stochastic Environmental Research and Risk Assessment **33**, pp. 451 – 463, 2019.
28. T. Nishino, Probabilistic urban cascading multi-hazard risk assessment methodology for ground shaking and post – earthquake fires, Natural Hazards **116**, pp. 3165 – 3200, 2023.
29. C.R. Jennings, Social and economic characteristics as determinants of residential fire risk in urban neighborhoods: A review of the literature, Fire safety Journal **62**, pp. 13-19, 2013.
30. D. Rohde, J. Corcoran, P. Chhetri, Spatial forecasting of residential urban fires: A Bayesian approach. Computers, Environment and Urban Systems **34**, pp. 58-69, 2010.
31. R. Ardianto, P. Chhetri, Modeling Spatial-Temporal Dynamic of Urban Residential Fire Risk Using a Markov Chain Technique, International Journal of Disaster Risk Science **10**, pp. 57-73, 2019.
32. 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.
33. Y. Maniatis, A. Doganis, M. Chatzigeorgiadis, Fire Risk Probability Mapping Using Machine Learning Tools and Multi – Criteria Decision Analysis in the GIS Environment: A Case Study in the National Park Forest Dadia-Lefkimi-Soufli, Greece, Applied Science **12**, 2938, 2022.
34. L. Breiman, Random forests, Machine learning **45**, pp. 5-32, 2001.
35. S.J. Rigatti, Random forest, Journal of Insurance Medicine **47**, pp. 31-39, 2017.
36. Chen, T., & Guestrin, C. (2016, August). Xgboost: A scalable tree boosting system. In Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining (pp. 785-794).
37. Dey, Arnab., Heger, Andrew., England, Darin. Urban Fire Station Planning using Predicted Demand and Service Quality Index. Springer Nature 2021.
38. Singh Walia, Bhavkaran., Hu, Qiany., Chen, Jeffrey., Chen, Fangyan., Lee, Jessica., et al. A Dynamic pipeline for Spatio - Temporal Fire Risk Prediction. Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining.2018. pp. 764–773.
39. G. Jin, Q. Wang, C. Zhu, Y. Feng, J. Huang, X. Hu, Urban Fire Situation Forecasting: Deep sequence learning with Spatio – temporal dynamics, Applied Soft Computing **97**, 2020.
40. A. Sahebi, B. Havasy, Y. Veisani, Predicting firefighting operation time in urban areas using machine learning: identifying key determinants for improved emergency response, Discover Applied Science **7**, 2025.
41. Yuan, Y., & Wylie, A. G. (2024). Comparing Machine Learning and Time Series Approaches in Predictive Modeling of Urban Fire Incidents: A Case Study of Austin, Texas. ISPRS International Journal of Geo-Information, 13(5), 149.
42. Zhou, Y., Lin, P., & Wang, N. (2025). A deep neural network approach for regional-scale 30-day accumulated urban fire occurrence forecast. Fire Safety Journal, 152, 104331.
43. Liu, Z., & Zhuang, Y. (2024). An investigation using resampling techniques and explainable machine learning to minimize fire losses in residential buildings. Journal of Building Engineering, 95, 110080.
44. Schmidt, A., Gemmil, E., & Hoskins, R. Machine Learning Based Risk Analysis and Predictive Modeling of Structure Fire Related Casualties. Available at SSRN 4757205.
45. Seo, M. S., Castillo-Osorio, E. E., & Yoo, H. H. (2023). Fire Risk Prediction Analysis Using Machine Learning Techniques. Sensors & Materials, 35.
46. MacQueen, J.: Some methods for classification and analysis of multivariate observations, in: Proceedings of the fifth Berkeley symposium on mathematical statistics and probability, Vol. 1, No. 14, pp. 281-297, 1967.

47. W. Lizhi, R. Aizhu, Urban Fire Risk Clustering Method Based on Fire Statistics, Tsinghua science & Technology **13**, pp. 418 – 422, 2008.
48. A.A. Ishola, D. Valles, Enhancing safety and Efficiency in Firefighting Operations via Deep Learning and Temperature Forecasting Modeling in Autonomous Unit., Sensors **23**, 4628, 2023.
49. Di Martino, T., Le Saux, B., Guinvarc'h, R., Thirion-Lefevre, L., & Colin, E. (2023). Detection of forest fires through deep unsupervised learning modeling of Sentinel-1 time series. ISPRS International Journal of Geo-Information, 12(8), 332.
50. Çiftçioğlu, A. Ö., & Naser, M. Z. (2022, February). Unsupervised Machine Learning for Fire Resistance Analysis. In International Conference on Science, Engineering Management and Information Technology (pp. 211-221). Cham: Springer Nature Switzerland.
51. Rahimi, I., Duarte, L., & Teodoro, A. C. Unsupervised Image Classification Algorithms Applied to Fire-Prone Area Detection.
52. M. O'Neill, C. Ryan, Grammatical evolution, IEEE Trans. Evol. Comput. **5**,pp. 349–358, 2001.
53. C. Bishop, Neural Networks for Pattern Recognition, Oxford University Press, 1995.
54. G. Cybenko, Approximation by superpositions of a sigmoidal function, Mathematics of Control Signals and Systems **2**, pp. 303-314, 1989.
55. Nawi, N. M., Atomi, W. H., & Rehman, M. Z. (2013). The effect of data pre-processing on optimized training of artificial neural networks. Procedia Technology, 11, 32-39.
56. J. W. Backus. The Syntax and Semantics of the Proposed International Algebraic Language of the Zurich ACM-GAMM Conference. Proceedings of the International Conference on Information Processing, UNESCO, 1959, pp.125-132.
57. C. Ryan, J. Collins, M. O'Neill, Grammatical evolution: Evolving programs for an arbitrary language. In: Banzhaf, W., Poli, R., Schoenauer, M., Fogarty, T.C. (eds) Genetic Programming. EuroGP 1998. Lecture Notes in Computer Science, vol 1391. Springer, Berlin, Heidelberg, 1998.
58. M. O'Neill, M., C. Ryan, Evolving Multi-line Compilable C Programs. In: Poli, R., Nordin, P., Langdon, W.B., Fogarty, T.C. (eds) Genetic Programming. EuroGP 1999. Lecture Notes in Computer Science, vol 1598. Springer, Berlin, Heidelberg, 1999.
59. C. Ryan, M. O'Neill, J.J. Collins, Grammatical evolution: Solving trigonometric identities, proceedings of Mendel. Vol. 98. 1998.
60. A.O. Puente, R. S. Alfonso, M. A. Moreno, Automatic composition of music by means of grammatical evolution, In: APL '02: Proceedings of the 2002 conference on APL: array processing languages: lore, problems, and applications July 2002 Pages 148–155.
61. Lídio Mauro Limade Campo, R. Célio Limã Oliveira,Mauro Roisenberg, Optimization of neural networks through grammatical evolution and a genetic algorithm, Expert Systems with Applications **56**, pp. 368-384, 2016.
62. K. Soltanian, A. Ebnenasir, M. Afsharchi, Modular Grammatical Evolution for the Generation of Artificial Neural Networks, Evolutionary Computation **30**, pp 291–327, 2022.
63. I. Dempsey, M.O' Neill, A. Brabazon, Constant creation in grammatical evolution, International Journal of Innovative Computing and Applications **1** , pp 23–38, 2007.
64. E. Galván-López, J.M. Swafford, M. O'Neill, A. Brabazon, Evolving a Ms. PacMan Controller Using Grammatical Evolution. In: , et al. Applications of Evolutionary Computation. EvoApplications 2010. Lecture Notes in Computer Science, vol 6024. Springer, Berlin, Heidelberg, 2010.
65. N. Shaker, M. Nicolau, G. N. Yannakakis, J. Togelius, M. O'Neill, Evolving levels for Super Mario Bros using grammatical evolution, 2012 IEEE Conference on Computational Intelligence and Games (CIG), 2012, pp. 304-31.
66. D. Martínez-Rodríguez, J. M. Colmenar, J. I. Hidalgo, R.J. Villanueva Micó, S. Salcedo-Sanz, Particle swarm grammatical evolution for energy demand estimation, Energy Science and Engineering **8**, pp. 1068-1079, 2020.
67. N. R. Sabar, M. Ayob, G. Kendall, R. Qu, Grammatical Evolution Hyper-Heuristic for Combinatorial Optimization Problems, IEEE Transactions on Evolutionary Computation **17**, pp. 840-861, 2013.
68. C. Ryan, M. Kshirsagar, G. Vaidya, G. et al. Design of a cryptographically secure pseudo random number generator with grammatical evolution. Sci Rep **12**, 8602, 2022.
69. I.G. Tsoulos, D. Gavrilis, E. Glavas, Neural network construction and training using grammatical evolution, Neurocomputing **72**, pp. 269-277, 2008.
70. G.V. Papamokos, I.G. Tsoulos, I.N. Demetropoulos, E. Glavas, Location of amide I mode of vibration in computed data utilizing constructed neural networks, Expert Systems with Applications **36**, pp. 12210-12213, 2009.
71. I.G. Tsoulos, D. Gavrilis, E. Glavas, Solving differential equations with constructed neural networks, Neurocomputing **72**, pp. 2385-2391, 2009.
72. Poli, R., & Langdon, W. B. (1998). Genetic programming with one-point crossover (pp. 180-189). Springer London.
73. Dimitris Gavrilis, Ioannis G. Tsoulos, Evangelos Dermatas, Selecting and constructing features using grammatical evolution, Pattern Recognition Letters **29**,pp. 1358-1365, 2008.
74. J. Park and I. W. Sandberg, Universal Approximation Using Radial-Basis-Function Networks, Neural Computation 3, pp. 246-257, 1991.
75. H. Yu, T. Xie, S. Paszczynski, B. M. Wilamowski, Advantages of Radial Basis Function Networks for Dynamic System Design, in IEEE Transactions on Industrial Electronics **58**, pp. 5438-5450, 2011.

76. Tsoulos, I.G. Creating classification rules using grammatical evolution. Int. J. Comput. Intell. Stud. 2020, 9, 161–171.
77. Anastasopoulos, N.; Tsoulos, I.G.; Tzallas, A. GenClass: A parallel tool for data classification based on Grammatical Evolution. SoftwareX 2021, 16, 100830.
78. 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.
79. Ben-Gal, I. (2008). Bayesian Networks. In Encyclopedia of Statistics in Quality and Reliability (eds F. Ruggeri, R.S. Kenett and F.W. Faltin).
80. Koski, T., & Noble, J. (2011). Bayesian networks: an introduction. John Wiley & Sons.
81. D.E. Rumelhart, G.E. Hinton and R.J. Williams, Learning representations by back-propagating errors, Nature **323**, pp. 533 - 536 , 1986.
82. T. Chen and S. Zhong, Privacy-Preserving Backpropagation Neural Network Learning, IEEE Transactions on Neural Networks **20**, , pp. 1554-1564, 2009.
83. Wilcoxon, F. (1945). Individual Comparisons by Ranking Methods. International Biometric Society: Vol. 1, No:6, pp.:80-83.