AI

MDPI

Article

Detection of Anomalies in Large-Scale Cyberattacks Using Fuzzy Neural Networks

Paulo Vitor de Campos Souza [1,*,†], Augusto Junio Guimarães [1,†], Thiago Silva Rezende [2], Vinicius Jonathan Silva Araujo [2] and Vanessa Souza Araujo [2]

[1] CEFET-MG- Av. Amazonas, 5.253, Nova Suiça, 30421-169 Belo Horizonte, Brazil; augustojunioguimaraes@gmail.com

[2] Faculty Una of Betim- Av. Gov. Valadares, 640-Centro, 32510-010 Betim, Brazil; silvarezendethiago@hotmail.com (T.S.R.); vinicius.j.s.a22@hotmail.com (V.J.S.A.); v.souzaaraujo@yahoo.com.br (V.S.A.)

* Correspondence: goldenpaul@informatica.esp.ufmg.br

† These authors contributed equally to this work.

Received: 27 December 2019; Accepted: 3 February 2020; Published: 7 February 2020

Abstract: The fuzzy neural networks are hybrid structures that can act in several contexts of the pattern classification, including the detection of failures and anomalous behaviors. This paper discusses the use of an artificial intelligence model based on the association between fuzzy logic and training of artificial neural networks to recognize anomalies in transactions involved in the context of computer networks and cyberattacks. In addition to verifying the accuracy of the model, fuzzy rules were obtained through knowledge from the massive datasets to form expert systems. The acquired rules allow the creation of intelligent systems in high-level languages with a robust level of identification of anomalies in Internet transactions, and the accuracy of the results of the test confirms that the fuzzy neural networks can act in anomaly detection in high-security attacks in computer networks.

Keywords: fuzzy neural network; anomaly detection; pattern classification problem

1. Introduction

The sheer volume of data these days has created new services, expectations, laws [1], and also precautions. The amount of essential data circulating in digital solutions enables strategic decision-making to be performed more efficiently, provided that they are treated consistently by its staff and managers [2]. The organic evolution of computing resources allows the change of several scenarios such as online shopping [3], service sales, business deployment, and even on the security and privacy of users' data on the internet [4].

Several companies have already visualized the power of the internet [4] and work with the capture of potential customers who engage in various interactions with online solutions, thus allowing them to create business opportunities or direct more focused marketing actions for this group of people. That happened recently in the US and Brazilian elections [5,6], where digital media was instrumental in mass dissemination of news and information sharing. In this context of a high number of interactions of customers, companies, and people on the Internet, Big Data comes [7], a concept capable of representing a large volume of data connected to several origins, mainly companies, business, and government. As the increasing use of computing resources generates a large amount of data traffic, many precautions and controls need to be taken so that companies and ordinary people do not have their data stolen by malicious people [8].

A fraudulent attack to fetch data from people through computer connections is considered a cyberattack. There are several techniques for stealing user data or information base for other purposes, not legal [9]. These techniques exploit security flaws in electronic devices and use human weaknesses in manipulating resources that use the internet to perform their daily tasks, such as smartphones and tablets [10,11].

Diverse smart systems are being developed to identify a cyberattack in a variety of contexts [9], but getting a template with an expected level of fraud matching is that unbalanced databases become a considerable challenge. In general, databases with more than 100000 records have as frauds less than 2% of these values. Within the computational context, this factor is considered an anomaly because it goes against the normal behavior of the system and does not happen so often [12].

There is a study area that deals with the detection of anomalies in different contexts. We can highlight the models that act in problems of the financial branch [13], of the field of the health [14], in the computation as in the paper of Fugate et al. [15], of the industry like the works of Hadeli et al. [16], Kumarage et al. [17], Dong et al. [18] among others [19]. In general, these studies are focused on specific elements that can generate significant financial losses or the physical integrity of the people. More recent work addresses the identification of anomalous behaviors using artificial intelligence to identify behaviors in vegetation [20], authorization logs [21], computer systems [22] and finally in modern industry [23].

However, this kind of context is not simple to be understood. In many cases, an anomaly specialist is rare and requires years of study and experience. Therefore, several studies are proposing the use of intelligent models based on knowledge to perform the detection of anomalies and at the same time, provide a knowledge base that can serve several purposes [24], such as training of people of a company. These expert systems represent the union of two techniques commonly used in artificial intelligence: fuzzy systems and artificial neural networks [25,26]. The first is capable of bringing interpretability to the results by bringing the responses of the social context through their linguistic and interpretable characteristics. The second is responsible for advances in intelligent models in simulating human reasoning through training [27]. These two concepts united are called fuzzy neural networks, which can act in diverse contexts such as the classification of binary patterns as in the work of Lin et al. [28] and Meesad et al. [29], models that operate in the financial market as proposed by Lin et al. [30] and Kuo et al. [31]. Elements related to the health area such as the models proposed by Wang et al. [32] and Cheng et al. [33], relevant aspects of the industry [34–36], even with the specific focus on cyberattacks such as Batista et al. [37], Gang Wanget al. [38] and Souza et al. [39]. Several models of different characteristics have been developed to work on the detection of cyberattacks such as Demertzis et al. [40–42] and Yusob et al. [43], which in turn may also represent an anomaly in the traditional behavior of the internet.

Therefore, this paper proposes the use of a fuzzy neural network model for the detection of anomalies in cyberattacks. The model is based on the concepts of neural networks with artificial neurons with the leaky ReLU function created by Maas et al. [44], besides fuzzification procedures based on the ANFIS technique [45], capable of generating similarly spaced membership functions to granularize the feature space. In the second layer are used fuzzy logical neurons capable of adding fuzzy inputs with numerical weights. The network seeks its simplicity through the training based on extreme learning machine [46]. A single neuron represents the neural network of aggregation responsible for detecting anomalies in the system [47]. Therefore, the use of this model, in addition to detecting anomalies in Internet transactions, aims to create a system of fuzzy rules capable of serving as a knowledge base on possible forms of attacks and, consequently, in the detection of anomalies. The database used to prove the approach is a set of data commonly used for the detection of cyberattacks, and many results obtained in the literature may corroborate the results obtained by the model.

The main contribution of this paper is to bring a hybrid model with a high degree of assertiveness in the prediction of anomalous behaviors and to transform this behavior into a set of rules that are

interpretable and capable of building an intelligent system. Knowledge gained from fuzzy rules in the training phase allows for the creation of expert systems for an audience that does not work directly with artificial intelligence concepts. Thus, neuro-fuzzy models will disseminate knowledge to a broader audience, especially those who are involved with access control over the internet. As the neural networks used in this paper use the division of the problem space into equally spaced pertinence functions, it is simpler to visualize the fuzzy relations, assigning them values as small, medium, and large, according to the dimension of the problem evaluated.

The paper is organized as follows: Section 2 presents the fundamentals that will support the paper, in Section 3, the main concepts about fuzzy neural networks and related works. In Section 4, we offer the training model that the algorithm will use to detect the web anomaly patterns. Section 5 presents the detection methodology proposed in this article. In Section 6, they show aspects of the database and the configurations of the tests, besides the obtained results. Finally, Section 7 discusses the conclusions and presents a new perspective to future works.

2. Literature Review

2.1. Large-Scale Problems

The evolution of the media and the growing use of computational resources produce large-scale data volumes. This new routine affects the behaviors of software developers, marketing people, managers, and several people involved in maintaining computerized systems and performing specific decisions [7]. Including a high volume of data, decision making becomes more complex and time-consuming, but this goes against the dynamics of today's days where people and companies need to make consistent decisions within a reasonable time so as not to miss out on great opportunities [7]. Some parameters are fundamental for the evaluation of a large-scale data set. They stand out as factors availability, reliability, performance, validation, and system parameterization. When dealing with cyber attacks, the main factors to be discussed are the evaluation, and the validation of the results, mainly if the target system contains information of high relevance.

Based on this massive data volume was created one of the concepts that would be part of the science and routine of developers and researchers data science area: the big data [7]. This concept reveals the existence of an extensive data stream in small time lapses. That happens, for example, when a large number of purchase requisitions are made at the same time by an online shopping site while thousands of other requests about searching for products on the same platform make the site performance slower. This type of experience on the new trends in the information market must ensure that decisions, analyze, and factors are checked promptly for decision-making. When reviewing a search trend for a specific product, the manager can choose to make a lightning-fast promotion, while the information security technicians must be aware of the site's integrity and possible transactions, mainly due to the significant number of malicious requests made by hackers [48].

In the Big Data context, several security factors cannot be overlooked by the technology team. Extensive data in a digital solution can generate an exacerbated profit for solution owners, just as it will also create a range of attacks on sensitive information [49]. When it comes to Big Data, several factors compete with the attention of those involved: processing data and information promptly, protecting the system from malicious attacks, understanding the needs of the target audience, and managing the performance of routine of the computerized resource so that it does not lose fundamental usability requirements for solution users [50].

Large-scale attacks can be viewed as a high number of requests to a server or service, in a short time. They can also result from requisitions with many characteristics to be evaluated by the systems. Cyber attacks, especially when small distortions are inserted in a large group of requests, can cause the

unavailability of systems relevant to society, such as security, online shopping, government services, among others. Figure 1 below presents features and challenges that involve the Big Data concepts for the present day.

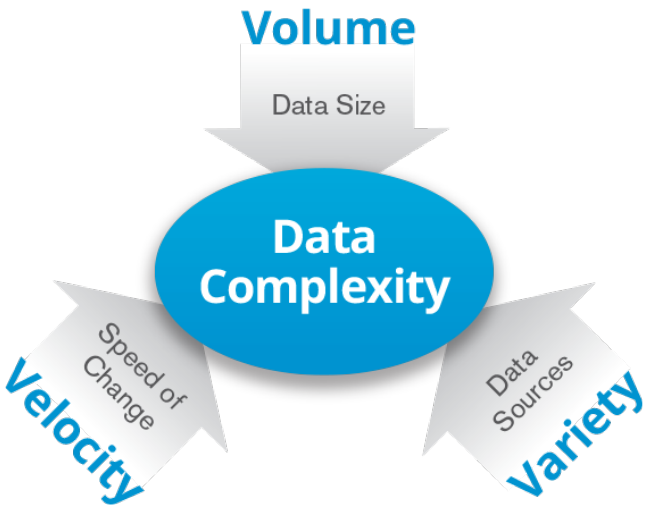

Figure 1. Big Data problems. Available in: https://dwbi.org/analysis/data-mining/177-what-is-big-data.

2.2. Cybernetic Invasions and Intrusion Detectors

Intelligent models can help identify patterns related to different contexts, especially concerning aspects of the internet. These criteria follow characteristics that were reported in simulations conducted by Lincoln Labs through the 1998 DARPA Intrusion Detection Evaluation Program. The task of learning the intrusion detector is to construct a predictive model (i.e., a classifier) that can distinguish between "bad" connections, called intrusions or attacks, and regular "good" connections. These behaviors can be qualified as anomalies because the correct pattern on the internet is a proper connection. The objective of this study was to investigate and evaluate intrusion detection research, in the same way, it can serve as a database to identify anomalies in Internet connections. A typical dataset to be audited, which includes a wide variety of simulated intrusions in a military network environment, was provided [51].

Figure 2 presents a possible step for simulating cyber-attack generating problems for individuals and businesses.

To simulate the database to be provided to the academic community, Lincoln Labs set up an environment in their labs to acquire nine weeks of raw TCP transmission data to a local area network (LAN), thus allowing them to simulate a typical LAN, present in a national security organization. In these tests, they operated the LAN as if it were a real Air Force environment, enabling simulations of multiple network attacks [51].

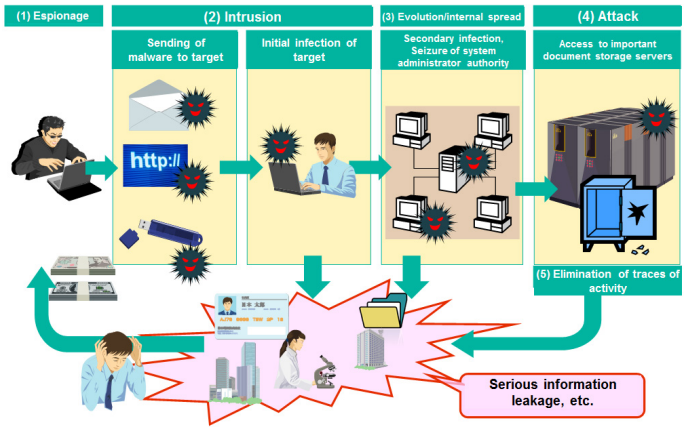

Figure 2. Simulation of a cyberattack. Available in: https://www.nec.com/en/global/solutions/safety/info_management/cyberattack.html.

The training data obtained takes up about four gigabytes of compressed binary TCP data from seven weeks of network traffic. This data was processed in approximately five million connection records. To collect the test data were received in two weeks, the records of connections with the evaluated network [51]. For this type of study, consider [51]:

- A connection is a sequence of TCP packets starting and ending at some well-defined times, between which data flows to and from a source IP address to a destination IP address in some well-defined protocol.
- Each connection is labeled regular or as an attack, with precisely one type of attack specific.
- Each connection record consists of about 100 bytes.

The attacks fall into four main categories [51]:

- DOS: denial of service;
- R2L: unauthorized access of a remote machine;
- U2R: unauthorized access to root superuser privileges;
- Poll: Surveillance and other polls.

As it is a realistic basis, several features in the test data are not present at the same intensity in the training data, helping the models to have characteristics of identifying anomalies according to the nature of the new data. The datasets contain a total of 24 training attack types, with an additional 14 types only in the test data. Stolfo et al. [51] have defined high-level features that help distinguish standard connections from attacks in a statistical way. In this approach, host resources only examine the connections in the last two seconds that have the same destination host as the current connection and calculate the statistics related to the behavior of the protocol, service. Thus it is possible to identify the action of the transaction statistically.

In another approach, similar features of the same service only examine the connections in the last two seconds that have the same function as the current connection. To facilitate the naming of propositions, "Same host" and "same service" are called time-based traffic resources of the connection records [51]. However, other techniques exploit attacks on computer networks using different methodological variations. Some probe attacks examine the hosts (or ports) using a time interval much longer than two seconds, for example, once per minute. Therefore, connection records were also sorted by the destination host, and features were constructed using a window of 100 connections to the same host instead of a time window as in the statistical approach. That produces a set of host-based traffic features [51].

When conducting studies on DOS attacks and probing, there appear to be no sequential patterns that are frequent in R2L and U2R attack records. That is because DOS and probe attacks involve many connections to some hosts in a short time, but R2L and U2R attacks are embedded in the packet data parts and usually require only a single connection [51].

In their study, Stolfo et al. [51] used domain knowledge to add features that look for suspicious behavior in data parts, such as the number of failed login attempts. These resources are called 'content' resources, so there are several techniques for working with anomaly detection according to the characteristics of connections made by people who want to break into a computer network. In this paper, the methods of detecting anomalies made by intelligent models will follow the standards defined by these studies.

2.3. Anomaly Detection

Anomalies detection consists of identifying patterns in dataset data with behavior different than expected. Humans can identify when something is not following a pattern because of their lived experience in the context, such as a security that identifies inappropriate behavior at the time of their night round or even a cook who checks that their cake is not growing as usual. These patterns are often referred to as anomalies, outliers, exceptions, aberrations, discordant observations, among others, varying according to context. The terms anomaly and outlier are the most used in the context of anomaly detection [52].

Anomaly Detection uses many advanced statistical techniques to determine whether an observation should be considered anomalous or not based on pre-established patterns. These techniques seek to evaluate behaviors different from the normality of a context. Consequently, the detection of anomalies is multidisciplinary and can serve different backgrounds [53]. They can be classified as punctual anomalies: when an individual data type can be considered as abnormal relative to the rest of the data, the class is designated as an anomaly punctual. This anomaly classification is the simplest and is the focus of most anomaly detection research. They are usually defined as the points out of normality, that one that differs from the central region where the other data are. As a real example, we have the detection of fraud in the use of a credit card [54]. In Figure 3, we can see different behaviors of the data sets evaluated. These outliers are considered behavioral anomalies.

In contextual anomalies, an instance of data can be considered anomalous in a specific context, but not otherwise in another type of evaluation. A temperature of 40 degrees may be average in tropical countries but would be unusual at the poles of the planet. To determine the context of an anomaly, the conceptual attributes are used. For example, in geographic data sets, the length and latitude of a location are the attributes of the context. In time-series data, time is a contextual attribute that determines the position of an instance throughout the sequence [52]. Another way to represent the attributes is to evaluate the behavioral form. In a set of geographic data describing the average rainfall across the world, the amount of precipitation at any location is a behavioral attribute.

Finally, the collective anomaly of data is considered when referring to the whole set of data. Specific data instances in a cumulative anomaly may not be anomalies by themselves, but their example in a collection, such as a collection, is anomalous [52].

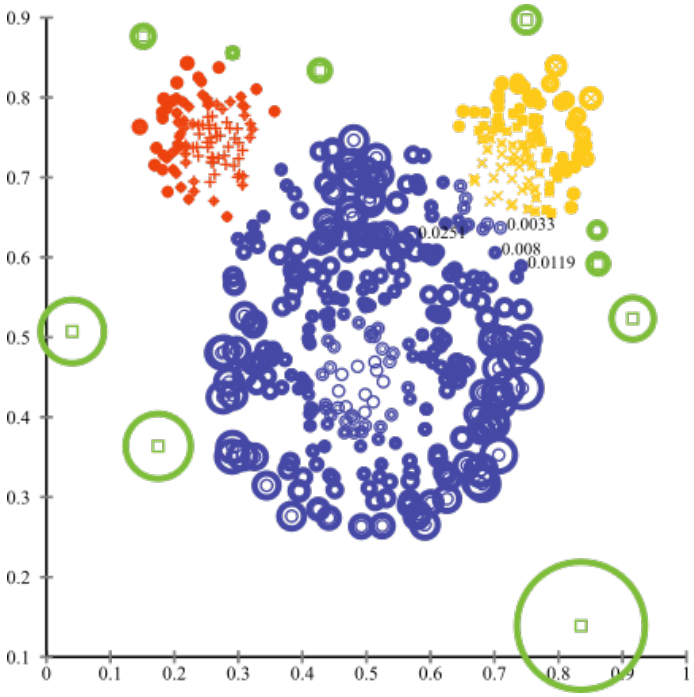

Figure 3. Anomaly detection performed by the k-means algorithm [55]. Available in: https://stats.stackexchange.com/questions/160260/anomaly-detection-based-on-clustering.

2.4. Related Work

As it is an area of interest in different contexts of science, the detection of anomalies has been the target of several researchers and fields of science. The work of [52] brings you a survey with the central regions of the detection of anomalies and the right jobs. Other papers also use reviews to collect the principal works of the literature addressing diverse topics on the discovery of anomalies like Chandola et al. [56], Ahmed et al. [57], Estevez-Tapiador et al. [58], and Patcha and Park [53].

Other works, such as Sabahi and Movaghar [59], address the detection of anomalies as intruders. The paper of Xie et al. [60] has dealt with anomalies in wireless sensors. The hyperspectral imagery animals were treated with algorithms in the work of Stein et al. [61]. A work, well known in the literature (Garcia et al. [62]), deals with anomalies in network invasion, another one uses the concepts of anomalies in web-based and networks attacks [63–67], IP address [68,69] and using call stack information [70]. Paper proposed by Lee and Xiang [71] has developed techniques to form a set of information to prevent more and more types of anomalies. Works were produced using the Markov chain to identify the anomalies [72], semi Markov chain [73], use immune-inspired algorithm [74], fuzzy judgment [75], and support vector machine's concepts [76].

Another branch of research where anomaly detection is well used is in the detection of credit card fraud such as Aleskerov et al. [77], monitoring smartphones [78], video anomalies [79], sensor of hardness recognition based on magnetic anomalies [80], quantum anomaly detection [81], energy anomalies [82], sonar imagery[83], wide area network [84] and fast anomaly detection in crowded scenes [85]. In recent paper the use of deep learning and large margin to video-based anomaly detection [86], time series anomalies [87], use noise binary search [88], classification and anomaly detection of side-channel signals uses deep-learning [89], authorization log [21] were used in the context of anomaly identification.

3. Fuzzy Neural Networks

Fuzzy neural networks (FNN) are intelligent hybrid models composed of artificial neural networks, their techniques of training and updates of parameters and fuzzy systems that have the characteristics of transforming the data of a problem into representations in the space of features, assigning them specific characteristics that can be interpreted as linguistic and interpretive terms. These models allow the union of the main benefits of neural network training techniques and the interpretability of fuzzy systems, allowing intelligent systems to update parameters and perform training with more interpretable results [27].

A wide range of applications can be attributed to these types of problems. Since the 1970s, they have been working synergistically in solving complex problems in various social contexts, industry, and science [90].

This intelligent model has as main characteristic the replacement of the artificial neurons commonly employed in artificial neural networks by fuzzy neurons, which perform the fuzzy input weighting with synaptic weights, which may or may not be fuzzy numbers. This approach uses elements ranging from the transformation of the input space into fuzzy characteristics (also called the fuzzification process) until the answers are obtained in the context to which they belong (defuzzification process). In the middle of these two processes, there is an extraction of knowledge from networks, where they can interpret the characteristics of the database and transform them into a representation through membership functions. These membership functions are constructed using techniques that measure the behavior of data in space. They may have methods that perform the comparison by grouping, similarity, density, or distance [27].

The method with which the inputs are manipulated can define parameters such as activation functions, fuzzy set membership functions, network topology, among others [91]. Fuzzy neural networks can use clustering methods, such as c-means and its fuzzy version, fuzzy c-means ([92,93]), methods based on data density called clouds [94], eClustering+ [95] and ePL [96], among others. These methods help define fundamental network structures and somehow allow the construction of fuzzy rules based on data contexts in the feature space [97].

In this context, the group of characteristics becomes represented by fuzzy elements, where these characteristics can receive qualitative labels (as definitions of the small, medium, large, or warm and warm cold). Thus, the fuzzy system is considered as an alternative to treating situations where the binary results are insufficient to represent the problems and that they require more factors to evaluate the problem. Fuzzy logic allows us to express data with linguistic labels, enabling a set of ages collected from patients in a clinic to qualify as new age, middle age, and older age. The amount of features depends on the nature of the problem being evaluated and the people who understand the target context [27].

The fuzzy neural networks present applicability in the obtaining of expert systems, classification of standards, linear regression, prediction of time series, besides problems applied in robotics, aspects of nature, health, education, and in the industry as can be visualized below.

Fuzzy Neural Networks and Their Practical Applications

The fuzzy neural network is used for anomalies detection in Han and Cho [98], Meneganti et al. [99], and learning rules can be visualized in Mahoney and Chan [100]. Expert systems have been approached as a form of knowledge transfer before the 1990s, as in Wiig [101] and in Gang Wang et al. [38] in 2010.

FNNs act on problem-solving with various types of complexity. They can work on simple pattern classification problems by using pruning approaches of their architecture [102] and at the same time, can generate patterns for the evaluation of nonlinear systems [103], time series forecasting [104] and linear and non-linear regression problems [102,105].

Hybrid models also stand out in finding women with breast cancer [106,107] according to characteristics informed by patients by clinical trials. In the same way, it assists in the detection of an autistic trait in children [108] and immunotherapy treatment [109].

More recent work involves the concepts of robot manipulation and control [110], prediction of chaotic series [103], effort forecast in software building [111], anomalies identification's in children and adolescents locomotion [112], absenteeism at work [113].

4. Fuzzy Neural Networks Model

4.1. Network Architecture

The fuzzy neural network described in this section is composed of three layers. Besides, it was derived from the work of Souza [47]. In the first layer, the fuzzification is realized through the concept of the ANFIS model [45] in its version that can generate membership functions equally spaced in the sample space. The association functions adopted in the first layer are of the Gaussian type created with the centers and the sigma values obtained by genfis1 in the generation of the input space granularization. Already in the second layer, the logical neurons of type andneuron [114], and the type proposed by [91] were used. These neurons have weights and activation functions determined randomly and through t-norms (calculated through the product) and s-norms (using the probabilistic sum concepts) for all neurons in the first layer. To define the weights that connect the second layer to the output layer, the idea of an extreme learning machine [46] is used to act on the neuron with a leaky ReLU type activation function. The fuzzy logical neurons are used to construct fuzzy neural networks in the second layer to solve problems of pattern recognition and to bring interpretability to the model through the extraction of fuzzy rules of type IF/THEN. Figure 5 illustrates the feedforward topology of the fuzzy neural networks considered in this article.

The first layer is composed of neurons whose activation functions are functions of association of fuzzy sets defined for the input variables using ANFIS in its approach called genfis1. For each input variable x_{ij}, L membership functions are defined A_{lj}, $l = 1 \ldots L$ whose association functions are the activation functions of the corresponding neurons. Thus, the outputs of the first layer are the degrees of association associated with the input values, i.e., $a_{jm} = \mu_m^A$ for $j = 1 \ldots N$, where N is the number of inputs and M is the number of fuzzy sets for each input result defined by ANFIS [115]. Therefore, this problem of the fuzzy neurons creation in the first layer. It is an exponential problem because the number of neurons is closely related to the relation of the number of functions of pertinence for each evaluated characteristic.

The second layer is composed by L fuzzy logic neurons. Each neuron performs a weighted aggregation of some of the first layer outputs. This aggregation is performed using the weights w_{il} (for $i = 1 \ldots N$ and $l = 1 \ldots L$). For each input variable j, only one first layer output a_{jl} is defined as input of the l-th neuron. So that $\mathbf{w}$ is sparse, each neuron of the second layer is associated with an input variable [47]. Finally, the output layer is composed of one neuron whose activation functions are leaky ReLu [44]. The output of the model is:

$$y = sign \sum_{j=0}^{l} f_{LeakyReLU}(z_l v_l) \tag{1}$$

where $z_0 = 1$, v_0 is the bias, and z_j and v_j, $j = 1, \ldots, l$ are the output of each fuzzy neuron of the second layer and their corresponding weight, respectively.

Leaky ReLu is an improved function of the ReLU function [116] because a small linear component is inserted at the input of the neuron. This type of change allows small changes to be noticed, and neurons that would be relevant to the model are not discarded. Its function is expressed by [44]:

$$f_{LeakyReLU}(x, \alpha) = max(\alpha x, x) \tag{2}$$

The logical neurons used in the second layer of the model are of the andneuron or unineuron type, where the input signals are individually combined with the weights and performed the subsequent global aggregation. The andneuron used in this work can be expressed as [117]:

$$z = AND(w; a) = T_{i=1}^{n}(w_i s a_i) \tag{3}$$

where T are *t-norms* (product), s is a *s-norms* (probabilistic sum).

The unineuron uses the concepts of uninorm [118] to perform more simplified operations according to the functions of activation of the fuzzy neurons.

Instead of values 0 and 1 for t-norm and s-norm respectively, the neutral element is allowed to assume values in the unit interval. One of the main characteristics of the uninorm is that it no longer has the so-called neutral element, now being called the entity element [119]. Through this identity element (e), the uninorms extend t-norms and s-norms by varying the value e in the interval between 0 and 1 allowing the alternation between an s-norm (e = 0) and t-norm (e = 1). The uninorm used in this work is expressed as follows [119]:

$$U(x, y) = \begin{cases} e\ T(\frac{x}{e}, \frac{y}{e}), if\ y\ \in [0, e] \\ e + (1 - e)\ S\ (\frac{x-e}{1-e}, \frac{y-e}{1-e}),\ if\ y\ \in\ (e, 1] \\ \varphi(x, y),\ otherwise \end{cases} \tag{4}$$

and

$$\varphi(x, y) = \begin{cases} max(x, y), & \text{if } e \in [0, 0.5] \\ \min(x, y), & \text{if } e \in (0.5, 1] \end{cases} \tag{5}$$

Its formatting allows the unineuron to use either concept of a neuron and, or a neuron or. [119] explain important concepts about a unineuron. The processing of neurons occurs at two levels. At the first level of L_1 locations, the input signals are combined individually with the weights. In the second, at a global level of L_2, a global aggregation operation is performed on the results of all first-level combinations.

Traditional logical neurons use t-norms and s-norms to perform the described operations.

1. ach pair *(a_i, w_i)* is transformed into a single value b_i = **h** *(a_i, w_i)*;
2. calculate the unified aggregation of the transformed values **U** *($b_1, b_2 \ldots b_n$)*, where n is the number of inputs.

The function p is responsible for transforming the inputs and corresponding weights into individual transformed values. A formulation for the p function can be described as [91]:

$$p(w, a, e) = wa + \bar{w}e \tag{6}$$

where $\bar{w}$ represents the complement of w. Using the weighted aggregation reported above the unineuron can be written as:

$$z = UNI(w; a; e) = U_{i=1}^{n}\ p\ (w_i,\ a_i, e) \tag{7}$$

Figure 4 shows how the formation of the logical neuron can vary according to the value of e.

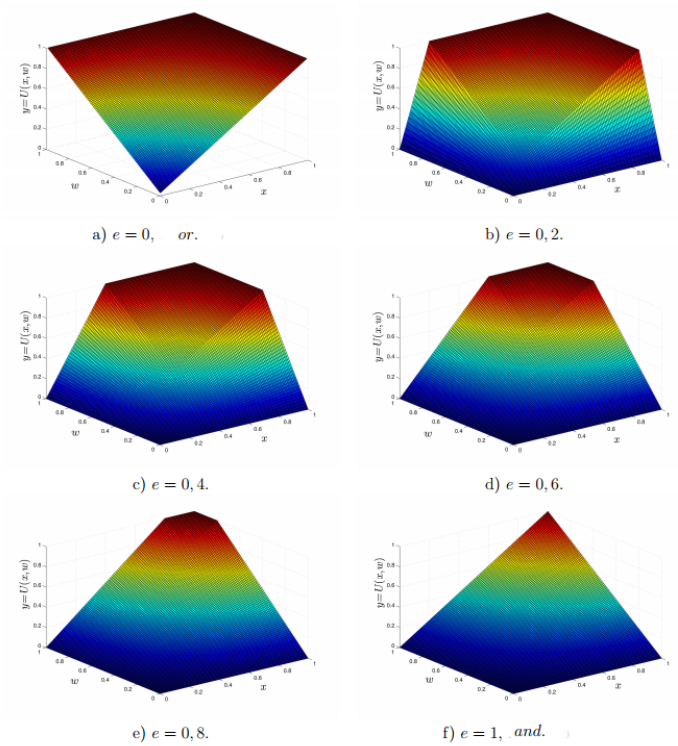

Figure 4. Example of Andneuron, OrNeuron and Unineuron [120].

Fuzzy rules can be extracted from andneurons according to the following example:

$$\begin{aligned} Rule_1: \ & If x_{i1} \text{ is } A_1^1 \text{ with certainty } w_{11} \ldots \\ & \text{and } x_{i2} \text{ is } A_1^2 \text{ with certainty } w_{21} \ldots \\ & \text{Then } y_1 \text{ is } v_1 \\ Rule_2: \ & If x_{i1} \text{ is } A_2^1 \text{ with certainty } w_{12} \ldots \\ & \text{and } x_{i2} \text{ is } A_2^2 \text{ with certainty } w_{22} \ldots \\ & \text{Then } y_2 \text{ is } v_2 \\ Rule_3: \ & If \ x_{il} \text{ is } A_3^1 \text{ with certainty } w_{13} \ldots \\ & \text{Then } y_3 \text{ is } v_3 \\ Rule_4: \ & If \ x_{i2} \text{ is } A_3^2 \text{ with certainty } w_{23} \ldots \\ & \text{Then } y_4 \text{ is } v_4 \end{aligned} \tag{8}$$

These rules allow the creation of a building base for expert systems [121]. Figure 5 presents an example of fuzzy neural network architecture.

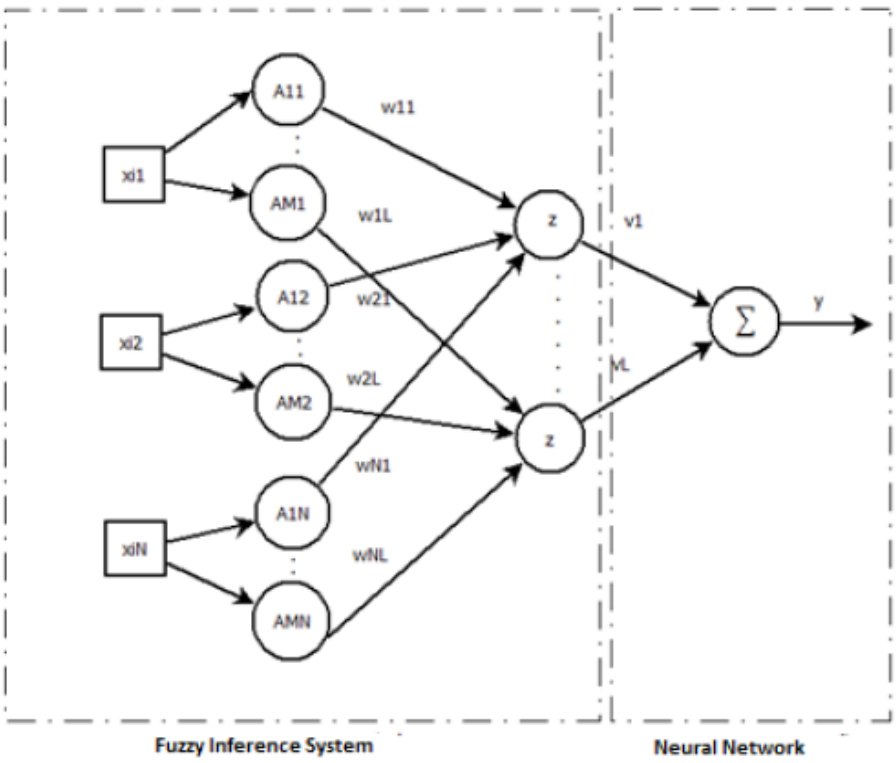

Figure 5. FNN architecture.

4.2. Training Fuzzy Neural Network

The membership functions in the first layer of the FNN are adopted as Gaussian. The number of neurons created with the input data partition is exponential between the number of membership functions and the number of features present in the problem database. The number of neurons L in the first layer is defined according to the input data and by the number of membership functions M), defined parametrically. The second layer performs the aggregation of the L neurons from the first layer through the andneurons.

After the construction of the L fuzzy logical neurons a filter select the 200 most significant neurons (called L_s) like in [47]. The final network architecture is defined through a feature extraction technique based on l_1 regularization and resampling. The learning algorithm assumes that the output hidden layer composed of the candidate neurons can be written as [115]:

$$f(x_i) = \sum_{i=0}^{L_\rho} v_i z_i(x_i) = sign(z(x_i)v) \tag{9}$$

where $\mathbf{v} = [v_0, v_1, v_2, \ldots, v_{L_\rho}]$ is the weight vector of the output layer and $\mathbf{z}\,(x_i) = [z_0, z_1(x_i), z_2(x_i) \ldots z_{L_\rho}(x_i)]$ the output vector of the second layer, for $z_0 = 1$ and sign is a step function that transforms values greater than zero into 1 and values smaller than zero into -1. In this context, $\mathbf{z}\,(x_i)$ is considered as the non-linear mapping of the input space for a space of fuzzy characteristics of dimension L_ρ [115]. The sign function is defined by:

$$sign = \begin{cases} -1, & z(x_i)v < 0 \\ 1, & z(x_i)v) \geq 0 \end{cases} \tag{10}$$

Subsequently, following the determination of the network topology, the predictions of the evaluation of the vector of weights' output layer are performed. In this paper, this vector is considered by the Moore-Penrose pseudo Inverse [115]:

$$v = Z^{+}y \tag{11}$$

Z^{+} is the Moore-Penrose pseudo-inverse [122] of z, which is the minimum norm of the least-squares solution for the output weights.

5. Proposed Detection of Cyber Invasions Through Detection of Anomalies Through Hybrid Models and the Creation of Expert Systems

The hybrid system proposes to use fuzzy neural networks and train them with the database that determines patterns of anomalies. Through these standards, the model learns the trends and characteristics of the database, allowing in addition to pattern classification, create an expert system based on fuzzy rules.

The model will have four dimensions (service, duration, bytes received, bytes sent) according to the formatting of bases for the detections of anomalies. These four features will be combined according to equally spaced membership functions. In an example with two pertinence functions for each input of the model, eight Gaussian neurons are generated in the first layer and consequently 16 fuzzy logical neurons in the second layer. Therefore $N = 4$, $M = 2$, $L = 16$. These 16 fuzzy neurons represent the union of fuzzy rules of the first layer.It follows then that $L = N^{M}$.

The methodology is synthesized as demonstrated in Algorithm 1. It has two parameter:

1. the number of membership functions, M;
2. the type of fuzzy logic neuron, *andneuron or unineuron*;

Algorithm 1: Fuzzy Neural Network for anomaly detection -FNN training

(1) Define the number of membership functions, M.
(2) Calculate M neurons for each characteristic in the first layer using ANFIS.
(3) Construct L fuzzy neurons with Gaussian membership functions constructed with center and σ values derived from ANFIS.
(4) Define the weights and bias of the fuzzy neurons randomly.
(5) Construct L fuzzy logical neurons with random weights and bias on the second layer of the network by welding the L fuzzy neurons of the first layer.
(6) **For all** K input do
(6.1) Calculate the mapping $z_k(x_k)$ using andneurons
end for
(7) Estimate the weights of the output layer using Equation (11).
(8) Calculate output **y** using leaky ReLU using Equatiojn (1).

6. Anomaly Detection Test

6.1. Dataset Uses

The dataset used for the experiments in this paper was originally provided in the KDD Cup 1999 and is currently available in the main data repository for machine learning. It contains 41 attributes (34 continuous and seven categorical). However, they are reduced to 4 attributes (service, duration, bytes received, and bytes sent) because these attributes are considered the most basic attributes where only the service is categorical. Using the service attribute, the data is divided into http, SMTP, FTP, FTP data, other

subsets. That allows distinct types of attacks to be verified by intelligent algorithms. Here, only HTTP service data is used. Since the values of the continuous attributes are concentrated around 0, we transform each value into a value far from 0, by $y = \log(x + 0.1)$. The original dataset has 3.925.651 attacks (80.1%) of 4.898.431 records. A smaller set is forged by having only 3.377 attacks (0.35%) of 976.157 records, where the logged-in an attribute is positive. From this forged dataset, 567.497 HTTP service data is used to construct the HTTP dataset [123–128].

The database was selected precisely with the main feature of a cyber attack: a large volume of requests with attacks entered together with them. Thus, protection systems are overloaded and often miss attacks that can compromise system integrity. Therefore, a system that acts dynamically in identifying these patterns, especially as assertively as possible, is necessary for maintaining system integrity. The database provided by the KDD Cup has the characteristics of large-scale attacks as the number of requests is exceptionally high. Moreover, in these requests, there are less than 2% of malicious attacks. Therefore, the database meets the anomaly detection criteria (when the database is hugely unbalanced about its labels) and the large scale criteria for having more than millions of requests.

6.2. Definitions and Models Used in the Tests

In preliminary tests were run using 10-k-fold and cross-validation (70 % for training and 30 % for the test) to find the best value of M between the interval [2, 5] (values defined by a specialist in problems). Another factor evaluated in the preliminary analysis was the logical neuron used. The test used logical neurons of the andneuron or unineuron type. After performing initial tests, the values of M and logical neuron type that maximize training accuracy and maintain the shortest execution time is $M = 3$ and the andneuron. Therefore they will be used for the final experiments of this paper. Simulations were performed on a Core (TM) 2 Duo CPU, 2.27 GHz with 3-GB RAM, and the model are implemented and executed in Matlab.

To perform the comparison with the results obtained by the fuzzy neural network, we used artificial neural network models that are provided by the tool in Java called WEKA [129]. For this experiment, the models commonly used in the detection of anomalies such as Multilayer Perceptron (MLP) [130] (batch size = 100, hidden layers = 1, learning rate = 0.3, momentum = 0.2, validation Threshold = 20.), Naive Bayes (NV) [131] (useKernelEstimator = false, debug = false, displayModelInOldFormat = false, doNotCheckCapabilities = false, useSupervisedDiscretization = false), Random Tree (RT) [132] (seed = 1, allowUnclassifiedInstances = false, debug = false, minNum = 1.0, numFolds = 0, doNotCheckCapabilities = false, maxDepth=0, minVarianceProp = 0.001, KValue = 0) and Support Vector Machine (SVM) [133] (c = 1.0, kernel = radialbasis function kernel coefficient = 1, features shrinking = true, tolerance = 0.001) were chosen.

In addition to tests with traditional approaches, the results will also be compared with other hybrid models of neural networks and fuzzy systems, where we can highlight an evolving fuzzy neural network model (EFNN) [39], one that works with incremental fuzzification (IFNN) [134] and one that works with self-organizing fuzzification (SFNN) [135]. All models use the extreme learning machine to define the weights of the output layer, have three layers, and are composed of unineurons in the second layer and Gaussian neurons in the first layer. All hyperparameters were defined using cross-validation in the interval between [3 and 6], mainly in the fuzzification stage.

In the final test of the model were evaluated the accuracy of training and test to verify if the model does not suffer from overfitting, time of execution (in seconds), sensitivity, specificity, and AUC (Area under Curve- Calculated by Matlab. These other factors are relevant to confirm if there are no false positives or false negatives present in the model results. The following equations were used to determine the test parameters.

$$accuracy = \frac{TP + TN}{TP + FN + TN + FP} \tag{12}$$

$$sensitivity = \frac{TP}{TP + FN} \tag{13}$$

$$specificity = \frac{TN}{TN + TP} \tag{14}$$

$$AUC = \frac{1}{2}\left(sensitivity + specificity\right) \tag{15}$$

where, TP = true positive, TN = true negative, FN = false negative and FP = false positive.

6.3. Results of anomaly detection

The Table 1 presents the results of detecting anomalies. It shows the training percentage, the number of neurons used in the test, the expected results, the time (in seconds). The values in parentheses represent the standard deviation of the 30 random repetitions performed.

Table 1. Data from the anomaly identification tests.

Model	Training	L_p	Sensitivity	Specificity	AUC	Test	Training and Test Time
FNN	**99.99 (0.01)**	27 (0.00)	**100 (0.00)**	**98.81 (0.35)**	**99.90 (0.01)**	**98.76 (1.67)**	2140.35 (59.02)
SFNN	98.18 (0.01)	27 (0.00)	99.92 (0.00)	98.14 (0.16)	99.03 (0.01)	97.25 (0.42)	162.31 (14.27)
EFNN	96.18 (0.16)	12.25 (4.07)	94.65 (0.29)	94.33 (0.19)	94.49 (0.20)	95.99 (0.84)	512.25 (21.17)
IFNN	97.99 (0.17)	28.68 (6.67)	**100 (0.00)**	97.89 (0.15)	98.94 (0.02)	96.76 (2.47)	1587.78 (28.02)
MLP	98.78 (0.00)	28.14 (0.78)	**100 (0.00)**	94.81 (1.35)	97.41 (0.46)	98.16 (1.67)	2090.13 (71.02)
NB	97.45 (0.35)	**8.64 (0.14)**	98.87 (0.05)	97.15 (0.75)	98.01 (0.12)	96.54 (2.05)	**67.45 (15.53)**
RT	98.02 (0.14)	32.00 (2.41)	97.87 (0.54)	96.81 (0.35)	97.34 (0.87)	97.34 (1.54)	134.36 (12.52)
SVM	99.15 (0.01)	56.98 (6.87)	98.14 (0.25)	92.17 (0.76)	95.15 (0.53)	98.29 (1.54)	434.36 (65.67)

The results obtained in the literature, highlighting the work of Tan et al. [128] present the results equivalent to the results obtained by the model, placing as the only unattended situation the execution time that was already expected to be high due to great sensitivity to solve problems with a large number of samples that ANFIS presents.

Despite being a disproportionate dataset, the model was very efficient in detecting the anomalies in the 30 tests performed. The standard deviation was very low in all parameters (except time), and the accuracy, sensitivity, and specificity corroborate that the model is a unique identifier of anomalies. What the proposal of this article has of differential are the fuzzy rules generated, and that can serve as the knowledge base for training and another type of dissemination of knowledge. The best results are presented in Table 1 of the accuracy for FNN, MLP, and SVM. However, it is noteworthy that the best results for an anomaly base were the model described in this paper. FNN specificity had better results (measures the proportion of actual anomalies that are correctly identified as such). The execution time of the algorithm has values close to MLP, but its results stand out as the best assertiveness. Regarding SVM, the results are also statistically similar, but their results are not interpretable.

Among the fuzzy neural network models, the model presented in the paper obtained the best accuracy results, despite spending more time in the training phase. That identifies that the FNN has some degree of accuracy, but it still needs adjustments to adapt to the time of obtaining ideal responses. Other models, such as the evolving fuzzy neural network, showed shorter training times, although the results of success were considerably below those of the other models.

The method proposed in this paper works efficiently to determine anomalies in an extensive data set as it has satisfactory results concerning its sensitivity, specificity, and AUC. When a model has high numbers in these three criteria, it means that it is an excellent model in determining class labels, especially when the imbalance is high. When a model is highly accurate but has low AUC numbers, it means that it has not been able to identify the smallest number of labels correctly. As can be seen in Table 1, the fuzzy neural network model proposed in this paper had the best AUC index, which, consequently, can be said to be the model that most identified anomalies in this large volume of requests.

The next topic will present the characteristics of the rules obtained.

6.4. Expert Systems in Detecting Anomalies in Cyberattacks Through Fuzzy Rules

The linguistic characteristics adopted for the formation of the rules were defined in consultation with experts in the field.

The Figure 6 shows the ANFIS structure formed with one of the results of the 30 replicates performed with the model. The three dimensions were shaped using equally spaced Gaussian membership functions. Here it can see the influence of fuzzy inputs for the fuzzy inference system that will generate rules based on dataset knowledge.

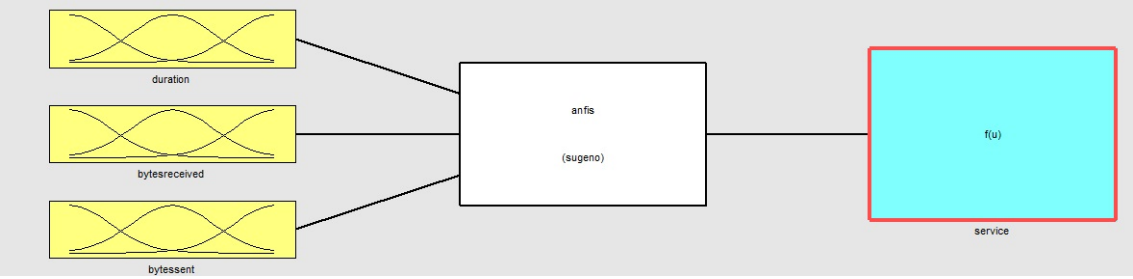

Figure 6. ANFIS Model.

The genfis1 model, which is the Matlab technique responsible for identifying the membership functions, is presented in Figure 7.

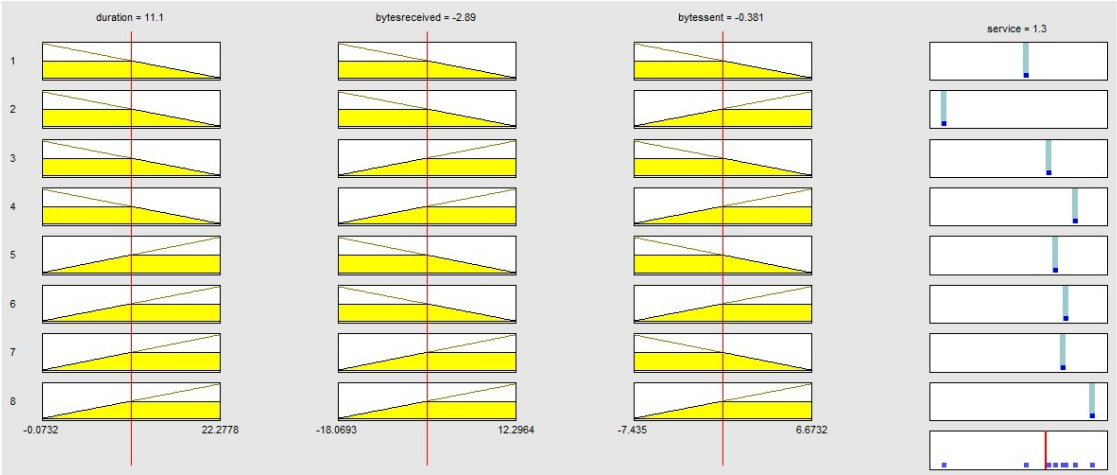

Figure 7. Graphical mode of membership functions.

The fuzzy rules generated were in 8 in total and are presented linguistically as a knowledge base for the formation of expert systems in the Figure 8.

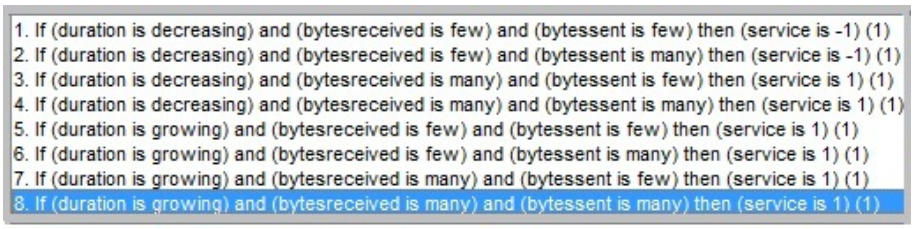

Figure 8. Fuzzy rules generated.

The model of neurons that represent the first layer and fuzzify the input space is described in Figure 9.

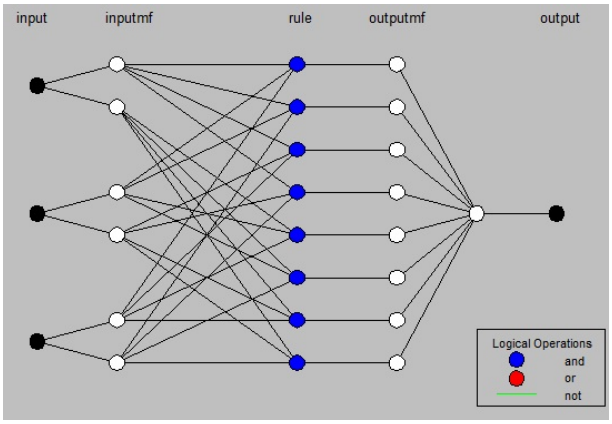

Figure 9. Fuzzy Model- ANFIS.

The decision space that assists in the identification of anomalies is presented in Figure 10.

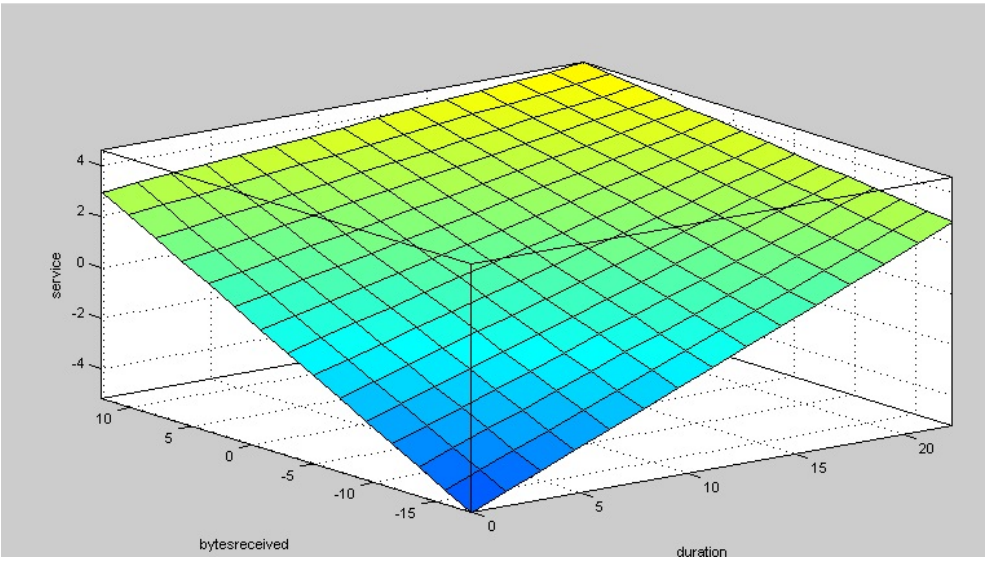

Figure 10. Decision space- ANFIS.

In fuzzy rules, it is possible to identify the decision space of the model by the duration of a request, the number of bytes received and sent. Large-scale cyber attacks work with the methodology of overloading data servers to make them more susceptible to attack. Thus, in Figure 7, it is possible to define the number of elements evaluated in each dimension of the problem for FNN decision making. Likewise, the graphic knowledge of a fuzzy neural network can be presented linguistically and relationally (Figure 8), allowing anyone interested in protecting computer systems to understand when a cyber-attack can occur, even those who are not profoundly knowledgeable. of artificial intelligence. So this is the most significant advantage of the model because it allows the clear and straightforward dissemination of implicit knowledge in a database. In these relationships obtained, it can be seen that the highest correlation between the identification of cyber-attacks is linked to the reduction of the duration of requests tied mostly to a low amount of bytes received.

7. Conclusions

After the presented results, we can conclude that the fuzzy neural networks used in this paper can act as unique identifiers of anomalies. Because it is an unbalanced problem or more than 99% of the samples are of one category, the model behaved efficiently to identify the anomalies, and in most of the trials, it has found all of them.

About Table 1, it is possible to analyze some aspects regarding the results obtained. All models submitted to the cyber attack classification test obtained excellent results with a balance between the training and test percentage, identifying that none of the models chosen in the test suffered from overfitting. The fuzzy neural network obtained the best training and test percentage, allowing the conclusion that the model has the best ability to identify cyber-attacks in this evaluation. However, it is noted that their training and testing time was longer compared to other models of the cyber invasion test. Because it is a high data volume, multi-layered networks take more time to solve problems (as can also be seen at MLP runtime). Models with the shortest runtimes were not as effective at detecting cyber attacks. This is because when a problem is so unbalanced (only 0.35% of the base contains attacks), the value of specificity (correct prediction of attacks) is one of the most important indices for defining model performance. Therefore, what obtained the best results was the FNN with results close to 99%.

However, it should be noted that even with the high execution time of the algorithm, the results of the model proposed in this paper were the best in the evaluation indexes, adding to the results of the possibility of obtaining knowledge about the attacks and improving the results. Protection devices that operate on cyber threat systems with knowledge extracted from the dataset. Fuzzy rules can be easily implemented in information systems that have logical programming, as can electronic devices that can also be programmable.

The model can be seen as an approach to knowledge management in Big Data since it can extract knowledge from a database and turn it into a set of linguistic rules, more accessible to interpret by people who do not are directly linked to the computer science area. This type of approach assists in the dissemination of intelligent techniques and can contribute to advances in science and the prevention of anomalies.

For future work, the challenge is to decrease the execution time of the algorithm while maintaining its ability to find the anomalies. Other techniques of fuzzification and training can be tested, as well as the comparison of other intelligent models so that comparisons can be made in different contexts of artificial intelligence. Another factor that can be taken into account for future extensions of this study is linked to the identification of anomalies in several types of cyberattacks, with more current databases, as reported in the work of Rupa Devi and Badugu [136].

Author Contributions: Conceptualization, P.V.d.C.S. and T.S.R.; methodology, A.J.G.; software, P.V.d.C.S. and A.J.G.; validation, V.S.A., T.S.R. and V.J.S.A.; formal analysis, P.V.d.C.S.; investigation, T.S.R.; resources, P.V.d.C.S.; data curation, P.V.d.C.S. and V.S.A.; writing–original draft preparation, P.V.C.S. and T.S.R.; writing–review and editing, P.V.C.S. and T.S.R.; visualization, A.J.G.; supervision, P.V.d.C.S.; project administration, P.V.d.C.S.; funding acquisition, P.V.d.C.S. and V.J.S.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: The thanks of this work are destined to CEFET-MG and UNA.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

MDPI	Multidisciplinary Digital Publishing Institute
DOAJ	Directory of open access journals
FNN	Fuzzy Neural Network
ANFIS	Adaptive neuro-fuzzy inference system
DOS	Disk Operating System
ReLU	Rectified Linear Activation

References

1. Duggal, P. # Cyberlaw: Global trends in 2014. *AI Soc.* **2014**. [CrossRef]
2. Jansen, C.; Jeschke, S. Mitigating risks of digitalization through managed industrial security services. *AI Soc.* **2018**, *33*, 163–173. [CrossRef]
3. Shoji, H.; Hori, K. S-Conart: An interaction method that facilitates concept articulation in shopping online. *AI Soc.* **2005**, *19*, 65–83. [CrossRef]
4. Gill, K.S. The transformation of the human dimension in the cyberspace. *AI Soc.* **2012**, *27*, 429–430. [CrossRef]
5. Enli, G. Twitter as arena for the authentic outsider: Exploring the social media campaigns of Trump and Clinton in the 2016 US presidential election. *Eur. J. Commun.* **2017**, *32*, 50–61. [CrossRef]
6. Green, J.N. Brazil faces its most important election. *Green Left Wkly.* **2018**, *1200*, 15.
7. Gandomi, A.; Haider, M. Beyond the hype: Big data concepts, methods, and analytics. *Int. J. Inf. Manag.* **2015**, *35*, 137–144. [CrossRef]
8. Wang, C.; Wang, Q.; Ren, K.; Lou, W. Privacy-preserving public auditing for data storage security in cloud computing. In Proceedings of the 2010 Proceedings IEEE INFOCOM, San Diego, CA, USA, 14–19 March 2010; pp. 1–9.
9. Pasqualetti, F.; Dörfler, F.; Bullo, F. Attack detection and identification in cyber-physical systems. *IEEE Trans. Autom. Control* **2013**, *58*, 2715–2729. [CrossRef]
10. Ball, J.; Borger, J.; Greenwald, G. Revealed: How US and UK spy agencies defeat internet privacy and security. *Guardian* **2013**, *6*, 2–8.
11. Greenwald, G. NSA collecting phone records of millions of Verizon customers daily. *Guardian* **2013**, *6*, 2013.
12. Yueai, Z.; Junjie, C. Application of unbalanced data approach to network intrusion detection. In Proceedings of the 2009 First International Workshop on Database Technology and Applications, Wuhan, China, 25–26 April 2009; pp. 140–143.
13. Ngai, E.W.; Hu, Y.; Wong, Y.; Chen, Y.; Sun, X. The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature. *Decis. Support Syst.* **2011**, *50*, 559–569. [CrossRef]
14. Friedberg, M.K.; Silverman, N.H.; Moon-Grady, A.J.; Tong, E.; Nourse, J.; Sorenson, B.; Lee, J.; Hornberger, L.K. Prenatal detection of congenital heart disease. *J. Pediatr.* **2009**, *155*, 26–31. [CrossRef]
15. Fugate, M.; Gattiker, J.R. Anomaly detection enhanced classification in computer intrusion detection. In *Pattern Recognition with Support Vector Machines*; Springer: New York, NY, USA, 2002; pp. 186–197.

16. Hadeli, H.; Schierholz, R.; Braendle, M.; Tuduce, C. Leveraging determinism in industrial control systems for advanced anomaly detection and reliable security configuration. In Proceedings of the IEEE Conference on Emerging Technologies & Factory Automation (ETFA 2009), Mallorca, Spain, 22–25 September 2009; pp. 1–8.
17. Kumarage, H.; Khalil, I.; Tari, Z.; Zomaya, A. Distributed anomaly detection for industrial wireless sensor networks based on fuzzy data modelling. *J. Parallel Distrib. Comput.* **2013**, *73*, 790–806. [CrossRef]
18. Zhou, D.-H.; Wei, M.-H.; Si, X.-S. A survey on anomaly detection, life prediction and maintenance decision for industrial processes. *Acta Autom. Sin.* **2013**, *39*, 711–722. [CrossRef]
19. Eskin, E. Anomaly detection over noisy data using learned probability distributions. In Proceedings of the International Conference on Machine Learning. Citeseer, San Francisco, CA, USA, 29 June–2 July 2000.
20. Meroni, M.; Fasbender, D.; Rembold, F.; Atzberger, C.; Klisch, A. Near real-time vegetation anomaly detection with MODIS NDVI: Timeliness vs. accuracy and effect of anomaly computation options. *Remote Sens. Environ.* **2019**, *221*, 508–521. [CrossRef]
21. Zamanian, Z.; Feizollah, A.; Anuar, N.B.; Kiah, L.B.M.; Srikanth, K.; Kumar, S. User Profiling in Anomaly Detection of Authorization Logs. In *Computational Science and Technology*; Springer: New York, NY, USA, 2019; pp. 59–65.
22. Borghesi, A.; Bartolini, A.; Lombardi, M.; Milano, M.; Benini, L. Anomaly detection using autoencoders in high performance computing systems. In Proceedings of the AAAI Conference on Artificial Intelligence, Honolulu, HI, USA 2019; Volume 33, pp. 9428–9433.
23. Wang, N.; Zhang, Z.; Zhao, X.; Miao, Q.; Ji, R.; Gao, Y. Exploring High-Order Correlations for Industry Anomaly Detection. *IEEE Trans. Ind. Electron.* **2019**, *66*, 9682–9691. [CrossRef]
24. Ben-Asher, N.; Gonzalez, C. Effects of cyber security knowledge on attack detection. *Comput. Hum. Behav.* **2015**, *48*, 51–61. [CrossRef]
25. Lin, C.T.; Lee, C.G.; Lin, C.T.; Lin, C. *Neural fUzzy Systems: A Neuro-Fuzzy Synergism to Intelligent Systems*; Prentice hall PTR: Upper Saddle River, NJ, USA, 1996; Volume 205.
26. Pal, S.K.; Mitra, S. *Neuro-Fuzzy Pattern Recognition: Methods in Soft Computing*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 1999.
27. Pedrycz, W.; Gomide, F. *Fuzzy Systems Engineering: Toward Human-Centric Computing*; John Wiley & Sons: Hoboken, NJ, USA, 2007.
28. Lin, C.T.; Yeh, C.M.; Liang, S.F.; Chung, J.F.; Kumar, N. Support-vector-based fuzzy neural network for pattern classification. *IEEE Trans. Fuzzy Syst.* **2006**, *14*, 31–41.
29. Meesad, P.; Yen, G.G. Pattern classification by a neurofuzzy network: application to vibration monitoring. *ISA Trans.* **2000**, *39*, 293–308. [CrossRef]
30. Lin, J.W.; Hwang, M.I.; Becker, J.D. A fuzzy neural network for assessing the risk of fraudulent financial reporting. *Manag. Audit. J.* **2003**, *18*, 657–665. [CrossRef]
31. Kuo, R.J.; Chen, C.; Hwang, Y. An intelligent stock trading decision support system through integration of genetic algorithm based fuzzy neural network and artificial neural network. *Fuzzy Sets Syst.* **2001**, *118*, 21–45. [CrossRef]
32. Wang, Y.; Zhu, Y.S.; Thakor, N.V.; Xu, Y.H. A short-time multifractal approach for arrhythmia detection based on fuzzy neural network. *IEEE Trans. Biomed. Eng.* **2001**, *48*, 989–995. [CrossRef] [PubMed]
33. Cheng, H.; Cui, M. Mass lesion detection with a fuzzy neural network. *Pattern Recognit.* **2004**, *37*, 1189–1200. [CrossRef]
34. Li, X.; Lim, B.; Zhou, J.; Huang, S.; Phua, S.; Shaw, K.; Er, M. Fuzzy neural network modelling for tool wear estimation in dry milling operation. In Proceedings of the Annual Conference of the Prognostics and Health Management Society, San Diego, CA, USA, 27 September–1 October 2009; pp. 1–11.
35. Chang, P.C.; Liu, C.H.; Fan, C.Y. Data clustering and fuzzy neural network for sales forecasting: A case study in printed circuit board industry. *Knowl.-Based Syst.* **2009**, *22*, 344–355. [CrossRef]
36. Hsiao, S.W.; Tsai, H.C. Applying a hybrid approach based on fuzzy neural network and genetic algorithm to product form design. *Int. J. Ind. Ergon.* **2005**, *35*, 411–428. [CrossRef]

37. Batista, L.O.; de Silva, G.A.; Araújo, V.S.; Araújo, V.J.S.; Rezende, T.S.; Guimarães, A.J.; Souza, P.V.d.C. Fuzzy neural networks to create an expert system for detecting attacks by SQL Injection. *Int. J. Forensic Comput. Sci.* **2018**, *13*, 8–21. [CrossRef]
38. Wang, G.; Hao, J.; Ma, J.; Huang, L. A new approach to intrusion detection using Artificial Neural Networks and fuzzy clustering. *Expert Syst. Appl.* **2010**, *37*, 6225–6232. [CrossRef]
39. de Campos Souza, P.V.; Rezende, T.S.; Guimaraes, A.J.; Araujo, V.S.; Batista, L.O.; da Silva, G.A.; Silva Araujo, V.J. Evolving fuzzy neural networks to aid in the construction of systems specialists in cyber attacks. *J. Intell. Fuzzy Syst.* **2019**, *36*, 6743–6763. [CrossRef]
40. Demertzis, K.; Iliadis, L. A hybrid network anomaly and intrusion detection approach based on evolving spiking neural network classification. In Proceedings of the International Conference on e-Democracy, Athens, Greece, 5–6 December 2013; pp. 11–23.
41. Demertzis, K.; Iliadis, L. Evolving smart URL filter in a zone-based policy firewall for detecting algorithmically generated malicious domains. In Proceedings of the International Symposium on Statistical Learning and Data Sciences, Egham, UK, 20–23 April 2015; pp. 223–233.
42. Demertzis, K.; Iliadis, L. Computational intelligence anti-malware framework for android OS. *Vietnam. J. Comput. Sci.* **2017**, *4*, 245–259. [CrossRef]
43. Yusob, B.; Mustaffa, Z.; Sulaiman, J. Anomaly Detection in Time Series Data Using Spiking Neural Network. *Adv. Sci. Lett.* **2018**, *24*, 7572–7576. [CrossRef]
44. Maas, A.L.; Hannun, A.Y.; Ng, A.Y. Rectifier nonlinearities improve neural network acoustic models. *Proc. ICML* **2013**, *30*, 3.
45. Jang, J.S. ANFIS: adaptive-network-based fuzzy inference system. *IEEE Trans. Syst. Man Cybern.* **1993**, *23*, 665–685. [CrossRef]
46. Huang, G.B.; Zhu, Q.Y.; Siew, C.K. Extreme learning machine: Theory and applications. *Neurocomputing* **2006**, *70*, 489–501. [CrossRef]
47. Souza, P.V.C. Regularized Fuzzy Neural Networks for Pattern Classification Problems. *Int. J. Appl. Eng. Res.* **2018**, *13*, 2985–2991.
48. Cukier, K.; Mayer-Schoenberger, V. The rise of big data: How it's changing the way we think about the world. *Foreign Aff.* **2013**, *92*, 28.
49. Tankard, C. Big data security. *Netw. Secur.* **2012**, *2012*, 5–8. [CrossRef]
50. Zikopoulos, P.; Eaton, C. *Understanding Big Data: Analytics for Enterprise Class Hadoop and Streaming Data*; McGraw-Hill Osborne Media: New York, NY, USA, 2011.
51. Stolfo, J.; Fan, W.; Lee, W.; Prodromidis, A.; Chan, P.K. Cost-based modeling and evaluation for data mining with application to fraud and intrusion detection. *Results Jam Proj. Salvatore* **2000**, 1–15.
52. Chandola, V.; Banerjee, A.; Kumar, V. Anomaly detection: A survey. *ACM Comput. Surv. (CSUR)* **2009**, *41*, 15. [CrossRef]
53. Patcha, A.; Park, J.M. An overview of anomaly detection techniques: Existing solutions and latest technological trends. *Comput. Netw.* **2007**, *51*, 3448–3470. [CrossRef]
54. Bhuyan, M.H.; Bhattacharyya, D.K.; Kalita, J.K. Network anomaly detection: methods, systems and tools. *IEEE Commun. Surv. Tutor.* **2014**, *16*, 303–336. [CrossRef]
55. Hartigan, J.A.; Wong, M.A. Algorithm AS 136: A k-means clustering algorithm. *J. R. Stat. Soc. Ser. C (Appl. Stat.)* **1979**, *28*, 100–108. [CrossRef]
56. Chandola, V.; Banerjee, A.; Kumar, V. Anomaly detection for discrete sequences: A survey. *IEEE Trans. Knowl. Data Eng.* **2012**, *24*, 823–839. [CrossRef]
57. Ahmed, M.; Mahmood, A.N.; Hu, J. A survey of network anomaly detection techniques. *J. Netw. Comput. Appl.* **2016**, *60*, 19–31. [CrossRef]
58. Estevez-Tapiador, J.M.; Garcia-Teodoro, P.; Diaz-Verdejo, J.E. Anomaly detection methods in wired networks: A survey and taxonomy. *Comput. Commun.* **2004**, *27*, 1569–1584. [CrossRef]
59. Sabahi, F.; Movaghar, A. Intrusion detection: A survey. In Proceedings of the 3rd International Conference on Systems and Networks Communications (ICSNC'08), Sliema, Malta, 26–31 October 2008; pp. 23–26.

60. Xie, M.; Han, S.; Tian, B.; Parvin, S. Anomaly detection in wireless sensor networks: A survey. *J. Netw. Comput. Appl.* **2011**, *34*, 1302–1325. [CrossRef]
61. Stein, D.W.; Beaven, S.G.; Hoff, L.E.; Winter, E.M.; Schaum, A.P.; Stocker, A.D. Anomaly detection from hyperspectral imagery. *IEEE Signal Process. Mag.* **2002**, *19*, 58–69. [CrossRef]
62. Garcia-Teodoro, P.; Diaz-Verdejo, J.; Maciá-Fernández, G.; Vázquez, E. Anomaly-based network intrusion detection: Techniques, systems and challenges. *Comput. Secur.* **2009**, *28*, 18–28. [CrossRef]
63. Kruegel, C.; Vigna, G. Anomaly detection of web-based attacks. In Proceedings of the 10th ACM Conference on Computer and Communications Security, Washington, DC, USA, 27–31 October 2003; pp. 251–261.
64. Sekar, R.; Gupta, A.; Frullo, J.; Shanbhag, T.; Tiwari, A.; Yang, H.; Zhou, S. Specification-based anomaly detection: A new approach for detecting network intrusions. In Proceedings of the 9th ACM Conference on Computer and Communications Security, Washington, DC, USA, 18–22 November 2002; pp. 265–274.
65. Noble, C.C.; Cook, D.J. Graph-based anomaly detection. In Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Washington, DC, USA, 24–27 August 2003; pp. 631–636.
66. Depren, O.; Topallar, M.; Anarim, E.; Ciliz, M.K. An intelligent intrusion detection system (IDS) for anomaly and misuse detection in computer networks. *Expert Syst. Appl.* **2005**, *29*, 713–722. [CrossRef]
67. Krügel, C.; Toth, T.; Kirda, E. Service specific anomaly detection for network intrusion detection. In Proceedings of the 2002 ACM symposium on Applied Computing, Madrid, Spain, 10–14 March 2002; pp. 201–208.
68. Thottan, M.; Ji, C. Anomaly detection in IP networks. *IEEE Trans. Signal Process.* **2003**, *51*, 2191–2204. [CrossRef]
69. Nychis, G.; Sekar, V.; Andersen, D.G.; Kim, H.; Zhang, H. An empirical evaluation of entropy-based traffic anomaly detection. In Proceedings of the 8th ACM SIGCOMM Conference on Internet Measurement, Vouliagmeni, Greece, 20–22 October 2008; pp. 151–156.
70. Feng, H.H.; Kolesnikov, O.M.; Fogla, P.; Lee, W.; Gong, W. Anomaly detection using call stack information. In Proceedings of the 2003 Symposium on Security and Privacy, Berkeley, CA, USA, 11–14 May 2003; pp. 62–75.
71. Lee, W.; Xiang, D. Information-theoretic measures for anomaly detection. In Proceedings of the 2001 IEEE Symposium on Security and Privacy (S&P 2001), Oakland, CA, USA, 14–16 May 2001; pp. 130–143.
72. Ye, N. A markov chain model of temporal behavior for anomaly detection. In Proceedings of the 2000 IEEE Systems, Man, and Cybernetics Information Assurance and Security Workshop. West Point, NY, USA, 6–7 June 2000; Volume 166, p. 169.
73. Xie, Y.; Yu, S.Z. A large-scale hidden semi-Markov model for anomaly detection on user browsing behaviors. *IEEE/ACM Trans. Netw.* **2009**, *17*, 54–65. [CrossRef]
74. Greensmith, J.; Aickelin, U.; Cayzer, S. Introducing dendritic cells as a novel immune-inspired algorithm for anomaly detection. In Proceedings of the International Conference on Artificial Immune Systems, Banff, AB, Canada, 14–17 August 2005; pp. 153–167.
75. Zhang, J.; Gong, J. An anomaly detection method based on fuzzy judgment. *J. Comput. Res. Dev.* **2003**, *40*, 776–783.
76. Hu, W.; Liao, Y.; Vemuri, V.R. Robust Support Vector Machines for Anomaly Detection in Computer Security. In Proceedings of the ICMLA, Los Angeles, CA, USA, 23–24 June 2003; pp. 168–174.
77. Aleskerov, E.; Freisleben, B.; Rao, B. Cardwatch: A neural network based database mining system for credit card fraud detection. In Proceedings of the IEEE/IAFE 1997 Computational Intelligence for Financial Engineering (CIFEr), New York City, NY, USA, 24–25 March 1997; pp. 220–226.
78. Schmidt, A.D.; Peters, F.; Lamour, F.; Scheel, C.; Çamtepe, S.A.; Albayrak, S. Monitoring smartphones for anomaly detection. *Mob. Netw. Appl.* **2009**, *14*, 92–106. [CrossRef]
79. Saligrama, V.; Chen, Z. Video anomaly detection based on local statistical aggregates. In Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Providence, RI, USA, 16–21 June 2012; pp. 2112–2119.
80. Xue, L.; Zhang, D.; Chen, Q.; Rao, H.; Xu, P. Tactile sensor of hardness recognition based on magnetic anomaly detection. In Proceedings of the Young Scientists Forum 2017, Shanghai, China, 5 March 2018; Volume 10710, p. 107102J.

81. Liu, N.; Rebentrost, P. Quantum machine learning for quantum anomaly detection. *Phys. Rev. A* **2018**, *97*, 042315. [CrossRef]
82. Capozzoli, A.; Piscitelli, M.S.; Brandi, S.; Grassi, D.; Chicco, G. Automated load pattern learning and anomaly detection for enhancing energy management in smart buildings. *Energy* **2018**, *157*, 336–352. [CrossRef]
83. Lyons, P.; Suen, D.; Galusha, A.; Zare, A.; Keller, J. Comparison of prescreening algorithms for target detection in synthetic aperture sonar imagery. In Proceedings of the Detection and Sensing of Mines, Explosive Objects, and Obscured Targets XXIII, Orlando, FL, USA, 16–18 April 2018; Volume 10628, p. 1062811.
84. Zhang, J.; Vukotic, I.; Gardner, R. Anomaly detection in wide area network mesh using two machine learning anomaly detection algorithms. *arXiv* **2018**, arXiv:1801.10094
85. Sabokrou, M.; Fayyaz, M.; Fathy, M.; Moayed, Z.; Klette, R. Deep-anomaly: Fully convolutional neural network for fast anomaly detection in crowded scenes. *Comput. Vis. Image Underst.* **2018**, *172*, 88–97. [CrossRef]
86. Min, R.; Song, D.; Cosatto, E. Large Margin High-Order Deep Learning With Auxiliary Tasks For Video-Based Anomaly Detection. US Patent 9,864,912, 9 January 2018.
87. Settanni, G.; Filzmoser, P. Time Series Analysis: Unsupervised Anomaly Detection Beyond Outlier Detection. *Inf. Secur. Pract. Exp.* **2018**, *11125*, 19.
88. Bittner, D.M.; Sarwate, A.D.; Wright, R.N. Using Noisy Binary Search for Differentially Private Anomaly Detection. In Proceedings of the International Symposium on Cyber Security Cryptography and Machine Learning, Beer-Sheva, Israel, 21–22 June 2018; pp. 20–37.
89. Wang, X.; Zhou, Q.; Harer, J.; Brown, G.; Qiu, S.; Dou, Z.; Wang, J.; Hinton, A.; Gonzalez, C.A.; Chin, P. Deep learning-based classification and anomaly detection of side-channel signals. *Cyber Sens.* **2018**, *10630*, 1063006.
90. Pedrycz, W. *Granular Computing: Analysis and Design of Intelligent Systems*; CRC Press: Boca Raton, FL, USA, 2016.
91. Lemos, A.; Caminhas, W.; Gomide, F. New uninorm-based neuron model and fuzzy neural networks. In Proceedings of the 2010 Annual Meeting of the North American Fuzzy Information Processing Society (NAFIPS), Toronto, ON, Canada, 12–14 July 2010; pp. 1–6.
92. Dunn, J.C. A Fuzzy Relative of the ISODATA Process and Its Use in Detecting Compact Well-Separated Clusters. *J. Cybern.* **1973**, *3*, 32–57. [CrossRef]
93. Bezdek, J.C. Objective function clustering. In *Pattern Recognition with Fuzzy Objective Function Algorithms*; Springer: New York, NY, USA, 1981; pp. 43–93.
94. Angelov, P.; Yager, R. A new type of simplified fuzzy rule-based system. *Int. J. Gen. Syst.* **2012**, *41*, 163–185. [CrossRef]
95. Angelov, P. Evolving Takagi-Sugeno Fuzzy Systems from Streaming Data (eTS+). In *Evolving Intelligent Systems: Methodology and Applications*; Wiley: Hoboken, NJ, USA, 2010; pp. 21–50.
96. Lima, E.; Gomide, F.; Ballini, R. Participatory evolving fuzzy modeling. In Proceedigs of the 2006 International Symposium on Evolving Fuzzy Systems, Ambleside, UK, 7–9 September 2006; pp. 36–41.
97. Jang, J.S. Structure determination in fuzzy modeling: A fuzzy CART approach. In Proceedings of the Third IEEE Conference on Fuzzy Systems, Orlando, FL, USA, 26–29 June 1994; pp. 480–485.
98. Han, S.J.; Cho, S.B. Evolutionary neural networks for anomaly detection based on the behavior of a program. *IEEE Trans. Syst. Man Cybern. Part B Cybern.* **2005**, *36*, 559–570.
99. Meneganti, M.; Saviello, F.S.; Tagliaferri, R. Fuzzy neural networks for classification and detection of anomalies. *IEEE Trans. Neural Netw.* **1998**, *9*, 848–861. [CrossRef] [PubMed]
100. Mahoney, M.V.; Chan, P.K. Learning rules for anomaly detection of hostile network traffic. In Proceedings of the Third IEEE International Conference on Data Mining, Melbourne, FL, USA, 22 November 2003; pp. 601–604. [CrossRef]
101. Wiig, K.M. Knowledge-based systems and issues of integration: A commercial perspective. *AI Soc.* **1988**, *2*, 209–233. [CrossRef]
102. de Campos Souza, P.V.; Guimaraes, A.J.; Araújo, V.S.; Rezende, T.S.; Araújo, V.J.S. Fuzzy Neural Networks based on Fuzzy Logic Neurons Regularized by Resampling Techniques and Regularization Theory for Regression Problems. *Intel. Artif.* **2018**, *21*, 114–133. [CrossRef]

103. Han, M.; Zhong, K.; Qiu, T.; Han, B. Interval Type-2 Fuzzy Neural Networks for Chaotic Time Series Prediction: A Concise Overview. *IEEE Trans. Cybern.* **2018**, *49*, 2720–2731. [CrossRef]
104. de Campos Souza, P.V.; Torres, L.C.B. Regularized fuzzy neural network based on or neuron for time series forecasting. In Proceedings of the North American Fuzzy Information Processing Society Annual Conference, Fortaleza, Brazil, 4–6 July 2018; pp. 13–23.
105. d. C. Souza, P.V.; Guimares, A.J.; Rezende, T.S.; Araujo, V.S.; Araujo, V.J.S.; Batista, L.O. Bayesian Fuzzy Clustering neural network for regression problems. In Proceedings of the 2019 IEEE International Conference on Systems, Man and Cybernetics (SMC), Bari, Italy, 6–9 October 2019; pp. 1492–1499. [CrossRef]
106. Guimarães, A.J.; Araújo, V.J.; de Oliveira Batista, L.; Souza, P.V.C.; Araújo, V.; Rezende, T.S. Using Fuzzy Neural Networks to Improve Prediction of Expert Systems for Detection of Breast Cancer. In *Anais do XV Encontro Nacional de Inteligência Artificial e Computacional*; SBC: Porto Alegre, Brasil, 2018; pp. 799–810. [CrossRef]
107. Silva Araújo, V.J.; Guimarães, A.J.; de Campos Souza, P.V.; Silva Rezende, T.; Souza Araújo, V. Using resistin, glucose, age and bmi and pruning fuzzy neural network for the construction of expert systems in the prediction of breast cancer. *Mach. Learn. Knowl. Extr.* **2019**, *1*, 466–482. [CrossRef]
108. de Campos Souza, P.V.; Guimaraes, A.J. Using fuzzy neural networks for improving the prediction of children with autism through mobile devices. In Proceedings of the 2018 IEEE Symposium on Computers and Communications (ISCC), Natal, Brazil, 25–28 June 2018; pp. 01086–01089.
109. Guimarães, A.J.; Araujo, V.J.S.; de Campos Souza, P.V.; Araujo, V.S.; Rezende, T.S. Using Fuzzy Neural Networks to the Prediction of Improvement in Expert Systems for Treatment of Immunotherapy. In Proceedings of the Ibero-American Conference on Artificial Intelligence, Trujillo, Peru, 13–16 November 2018; pp. 229–240.
110. Camci, E.; Kripalani, D.R.; Ma, L.; Kayacan, E.; Khanesar, M.A. An aerial robot for rice farm quality inspection with type-2 fuzzy neural networks tuned by particle swarm optimization-sliding mode control hybrid algorithm. *Swarm Evol. Comput.* **2018**, *41*, 1–8. [CrossRef]
111. Souza, P.V.d.C.; Guimaraes, A.J.; Araujo, V.S.; Rezende, T.S.; Araujo, V.J.S. Regularized Fuzzy Neural Networks to Aid Effort Forecasting in the Construction and Software Development. *arXiv* **2018**, arXiv:1812.01351
112. Souza, P.V.C.; dos Reis, A.G.; Marques, G.R.R.; Guimaraes, A.J.; Araujo, V.J.S.; Araujo, V.S.; Rezende, T.S.; Batista, L.O.; da Silva, G.A. Using hybrid systems in the construction of expert systems in the identification of cognitive and motor problems in children and young people. In Proceedings of the 2019 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), New Orleans, LA, USA, 23–26 June 2019; pp. 1–6. [CrossRef]
113. Araujo, V.S.; Rezende, T.S.; Guimarães, A.J.; Araujo, V.J.S.; de Campos Souza, P.V. A hybrid approach of intelligent systems to help predict absenteeism at work in companies. *SN Appl. Sci.* **2019**, *1*, 536. [CrossRef]
114. Hirota, K.; Pedrycz, W. OR/AND neuron in modeling fuzzy set connectives. *IEEE Trans. Fuzzy Syst.* **1994**, *2*, 151–161. [CrossRef]
115. de Campos Souza, P.V.; Silva, G.R.L.; Torres, L.C.B. Uninorm based regularized fuzzy neural networks. In Proceedings of the 2018 IEEE Conference on Evolving and Adaptive Intelligent Systems (EAIS), Rhodes, Greece, 25–27 May 2018; pp. 1–8.
116. Nair, V.; Hinton, G.E. Rectified linear units improve restricted boltzmann machines. In Proceedings of the 27th International Conference on Machine Learning (ICML-10), Haifa, Israel, 21–24 June 2010; pp. 807–814.
117. Pedrycz, W. Neurocomputations in relational systems. *IEEE Trans. Pattern Anal. Mach. Intell.* **1991**, *13*, 289–297. [CrossRef]
118. Yager, R.R.; Rybalov, A. Uninorm aggregation operators. *Fuzzy Sets Syst.* **1996**, *80*, 111–120. [CrossRef]
119. Lemos, A.P.; Caminhas, W.; Gomide, F. A fast learning algorithm for uninorm-based fuzzy neural networks. In Proceedings of the 2012 Annual Meeting of the North American Fuzzy Information Processing Society (NAFIPS), Berkeley, CA, USA, 6–8 August 2012; pp. 1–6.
120. Rosa, R.; Gomide, F.; Ballini, R. Evolving hybrid neural fuzzy network for system modeling and time series forecasting. In Proceedings of the 2013 12th International Conference on Machine Learning and Applications, Miami, FL, USA, 4–7 December 2013; Volume 2, pp. 378–383.
121. Wang, L.X.; Mendel, J.M. Generating fuzzy rules by learning from examples. *IEEE Trans. Syst. Man, Cybern.* **1992**, *22*, 1414–1427. [CrossRef]

122. Albert, A. *Regression and the Moore-Penrose Pseudoinverse*; Elsevier: Amsterdam, The Netherlands, 1972.
123. Yamanishi, K.; Takeuchi, J.I.; Williams, G.; Milne, P. On-line unsupervised outlier detection using finite mixtures with discounting learning algorithms. *Data Min. Knowl. Discov.* **2004**, *8*, 275–300. [CrossRef]
124. Williams, G.; Baxter, R.; He, H.; Hawkins, S.; Gu, L. A comparative study of RNN for outlier detection in data mining. In Proceedings of the 2002 IEEE International Conference on Data Mining, Maebashi City, Japan, 9–12 December 2002; pp. 709–712.
125. Liu, F.T.; Ting, K.M.; Zhou, Z.H. Isolation forest. In Proceedings of the 2008 Eighth IEEE International Conference on Data Mining, Pisa, Italy, 15–19 December 2008; pp. 413–422.
126. Ting, K.; Tan, S.; Liu, F. *Mass: A New Ranking Measure for Anomaly Detection*; Monash University: Melbourne, Australia, 2009.
127. Ting, K.M.; Zhou, G.T.; Liu, F.T.; Tan, J.S.C. Mass estimation and its applications. In Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Washington, DC, USA, 24–28 July 2010; pp. 989–998.
128. Tan, S.C.; Ting, K.M.; Liu, T.F. Fast anomaly detection for streaming data. In Proceedings of the IJCAI Proceedings-International Joint Conference on Artificial Intelligence, Barcelona, Spain, 16–22 July 2011; Volume 22, p. 1511.
129. Hall, M.; Frank, E.; Holmes, G.; Pfahringer, B.; Reutemann, P.; Witten, I.H. The WEKA Data Mining Software: An Update. *SIGKDD Explor.* **2009**, *11*, 10–18. [CrossRef]
130. Rumelhart, D.E.; Hinton, G.E.; Williams, R.J. *Learning Internal Representations by Error Propagation*; Technical Report; California Univ San Diego La Jolla Inst for Cognitive Science: La Jolla, CA, USA, 1985.
131. Lewis, D.D. Naive (Bayes) at forty: The independence assumption in information retrieval. In Proceeedings of the European Conference on Machine Learning, Chemnitz, Germany, 21–23 April 1998; pp. 4–15.
132. Aldous, D. The continuum random tree III. *Ann. Probab.* **1993**, *21*, 248–289. [CrossRef]
133. Cortes, C.; Vapnik, V. Support-vector networks. *Mach. Learn.* **1995**, *20*, 273–297. [CrossRef]
134. de Campos Souza, P.V.; Guimaraes, A.J.; Araujo, V.S.; Rezende, T.S.; Araujo, V.J.S. Incremental regularized Data Density-Based Clustering neural networks to aid in the construction of effort forecasting systems in software development. *Appl. Intell.* **2019**, *49*, 3221–3234. [CrossRef]
135. de Campos Souza, P.V.; Guimaraes Nunes, C.F.; Guimares, A.J.; Silva Rezende, T.; Araujo, V.S.; Silva Arajuo, V.J. Self-organized direction aware for regularized fuzzy neural networks. *Evol. Syst.* **2019**.10.1007/s12530-019-09278-5. [CrossRef]
136. Rupa Devi, T.; Badugu, S. A Review on Network Intrusion Detection System Using Machine Learning. In *Advances in Decision Sciences, Image Processing, Security and Computer Vision*; Satapathy, S.C., Raju, K.S., Shyamala, K., Krishna, D.R., Favorskaya, M.N., Eds.; Springer International Publishing: Cham, Switzerland, 2020; pp. 598–607.