Int. J. Nonlinear Anal. Appl. 14 (2023) 6, 221–248
ISSN: 2008-6822 (electronic)
http://dx.doi.org/10.22075/ijnaa.2022.28330.3860

A recurrent rough polynomial artificial neural network and its biomedical application to the classification of a cardiac patient

Alireza Mehrankia, Mohammad Reza MollaKhalili Meybodi*, Kamal Mirzaie

Department of Computer Engineering, Maybod Branch, Islamic Azad University, Maybod, Iran

(Communicated by Abdolrahman Razani)

Abstract

Since the prevalence of cardiovascular disease and consequent mortality has increased, accurate prediction of the disease status of individuals is of great importance. Therefore, models that have the least error and maximum reliability should be used. In recent years, the use of intelligent systems in engineering sciences and mechanics, especially in the prediction of diseases has increased dramatically. The diagnosis of disease by computer systems has become one of the important fields of study for researchers in this field. Diagnosis of heart disease is an attractive and challenging field of research because of the high sensitivity of communities to the death and life of the patient. The use of medical information such as age, gender, blood pressure, blood glucose level, weight, blood cholesterol levels, bio-signal of electrocardiogram, etc. can help physicians in predicting heart disease to prevent the progression of the disease, recurrent heart attacks, and consequently, reduce mortality. This data should be collected in an organized manner and used to integrate the disease prevention and diagnosis system. Evaluating these data and obtaining useful results and patterns in relation to it using data mining techniques and neural networks help to predict the early detection of this disease. This study presents a method of investigating factors influencing heart attacks using the recurrent rough group model of data handling (RRGMDH) neural network. We also compare the results of the proposed method to the results of five model of data handling neural network models, namely long short-term memory (LSTM), gated recurrent unit (GRU), redial basis function (RBF), probabilistic neural network (PNN) and recurrent group model of data handling (GMDH). The results indicate that the proposed method outperforms the five other methods.

Keywords: Analysis of Data, Recurrent Rough GMDH neural network, Diagnosis, Heart attack
2020 MSC: 68T09, 68T07

1 Introduction

In the medical world, the data related to the symptoms of various diseases and assistive methods are so vast and extensive for the diagnosis of diseases, as far as in the normal conditions, it seems difficult to analyze, examine and consider all the factors causing a disease by a person. In the contemporary world, new solutions are offered called data mining using the information of experts in any field and combining this information with computer. In the field of data mining, we seek to identify unknown patterns or predict changes in the future; therefore, it is different from statistics in general definition. According to the range of diseases, medical centers and laboratories collect data on

*Corresponding author
Email addresses: ar.mehrankia@gmail.com (Alireza Mehrankia), mollakhalili@maybodiau.ac.ir (Mohammad Reza MollaKhalili Meybodi), k.mirzaie@maybodiau.ac.ir (Kamal Mirzaie)

Received: July 2022 *Accepted:* September 2022

various diseases and use data mining techniques as a suitable solution to analyze the data and discover patterns of this large amount of data. Extracting knowledge from a very large amount of data related to patient history and medical records using data mining processes can lead to the detection of the type of disease as quickly as possible and it provides valuable information to health professionals and practitioners for the prediction and treatment of diseases according to the environmental factors. At the beginning of the twentieth century, about 10% of all deaths in the world were due to cardiovascular diseases. At the end of the twentieth century, deaths caused by heart disease increased about twenty-five percent. It is predicted that according to indicators such as air pollution and the like, the death rate will have been increased to 60% of total deaths by 2025 [16, 45, 52]. The quality of health is very important in the community and in the health industry. The most important pillar of the health industry is health data, it can be stated that data is the heart of the system. System analysis also requires the collection and proper analysis of this data. The dramatic growth of diseases and their effects and complications and the high costs to society led the medical community to seek programs for further investigation, prevention, early detection and effective treatment of diseases. In recent years, the use of data mining methods on large volumes of data for producing predictive models and patterns has become common in the various fields of medicine. Numerous and emphatic studies indicate that data mining techniques and their combination with neural networks provide an effective tool for identifying important patterns of health from medical records. Computer health records are rich in knowledge by containing a set of data about diagnosis, treatment, laboratory and potentially pharmacological measures. Although, it is not impossible for humans to discover knowledge from large amounts of data, it is very difficult. Data mining and combining it with neural networks is the best way to solve this challenge. Therefore, the concepts related to heart diseases and prediction methods introduced in the past and present will be examined in this chapter [28, 39, 53, 58]. In other words, the increase of cardiovascular diseases, the effects, complications and high costs to society led the medical community to seek programs for prevention, early detection and effective treatment of this disease according to the studies. Various research centers around the world have presented plans to study the effective factors on this disease and its control. Valuable data obtained from these projects will help the medical community around the world to obtain useful models for the early diagnosis and treatment of this disease. The use of statistical methods in investigating data and their combination with artificial neural networks in various fields has prospered, especially disease prediction, so that most of the analysis in the field of medicine has been performed by statistical methods and their combination with artificial neural networks. Statistical methods are generally used to prove the hypothesis so that, a theory is proposed and then evaluated by its correct statistical methods. The collaboration of experts in the field of computer science and medicine offers a new way for analyzing data and obtaining useful and practical models, which is the same as data mining and its combination with artificial neural networks.

Predicting the occurrence of heart diseases is very important in medical science, since it requires extremely high accuracy. This issue can be done by being aware of the patient's conditions such as age, gender, blood pressure, cholesterol, electrical signal of the heart, heart rate, blood sugar, etc.

According to the need for early and timely diagnosis of heart disease, a new method has been proposed based on the intelligent system of artificial neural networks. Many researchers tend to use this tool; however, they should be careful when using neural networks to consider two cases, which are:

1. Training data must be noise-free.
2. Try to use neural networks with deep learning so that the diagnosis can be done with high accuracy.

The necessities of this study are summarized as follows:

- Using the recursive Rough GMDH layered neural network instead of the non-recursive multi-layer perceptron neural network that results in deeper learning in the neural network.
- Improving the accuracy of the prediction model using the recursive Rough GMDH layered neural network.
- Modelling and constructing various recursive Rough GMDH layered neural networks to match them with different datasets.
- Desired output of the proposed method despite data uncertainty (noise).

The main inventions of this study are as follows:

- Presenting a novel recursive neural network based on Rough GMDH neurons.

- Presenting models of the recursive Rough GMDH layered neural network that provide more desired outputs using noisy data.
- Deep learning of the proposed neural network based on Rough GMDH neurons.
- Results of the proposed algorithm are reported compared to PNN, RBF, GRU, Recurrent GMDH, and LSTM methods on two datasets.

This paper is organized in the following sections: Data mining description and the reasons for its use in data modeling, data mining application in medicine, artificial neural networks, primary concepts, recurrent artificial neural networks, LSTM artificial neural networks, GRU artificial neural networks, PNN artificial neural networks, RBF artificial neural networks, polynomial artificial neural networks, recurrent polynomial artificial neural networks, Rough artificial neural networks, Rough polynomial artificial neural networks, recurrent Rough polynomial artificial neural networks and finally the previous study conducted in predicting heart diseases using artificial neural networks are described in Section 2. Data mining methods and recurrent rough polynomial artificial neural networks are combined to predict heart diseases more accurately, which includes two database with 13 features respectively with 964 people and 270 people, which finally the proposed model is evaluated using standard evaluation methods in Section 3. And its comparison to the evaluation results using artificial neural networks of LSTM and GRU and RBF and PNN and recurrent GMDH is presented in Section 4. Finally, the paper is concluded, and future suggestions are presented in Section 5.

2 Background

2.1 Data mining

The need for fast and accurate methods of discovering and extracting knowledge and information was increased by expanding information storage in database. To respond to this need, various techniques and methods was created and gradually introduced as data mining in this field.

There are different definitions for data mining. In this section, we refer to some of these definitions [11, 15, 20], which are:

1. Data mining is a process in which it discovers meaningful relationships, patterns and process in a large amount of data [20].
2. Data mining is a process analyzing a set of data to find the relationships between them [20].
3. Data mining is a method of summarizing data, in addition to making it easier to understand the data, it can use the knowledge contained in the data [20].
4. Data mining is the analysis of the observed data to discover the unpredictable relationships and to summarize the data in a new way, so that it is understandable and useful for data owners [12].
5. Data mining is an interdisciplinary field including machine learning techniques, pattern recognition and statistical techniques to discover information from a large dataset [12].
6. Data mining is the process of finding knowledge from large amounts of data stored in a database or data warehouse [12].

The purpose of data mining is to discover knowledge as well as data mining is created to benefit from the knowledge hidden in the data. However, data mining should not be considered the same as the process of knowledge discovery; data mining is a step in the process of knowledge discovery. In other words, data mining is a process extracting knowledge from a set of data using intelligent techniques. The extracted knowledge is presented in the form of models, patterns or rules. This knowledge can be the criterion for future decisions, next actions or necessary changes in the system. Many studies indicated that data mining methods are more accurate in modeling compared to traditional statistical methods. In addition, the advantage of such methods over traditional methods is that data mining methods use existing data for modeling [6, 49].

Data mining root can be sought to four families of sciences, which are:

Classical Statistics: Statistics is the basis of most technologies on which data mining is made. Classical statistics include concepts such as regression analysis, standard distribution, standard deviation, variance, cluster analysis, and confidence intervals, all of which are for data study and the relation between data. Classical statistical analysis plays an essential role in data mining techniques [20].

Artificial intelligence: Artificial intelligence is based on innovative methods and it is opposed to statistics, trying to use a process such as human thought to solve statistical problems. Since this approach requires high computational power, it was not implemented until the early 1980s. Artificial intelligence found few applications in the scientific and governmental fields. The need to use large computers made it impossible for everyone to use the techniques provided, however, it has found its position more and more in all areas by the advancement of technology [20].

Machine learning: The main concept of machine learning is the community of statistics and artificial intelligence. Machine learning was mentioned as a change in artificial intelligence, since it is a mix of innovative methods of artificial intelligence with advanced statistical analysis. Machine learning allows computer programs to learn about the data reading, such as programs making different decisions based on the quality of the data being studied, and use statistics for their basic concepts. In addition, they use heuristic algorithms and methods of artificial intelligence to achieve the goal [20].

Database: A large collection that is the main reservoir of all present and past data of an organization, and it is always available to managers to perform reporting and analysis operations. In other words, for a very large and complex set of data that the use of existing database management tools is very difficult to be processed, a set of new methods and tools capable of intelligently processing information is required. In this case, data mining is presented [20]. Figure 1 show a schematic of the explanation given above [20].

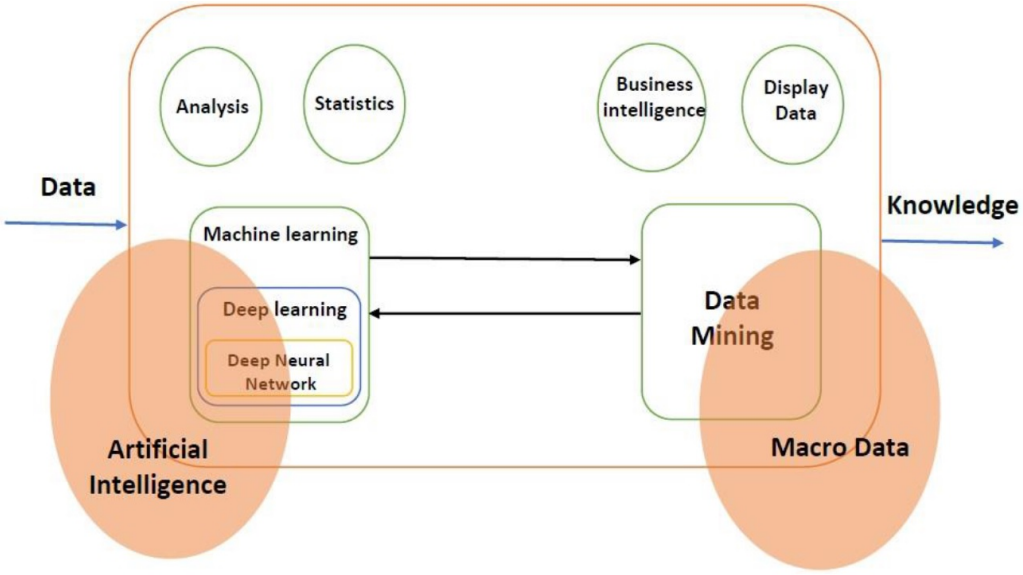

Figure 1: The position of data mining in various sciences

There are important reasons for the need to use data mining to solve data problems, some of which are listed below.

1. **High volume of data:** As the volume of data increases, the efficiency of traditional methods decreases. In many cases, traditional methods do not have efficiency. Data mining can obtain information and ultimate knowledge from large volumes of data.
2. **High data dimension:** Data dimension means the number of features available for each record. The more the number of features is higher; the data mining shows its ability better.
3. **Heterogeneity of data:** The difference in the type of features in a data set is called data heterogeneity, meaning numerical, sequential, interval values and....
4. **Dispersion of data:** Data are collected from different parts. This problem has been partially solved by forming data warehouse.
5. **Availability of computers:** Computers have become cheaper, more accessible, and much more powerful than ever.

One of the most widely used data mining areas is the diagnosis of various diseases in medicine, which can help physicians to predict patients' survival. Data mining techniques and neural networks can be combined to invent intelligent systems that require less physician supervision and can help people to understand medicine automatically and help physicians to diagnose the diseases as accurately as possible by providing useful information [40, 56, 58].

2.2 Applying data mining in medicine

The rapid increase of medical databases in developed countries caused medical researchers to be motivated to use data mining techniques to extract knowledge from these databases. The use of data mining techniques and neural networks in medical and health data has been one of the challenging areas. Medical data mining has its own characteristics distinguishing it from data mining in other issues and applications. Researchers doing data mining in other fields may encounter limitations and problems such as exploring heterogeneous, confidential or private, bulk and distributed data, and less ethical, legal, or security aspects [4, 45].

Some medical centers issue the bill of the patient, order medication, and manage inventory using patient decision systems, and they are able to generate simple statistics, however their maneuverability is limited. They are able to answer simple questions such as "What is the average age of patients with heart attacks?" However, according to the information obtained from the patient and according to the previous experiences, limited predictions can be made. Clinical decision-making is often based on the physician's observations and experience leading to unwanted biases, errors, and high medical costs; consequently, it has a direct impact on the quality of services to the patient. If we integrate clinical decision making with computer-based patient history, we can reduce mistakes in the correct diagnosis of the disease and, as a result, improve the patient as quickly as possible; it must be done by modeling and analyzing the same data mining. Data mining can have many and varied applications in the field of medicine and health [46, 54].

In addition, the existence of a system to diagnose and select the appropriate type of treatment for patients can be used by physicians that it will increase accuracy and efficiency while reducing costs [4, 42, 45, 58].

2.3 Artificial Neural Network

Artificial neural networks can be called an electronic model of the neural structure of the human brain. The mechanism of learning and training of the human brain is basically based on experience. The electronic model of neural networks is based on the same model. The way of dealing with such models with problems differs from the computational methods typically used by computer systems. We know that even the simplest animal brains can solve problems that, at least modern computers face with problems in solving them. For example, various problems of pattern recognition are examples of cases in which conventional computational methods fail to obtain the desired result to solve them. While the brain of the simplest animals can easily handle such problems, studies in this field have indicated that the brain stores information as patterns. The process of storing information as a pattern and analyzing it form the basis of a new computational method. This field of computational knowledge does not use traditional programming methods in any way, and instead, it uses large networks arranged and trained in parallel [44].

Among the advantages of artificial neural networks, it can refer to the following which are [41]:

1. **Adaptive learning:** The capability to learn how to perform tasks based on information provided for practice and introductory experiences.
2. **Automated organization:** An artificial neural network can organize the information receiving during the learning period.
3. **Real-time performance:** Calculations of artificial neural network can be performed in parallel, and special hardware has been designed and built that can use this capability.
4. **Error tolerance without interrupting when coding information:** A minor failure of a network leads to a performance degradation corresponding to it, although some network capabilities may remain intact, even by great damages.

Artificial neural networks have a high potential for nonlinear mapping of inputs to outputs. Therefore, they have been used in many branches of science and engineering. The structure of artificial neural networks is designed appropriate to the type of application, available information and specific needs of the problem, accordingly, different types of neural networks with specific efficiency and features have been invented. These networks are an information processing system and, as mentioned, they have functional properties similar to the neural network of living organisms. Artificial neural networks have been developed and extended from mathematical modeling of neural networks of living organisms based on the following hypotheses, which are [2, 14, 17, 40]:

1. Information processing is done in simple components called nerve cells or neurons.
2. Information is exchanged between neurons through the connections between them; the process takes place within the nervous system.
3. These connections have a weight multiplied by the information transmitted from one neuron to another; the weights represent the information needed to solve the problem.

4. Neurons apply a stimulus function to its inputs to calculate its output.
5. Neuron in an artificial neural network is a type of information processing unit, figure 2 shows mathematical model as well as an example of a three-layer perceptron neural network.

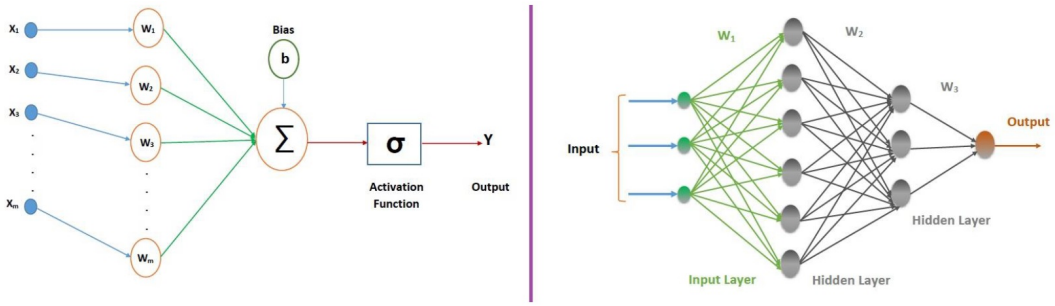

Figure 2: Scheme of the neurons and three-layer neural network structure.

2.4 Recurrent Artificial Neural Network

Recursive artificial neural network, called recurrent neural network, is a type of artificial neural network used in speech recognition, natural language processing, as well as in predictive sequential data processing. Many deep networks, such as convolutional artificial neural networks, are leading networks, in which the signal moves only in one direction from the input layer to the hidden layers and then to the output layer and previous data is not stored in memory. However, in the recursive artificial neural networks, they have a feedback layer in which the network output, along with the next input, is returned to the network. Recursive artificial neural networks can remember their previous input due to internal memory and use this memory to process a sequence of inputs. In other words, recursive neural networks consist of a recursive loop caused the information gained from previous moments not to be lost and to be remained in the network. These networks are designed to process continuous signals. In a typical neural network, all inputs and outputs are independent from each other, but in many cases, this idea may not be appropriate. Consider that you want to predict the next word in a sentence; if the network doesn't understand the relationships between the words, it certainly cannot predict the next word correctly [36].

One of the attractions of recurrent neural networks is that they may be able to relate previously observed information to the current work. These networks are designed to process continuous signals. In a typical neural network, all inputs and outputs are independent of each other, but in many cases, the idea can be terrible. Consider you want to predict the next word in a sentence. If the network cannot understand the relationships between words, it certainly cannot predict the next word correctly. These networks have a type of memory and record the information they have seen thus far. Theoretically, these networks appear to be able to capture and use information in a long sequence. However, in practice, it is not so, and they are constrained. In other words, they only record the information in the last few steps [48]. Figure 3 show a Schematic overview of the artificial recurrent neural network structure [48].

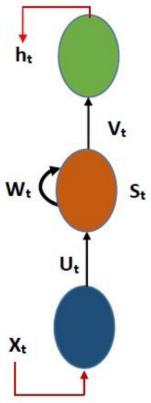

Figure 3: Schematic overview of the artificial recurrent neural network structure.

In Fig. 3, x_t is the input at time step t, and S_t is the hidden state at time step t. In the recurrent neural network structure, S_t changes according to the previous hidden state and the current input. The output is taken at time step t. Unlike conventional neural networks using different parameters in each layer, an artificial recurrent neural network shares similar parameters across all time steps, indicating that we perform the same operations at each time step, and that only the inputs are different. This technique reduces the total number of parameters that the network must learn dramatically [5]. The main feature of the artificial recurrent neural network is its hidden state that stores the information of a sequence. In addition, we do not need to have the same output/input at all-time steps. This might change depending on the desired work. The artificial recurrent neural network is called so, since the output of each layer depends on computing the previous layers. In other words, these networks have a memory that stores information about the seen data. At the first glance, it may sound slightly strange, but these networks are in fact duplicates of regular neural networks, each transmitting a message to each other. In the structure of these networks, there is at least one feedback loop [42]. Generally, the processes of a recurrent artificial neural network occur in components called neurons, and information is exchanged between neurons via their connections, where each one has a weight that is transferred via the information transmitted by a neuron. It is multiplied by the other neurons, and the weights represent the information needed to solve the problem. Each neuron applies a stimulus function to its inputs to calculate its output, the output of each neuron in recurrent artificial neural networks can be assigned to the input of the same feedback neuron, which forms the self-feedback loop. Then, it can be given to the inputs of the hidden layer or previous layers, and fed back to the recurrent neural network input [8]. Figure 4 show a Schematic of the three-layer recurrent neural network structure.

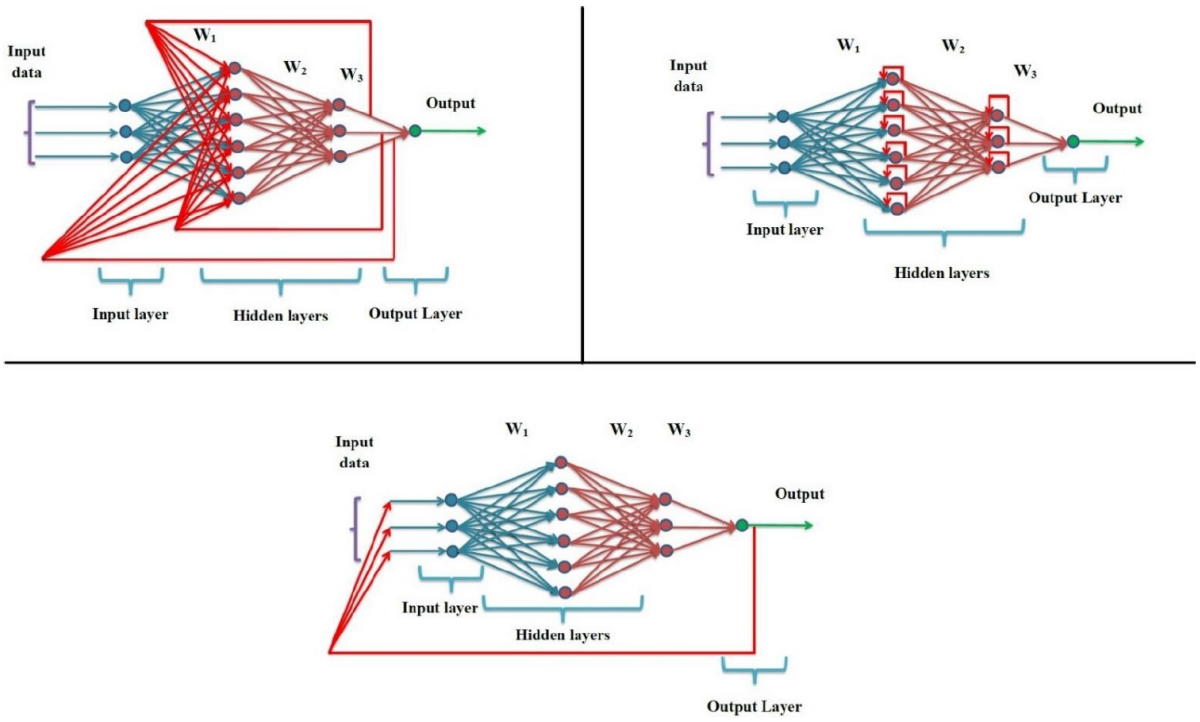

Figure 4: Schematic of the three-layer recurrent neural network structure.

2.5 LSTM Network

The LSTM recursive artificial neural network was first introduced by Hochreiter and Schmidhuber in [21] for the development of recursive artificial neural networks [49], and many researchers have played a significant role in improving these networks. Simple recursive artificial neural networks are very powerful tools; however, they suffer from some problems such as fading gradient and lack of learning long sequences. A simple recursive network encounters difficulty in remembering long sequences and it is unable to store information for long [49].

These problems led to the use of extended versions of this network. Consequently, LSTM artificial neural networks are a special type of recursive artificial neural networks having the ability to learn long-term dependencies. In addition, this network was created to solve the problem of fading gradient in the recursive neural networks. The main change

is to replace the middle layer of the above network with a block called LSTM block [37, 57, 59], Figures 5 and 6 show schematics of the overall structure of the LSTM network.

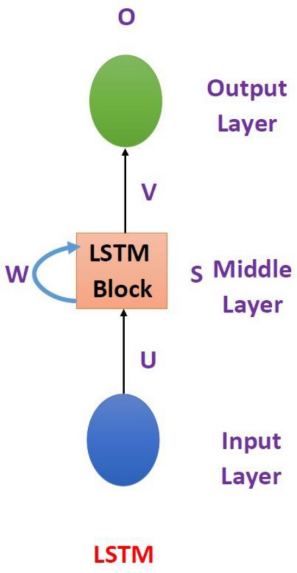

Figure 5: Schematic of the LSTM network structure.

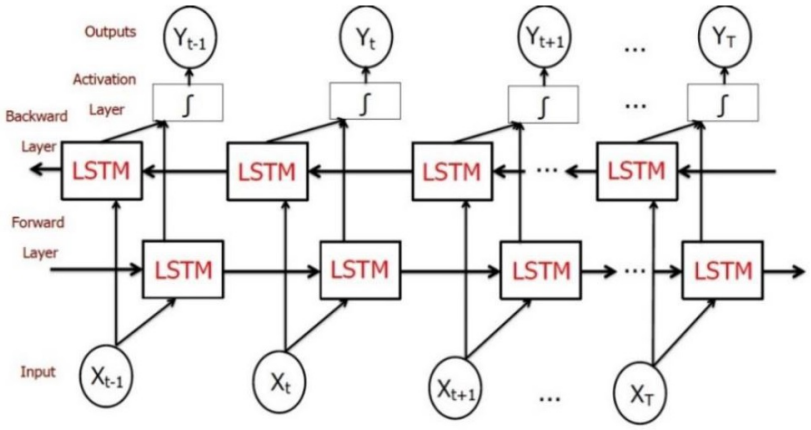

Figure 6: Schematic of the LSTM network structure.

2.6 GRU Network

The GRU recursive artificial neural network is an improved version of the LSTM recursive artificial neural network, since both use the same design and in some cases, provide excellent results. The GRU recursive artificial neural network was introduced in [25], this network is a special type of recursive neural networks having the ability to learn long-term dependencies. The GRU recursive artificial neural network acts like LSTM mechanism, its limitation is that it uses two gates instead of three gates, so it will improve well in terms of speed [24, 25, 54].

This network is presented to remove the shortcomings of the traditional recursive neural network, such as the problem of fading gradient as well as reducing the overhead in the LSTM architecture. This type of architecture uses concepts called update gate and reset gate. These two gates are basically two vectors, it is decided using them what information is transmitted to the output. The special thing about these gates is that they can be trained to retain information about long time steps without changing over time [18, 25]. Figure 7 shows the structure of GRU artificial neural network [25].

GRU recursive artificial neural network can use the memory unit in a network to deal with any input data sequence, thus, having the ability to learn time series. The network not only can browse long time series, but also it automatically

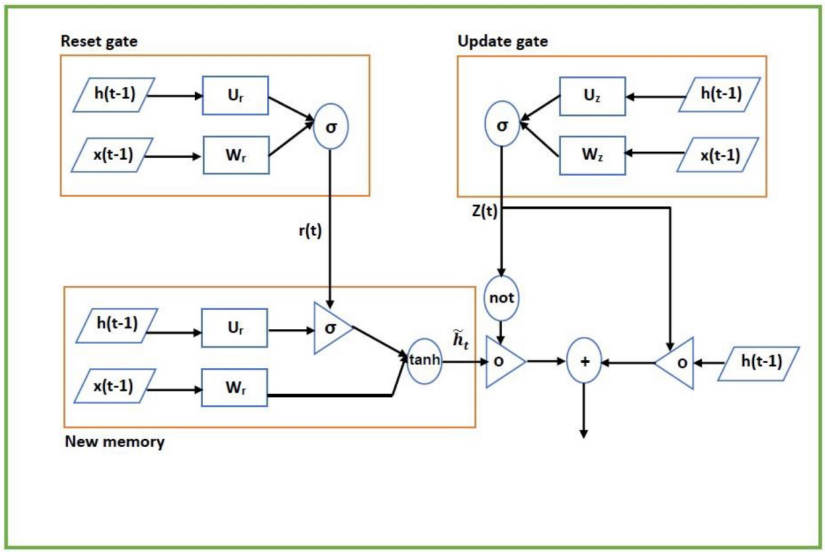

Figure 7: Schematic of the GRU network structure.

determines the optimal time balance for prediction. In recent years, GRU artificial neural network has been used by success in handwriting recognition, human movement detection, robot control, and so on [50].

2.7 PNN Network

PNN is a network that is used to identify and classify patterns, generally learns patterns in training data, and then categorizes new samples based on these patterns. PNN is a type of leading neural network derived from Bayesian networks and statistical algorithms known as Fisher's analysis used to identify and classify patterns. The PNN probabilistic neural network uses data-transit one-time learning algorithms, and this is considered one of its biggest advantages. Since methods using subjective approaches to obtain patterns in data usually require large amounts of small changes in the network parameters to gradually improve their efficiency, which means a long learning time. In addition, it is guaranteed in the learning algorithm of probabilistic neural networks that decision levels tend to optimal decision boundaries by increasing the number of training samples. Moreover, by selecting the smoothing factor, decision levels can be adjusted according to very simple or very complex needs. Probabilistic neural networks have a major drawback, it is the need to store all training samples to classify new patterns; consequently, a large amount of memory is occupied.

This type of neural networks has the simplicity and transparency of traditional models of statistical classification; in the meantime, it has the flexibility and computational power of error back-propagation neural networks. PNN probabilistic neural network is a non-parametric nonlinear modeling technique of model recognition proposed first by [10, 49].

This network is one of the most successful and practical artificial neural networks that is a powerful tool for identifying and classifying patterns with maximum probability of success. The PNN neural network is one of the most widely used neural networks in classifying input patterns. This network uses non-parametric estimation methods for classification, unlike leading networks; it does not have a local minimum problem. By replacing the adder function used in neural networks with an exponential function, a PNN neural network capable of determining nonlinear decision boundaries is obtained using Bayesian optimization policy. Bayesian strategy represents a set of rules and strategies minimizing the expected risk for classifying patterns. The PNN neural network is a four-layer neural network that can map any input scheme to any number of classes and it can quickly provide decision boundaries for new data if the old data changes [7, 10, 31, 47].

One of the advantages of using PNN probabilistic neural network is that it is faster than recursive artificial neural network for problems where incremental recursive adaptation time is an important part of the total computation time. The actual speed of training varies depending on the number of training samples, the number of hidden layers of the network, the number of elements in each hidden layer, the degree of nonlinearity of the required decision range, and the methods used for validation. The degree required for nonlinearity affects the number of training phases required for the recursive artificial neural network, however, it does not affect the time required for the probabilistic neural

network [49]. Figure 8 shows the structure of PNN artificial neural network [25].

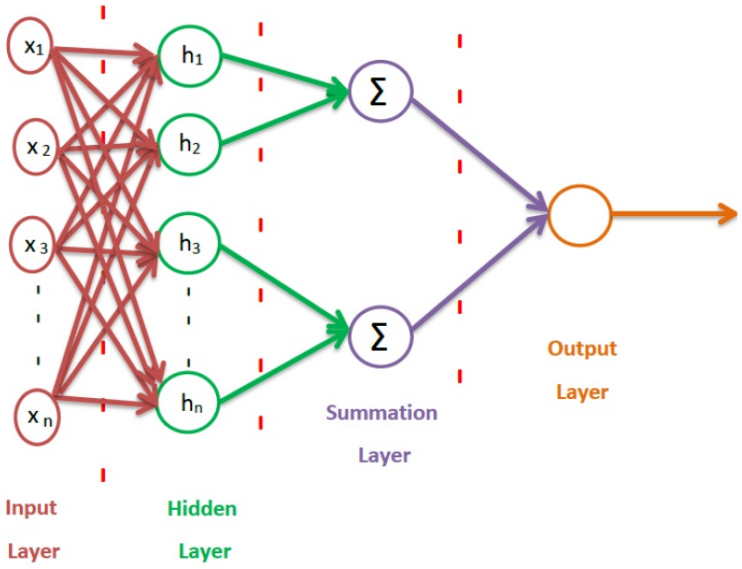

Figure 8: Schematic of the PNN network structure.

Other potential advantages of neural network include the following [47]:

- Probabilistic neural networks do not need any assumptions about the distribution of random variables used for classification. These variables can have multiple distributions.
- Probabilistic neural networks are trained well and even better than multilayer perceptron artificial neural networks.
- Probabilistic neural networks have the capability to generate mathematically levels of confidence.
- Probabilistic neural networks use all available data for modeling.

2.8 Redial Basis Function (RBF) Network

Another type of artificial neural network is the neural network based on radial basis functions, which is a feedforward network with a structure similar to the multilayer perceptron network. The function used in this network is a type of function called radial function. This type of function was first introduced by Powell in [38] as a solution to multivariate interpolation problems. Bromhead and Lowe in [13], as well as Moody and Darken in [32], were the first to use the radial-based function in the design of neural networks. This network is widely used in estimating unknown connections [3, 50].

The RBF artificial neural network has three layers: input layer, hidden layer and output layer. The RBF network requires more neurons compared to other networks, however it is designed in a fraction of time required to train other networks.

The main advantage of this network is to reduce the error to zero on the input data. In other words, there is no need to determine the number of hidden layer neurons in this network, since the number of neurons is equal to the number of input vectors. Another advantage of the RBF artificial neural network is the simplicity of the algorithm and its speed, the non-duplication of the training algorithm and, finally, the lack of convergence problems. In the RBF artificial neural network, the nodes of the input layer receive information and transmit it to the hidden layer. The transition from the input space to the middle space is nonlinear and the transition from the middle space to the output is linear. The nonlinear transition is created by the radial transition function and the linear transition function by w_1. The number of hidden layer nodes is one of the key parameters of RBF network and it should be optimized, however there is no easy way to get the optimal number of hidden layer nodes in RBF network and this number can be obtained only via experiment. The network is calculated with a different number of nodes and the best answer is considered as the optimal number of nodes. Each hidden layer node contains a function with radial base, as a

transfer function. Radial functions have parameters such as center and width shown by c and σ, respectively. These parameters are determined in the network training phase. Accurate determination of c and σ plays an important role in the efficiency of these networks. The schematic structure of an RBF neural network can be seen in Figure 9 [3, 23, 29, 50].

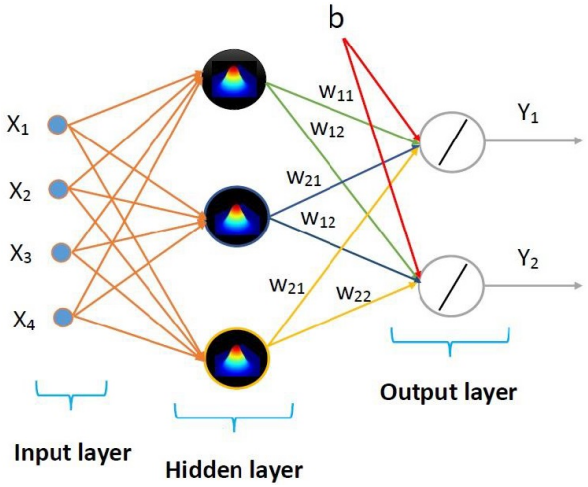

Figure 9: Schematic of the RBF network structure.

In each hidden layer node, the Euclidean distance between the center of each function and the input vectors must be calculated. Euclidean distance is the direct distance between two points in space. The transition function is applied to this distance φ and the output of the hidden layer is obtained. One of the radial functions widely used in RBF artificial neural network is Gaussian function. Figure 10 shows the diagram of this function [3, 50].

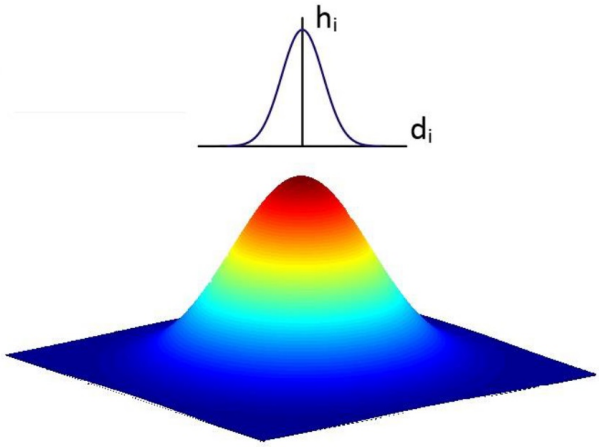

Figure 10: Schematic of the RBF network structure.

2.9 Polynomial Artificial Neural Network

Numerical data classification method is a network statistical training technology obtained from the cybernetic research including self-organizing systems, information theory, and control and computer science. GMDH is not a classical statistical modeling method. This method is a regular process to overcome statistical weaknesses and neural networks.

The polynomial neural network is formed from the GMDH algorithm. The GMDH algorithm is an inductive algorithm that mathematically models multi-parameter datasets with automatic structural and parametric optimization properties of the model. This algorithm is used in neural networks and is a good alternative to standard regressions with a multiplicative form. One of the most important features of a polynomial algorithm is its ability to identify and remove redundant variables. This algorithm was devised to solve the complexity of calculations and the problem of the

linear dependence of regressions. In fact, this algorithm is a technique to construct a high-order extended polynomial. The polynomial model with multiple inputs and one output is a subset of the basic functional components that can be seen in Eq. (2.1) [34, 43].

$$Y(x_1, ..., x_n) = a_0 \sum_{i=1}^{m} (a_i - f_{ij}) \tag{2.1}$$

f represents the elementary function that depends on different inputs sets, m is the number of basic performance components, and a represents the coefficients. Similarly, a polynomial neural network is a one-sided and self-organized network comprised of several layers, where each layer is comprised of multiple neurons. The structure of all neurons is similar, with two inputs and one output. Each neuron has one bias and five weights, and processes the relationship between the input and output data. The weights are obtained using the least squares error method. Then, they are inserted into each neuron as specific fixed values. A prominent feature of this network is that the previous neurons or layers are the factor and generator of the new neurons with the number obtained in Eq. (2.2) [34, 35, 43].

$$\binom{m}{2} = \frac{(m(m-1))}{2} \tag{2.2}$$

Some of the produced neurons have been removed to prevent the network from diverging. The remained neurons extending the network, may be removed to make the network converge and to disassociate from the bottom layer neurons. These neurons are known as inactive neurons. Therefore, the polynomial neural network is a self-organizing model gradually increasing the partial models to find the best solution, so that a structural model of optimal complexity could be found. Evolutionary methods, including genetic algorithms, are widely used in various stages of neural network design. In addition, they have unique capabilities in finding optimal quantities and in exploring in an unpredictable space. Interval polynomial neural network also uses this method [33, 35].

2.10 Polynomial artificial neural network design

In this section, we develop a GMDH neural network based on the Ivakhnenko polynomial with a combination of N-adrenal neurons. Each neuron in this neural structure, according to Fig. 11, has two inputs and one output that produce a second-order polynomial of Ivakhnenko [34].

$$net = x_1 w_1 + x_1^2 w_2 + x_1 x_2 w_3 + x_2^2 w_4 + x_2 w_5 + w_0 \tag{2.3}$$

$$\hat{y}(k) = f(net(k)) \tag{2.4}$$

where $\hat{y}(k)$ is the estimated amount. This neuron can also be used for higher-order polynomials. Each neuron has two inputs and six weight parameters whose output equation is in accordance with Eq. (2.3) and Eq. (2.4) [34].

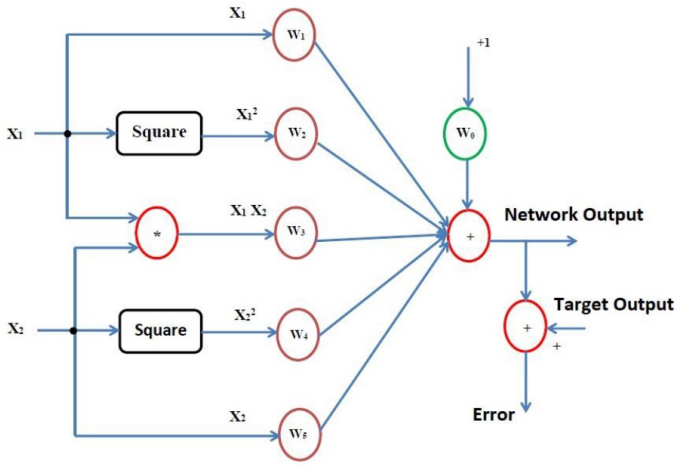

Figure 11: Polynomial network neuron structure.

Now, as explained, we show the GMDH neural network with five inputs, four activating layers, and one output in Fig. 12.

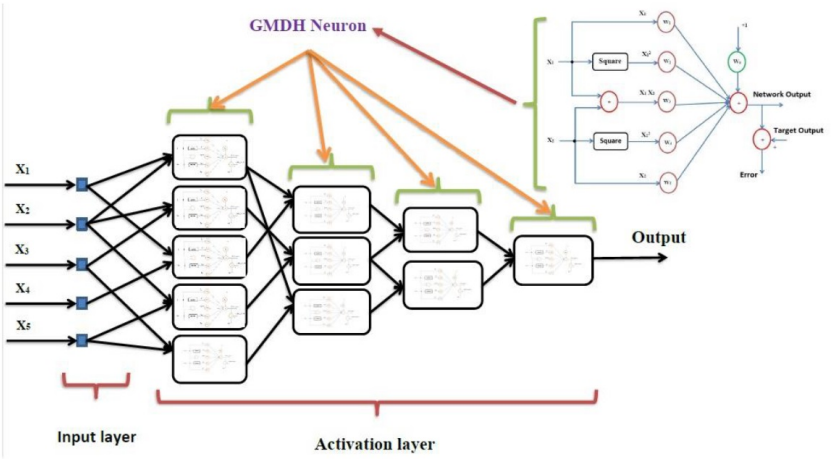

Figure 12: Polynomial network with five inputs, four active layers, and one output.

2.11 Recurrent polynomial network

Now, according to the explanations provided in Sections 2.3, 2.4, 2.9, it is possible to design a polynomial neural network based on the polynomial network and the properties of recurrent neural networks. These are described in the next section with an image [8, 30, 33, 34, 35].

2.12 Recurrent polynomial network design

In this section, we design and implement various recurrent polynomial neural networks according to the explanations presented in the previous sections. Figure 13, 14, 15 show a Schematics of a recurrent polynomial network with five inputs, five active layers, and one output.

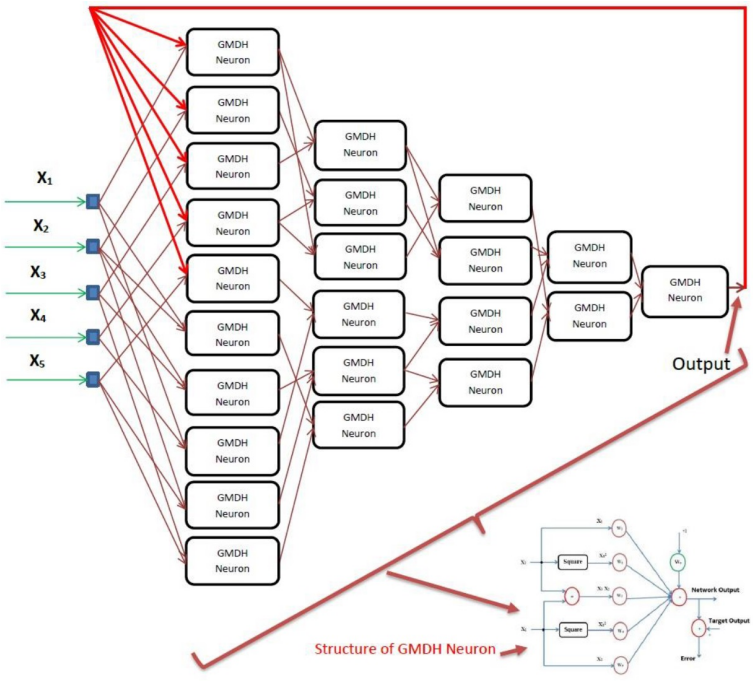

Figure 13: Schematic of a recurrent polynomial network with five inputs, five active layers, and one output.

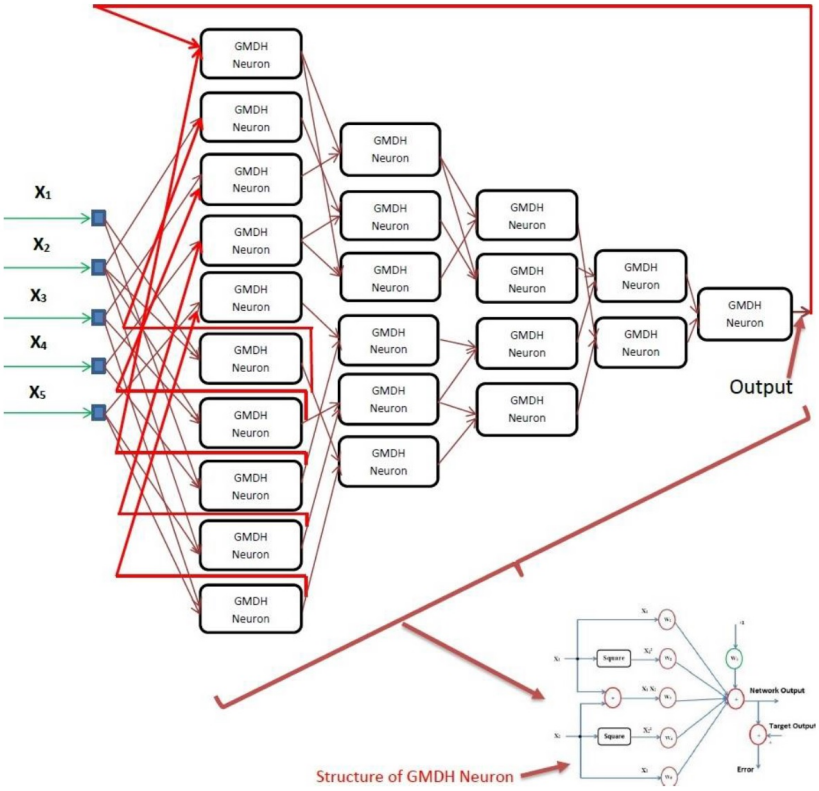

Figure 14: Schematic of a recurrent polynomial network with five inputs, five active layers, and one output.

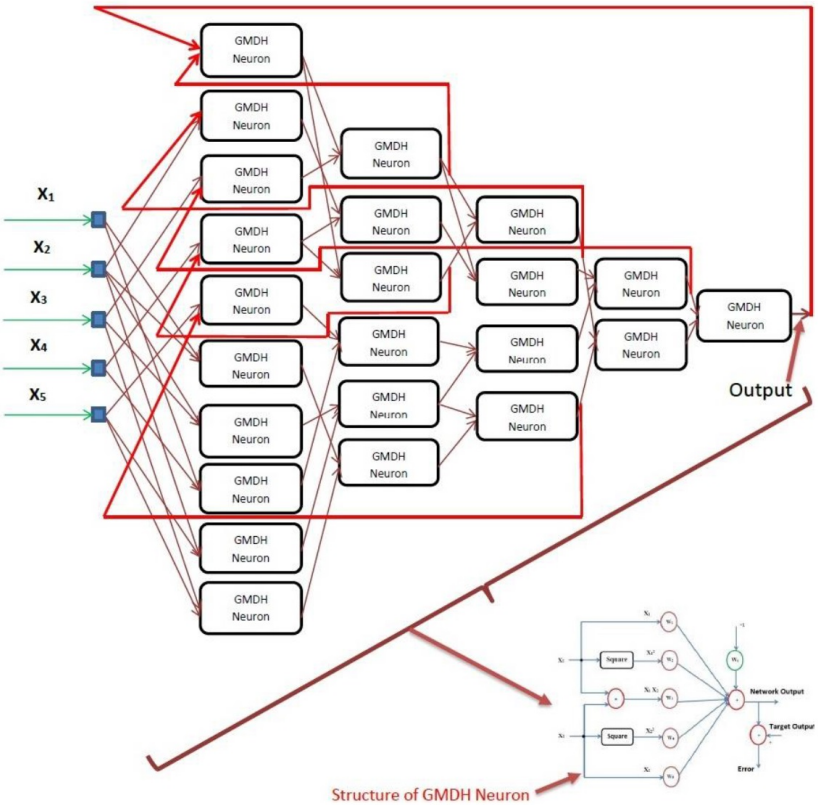

Figure 15: Schematic of a recurrent polynomial network with five inputs, five active layers, and one output.

2.13 Rough Artificial Neural Network

Rough artificial neural networks are called resistant networks against uncertainty. Rough artificial neural networks, unlike back-propagation error multilayer neural networks, perform better in the presence of noise-impregnated data. The main difference between these networks and back-propagation error multilayer neural networks is in neurons and communication weights. In these networks, rough weight is used for uncertainty instead of considering definite weights and rough neuron is used instead of definite neuron. Rough neuron can be used in all network layers. Figure 16 shows the structure of rough neuron well [16, 52].

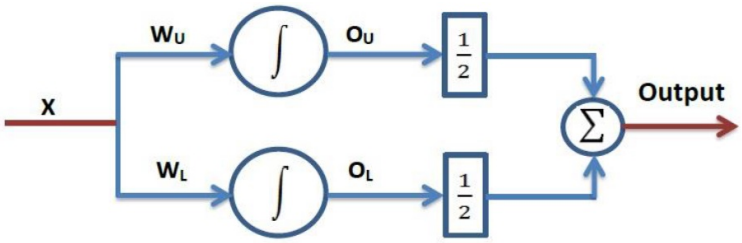

Figure 16: Schematic of a rough neuron structure.

In Figure 16, U represents the upper part of the neuron or the upper limit neuron and L represents the lower part of the neuron or the lower limit neuron. In the image above, O_U and O_L are the high and low limit neuron output. The final output obtained from the average of the two O_U and O_L outputs [16, 52].

2.14 Rough Artificial Neural Network design

According to the explanations in section 2.13, the rough neural network is designed with 5 inputs, 4 activation layers and one output, which can be seen in Figure 17.

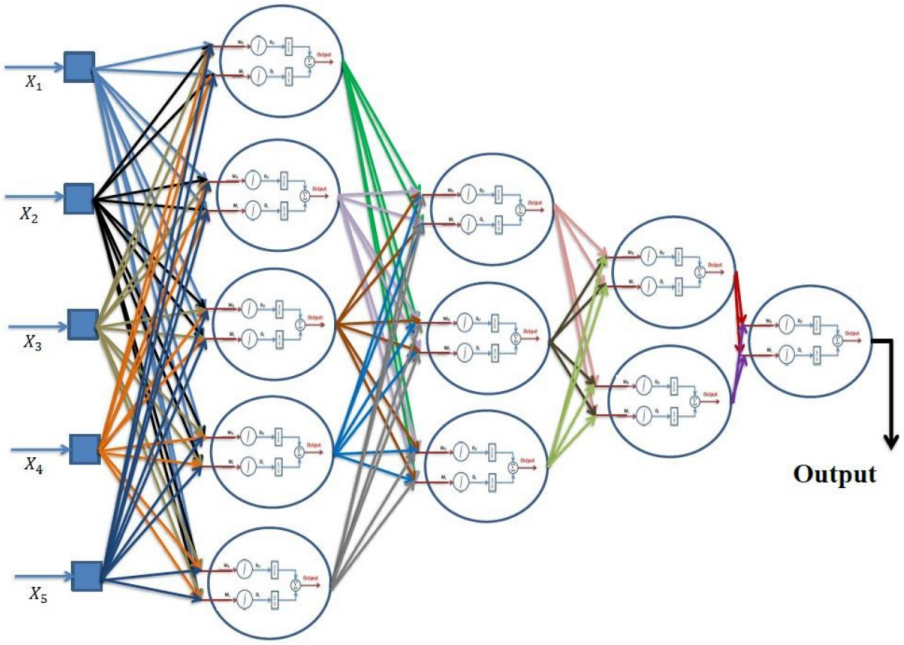

Figure 17: Rough network with five inputs, four active layers, and one output.

2.15 Rough Polynomial artificial neural network

According to the explanations about the structure of polynomial artificial neural networks as well as rough artificial neural networks in sections 2.13 and 2.14, we want to design the structure of GMDH Rough Neural Network.

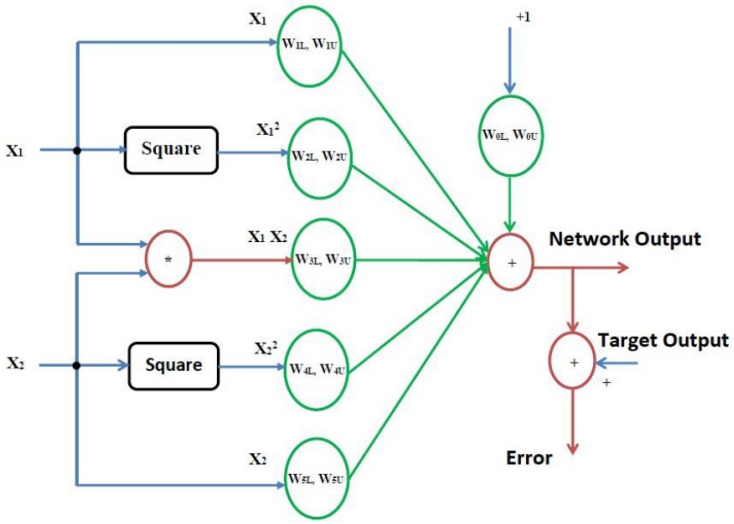

Figure 18: Rough Polynomial network neuron structure.

Each neuron has two inputs and one output in this neural structure according to Figure 18. Equation (2.5) shows the polynomial with Ivakhnenko's second-order combination in Section 2.10 and the concepts of neurons in the rough artificial neural network described in Section 2.13.

$$\frac{X_1W_{1L}+X_1W_{1U}}{2}+\frac{X_1^2W_{2L}+X_1^2W_{2U}}{2}+\frac{X_1X_2W_{3L}+X_1X_2W_{3U}}{2}+\frac{X_2^2W_{4L}+X_2^2W_{4U}}{2}+\frac{X_2W_{5L}+X_2W_{5U}}{2}+\frac{W_{0L}+W_{0U}}{2} \tag{2.5}$$

2.16 Rough Polynomial artificial neural network design

According to the explanations in sections 2.13 and 2.15, the polynomial rough artificial neural network with 5 inputs, 4 activator layers and one output can be seen in Figure 19.

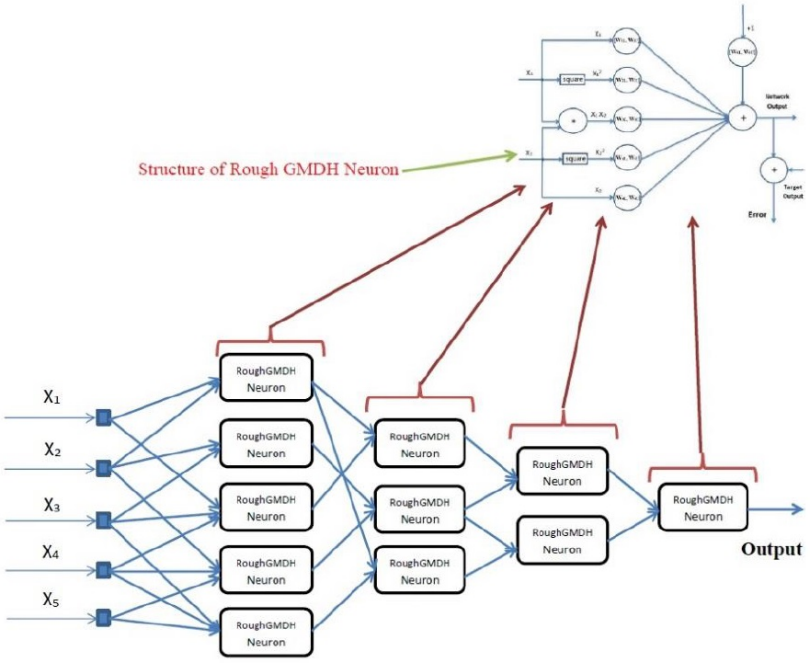

Figure 19: Schematic of a Rough polynomial network with five inputs, four active layers, and one output.

2.17 Recurrent Rough Polynomial artificial neural network

According to the explanations in sections 2.13 and 2.15, it is possible to design a recurrent polynomial neural network based on rough polynomial network and the characteristics of feedforward neural networks described in 2.18.

2.18 Recurrent Rough Polynomial artificial neural network design

According to the explanations in sections 2.4, 2.9, and 2.13, it is possible to design a recurrent rough polynomial neural network based on rough polynomial network. Figures 20, 21, and 22 show schematics of the proposed structure of the feedback rough polynomial neural network.

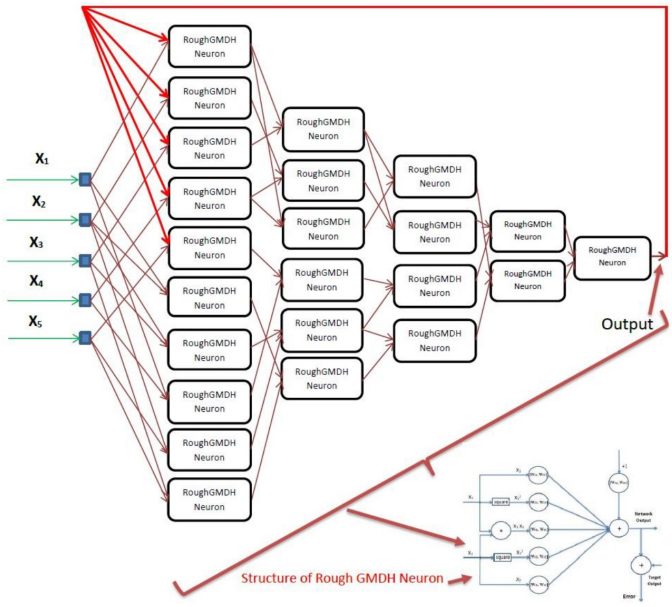

Figure 20: Schematic of a recurrent Rough polynomial network with five inputs, five active layers, and one output.

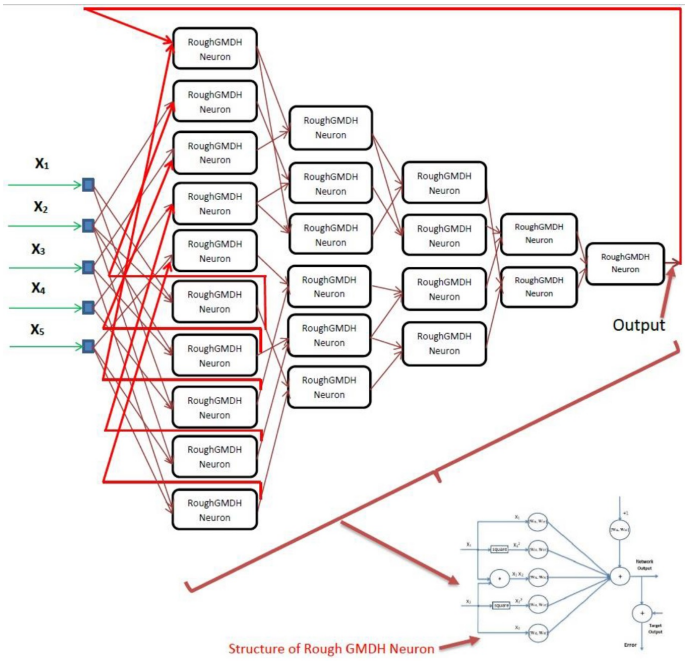

Figure 21: Schematic of a recurrent Rough polynomial network with five inputs, five active layers, and one output.

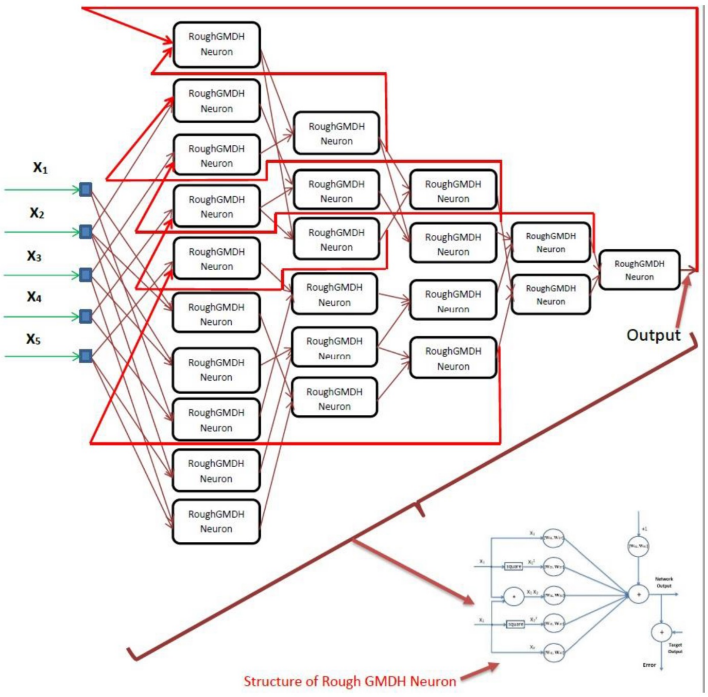

Figure 22: Schematic of a recurrent Rough polynomial network with five inputs, five active layers, and one output.

Table 1: Related Works

Author	Year	Algorithm(s) in use	Type of disease	Model accuracy	Results	Predictive variables
[27]	1995	Regression, Classification tree	Myocardial infarction	83%	Better decision tree performance in predicting myocardial infarction	Age, family history of heart diseases, cigarette smoking, chest pain, hypertension, diabetes, night sweats, sex
[29]	2009	Multilayer perceptron neural network	Congestive heart failure	93%	Neural network with better performance	Age, sex, palpitations, shortness of breath, chest pain, shortness of breath, swelling, fatigue, cough, sputum, and weight loss
[28]	2011	Bayesian network, Decision tree and artificial neural network	Heart diseases	78%	Creating rules to better communicate between variables	Age, smoking, chest pain, cholesterol, blood glucose, gender, and body mass index
[51]	2017	decision tree algorithm	Coronary heart diseases	88.4%	Correct classification of patients in need of angiography and drug therapy	Body mass index, creatinine, cholesterol, triglycerides, cigarette smoking, blood pressure, diabetes, exercise test, and history of heart diseases.
[19]	2010	Artificial neural network	Heart diseases	80.4%	Better performance of the neural network	Age, sex, hypertension, cholesterol, diabetes, chest pain, and ECG waves
[9]	2020	Classification	Heart disease	80.1%	Heart disease classification	Age, cigarette smoking, chest pain, hypertension, diabetes, night sweats, sex
[40]	2019	Hybrid genetic algorithm and fuzzy logic classifier	Heart diseases	90%	Heart disease classification	Age, sex, type of chest pain, cholesterol, blood pressure while resting, chest pain, and ECG waves

[22]	2020	Genetic algorithm and clustering	Heart diseases	94%	Heart disease classification	Age, smoking, chest pain, , hypertension, diabetes, night sweats, sex, Body mass index
[26]	2014	Decision tree, artificial neural network and Support vector machine	Heart diseases	74.19%	Better performance of the neural network and Support vector machine	Body mass index, cholesterol, cholesterol, cigarette smoking, blood pressure, diabetes, exercise test and body mass index

3 The proposed method

A large volume of information is required to diagnose diseases accurately. Attributes, or in other words, data, can be used to diagnose heart attacks accurately. Tables 2 and 3 present the details of approximately 964 people and 270 people extracted from the website (https://archive.ics.uci.edu/ml) and website (http://dataheart.ir/).

Table 2: Inputs applied to model their definitions and their range of values.

Feature	**Reference Value**
Sex	(Male=1,female=0)
Age	[29-78]
Chest pain	[1-4]
Test blood pressure stays stationary	[94-200]
Cholesterol	[126-564]
Fasting blood sugar	[0-1]
Rest electro cardio graph	[0-2]
Maximum heart rate	[71-202]
Exercise increase your angina	[0-1]
Old peak	[0-6.2]
Slope	[1-3]
Cardiac arteries	[0-3]
Represents angiographic result	[3,6,7]
Class of heart disease	Absence (1) or presence (2) of heart disease

Table 3: Inputs applied to model their definitions and their range of values.

Feature	**Reference Value**
Sex	(Male=1,female=0)
Age	[29-77]
Chest pain	[1-4]
Test blood pressure stays stationary	[94-200]
Cholesterol	[126-564]
Weight	[48-120]
Rest electro cardio graph	[0-2]
BMI (Body massing index (kg/m^2))	[18-41]
Exercise increase your angina	[0-1]
Old peak	[0-6.2]
Slope	[1-3]
PR (Pulse rate)(ppm)	[50-110]
Represents angiographic result	[3,6,7]
Class of heart disease	Absence (1) or presence (2) of heart disease

The proposed method considers thirteen features for diagnosing heart attacks.

3.1 Algorithm

The proposed method has two parts, the first part prepares data and separates training and test data, and the second part predicts if a heart attack has occurred using the recursive polynomial neural network.

3.2 Data separation

When data are not in a domain, and we intend to put them in a specific domain, we use normalization methods. This process may be useful, when it can affect the function or lengthen the checking and computation time by examining data with similar properties and different ranges. This operation is also referred to data preprocessing, being useful in the decision-making process [1, 9].

Before data separation, we should normalize it by Eq. (3.1) [1].

$$\text{Data normalization} = \frac{\text{Data} - \text{Data}_{\min}}{\text{Data}_{\max} - \text{Data}_{\min}} \tag{3.1}$$

To evaluate the performance of the proposed method, test and training data must be separated. We divide the data into two parts: training data, which contains 75% of the total data, and test data, which contains 25% of the total data.

Above, the input data include each individual's clinical symptoms, and the output data includes the prediction of a heart attack or no heart attack [1, 51].

3.3 Neural network

At this point, after preparing the data, we define the proposed neural network with 13 inputs and one output (a heart attack or no heart attack). Then, we train the recurrent polynomial neural network based on the known data and extract the output from the network as a result. Fig. 23 and Fig. 24 presents the details of the proposed recurrent rough polynomial neural network.

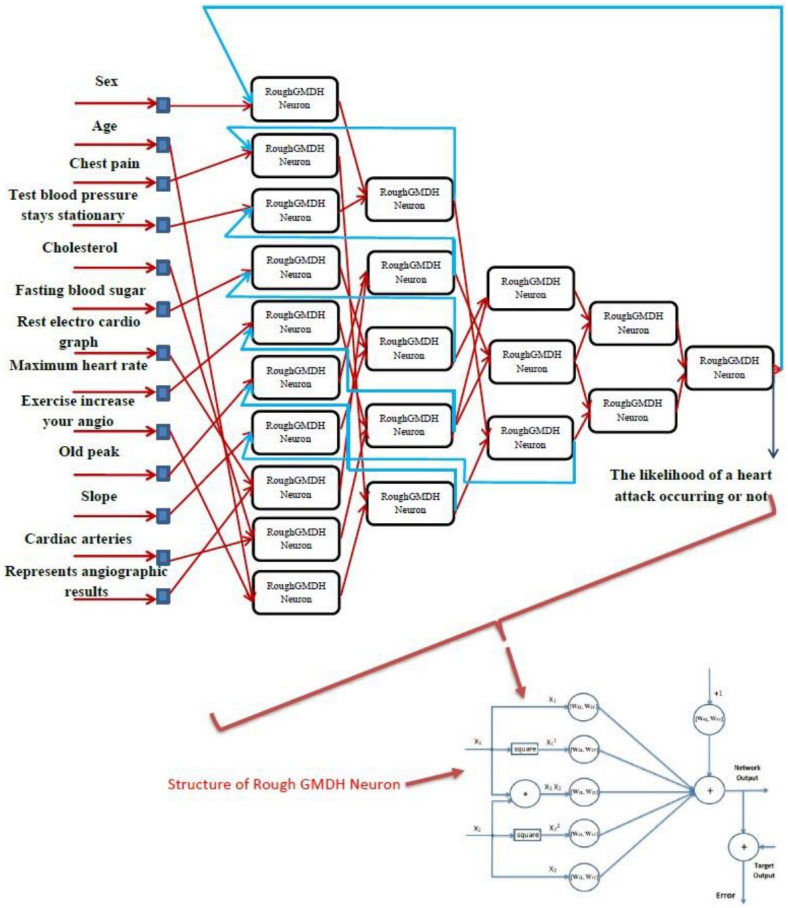

Figure 23: Proposed recurrent interval rough polynomial network structure base on the Table 2.

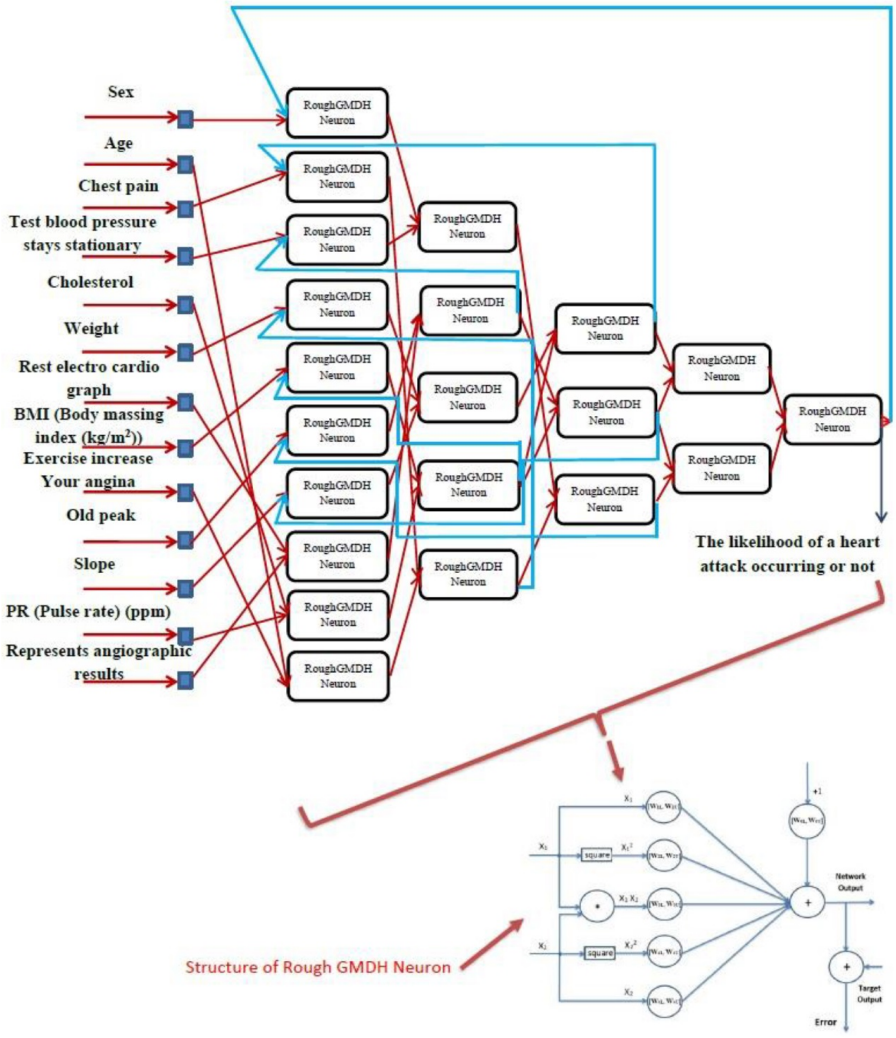

Figure 24: Proposed recurrent interval rough polynomial network structure base on the Table 3.

3.4 Evaluation of the proposed model

Accuracy of the test data classification is the models' measure of validity and accuracy. This method has features, ensuring that the trained network works well for unseen data. The most important criterion for determining the efficiency of this algorithm is the accuracy rate. This criterion is obtained as [1, 28].

$$\text{Accuracy Rate} = \frac{TN + TP}{TN + FN + TP + FP} \tag{3.2}$$

This equation states that the values of TN and TP are the most important values to be maximized.

The factors in the above equation have the following meaning:

- TN (True Negative): represents the number of records whose real batch is negative, and the classification algorithm correctly classifies them as negative.
- FN (False Negative): represents the number of records whose real batch is negative, and the classification algorithm incorrectly classifies them as negative.
- TP (True Positive): represents the number of records whose real batch is positive, and the classification algorithm correctly classifies them as positive.
- FP (False Positive): represents the number of records whose real batch is positive, and the classification algorithm incorrectly classifies them as positive.

In real-world problems, the classification accuracy criterion alone is not sufficient to evaluate the efficiency of the classification algorithm. Therefore, one has to use other criterions, such as precision and recall, as shown in Eq. (3.3) and Eq. (3.4), respectively. These criteria, which pay more attention to the positive category, explain the ability of the category to identify the positive class [1, 28, 56]. Figure 25 show a Schematic of Confusion matrix [56].

$$\text{Precision} = \frac{TP}{TP + FP} \tag{3.3}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{3.4}$$

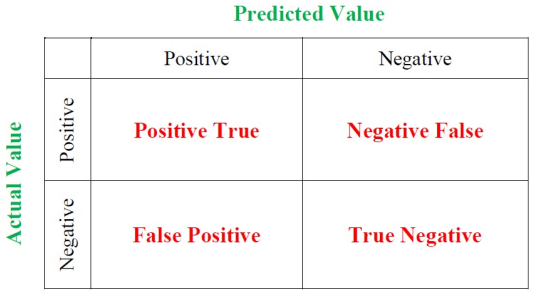

Figure 25: Confusion matrix

Precision indicates the accuracy of distinguishing the positive category for the correctly predicted positive category (TP) from the negative category for the incorrectly predicted positive (FP); therefore, the smaller the FP, the accuracy criterion would tend to 100%.

Performance criteria accurately represent the positive category for the correctly predicted positive category (TP) and the positive category for the incorrectly predicted negative (FN); thus, the smaller the FN, the more efficient the performance measure would be [1].

4 Experimental results

To this stage, heart attack prediction using the proposed neural network is explained. A total of 13 features were obtained on the first dataset with 964 records, and 13 features were obtained on the second dataset with 270 records, and the final results are compared to other conventional methods.

Table 4: Results of different methods - Training data.

Algorithm	TP	TN	FP	FN	Accuracy	Precision	Recall	Time–Consuming (second)
LSTM	304	312	57	52	0.849	0.842	0.853	8.4
GRU	326	343	21	35	0.922	0.939	0.903	7.9
RBF	310	366	29	20	0.932	0.914	0.939	14.2
PNN	317	371	16	21	0.948	0.951	0.937	10.6
GMDH	344	356	13	12	0.965	0.963	0.966	9.6
Proposed method	360	359	4	2	0.991	0.989	0.994	10.3

Table 5: Results of different methods - Test data.

Algorithm	TP	TN	FP	FN	Accuracy	Precision	Recall	Time–Consuming (second)
LSTM	90	110	10	29	0.836	0.900	0.756	2.8
GRU	105	120	6	8	0.941	0.945	0.929	2.6
RBF	100	111	9	19	0.882	0.917	0.840	4.6
PNN	112	116	6	5	0.953	0.949	0.957	3.5
GMDH	118	118	1	2	0.987	0.991	0.983	3.1
Proposed method	120	117	1	1	0.991	0.991	0.991	3.4

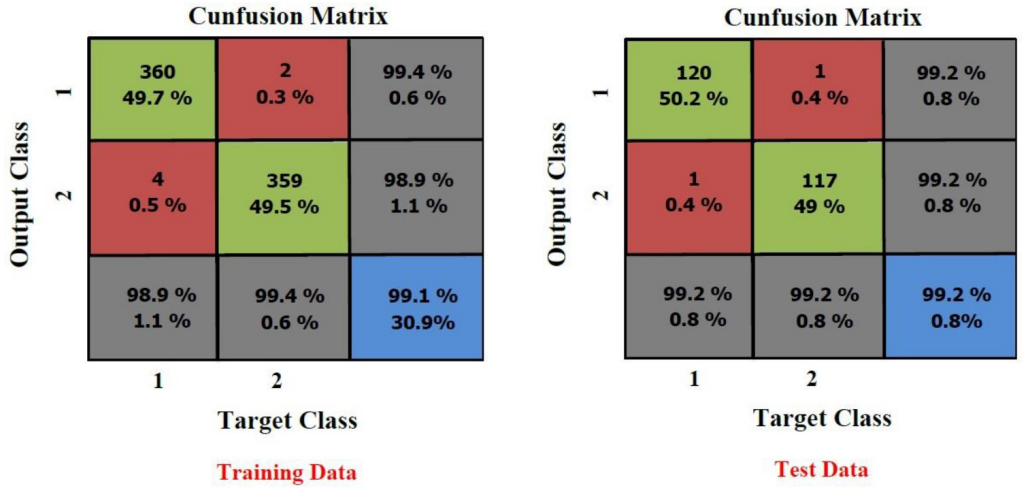

Figure 26: Confusion Matrix for training and test data.

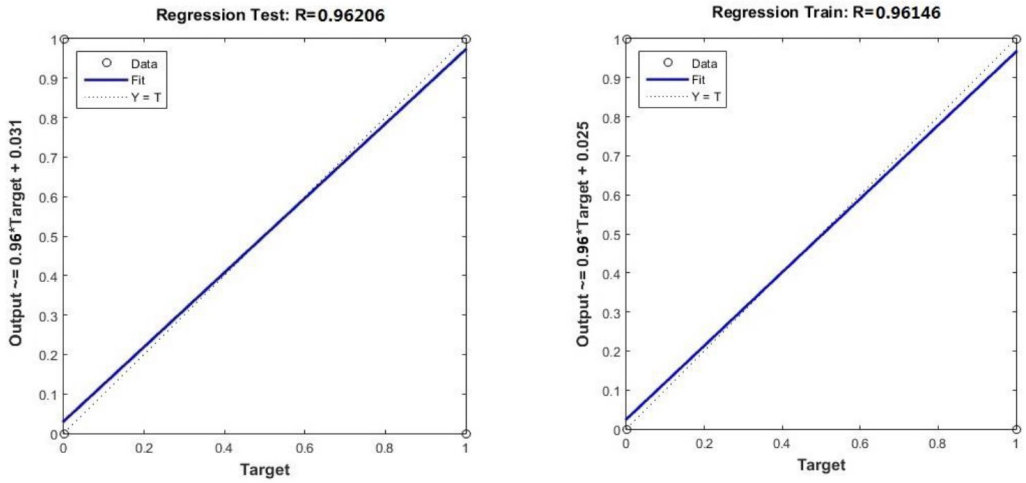

Figure 27: Regression of proposed method for training and test data.

Table 6: Results of different methods - Training data.

Algorithm	TP	TN	FP	FN	Accuracy	Precision	Recall	Time–Consuming (second)
LSTM	85	87	16	15	0.847	0.841	0.850	2.6
GRU	88	96	3	16	0.906	0.967	0.846	2.4
RBF	88	100	10	5	0.926	0.897	0.946	4.1
PNN	89	103	4	7	0.945	0.956	0.927	3.2
GMDH	96	101	3	3	0.970	0.969	0.969	2.9
Proposed method	103	97	2	1	0.985	0.980	0.990	3.3

Table 7: Results of different methods - Test data.

Algorithm	TP	TN	FP	FN	Accuracy	Precision	Recall	Time–Consuming (second)
LSTM	25	30	3	9	0.820	0.892	0.735	1.0
GRU	29	34	2	2	0.940	0.935	0.935	1.3
RBF	28	31	3	5	0.880	0.903	0.848	1.8
PNN	30	32	3	2	0.925	0.909	0.937	1.4
GMDH	33	32	1	1	0.970	0.970	0.970	1.2
Proposed method	34	31	1	1	0.985	0.971	0.970	1.5

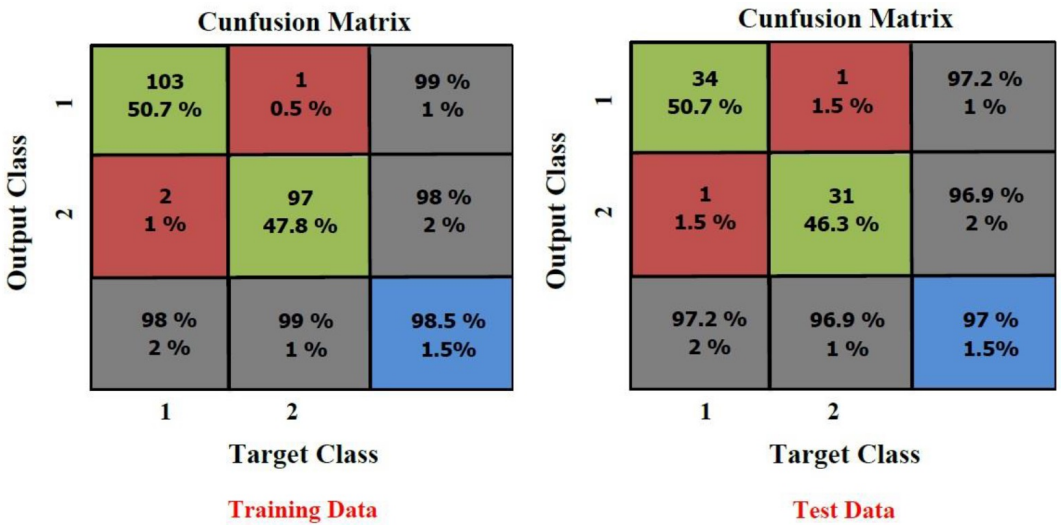

Figure 28: Confusion Matrix for training and test data.

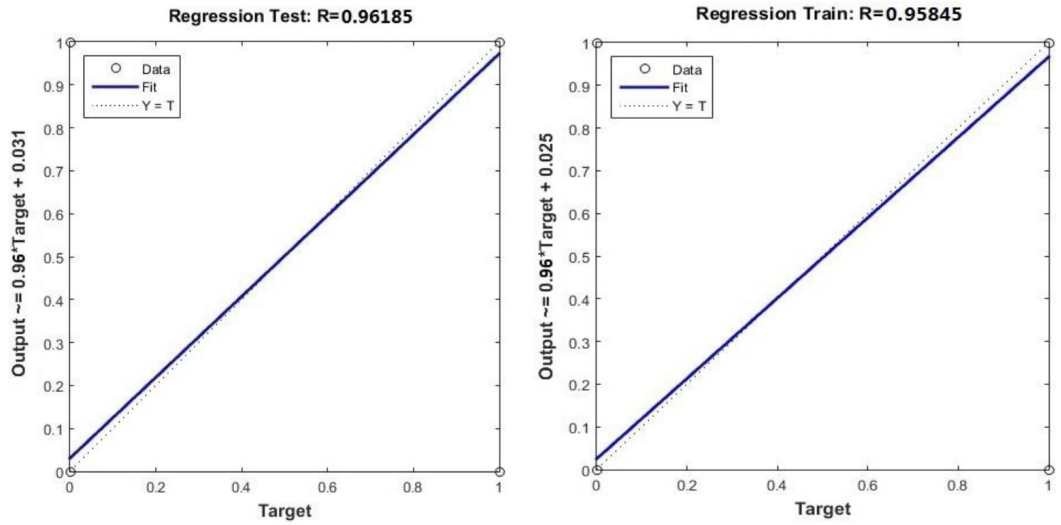

Figure 29: Regression of proposed method for training and test data.

Tables 4 and 5 as well as tables 6 and 7 present the results obtained based on the proposed method and comparison to GMDH, PNN, RBF, GRU, and LSTM for the first and second datasets, respectively. To ensure performance of the proposed method on different data considering accuracy, precision, recall, and time-consuming measures on the first dataset, test and training data in tables 4 and 5 indicate that the proposed method has better performance than other compared methods. Furthermore, performance on the second dataset based on test and training data shows that the proposed method performs better than other conventional methods do in terms of accuracy, precision, recall, and time-consuming according to tables 6 and 7.

5 Conclusion

Heart disease diagnosis is a significant and tedious task in medical science; however it is a complex task that must be done accurately and efficiently. Accordingly, there are tools for analyzing and extracting data that the availability of this huge collection of medical data led to a correct analysis in this area. It can increase the likelihood of predicting heart disease using medical information such as age, sex, blood pressure and blood sugar. This data must be collected in an organized manner, which can be used to integrate the prevention system. One of the methods used to diagnose heart disease is to use data mining methods and combine them with neural networks.

Data mining techniques and their combination with artificial neural networks and their application to medical

data is very important and it helps us to design systems to diagnose the type of disease or choose the appropriate treatment to save human lives. According to the sensitivity of heart disease and being unforgivable of misdiagnosis of heart disease, researchers attempt to be able to diagnose this disease with the least possible error. In this study, the factors affecting heart attacks were investigated. Previous studies have used a variety of methods to diagnose diseases, including heart disease, the most important of which are classification techniques, Bayesian network, decision tree, multilayer perceptron neural network, and regression tree. Artificial neural networks are one of the techniques used in data mining to detect patterns and diagnose diseases. According to cardiologists, the risk parameters such as old age, blood pressure, high blood fats and smoking have the greatest impact on cardiovascular disease. In the proposed method, all of these parameters to predict the cardiovascular disease have been used. These parameters indicate their importance in the treatment method. In this study, the proposed method with the best performance provides a suitable treatment method for patients with cardiovascular disease, which is the prediction with the most accuracy of the disease.

The current study handles the following constraints:

- The neural network neurons were selected to return to the previous layer using a random method in the selection of neurons, it is better to provide an automated method for the selection of recursive neurons.
- Hardware limitations, especially CPU and cache, which increase the forecast time, when the data set size is large, the processing speed is reduced, consequently increases the forecast time.
- There are legal limitations on obtaining up-to-date datasets from reputable medical centers.
- Full and long-term access to information of major treatment centers related to cardiovascular disease with a large number of patients may help to better and more efficiently design of the proposed network.

Accuracy of this method is tested on two datasets for 964 individuals and 270 individuals, respectively. The results of the proposed algorithm indicate that the prediction accuracy of the studied data is higher than other widely used and popular methods. Other combinatorial models can be discussed in future studies.

References

[1] M. Abdar, S.R. NiakanKalhori, T. Sutikno, I. Subroto and G. Arji, *Comparing performance of data mining algorithms in prediction heart diseases*, Int. J. Electr. Comput. Eng. **5** (2015), 1569–1576.

[2] I. Abiodun, A. Jantan, E. Omolara, V. Dada and H. Arshad, *State-of-the-art in artificial neural network applications: A survey*, Heliyon **4** (2018).

[3] M. Alexandridis and E. Chondrodima, *A medical diagnostic tool based on radial basis function classifiers and evolutionary simulated annealing*, J. Biomed. Inf. **49** (2014), 61–72.

[4] A.K. Arslan, C. Colak and M.E. Sarihan, *Different medical data mining approaches based prediction of ischemic stroke*, Comput. Meth. Prog. Biomed. **130** (2016), 87–92.

[5] M. Assaad, R. Bone and H. Cardot, *A new boosting algorithm for improved time-series forecasting with recurrent neural networks*, Inf. Fusion **354** (2008), 506–526.

[6] P.C. Austin, J.V. Tu, J.E. Ho, D. Levy and D.S. Lee, *Using methods from the data-mining and machine-learning literature for disease classification and prediction: A case study examining classification of heart failure subtypes*, J. Clinic. Epidem. **66** (2013), 398–407.

[7] M. Bascil and H. Oztekin, *A study on hepatitis disease diagnosis using probabilistic neural network*, J. Med. Syst. **36** (2012), no. 3, 1603–1606.

[8] R. Bekhet, Y. Wu, N. Wang, X. Geng and D. Zhi, *A study of generalizability of recurrent neural network-based predictive models for heart failure onset risk using a large and heterogeneous EHR data set*, J. Biomed. Inf. **84** (2018), 11–16.

[9] H. Benhar and A. Idri, *Data preprocessing for heart disease classification: A systematic literature review*, Comput. Meth. Prog. Biomed. 195 (2020), 197–210.

[10] F. Beritelli, G. Capizzi, G. Lo Sciuto, C. Napoli and F. Scaglione, *Automatic heart activity diagnosis based on Gram polynomials and probabilistic neural networks*, Biomed. Eng. Lett. **8** (2018), no. 1, 77–85.

[11] C. Beulah and S. Carolin, *Improving the accuracy of prediction of heart disease risk based on ensemble classification techniques*, Inf. Med. Unlocked **16** (2019).

[12] M. Bramer, *Principles of data mining*, Undergraduate Topics in Computer Science, 2007.

[13] D.S. Bromhead and D. Lowe, *Multivariate functional interpolation and adaptive networks*, Complex Syst. **2** (1988), 321–355.

[14] S. Chliebs and N. Kasabov, *Evolving spiking neural network–A survey*, Evolv. Syst. **4** (2013), no. 2, 87–98.

[15] A. Choudhary, J. Harding and M. Tiwari, *Data mining in manufacturing: A review based on the kind of knowledge*, J. Intell. Manufact. **20** (2009), no. 5, 501–521.

[16] S. Ding and J. Chen, *Rough neural networks: A review*, J. Comput. Info. Syst. **7** (2011), 2338–2346.

[17] S. Ferrari and F. Stengel, *Smooth function approximation using neural networks*, IEEE Trans. Neural Networks **16** (2005), 24–38.

[18] R. Fu, Z. Zhang and L. Li, *Using LSTM and GRU neural network methods for traffic flow prediction*, Youth Acad. Ann. Conf. Chinese Assoc. Automation, 2017, pp. 324–328.

[19] M. Gudadhe, K. Wankhade and S. Dongre, *Decision support system for heart disease based on support vector machine and artificial neural network*, Int. Conf. Comput. Commun. Tech. IEEE, 2010, pp. 741–745.

[20] J. Han, M. Kamber and J. Pei, *Data mining concepts and techniques*, Morgan Kaufmann, 2011.

[21] S. Hochreiter and J. Schmidhuber, *Long short-term memory*, Neural Comput. **9** (1997), 1735–1780.

[22] M.T. Islam, S.R. Rafa and M.G. Kibria, *Early prediction of heart disease using PCA and hybrid genetic algorithm with K-means*, Int. Conf. Comput. Inf. Tech. 2016, pp. 1–6.

[23] M. Korurek and B. Dogan, *ECG beat classification using particle swarm optimization and radial basis function neural network*, Expert Syst. Appl. **37** (2010), 7563–7569.

[24] B. Liu, C. Fu, A. Bielefield and Q. Liu, *Forecasting of Chinese primary energy consumption in 2021 with GRU artificial neural network*, Energies **10** (2017), no. 10, p. 1453.

[25] H.M. Lynn, S.B. Pan and P. Kim, *A deep bidirectional GRU network model for biometric electrocardiogram classification based on recurrent neural networks*, Nat. Res. Found. **7** (2019), 145395–145405.

[26] I. Mahmoudi, R. moghadam, R. Moazzam and S. Saeghian, *Identify models predict coronary artery disease using Neural networks and variable selection based on classification and regression tree*, J. Shahrekord Univ. Med. Sci. **15** (2014), 47–56.

[27] J. Mair, Puschendorf, J. Smidt, P. Lechleitner and F. Dienst, *A decision tree for the early diagnosis of acute myocardial infarction in nontraumatic chest pain patients at hospital admission*, CHEST **108** (1995), 1502–1509.

[28] K. Mathan, P. Malarvizh, G. Manogaran and R. Varadharajan, *A novel Gini index decision tree data mining method with neural network classifiers for prediction of heart disease*, Design Autom. Embed. Syst. **22** (2018), 225–242.

[29] S. Mehrabi, M. Maghsoudloo, H. Arabalibeik, R. Noormand and Y. Nozari, *Application of multilayer perceptron and radial basis function neural networks in differentiating between chronic obstructive pulmonary and congestive heart failure diseases*, Expert Syst. Appl. **36** (2009), 6956–6959.

[30] A. Mehrankia, M. Mollakhalili and K. Mirzaie, *Prediction of heart attacks using biological signals based on recurrent rough GMDH neural network*, J. Neural Process. Lett. **49** (2021), no. 2, 987–1008.

[31] F. Modaresi and S. Araghinejad, *A comparative assessment of support vector machines, probabilistic neural networks, and K-nearest neighbor algorithms for water quality classification*, Water Resources Manag. **28** (2014), 4095–4111.

[32] J. Moody and C.J. Darken, *Fast learning in networks of locally-tuned processing units*, Neural Comput. **1** (1989), 281–294.

[33] H. Naderpour, D. Rezazadeh, P. Fakharian, A.H. Rafiean and M. Kalantari, *A new proposed approach for moment capacity estimation of ferrocement members using group method of data handling*, Eng. Sci. Tech. Int. J. **23** (2020),

382–391.

[34] N. Narimanzadeh, A. Darvizeh, M. Darvizeh and H. Gharababaei, *Modelling of explosive cutting process of plates using GMDH-type neural network and singular value decomposition*, Mater. Process. Technol. **128** (2002), 80–87.

[35] T.N. Nguyen, S. Lee, H. Nguyen-Xuan and J. Lee, *A novel analysis-prediction approach for geometrically nonlinearproblems using group method of data handling*, Comput. Meth. Appl. Mech. Engin. **354** (2019), 506–526.

[36] F. Nordin and F. Nagi, *Layer-recurrent network in identifying a nonlinear system*, Int. Conf. Control Autom. Syst., 2008, pp. 387–391.

[37] L. Oh, K. Ng, S. Tan and R. Acharya, *Automated diagnosis of arrhythmia using combination of CNN and LSTM techniques with variable length heart beats*, Comput. Bio. Med. **102** (2018), 278–287.

[38] M.J.D. Powell, *Radial basis functions for multi-variable interpolation: A review*, Algorithms Approximation, Clarendon Press, 1987.

[39] A. Rady and S. Anwar, *Prediction of kidney disease stages using data mining algorithms*, Inf. Med. Unlocked **15** (2019).

[40] G.T. Reddy, M.P.K. Reddy, K. Lakshmanna, D.S. Rajput, R. Kaluri and G. Srivastava, *Hybrid genetic algorithm and a fuzzy logic classifier for heart disease diagnosis*, Evolut. Intell. **13** (2020), 185–196.

[41] C. Romero and S. Ventura, *Educational data mining: A survey from 1995 to 2005*, Expert Syst. Appl. **33** (2007), 135–146.

[42] J.A. Sanz, M. Galar, A. Jurio, A. Brugos, M. Pagola and H. Bustince, *Medical diagnosis of cardiovascular diseases using an interval-valued fuzzy rule-based classification system*, Appl. Soft Comput. **20** (2014), 103–111.

[43] S. Shaghaghi, H. Bonakdari, A. Gholami, I. Ebtehaj and M. Zeinolabedini, *Comparative analysis of GMDH neural network based on genetic algorithm and particle swarm optimization in stable channel design*, Appl. Math. Comput. **313** (2017), 271–286.

[44] S. Shanmuganathan, *Artificial neural network modelling: An introduction*, Artific. Neural Network Model. **628** (2016), 1–14.

[45] A. Sharma and P. Mansotra, *Emerging applications of data mining for healthcare management–A critical review*, Int. Conf. Comput. Sustain. Glob. Dev., 2014, pp. 1–6.

[46] E. Shao, D. Hou and C. Chiu, *Hybrid intelligent modeling schemes for heart disease classification*, Appl. Soft Comput. **14** (2014), 47–52.

[47] N. Shree and T. Kumar, *Identification and classification of brain tumor MRI images with feature extraction using DWT and probabilistic neural network*, Brain Inf. **5** (2018), no. 1, 23–30.

[48] S. Singh, S.K. Pandey, U. Pawar and R. Janghel, *Classification of ECG arrhythmia using recurrent neural networks*, Proc. Comput. Sci. **132** (2018), 1290–1297.

[49] D. Specht, *Probabilistic neural networks*, Neural Networks, 3 (1990), 109–118.

[50] P. Strumiłło and W. Kamiński, *Radial basis function neural networks: Theory and applications*, Neural Networks Soft Comput. Physica, Heidelberg, 2003, pp. 107–119.

[51] M. Tayefi, M. Tajfard, S. Saffar, P. Hanachi and M. Moohebati, *HS-CRP is strongly associated with coronary heart disease (CHD): A data mining approach using decision tree algorithm*, Comput. Meth. Prog. Biomed. **141** (2017), 105–109.

[52] M. Teshnelab and P. Jafari, *Neural networks and neural controllers*, Khajeh Nasir al-Din Tusi University of Technology, 2018.

[53] E. Williams, H. Mosley, J. Kop, J. Couper, L. Welch and D. Rosamond, *Vital exhaustion as a risk factor for adverse cardiac events (from the atherosclerosis risk in communities [ARIC] study)*, Amer. J. Cardiol. **105** (2010), 1661–1665.

[54] Z. Wu and S. King, *Investigating gated recurrent networks for speech synthesis*, Int. Conf. Acoust. Speech Signal Process., 2016, pp. 5140–5144.

[55] D. Wu, K. Warwick, M.N. Gasson and J.G. Burgess, S. Pan and T.Z. Aziz, *Prediition of Parkinson's disease tremor onset using a radial basis function neural network based on particle swarm optimization*, Int. J. Neural Syst. **20** (2010), no. 2.

[56] H. Wu, S. Yang, Z. Huang, J. He and X. Wang, *Type 2 diabetes mellitus prediction model based on data mining*, Inf. Med. Unlocked **10** (2018), 100–107.

[57] J. Xie, B. Chen, X. Gu, F. Liang and X. Xu, *Self-attention-based BiLSTM model for short text fine-grained segment classification*, Comput. Inf. Process. Comput. Artif. Intel. IEEE **7** (2019), 180558–180570.

[58] H. Yang and Y. PhoebeChen, *Data mining in lung cancer pathologic staging diagnosis: Correlation between clinical and pathology information*, Expert Syst. Appl. **42** (2015), 6168–6176.

[59] O. Yildirim, U.B. Baloglu, S. Tan, E.J. Ciaccio and U.R. Acharya, *A new approach for arrhythmia classification using deep coded features and LSTM networks*, Comput. Meth. Prog. Biomed. **176** (2019), 121–133.