Selection of Artificial Neural Networks for Disease Prediction

Iryna Kyrychenko[1], Oleksii Nazarov[1], Nural Huliiev[1] and Oleksii Avdieiev[1]

[1] *Kharkiv National University of Radio Electronics, Nauky Ave, 14, Kharkiv, 61166, Ukraine*

Abstract
This work talks about the relevance of the use of neural networks in the medical field, because their contribution cannot be fully assessed. Due to the large amount of information, symptoms and independent risk factors for the occurrence and development of diseases in all people, it is extremely difficult to diagnose and predict diseases. Therefore, in order to preserve and improve health, medical professionals can use expert systems to solve problems with patients' conditions or prevent them from occurring altogether. The article is also devoted to the topic of the benefit of neural networks for medicine and choosing the best one for the given task.

Keywords
Artificial neural network, medicine, multilayer network, direct propagation, prediction, diagnosis, perceptron.

1. Introduction

In the modern world, neural networks have rapidly gained great popularity. All over the world, the models are engaged in their development for various purposes. With the help of artificial neural networks, it is possible to process information in a similar way to the work of the brain: analyze, process and refine. Therefore, it is not surprising that they are used in many areas: economy, medicine, in security systems.

Volumes of information, medical and diagnostic technologies require new ways of applying data. Similarly, in medical research, the influence of statistical methods is not significant. And medicine is rapidly developing every day, which requires flexibility and the ability to process a large amount of information.

The solution to the problem can be new expert systems aimed at considering both general and purely individual cases. But it also happens that the possibility of setting up such models cannot always be for an individual consumer, so the system should be better universal than built for some situational moments, for which it is necessary to look for data for conducting observations, relationships between input data and the expected result, creation of a knowledge base – determination of the rules by which the system will work, testing and verification of the operation of the programmed algorithm [1].

Disagreements and misunderstandings when the question of diagnosis or prognosis arises are removed by neural information technologies that are able to fulfill the conditions indicated above. Therefore, the design of a universal model, which should have an optimization architecture, theoretical and practical foundations of the functioning of neural networks in the processing of biological information, is relevant.

In the 1950s and 1960s, there were attempts to combine existing at that time approaches to creating neural networks, using which the possibilities of calculations related to human intelligence, memorization, analysis and processing of information, resembling the work of the human brain, appeared.

COLINS-2023: 7th International Conference on Computational Linguistics and Intelligent Systems, April 20–21, 2023, Kharkiv, Ukraine
EMAIL: Iryna/kyrychenko@nure.ua (I. Kyrychenko); oleksii.nazarov1@nure.ua (O. Nazarov); nural.huliiev@nure.ua (N. Huliiev); oleksii.avdieiev@nure.ua (O. Avdieiev);
ORCID: 0000-0002-7686-6439 (I. Kyrychenko); 0000-0001-8682-5000 (O. Nazarov); 0000-0003-2123-0377 (N. Huliiev); 0009-0008-7806-4991 (O. Avdieiev)

Medicine is one of the main areas of human activity, where neural networks are gaining more and more popularity every year. In this field, they are used mainly in the diagnosis of diseases.

In this field, they are used to diagnose diseases. An example is the cardiac diagnostics software package developed by R Informati. Similar systems are used in hospitals in England to detect heart attacks and other cardiovascular diseases.

Deep neural networks help identify diseases and pathologies from scans, electrocardiograms, endoscopy. Already today, Google developers have built a neural network capable of recognizing diseases of the heart and blood vessels in the retina, the accuracy of which is 70%. The sample consisted of 300,000 retinal views, but it is not enough, so the accuracy is not as high as it should be. It is worth noting that this model predicts the risk of diabetes.

Each model is distinguished from others by the research data, their quantity, homogeneity and heterogeneity, the accuracy of the results depends on the amount of data and dependent attributes. Of course, the more records, the more possible values are tested and applied for further processing, but the more difficult it is to build a more accurate model. Therefore, it is difficult to choose the neural network that will bring the desired result, which can be relied on or at least taken into account during diagnosis and decision-making in the treatment of patients.

The purpose of the study is the application of neural networks in medicine and the selection of the necessary network, which must meet the following requirements:

- high accuracy of results;
- fast learning;
- using a small amount of memory.

2. Related Works

Over the last decades, the diagnosis of diseases began to use intelligent data analysis, the basis of which are algorithms and mathematical approaches. It is neural networks that are effective in forecasting and diagnosis. There are many examples of their use in observing the detection and development of various types of diseases.

Machine learning helps predict drug costs using cost data and healthcare parameters. For example, researchers have designed multi-layer perceptron (MLP), long-short-term memory (LSTM), and convolutional neural network (CNN) to predict health outcomes [3].

Also, neural networks are widespread among the classifications of the choice of treatment for degenerative diseases [4]. In 2050, the number of elderly population may reach 426 million people, compared to 143 million in 2019 [5]. Unfortunately, due to age, elderly people can have degenerative diseases, for example, osteoarthritis. At the moment, it affects more than 130 million [6]: patients experience pain in the knees that have impaired mobility.

In such cases, doctors suggest taking painkillers, exercise and physiotherapy, but not all patients can perform exercises correctly, which reduces the likelihood of improvement in the patient's condition. When they undergo rehabilitation, doctors use a goniometer to monitor the movement of the knees, but this device cannot easily record data at home or in the hospital, so in this case, it is better to apply a deep learning algorithm to monitor the person's posture. A lack of understanding of the exact problem and the treatment required using an empirical analysis based on feature extraction that will classify foot imaging for knee rehabilitation. For this purpose, histograms of oriented gradients (HOG) [7] and local binary pattern (LBP) algorithms [8] were used, and features were compared using support vector methods (SVM) [9], k-nearest neighbor (kNN) [10] and multilayer perceptron (MLP) [11].

Consider the information from the article on diagnosing dermatological problems using neural networks [12]. Nowadays, smartphones are used to detect or diagnose skin diseases [13–16], such as Hand, Foot, and Mouth Disease (HFMD), but it is quite difficult to do it, so the accuracy of this method is lower than 90%, which requires better research. First of all, with such a disease, it is necessary to pay attention to the symptoms [17]. Smartphone prediction is resource-constrained and determined in the absence of a DL or ML architecture capable of handling mixed data.

For example, Google researchers designed a DL architecture capable of diagnosing such diseases [18], but it is not for smartphones and is not intended for detecting HFMD, so a new Hybrid Deep Neural Network was developed, which takes into account integrated clinical data and images when calculating the probability of skin lesions. This model consists of a multilayer perceptron (MLP) [19], which is responsible for processing clinical data, and a modified CNN model, which extracts features from images. The results of cross-validation show that the constructed hybrid neural network calculates the probability of HFMD development with an accuracy of more than 99%.

Neural networks also play an important role in predicting heart diseases [20]. It is known that 2–4% of people have mitral valve prolapse (MVP) [21–25]. It occurs in Marfan and Ehlers-Danlos syndromes [26–28]. A study by Tison et al. applied the Gradient Boost Machine (GBM) classifier to a range of ECG features to predict MVP in a large elderly population [29], the results indicated that the area under the curve (AUC) was suboptimal at 77%.

To date, many neural networks are used, so there is a problem of choosing the one necessary for the set goals with the help of many criteria, which are important and necessary in their own way. Therefore, it is necessary to conduct research and experimentally solve a multi-criteria problem.

3. Methods and Materials

The multi-criteria problem of selecting a neural network for medical observations is considered as a set of alternative neural networks with different values of their characteristics, and it is necessary to find the best option among the proposed ones.

When solving such problems, difficulties may arise due to the ambiguity of the choice. In such cases, methods from two groups are used: the first is designed to reduce the number of evaluation criteria, while making assumptions for ranking the values of characteristics and comparing all options, and the second is aimed at eliminating bad alternatives before the comparison algorithm begins.

For this study, the preferred method is one of the methods of the first group, which includes the convolution method, the method of boundary criteria, the distance method, and the method of the main criterion.

The convolution method means that all criteria of alternatives become one common one. The most commonly used are additive, multiplicative, and max-min convolution.

The additive one is presented as follows:

$$K(x) = \sum^{n}{}_{j=1} a_j K_j(x), \quad (1)$$

where *K(x)* – general criterion for the alternative $x \in X$;

$(K_1(x), \dots K_j(x), \dots\ K_n(x))$ – a set of initial criteria;

n – number of initial criteria;

a_j – normalizing factor, weight of alternative characteristics.

The best alternative is calculated as follows:

$$x^* = arg\max_{x \in x} K\,(x)\ . \quad (2)$$

That is, the solution is the largest value calculated using convolution.

The multiplicative convolution is calculated by the formula:

$$K(x) = M^{n}_{j=1}\ K^{a_j}_{j}(x). \quad (3)$$

Maximal convolution formula:

$$K(x) = \min_{j} a_j K_j(x). \quad (4)$$

The best solutions of the multiplicative and max-min convolutions are also calculated using formula 2.

The method of boundary criteria is necessary for solving planning and design problems when trial values of the criteria are set $k_j(x) \geq k_{j^0};\ j = 1 \dots, n$. The formula is:

$$K(x) = \min_{j}\ \left(\frac{K_j(x)}{K_{j0}(x)}\right). \quad (5)$$

The best solution is calculated using formula 2.

The distance method uses an additional metric called distance. Suppose the available information $(K_0,\ .\ .\ .\ ,\ K_{0n})$ is sufficient to select the ideal solution. Let's calculate for each alternative the distance to the maximum value d(x). In this case, the best alternative is found by the formula:

$$x^* = \arg\min_{x \in x} d\,(x). \tag{6}$$

The bits of the Minkowski and Mahalanobis functions can be used as a distance metric.

The main criterion method replaces the one-criterion method in a multi-criteria problem, but with restrictions, and it also requires knowing the thresholds of the non-main criteria.

Suppose we have enough information to identify the main criterion, which is more important than the others, then the best solution is calculated as follows:

$$x^* = \arg\max_{x \in x} K_0\,(x), \tag{7}$$

provided that the values of the other criteria are below the thresholds.

It is also worth noting that, along with the methods of the first group, methods from the second group are used, in particular the Pareto principle. When solving multi-criteria problems, this method is used: the best alternative for conflicting criteria is chosen among all possible ones determined by the Pareto set. That is, the method from the first group works with a narrowed set of alternatives, which includes only those that are difficult to compare using the values of the criteria: one alternative has a higher score for one of the criteria, but a lower score for the other.

Related to the Pareto principle is the principle of equilibrium, also known as the Nash principle, which allows you to reduce the set of alternatives when you are not determining the best solution, but a collective solution supported by each of the solutions, which means that they are inferior to the scores of their criteria.

Unfortunately, there are situations when it is difficult to find solutions to multi-criteria problems due to uncontrollable and unexpected influencing parameters that may come from the environment.

In this case, the guaranteed outcome method can be useful, which consists in determining the worst-case reaction, so it is impossible to find the best solution, but perhaps a guaranteed one, the probability of which is quite high [30].

After analyzing the methods of the first and second groups, you should choose the one you need to conduct your research. All possible alternatives cannot be excluded at once, so the methods of the second group are not suitable. But since it is difficult to find the thresholds of the criteria, it is better to use one of the convolutions.

Let's choose the additive one, because the multiplicative one requires normalization of values from 0 to 1, in which case we will have 0 at 0, despite other priorities of the criteria.

First, it is necessary to determine the criteria by which all proposed alternatives will be evaluated, and then weights are calculated for each of them – their importance in deciding which one is better and more suitable.

Each criterion can have different values: both qualitative and quantitative. This method works with the latter, so if you have the former, you should replace them with the appropriate values of the latter type.

Once the quantitative data are ready, you can eliminate some of the alternatives using the Pareto principle if there are those that are worse than others in all criteria, and then normalization is necessary. The scores for the criteria differ in the scales of their values, i.e., mass is measured in kilograms, speed in m/s or in s, so to properly assess the values, you should normalize them in the range from 0 to 1. Often, the higher the value, the better, but of course, the opposite is also true, depending on the task.

One way to normalize is to divide the criterion value by the maximum value, which is used in the case of maximizing values, and in the opposite case, one is divided by this value.

The next step is to determine the weighting factors for ranking the criteria. A weighting factor is a multiplier that determines the importance of how much a given criterion can influence the final choice.

We have n criteria: the most powerful one will have the value of the parameter n divided by n, the less important one will have *n-1* divided by n, and so on. Or you can do it another way: sum the values of each criterion and divide one by this sum.

The only thing left to do is to calculate the convolution values for all alternatives and then compare them: summarize for each alternative the sum of the products of all the criteria values by their weighting factors. Let's conduct the following experiment to select the best neural network for the task at hand.

4. Experiment

Let's model and solve a vector optimization problem:

We need to choose the best possible neural network with fast training and a high percentage of confidence in the results.

We will prepare the decision-making process for choosing the right network (we will analyze the most common neural networks):

1. describe the set of alternatives:
   - Kohonen neural network;
   - multilayer perceptron (MLP);
   - hybrid neural network;
   - adaptive resonance theory networks (ART-2);
   - Fuzzy-ART.
2. describe the selection criteria:
   - training speed – the time required to train a neural network based on the analysis and processing of data for training;
   - reliability of control – the reliability of operation and diagnostics of the formation of result classes;
   - memory size – the amount of memory required to store the neural network;
   - training – the type of network training: with or without a teacher;
   - application principle – types of analyzes for which the neural network is intended.
3. describe the rating scales by criteria:
   - learning speed, control reliability, and memory capacity are already known quantitative indicators derived from research;
   - training of a neural network can be of three types: with a teacher, without a teacher, or of a mixed type; we will present these qualitative indicators as follows:
     - learning without a teacher – 3 points;
     - mixed learning – 2 points;
     - learning with a teacher – 1 point.
   - the principle of application should include the capabilities of cluster analysis, prediction of the desired phenomenon, work with linearly inseparable hyperplanes of data, and dynamic expansion of the knowledge base of experimental objects.

We present the model in the form of a table with known data in Table 1 [31].

Table 1

Signs of neural networks

Indicator	Type of neural network				
	Kohonen's neural network	Multilayer Perceptron (MLP)	Hybrid neural network	Networks of adaptive resonance theory	Fuzzy-ART
Learning speed, c	5.3	0.2	5.7	1.5	2.1
Validity of control	0.88...0.93	0.98...0.99	0.96...0.98	0.98...0.99	0.97...0.99
Memory capacity, Kb	64	81	219	203	115
Teaching	Without a teacher	With the teacher	Mixed	Without a teacher	Without a teacher

Principle of application	Cluster analysis; Forecasting; Dynamically expands the knowledge base of objects.	Cluster analysis; Forecasting; Works with linear non-separating planes.	Cluster analysis; Forecasting; Works with linear non-separating planes; Dynamically expands the knowledge base of objects.	Cluster analysis; Forecasting; Works with linear non-separating planes; Dynamically expands the knowledge base of objects.	Cluster analysis; Forecasting; Works with linear non-separating planes; Dynamicall y expands the knowledge base of objects.

Let's change the problem model for further analysis as follows:

- learning rate and memory capacity are values that should be as small as possible, unlike other indicators, so we minimize them – we will present the inverse values to support the maximization of all values when comparing neural networks – when finding the best one;
- we will change the values of control reliability to the smallest of the specified ranges for a clearer and more rigorous analysis;
- we will change the training indicators to quantitative ones, as described above;
- the values of the application principle will depend on the number of functioning of a particular neural network required for the task of the master's research, where the maximum value is 4 and the minimum value is 0.

The modified table is shown below in Table 2.

Table 2
Quantitative indicators of neural networks

Indicator	Type of neural network				
	Kohonen's neural network	Multilayer Perceptron (MLP)	Hybrid neural network	Networks of adaptive resonance theory	Fuzzy-ART
Learning speed, c	0.19	5	0.175	0.66	0.48
Validity of control	0.88	0.98	0.96	0.98	0.97
Memory capacity, Kb	0.016	0.012	0.005	0.049	0.087
Teaching	3	1	2	3	3
Principle of application	3	3	4	4	4

At this stage, the networks are not comparable according to the Pareto principle, so we will perform linear additive convolution with normalizing factors for the entire problem model and consider the results of the experiment.

5. Results

Results of comparison of neural networks by convolution are presented in Table 3.

Table 3
Results of comparison of neural networks by convolution

Indicator	Type of neural network					
	Normalizing factor	Kohonen's neural network	Multilayer Perceptron (MLP)	Hybrid neural network	Networks of adaptive resonance theory	Fuzzy-ART
Learning speed, c	0.154	0.19	5	0.175	0.66	0.48
Validity of control	0.2	0.88	0.98	0.96	0.98	0.97
Memory capacity, Kb	5.9	0.016	0.012	0.005	0.049	0.087
Teaching	0.083	3	1	2	3	3
Principle of application	0.055	3	3	4	4	4
Convolution Result		0.7137	1.2848	0.6345	1.0557	1.252

As you can see, the multilayer perceptron (BSHP) [32] is the best for our requirements, which really has a fairly high reliability of the results, takes a fairly normal amount of memory and works very quickly. Let's analyze the calculated choice.

The multilayer perceptron is the best in terms of the learning speed of neural networks, as well as in terms of the reliability of the obtained results.

This neural network really does not require a lot of memory space to function.

The algorithm is trained by a teacher, which means that a multilayer perceptron needs a complete data set for model training, which consists of neuron characteristics (unique values for all parameters) and expected results. Of course, the more of them, the more accurate the algorithm makes predictions [33].

The network does not have the ability to dynamically increase the knowledge base of objects, but this is explained by the type of its training: this principle of application is not necessary for its operation. Its main advantage is the processing of linear non-separable data, which is always present in medicine and creates many problems of multivariability in diagnosing and predicting possible diseases of patients.

It is worth noting that according to the results of linear additive convolution, the Fuzzy-ART neural network is slightly worse than the multilayer perceptron. Let's compare these two models separately.

Figure 1 shows two histograms showing that the speed of the multilayer perceptron is still higher than the Fuzzy-ART neural network.

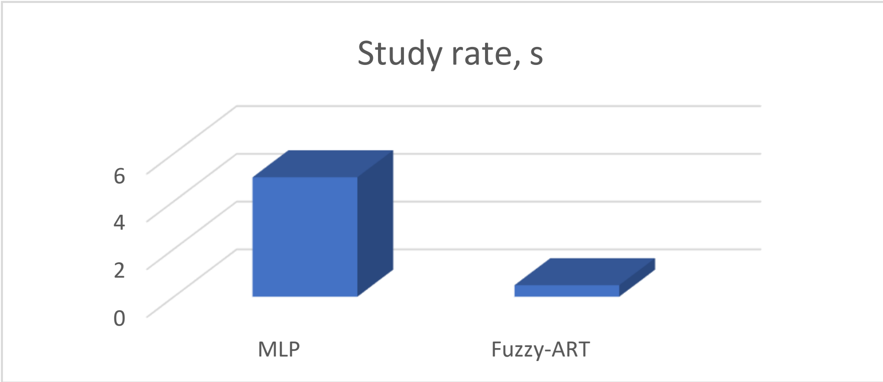

Figure 1: Speed histograms

Below in Figure 2, it can be seen that Fuzzy-ART is inferior to Perceptron in terms of reliability of results.

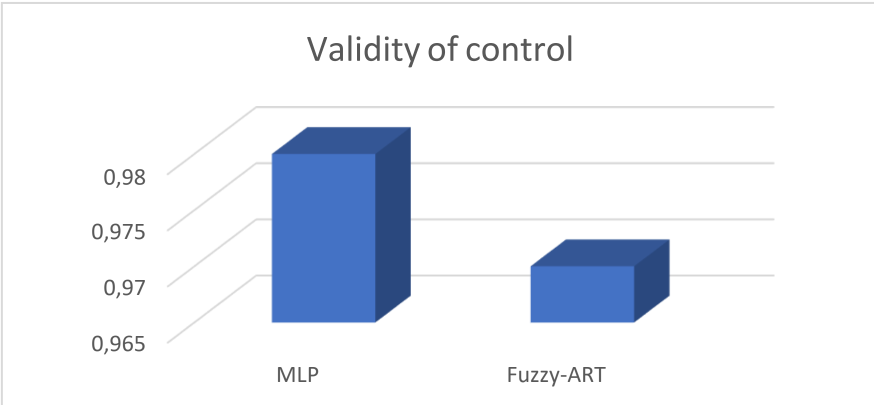

Figure 2: Control reliability histograms

As can be seen in Figure 3, Fuzzy-ART takes up more memory, so it is better to choose a multilayer perceptron here as well.

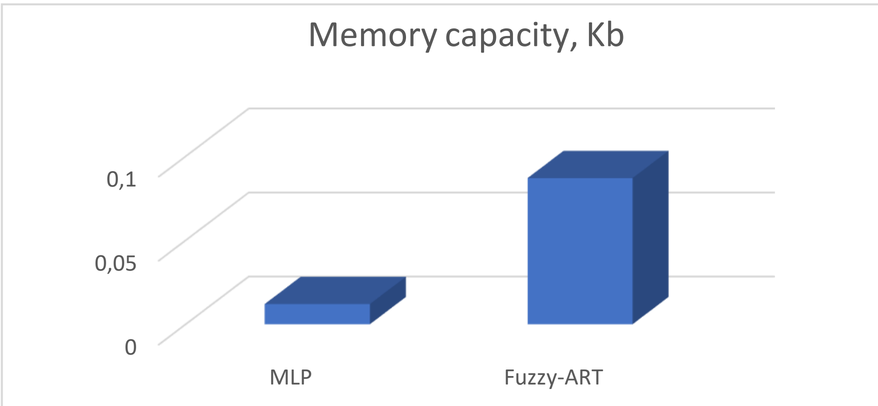

Figure 3: Histograms of memory volumes

But the Fuzzy-ART training is larger (3 – mixed form), when the perceptron has only 1 (training with a teacher), which is reflected in Figure 4. In this case, of course, the ideal option is the Fuzzy-ART network.

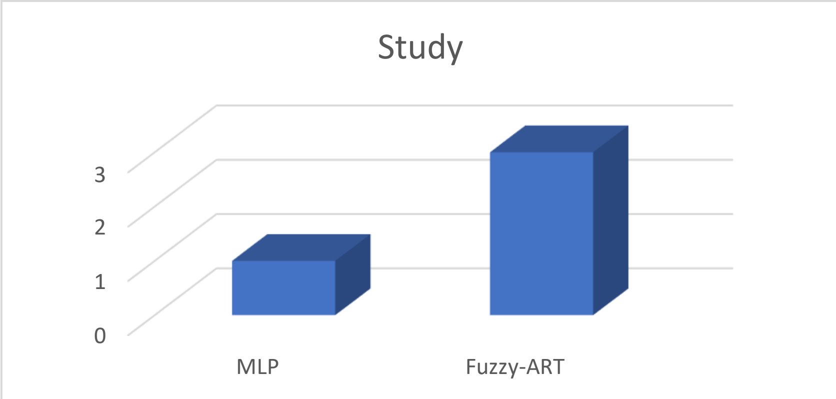

Figure 4: Histograms of types of training

Fuzzy-ART application principles, in addition to the main tasks, can also dynamically increase the database of observations.

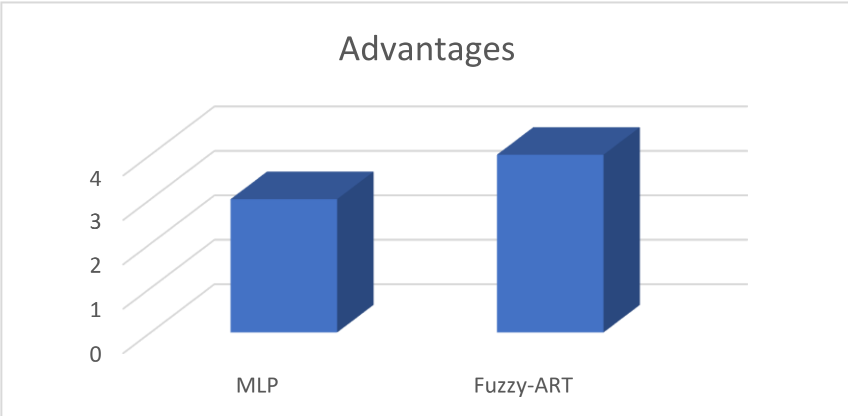

Figure 5: Histogram of application principles

But in the case of a multi-layered perceptron intended for disease prediction or diagnosis, this functionality is superfluous, because its efficiency is sufficient for the needs of the user, and it continues its learning in the process of application [34].

Comparisons confirm that the multilayer perceptron is the best choice for the development of neural networks in medicine.

6. Discussions

A single-layer perceptron is unable to separate the data sets linearly due to a number of limitations that can be overcome by adding additional layers.

A multilayer perceptron is a nonlinear system that allows for better classification and prediction than usual linear methods. It works with a large number of parameters, the impact of which on the expected results is difficult to assess, but it can find patterns when learning from examples for further decision-making. A special feature of its application is the ability to separate linearly inseparable data.

The choice of building this neural network is appropriate because its use in the medical field is already known for its merits, since its accuracy is due to the fact that it uses an inverse error propagation algorithm that minimizes the root mean square error when training: knowing the correct result, you can calculate the difference to understand what caused this value. Observations confirm the effectiveness of its work [35, 36].

For example, today there is a problem of long queues in cardiac surgery departments due to the small number of intensive care units, and the idea of increasing them requires a lot of money for intensive care, so the only way out of this situation is to use services efficiently and optimize the process.

Let's say that the condition of the operated patients is very serious, requiring additional observation and care, so they remain in the wards and at this time surgeons cannot accept new patients and operate on them, so it is possible to perform such operations at the beginning of the week for those who will be cared for for no more than two days, and before weekends or holidays, when they do not work and patients can rest. But now it is unclear which doctors need to stay after the operation and which do not.

Jack Tu and Michael Guerrier, colleagues at St. Michael's University Hospital in Toronto, decided to use neural networks to make such a prediction. The data were the preoperative indicators of patients: gender, age, diseases, and the complexity of the upcoming surgery. They trained a two-layer perceptron to classify patients into three categories: 16% of low-risk patients had been in their rooms for more than two days, 24% were expected to be in the hospital, and more than 60% of high-risk patients had no confirmed adverse scenario.

Consider a study conducted in Sarawak, Malaysia, a city known for the prevalence of hypertension. A total of 2461 data records were collected, of which 741 were sick (30.1%) and 69.9% were healthy. Five parameters were evaluated: weight, gender, age, ethnicity, and weight-to-height ratio. The developed multilayer perceptron had the following indicators according to neural network metrics

- sensitivity 0.41;
- specificity 0.91;
- F-score 0.50;
- accuracy 0.76.

When the prevalence of hypertension in the population was 30%, using Bayes' theorem, the probability that adolescents had hypertension was 66.2%. However, when the prevalence of the disease increased to 50%, the probability value was 82%. After a while, when hypertension in the city spread to only 10% of people, the model predicted only a 33.6% probability.

After increasing the sensitivity of the neural model to 65% and 90%, while maintaining the specificity value, the probability of developing hypertension in adolescents according to Bayes' theorem was 75.7% and 81.2%. The study confirms that socio-demographic data can be useful for detecting hypertension in children when using the developed multilayer perceptron.

Another study shows how an effective arrhythmia classification algorithm is used for different heart rate variability (HRV) signals. This method is based on a classifier built by a multilayer perceptron and reduced discriminant analysis (GDA) values. Nine linear and nonlinear signals are extracted from the HRV signals, of which GDA uses only three. This proposed method is applied to the input signals obtained from the MIT-BIH databases. The neural network helps to distinguish between such types of cardiac arrhythmias as first-degree heart block, ventricular activation, left bundle branch block, and supraventricular tachyarrhythmia with an accuracy of 95%, which is approximately the same as the value of the reliability of the results of this model from Table 1 [37].

As you can see, the contribution of this expert system to medicine is very important, but it is worth noting that it still has drawbacks. If the number of elements in the hidden layers is fixed, the perceptual system loses its ability to find solutions to complex problems using simple methods, so it is necessary that the number of elements in the hidden layers grows exponentially, which complicates the algorithm and increases its running time, so the perceptual system loses its ability to find solutions to complex problems using simple methods.

However, the results of the studies under consideration emphasize the peculiarity of this neural network – its important contribution to medicine, which refutes the work of specialists, but most importantly, the final decision is theirs, because the algorithm's task is to search for patterns among all patient indicators to predict a possible disease or confirm their assessment.

7. Conclusions

Neural networks are the key to new discoveries in our time, mainly in medicine, where they are used to diagnose and predict the occurrence of possible diseases.

To date, many observations have been made on the eternal problem of mankind – to invent triggers for the development of diseases, for which various neural networks are used, which was also discussed in this paper.

The article deals with the problem of choosing a neural network for predicting the occurrence and development of diseases. Journals and electronic sources that reflect the use of machine learning in solving problems in the medical field to reduce the likelihood of disease occurrence and process large amounts of medical information were reviewed [38–40]. The paper reflects the potential application of neural networks in medicine and solves the problem of multi-criteria selection of the required model, which should be accurate in prediction, fast for training and not take up much memory space, using a linear additive convolution algorithm that considered five neural networks:

- Kohonen neural network;
- multilayer perceptron (MLP);
- hybrid neural network;
- Adaptive resonance theory networks (APT-2);
- Fuzzy-ART.

In the course of the analysis, the alternatives were investigated according to the following criteria:

- learning speed;
- reliability of control;
- memory capacity;
- training;
- principle of application.

The study determined that the best model for this task is the multilayer perceptron (MLP), which stands out among other similar systems for its speed, use of a small amount of memory, teacher-assisted learning, and its ability to separate linearly inseparable data planes. It is not capable of dynamically increasing the knowledge base, which continues to learn even during use.

It has been established that the basis for the use of neural networks in medicine was often the construction of a multilayer perceptron, as described earlier. Also, for example, the variable accuracy of a neural network for diagnosing cardiovascular diseases ranged from 64 to 94%. These were models of a multilayer perceptron with two hidden layers with an accuracy of more than 90%, which were trained using genetic algorithms [41].

The main advantages of neural networks in medicine are

- the ability to search for relationships in very complex situations when it seems impossible or difficult to notice when assessing the situation;
- due to their ability to learn, they are able to find solutions to problems even in the absence of a priori knowledge of the input data, the development of the phenomenon under study, dependence on parameters, input data and expected results;
- the accuracy of forecasts does not depend on the availability of different types of less informative or omitted data [42, 43].

However, despite the usefulness of neural networks, they have a number of disadvantages: training takes some time, sometimes the neural network has to go through retraining stages during repeated use, and the larger the input value, the more time is required, so it is suggested that in such cases, computations should be performed in parallel; sometimes implementation requires the use of appropriate software.

Perceptron is a neural network capable of calculating the likelihood of developing diseases, but the final decision should always be made by a specialist in the relevant department at the hospital.

References

[1] Conferences of the State University "Zhytomyr Polytechnic". URL: https://conf.ztu.edu.ua/wp-content/uploads/2019/12/5.pdf.

[2] The use of neural networks is a promising area of science and society | Scientific conferences. December / December 2013 | Scientific conferences. URL: http://oldconf.neasmo.org.ua/node/139.

[3] S. Kaushik, A. Choudhury, S. Natarajan, L. A. Pickett and V. Dutt, "Medicine Expenditure Prediction via a Variance- Based Generative Adversarial Network," in IEEE Access, vol. 8, pp. 110947-110958, 2020, doi: 10.1109/ACCESS.2020.3002346.

[4] J. Jaruenpunyasak and R. Duangsoithong, "Empirical Analysis of Feature Reduction in Deep Learning and Conventional Methods for Foot Image Classification," in IEEE Access, vol. 9, pp. 53133-53145, 2021, doi: 10.1109/ACCESS.2021.3069625.
[5] Ageing, Apr. 2020, [online] Available: https://www.un.org/en/global-issues/ageing/.
[6] Osteoarthritis, Apr. 2020, [online] Available: https://www.who.int/medicines/areas/priority_medicines/Ch6_12Osteo.pdf.
[7] [N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection", Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR), pp. 886-893, Jun. 2005.
[8] C. Kalyoncu and, "GTCLC: Leaf classification method using multiple descriptors", IET Comput. Vis., vol. 10, no. 7, pp. 700-708, Jun. 2016.
[9] Y. Qi and G. Zhang, "Strategy of active learning support vector machine for image retrieval", IET Comput. Vis., vol. 10, no. 1, pp. 87-94, Feb. 2016.
[10] T. George, S. P. Potty and S. Jose, "Smile detection from still images using KNN algorithm", Proc. Int. Conf. Control Instrum. Commun. Comput. Technol. (ICCICCT), pp. 461-465, Jul. 2014.
[11] G. Khan, A. Siddiqi, G. Khan, S. Q. Wahla and S. Samyan, "Geometric positions and optical flow based emotion detection using MLP and reduced dimensions", IET Image Process., vol. 13, no. 4, pp. 634-643, Mar. 2019.
[12] S. Verma, M. A. Razzaque, U. Sangtongdee, C. Arpnikanondt, B. Tassaneetrithep and A. Hossain, "Digital Diagnosis of Hand, Foot, and Mouth Disease Using Hybrid Deep Neural Networks," in IEEE Access, vol. 9, pp. 143481-143494, 2021, doi: 10.1109/ACCESS.2021.3120199.
[13] N. Alamdari, K. Tavakolian, M. Alhashim and R. Fazel-Rezai, "Detection and classification of acne lesions in acne patients: A mobile application", Proc. IEEE Int. Conf. Electro Inf. Technol. (EIT), pp. 0739-0743, May 2016.
[14] S. Abdul-Rahman, A. Khairil Norhan, M. Yusoff, A. Mohamed and S. Mutalib, "Dermatology diagnosis with feature selection methods and artificial neural network", Proc. IEEE-EMBS Conf. Biomed. Eng. Sci., pp. 371-376, Dec. 2012.
[15] T. A. Rimi, N. Sultana and M. F. A. Foysal, "Derm-NN: Skin diseases detection using convolutional neural network", Proc. 4th Int. Conf. Intell. Comput. Control Syst. (ICICCS), pp. 1205-1209, May 2020.
[16] M. F. Aryan, W. Krathu, C. Arpnikanondt and B. Tassaneetrithep, "Image recognition for detecting hand foot and mouth disease", Proc. 11th Int. Conf. Adv. Inf. Technol., pp. 1-11, Jul. 2020.
[17] When Should I Suspect Hand Foot and Mouth Disease, Aug. 2020, [online] Available: https://cks.nice.org.uk/topics/hand-foot-mouth-disease/diagnosis/when-t%o-suspect-hand-foot-mouth-disease/.
[18] Y. Liu, A. Jain, C. Eng, D. H. Way, K. Lee, P. Bui, et al., "A deep learning system for differential diagnosis of skin diseases", Nature Med., vol. 26, no. 6, pp. 900-908, 2020.
[19] A. Taravat, S. Proud, S. Peronaci, F. D. Frate and N. Oppelt, "Multilayer perceptron neural networks model for meteosat second generation SEVIRI daytime cloud masking", Remote Sens., vol. 7, no. 2, pp. 1529-1539, 2015.
[20] G. -M. Lin and H. -C. Zeng, "Electrocardiographic Machine Learning to Predict Mitral Valve Prolapse in Young Adults," in IEEE Access, vol. 9, pp. 103132-103140, 2021, doi: 10.1109/ACCESS.2021.3098039.
[21] L. A. Freed, D. Levy, R. A. Levine, M. G. Larson, J. C. Evans, D. L. Fuller, et al., "Prevalence and clinical outcome of mitral-valve prolapse", New England J. Med., vol. 341, no. 1, pp. 1-7, Jul. 1999.
[22] R. B. Devereux, E. C. Jones, M. J. Roman, B. V. Howard, R. R. Fabsitz, J. E. Liu, et al., "Prevalence and correlates of mitral valve prolapse in a population-based sample of American Indians: The strong heart study", Amer. J. Med., vol. 111, no. 9, pp. 679-685, Dec. 2001.
[23] J. M. Flack, J. H. Kvasnicka, J. M. Gardin, S. S. Gidding, T. A. Manolio and D. R. Jacobs, "Anthropometric and physiologic correlates of mitral valve prolapse in a biethnic cohort of young adults: The CARDIA study", Amer. Heart J., vol. 138, no. 3, pp. 486-492, Sep. 1999.
[24] R. B. Devereux, I. Hawkins, R. Kramer-Fox, E. M. Lutas, I. W. Hammond, M. C. Spitzer, et al., "Complications of mitral valve prolapse. Disproportionate occurrence in men and older patients", Am. J. Med., vol. 81, no. 5, pp. 751-758, Nov. 1986.

[25] P.-Y. Liu, K.-Z. Tsai, Y.-P. Lin, C.-S. Lin, H.-C. Zeng, E. Takimoto, et al., "Prevalence and characteristics of mitral valve prolapse in military young adults in Taiwan of the CHIEF heart study", Sci. Rep., vol. 11, no. 1, pp. 2719, Feb. 2021.
[26] R. A. Levine et al., "Mitral valve disease–morphology and mechanisms", Nature Rev. Cardiol., vol. 12, no. 12, pp. 689-710, Dec. 2015.
[27] C. V. Leier, T. D. Call, P. K. Fulkerson and C. F. Wooley, "The spectrum of cardiac defects in the Ehlers-Danlos syndrome types I and III", Ann. Int. Med., vol. 92, pp. 171-178, Feb. 1980.
[28] K. Hirata, F. Triposkiadis, E. Sparks, J. Bowen, H. Boudoulas and C. F. Wooley, "The Marfan syndrome: Cardiovascular physical findings and diagnostic correlates", Amer. Heart J., vol. 123, no. 3, pp. 743-752, Mar. 1992.
[29] G. H. Tison, J. Zhang, F. N. Delling and R. C. Deo, "Automated and interpretable patient ECG profiles for disease detection tracking and discovery", Circulation Cardiovascular Qual. Outcomes, vol. 12, no. 9, Sep. 2019.
[30] ElAr :: Holovna. URL: https://openarchive.nure.ua/server/api/core/bitstreams/5a6f6f03-7222-4b95-8119-83137430b1ef/content (date of appeal: 02.03.2023).
[31] CORE – Aggregating the world’s open access research papers. URL: https://core.ac.uk/download/pdf/322991189.pdf (date of the application: 14.12.2022).
[32] S. Trenn, "Multilayer perceptrons: Approximation order and necessary number of hidden units", IEEE Trans. Neural Netw., vol. 19, no. 5, pp. 836-844, May 2008.
[33] E. Romero and J. M. Sopena, "Performing feature selection with multilayer perceptrons", IEEE Trans. Neural Netw., vol. 19, no. 3, pp. 431-441, Mar. 2008.
[34] H. Seo and D. -H. Cho, "Cancer-Related Gene Signature Selection Based on Boosted Regression for Multilayer Perceptron," in IEEE Access, vol. 8, pp. 64992-65004, 2020, doi: 10.1109/ACCESS.2020.2985414.
[35] Al.Hak, L. A. (2022). Diabetes Prediction Using Binary Grey Wolf Optimization and Decision Tree. International Journal of Computing, 21(4), 489-494. https://doi.org/10.47839/ijc.21.4.2785.
[36] Mienye, I. D., Sun, Y., & Wang, Z. (2020). IMPROVED PREDICTIVE SPARSE DECOMPOSITION METHOD WITH DENSENET FOR PREDICTION OF LUNG CANCER. International Journal of Computing, 19(4), 533-541. https://doi.org/10.47839/ijc.19.4.1986
[37] Qeethara Kadhim Al-Shayea and Itedal S.H. Bahia, “Urinarysystem Dieseases Diagnosis Using Artificial neural networks”,IJCSNS, Vol.10, No.7, July 2010.
[38] Sharonova, N., Kyrychenko, I., Gruzdo, I., Tereshchenko, G., “Generalized Semantic Analysis Algorithm of Natural Language Texts for Various Functional Style Types”, 2022 6th International Conference on Computational Linguistics and Intelligent Systems (COLINS-2022), 2022. – CEUR-WS 3171, 2022, ISSN 16130073. – Volume I: Main, PP. 16-26.
[39] Sharonova, N., Kyrychenko, I., Tereshchenko, G., “Application of big data methods in E-learning systems”, 2021 5th International Conference on Computational Linguistics and Intelligent Systems (COLINS-2021), 2021. – CEUR-WS, 2021, ISSN 16130073. – Volume 2870, PP. 1302-1311.
[40] Gruzdo, I., Kyrychenko, I., Tereshchenko, G., Shanidze, N. Metrics applicable for evaluating software at the design stage. 2021 5th International Conference on Computational Linguistics and Intelligent Systems (COLINS-2021), 2021. – CEUR-WS, 2021, ISSN 16130073. – Volume 2870, PP. 916-936.
[41] K. Smelyakov, A. Chupryna, D. Sandrkin and M. Kolisnyk, "Search by Image Engine for Big Data Warehouse," 2020 IEEE Open Conference of Electrical, Electronic and Information Sciences (eStream), Vilnius, Lithuania, 2020, pp. 1-4, doi: 10.1109/eStream50540.2020.9108782.
[42] K. Smelyakov, A. Datsenko, V. Skrypka and A. Akhundov, "The Efficiency of Images Reduction Algorithms with Small-Sized and Linear Details," 2019 IEEE International Scientific-Practical Conference Problems of Infocommunications, Science and Technology (PIC S&T), 2019, pp. 745-750, doi: 10.1109/PICST47496.2019.9061250.
[43] Danilenko V.D., Nazarov O.S. Research of forecasting methods for the needs of an automobile enterprise to ensure material reserves. "Modern trends in the development of information and communication technologies and management tools": Twelfth International Scientific and Technical Conference, April 27. 2022.