

UDC 004.032.26

## 8. THE ROLE OF NEURAL NETWORKS IN DATA MINING

*Kupreichyk A.S., Master's degree student, group 376701*

*Belarusian State University of Informatics and Radioelectronics, Minsk, Republic of Belarus*

*Liakh Y.V. – Senior Lecturer*

**Annotation.** This article describes the role of neural networks in data processing and analysis, their advantages and limitations. The trends in the development of neural networks in this field and their impact on the future of data analysis are presented.

**Keywords.** Neural networks, data processing, data analysis, function approximation, backpropagation, classification, regression, clustering.

Neural networks are a powerful tool in the field of data processing and analysis that has gained wide recognition in recent decades. Neural networks are models based on the functionality of brain neurons that can process and analyse complex data, recognise patterns, and learn from experience.

The role of neural networks in data processing and analysis is to automatically detect and extract information from large volumes of data. They are used for various tasks such as classification, regression, clustering, and forecasting [1].

One of the main advantages of neural networks is their ability to learn from existing data. They can recognise complex hierarchical structures and uncover hidden patterns, enabling them to tackle tasks that were previously considered difficult or impossible to automate. Neural networks also provide flexibility and adaptability in working with different types of data. They can be applied to text, audio and video files, as well as numerical and categorical data. Due to this flexibility, neural networks have found applications in many industries, including healthcare, finance, marketing, and manufacturing.

However, it is important to note that neural networks also have their limitations. Their computational requirements can be resource-intensive, especially when using deep neural networks. Additionally, training and interpreting the results of neural networks can be complex and require specific expertise.

Neural networks are composed of interconnected neurons that work together to process and analyse data. They have their own architecture and components that perform specific tasks. The fundamental principles of neural networks will be analyzed below, including their architecture, function approximation, and the learning process using back propagation [2].

1. Architecture of neural networks. Neural networks consist of layers of neurons, which are the main building blocks of the network. They are classified into input, hidden, and output layers. The input layer receives data into the network, while the output layer sends the processing results. Hidden layers reside between the input and output layers and perform computations for data processing.

2. Function approximation principle. The fundamental principle of neural networks lies in their ability to approximate complex functions. Each neuron in the network takes its input data, applies weights and an activation function to them, and then passes the resulting output to the next neurons. This allows the neural network to build nonlinear models and process complex data.

3. Backpropagation. The learning process of neural networks is achieved through backpropagation. Initially, the network is initialised with random weights. Then, data is fed into the input of the network, and the results are compared with the expected values. The error, or the difference between the actual and expected values, is calculated and propagated back through the network. The neuron weights are updated using gradient descent to minimise the error [1].

Backpropagation enables the neural network to adjust its weights and tune its parameters to achieve better performance and accuracy in data analysis. This process takes place over multiple training iterations until the network reaches a certain level of accuracy or obtains satisfactory quality of results.

Thus, the fundamental principles of neural networks include their architecture, function approximation, and the learning process using backpropagation. These principles allow neural networks to process complex data, tune their parameters, and achieve high performance in data analysis.

Neural networks are one of the most popular and effective methods for data analysis, thanks to their ability to uncover complex patterns and process large volumes of information. Consider the various tasks in which neural networks are used to analyze data. The main objective is to describe the essential methods for data representation and preprocessing, while also providing examples of successful usage in different domains [3].

### 1. Different Data Analysis Tasks:

– classification. Neural networks can be used to classify objects into different categories. For instance, they can classify emails as spam or non-spam, recognise images for automated quality control, or identify types of cancer cells in medical images;

- regression. Neural networks can predict numerical values based on existing data. For example, they can predict real estate prices based on house characteristics or forecast sales based on historical data;
- clustering. Neural networks can be used to group objects based on their similarity. For example, they can identify similar consumer groups based on their shopping behavior or cluster patients based on their medical characteristics;
- forecasting. Neural networks can forecast future values based on historical data. For example, they can predict time series for sales forecasting, weather forecasting, or changes in financial markets.

#### 2. Data Representation and Preprocessing:

before using a neural network, data requires specific representation and preprocessing. For example, numerical data may be normalised or standardised, categorical data may be transformed into binary features, and text data may be vectorised or transformed into embeddings. Additionally, data may be split into training, validation, and testing sets to evaluate model performance.

#### 3. Examples of Successful Neural Network Applications:

- medicine. Neural networks are used to analyse medical data, such as scanning images to detect cancer cells or recognise rare diseases;
- finance. Neural networks are employed for analysing financial data, including market trend forecasting, risk assessment, and detecting fraudulent transactions;
- marketing. Neural networks are applied to consumer data analysis like social media, purchasing behavior, and preferences for personalised marketing campaigns and demand prediction;
- industry. Neural networks are used for analysing data from manufacturing processes, monitoring and preventing equipment failures, optimising production processes, and enhancing product quality.

The application of neural networks in data analysis covers a variety of tasks and fields, with successful examples found in many sectors including medicine, finance, marketing, and industry. This illustrates the significance of neural networks in data processing and analysis, showcasing their ability to help uncover hidden patterns and make important decisions [4].

Neural networks serve as a powerful tool for data analysis; however, they come with their advantages and limitations.

#### 1. Advantages of Using Neural Networks:

- ability to handle complex data: neural networks can effectively process large volumes of information and complex data types like text, sound, images, and video;
- detection of hidden patterns: leveraging their capability to capture intricate data relationships, neural networks can uncover hidden patterns and subtle dependencies that may be missed by other methods;
- adaptability and trainability: neural networks can learn from existing data and adapt to data changes, allowing them to enhance performance over time;
- flexibility in solving various tasks: neural networks can be applied to diverse tasks such as classification, regression, clustering, and forecasting. They can be customised for specific requirements and objectives.

#### 2. Limitations of Neural Networks and Special Application Conditions:

- high computational resource requirements: training and utilising complex neural networks can demand significant computational resources, including high-performance processors and graphics accelerators;
- need for large amounts of data: neural networks deliver good results only with sufficient training data. In some fields, limited data volumes can hinder the use of neural networks;
- risk of overfitting: neural networks can overfit when they excessively fit the training data and struggle to generalise to new data, resulting in low model generalisation capability.

#### 3. Issues with Result Interpretation and Model Explainability:

- interpreting deep neural networks: neural networks, especially deep models, can be challenging to interpret in terms of results. sometimes they act like "black boxes", making it difficult to explain or decode their decisions and outputs;
- processing large data volumes: employing neural networks for analysing large data sets can be labor-intensive and requires scalability and computational optimisation.

In essence, neural networks are potent tools for data analysis; however, their application may be constrained by computational resource demands, data volume prerequisites, and the need for result explainability. Considering these factors, it is crucial to carefully assess the applicability of neural networks in each specific scenario and select an appropriate data analysis method based on the task at hand and available resources [5].

#### 1. Forecasting the Future Development of Neural Networks in Data Analysis:

- Increased Efficiency and Accuracy: With advancements in computational technology and learning algorithms, neural networks will become more efficient and accurate in data analysis. This will enable them to tackle more complex tasks and achieve higher quality results.

– Integration with Other Data Analysis Methods: Neural networks will increasingly be used in combination with other data analysis methods, such as statistical methods or cluster analysis, to gain a more comprehensive and accurate understanding of data.

## 2. New Methods of Training Neural Networks:

– Deep Learning: Deep learning is becoming increasingly popular in data analysis, allowing neural networks to process and analyse highly complex data and achieve outstanding results.

– Convolutional Neural Networks: Convolutional neural networks are particularly useful in image and video analysis, as well as natural language processing, opening up new possibilities in areas such as computer vision and speech recognition.

## 3. Potential Future Applications of Neural Networks:

– autonomous transport systems: neural networks will play a crucial role in the development and safety of autonomous transport systems, enabling them to process and analyse data from sensors for making quick and accurate decisions.

– medical diagnostics: neural networks can be applied for automatic diagnosis of various diseases and detection of pathologies based on medical images and data, improving the accuracy and efficiency of diagnostic procedures.

– social media analytics: neural networks will play a pivotal role in the analysis of social media data, empowering companies to gain insights into user sentiments, trends, and preferences. This understanding enables the enhancement of marketing campaigns and enables the delivery of personalised services.

Developments in neural networks and the use of large volumes of data raise important questions related to ethics, data confidentiality, and security. Neural networks, particularly when used in the field of data analysis, can have potential ethical and social implications that need to be considered and analysed.

One of the main issues that can arise when using neural networks is algorithmic bias. This refers to situations where a neural network model makes incorrect or biased conclusions based on insufficiently diverse data or the presence of biases in the input data. Algorithmic bias can lead to adverse or unfair outcomes, such as discrimination in employment decisions, loan applications, or judicial proceedings.

Data confidentiality is also a critical aspect when working with neural networks. Often, large datasets are required for training models, which may contain personal information about individuals. It is important to ensure effective mechanisms for data protection to prevent unauthorized access or leakage of confidential information [6].

Data security is also a concern. Neural networks can be targeted by attacks aimed at altering model outputs or injecting malicious code. Ensuring the security of neural networks, including defense against attacks and maintaining data integrity, becomes increasingly important.

In addition to algorithmic bias, data confidentiality and security, ethical considerations related to transparency and explainability of models are crucial. Neural networks, particularly deep models, can be difficult to interpret, leading to challenges in explaining the reasoning behind model decisions. This can raise concerns about trust in the results or contradictions regarding the decisions made by the model. It is important to continue research in the area of model explainability and develop methods that allow for clear understanding of causal relationships and model logic.

This article has explored the role of neural networks in data processing and analysis. They are powerful tools capable of handling and analysing complex data, uncovering hidden patterns, and predicting future values. The fundamental principles of neural networks, their architecture and components have been discussed, explaining their role as function approximators, and describing the training process using backpropagation of error.

The application of neural networks in data analysis spans diverse tasks, including classification, regression, clustering, and forecasting. The methods of data representation and preprocessing before using neural networks have been studied, along with examples of successful applications in various fields like medicine, finance, marketing, and industry. The future of neural networks will focus on improving model efficiency, accuracy, interpretability, and their application in new fields and tasks. Neural networks remain a vital tool for data analysis, evolving to meet the demands of the modern information society.

### References:

1. Gevorgyan, M. N. Pseudo-random number generator based on neural network / V. N. Gevorgyan, A. V. Demidova, A. V. Korolkova, D. S. Kulyabov, L. A. Sevastianov, I. M. Gostev, – GRID 2018, p. 568-572, Dubna, Moscow region, Russia, September 10 - 14, 2018.
2. Shymkovych, V. Hardware implementation of radial-basis neural networks with Gaussian activation functions on FPGA / V. Shymkovych, S. Telenyk, P. Kravets — *Neural Computing and Applications*, 2021. — 9467–9479 pp.
3. Sahin, I. Design and Implementation of Neural Networks Neurons with RadBas, LogSig, and TanSig Activation Functions on FPGA / I. Sahin, I. Koyuncu. — 120 p.
4. Machine learning, explained [Electronic resource]. – Mode of Access: <https://mitsloan.mit.edu/ideas-made-to-matter/machine-learning-explained>. – Date of Access: 03.03.24
5. Getting Started with Neural Networks in Python [Electronic resource]. – Mode of Access: <https://ai.plainenglish.io/getting-started-with-neural-networks-in-python-b4efc76ef91>. – Date of Access: 02.03.24.
6. Lezhnev, E. V. Development of the concept of hardware implementation of a neural network based on FPGA / E. V. Lezhnev, A. A. Amerikanov, 2017. — 214 – 215 pp