UDK 004.032.26; 657.6
DOI: 10.31891/CSIT-2021-4-3

TETIANA NESKORODIEVA, EUGENE FEDOROV,
ANASTASIIA NESKORODIEVA, TETIANA SICHKO, PAVLO RYMAR
Vasyl' Stus Donetsk National University

NEURAL NETWORK DETECTION METHOD OF DATA ANOMALIES OF WASTE-FREE PRODUCTION AUDIT

The paper presents a method for the detection of anomalies in waste-free production audit data based on the neural network model of Gauss-Bernoulli of the forward only restricted Cauchy machine (FORCM). The purpose of the work is to increase the efficiency of audit data analysis of waste-free production on the basis of the neural network model of anomalies detection without the use of the marked data that simplifies audit.

To achieve this goal, the following tasks have been set and solved: offered model of generalized multiple transformations of audit data in the form of a two-layer neural network. Cauchy offered neural network model of Gauss-Bernoulli of the forward only restricted Cauchy machine possesses a heteroassociative memory; works real data; has no restrictions for storage capacity; provide high accuracy of detection of anomalies; uses Cauchy's distribution that increases the speed of convergence of a method of parametrical identification. To increase the speed of Gauss-Bernoulli parametric identification of a forward only restricted Cauchy machine, a parametric identification algorithm was developed to be implemented on a GPU using CUDA technology. The offered algorithm allows increasing training speed by approximately proportional to the product of numbers of neurons in the hidden layer and power of a training set.

The made experiments confirmed the operability of the developed software and allow to recommend it for use in practice in a subsystem of the automated analysis of DSS of audit for detection of anomalies.

Keywords: audit, mapping by neural network, neural network model of Gauss-Bernoulli, forward only restricted Cauchy machine, detection of anomalies, audit of waste-free production.

ТЕТЯНА НЕСКОРОДЄВА, ЄВГЕН ФЕДОРОВ,
АНАСТАСІЯ НЕСКОРОДЄВА, ТЕТЯНА СІЧКО, ПАВЛО РИМАР
Донецький національній університет імені Василя Стуса, Вінниця, Україна

МЕТОД НЕЙРОМЕРЕЖЕВОГО ВИЯВЛЕННЯ АНОМАЛІЙ ДАНИХ АУДИТУ БЕЗВІДХОДНОГО ВИРОБНИЦТВА

В роботі представлено метод нейромережевого виявлення аномалій даних аудиту безвідходного виробництва по моделі Гауса- Бернуллі односпрямованої обмеженої машини Коші (FORCM). Метою роботи є підвищення ефективності аналізу даних аудиту безвідходного виробництва на основі нейромережевої моделі виявлення аномалій без використання розмічених даних, що спрощує аудит. Для досягнення цієї мети були поставлені і вирішені наступні завдання: запропонована модель узагальнених множинних перетворень даних аудиту у вигляді дворівневої нейронної мережі та модель виявлення аномалій у вигляді нейромережевої моделі Гауса-Бернуллі односпрямованої обмеженої машини Коші; обраний критерій оцінки ефективності нейромережевої моделі виявлення аномалій; запропонований метод параметричної ідентифікації нейромережевої моделі виявлення аномалій; проведені чисельні дослідження.

Запропонована нейромережева модель Гауса-Бернуллі односпрямованої обмеженої машини Коші має гетероасоціативну пам'ять; працює з дійсними даними; не має обмежень за обсягом зберігання; забезпечує високу точність виявлення аномалій; використовує розподіл Коші, що збільшує швидкість збіжності методу параметричної ідентифікації.

Для збільшення швидкості параметричної ідентифікації моделі Гауса-Бернуллі односпрямованої обмеженої машини Коші був розроблений алгоритм параметричної ідентифікації, який реалізований на графічному процесорі з використанням технології CUDA. Запропонований алгоритм дозволяє збільшити швидкість навчання приблизно пропорційно добутку кількості нейронів в прихованому шарі і потужності навчальної вибірки.

Проведені експерименти підтвердили працездатність розробленого програмного забезпечення і дозволяють рекомендувати його для використання на практиці в підсистемі автоматизованого аналізу DSS аудиту для виявлення аномалій. Перспективи подальших досліджень полягають у перевірці запропонованих методів на більш широкому наборі тестових баз даних.

Таким чином, актуальне завдання підвищення ефективності виявлення аномалій в даних аудиту безвідходного виробництва була вирішена за допомогою нейромережевої моделі Гауса-Бернуллі односпрямованої обмеженої машини Коші.

Ключові слова: аудит, відображення за допомогою нейронної мережі, модель нейронної мережі Гауса-Бернуллі, односпрямована обмежена машина Коші, виявлення аномалій, аудит безвідходного виробництва.

Introduction

Nowadays the scientific and technical issue of the modern information technologies in financial and economic sphere is creation methodology forming of the decision support systems (DSS) at the enterprises audit in the conditions of IT application on enterprises and with the use of information technologies. Modern automated DSS audit are based on the automated analysis of the large volumes of data about financial and economic activity and states of enterprises with the multi-level hierarchical structure of heterogeneous, multivariable, multifunction connections, intercommunications and cooperation of objects of audit. The tasks automated DSS audit are expansion of functional possibilities, increase of efficiency and universality of IT-audit [1].

Related works

Currently, the analytical procedures used during the audit are based on data mining techniques [2, 3]. Automated DSS audit means the automatic forming of recommendable decisions, based on the results of the automated analysis of data, that improves quality process of audit. Unlike the traditional approach, computer technologies of analysis of data in the system of audit accelerate and promote the process accuracy of audit, that extremely critical in the conditions of plenty of associate tasks on lower and middle levels, and also amounts of indexes and supervisions in every task.

The development of methods of estimation and prediction [4,5], formation of generalized associative relationships [6] are described in the works of the authors of this article. The goals of creating these methods: reducing the computational complexity for simple tasks (a single mapping of elements or sub-elements of the audit subject area), automatic structural identification, increasing the accuracy for complex tasks (compositions of mappings of elements or sub-elements of the audit subject area) and the possibility of applying these methods for the generalized analysis of elements and sub-elements of the audit subject area (Table 1).

Table 1

Comparative analysis of intelligent analysis methods in audit tasks

The economic content of the display	Model of processing elements of the subject area, Features of the model or method	Purpose of processing elements of the subject area	Advantages disadvantages of the model or method
Payment - delivery of raw materials	Modified Liquid State Machine, one-dimensional hidden layer, parameter identification based on matrix pseudoreversion [1]	Evaluation and prediction of indicators of raw material supplies (by type) based on the values of payment indicators in a direct check of the display	Reducing computational complexity, improving the forecast accuracy
Settlements with suppliers-customer settlements	A neural network model based on a gateway recurrent unit. For parametric identification of this model, adaptive cross entropy (a combination of random and directional search) is faster to learn but less accurate than in [1] because the pseudoreversal is not paralleled	Evaluation of indicators of settlements with customers on the basis of values of indicators of settlements with suppliers in a direct verification of mapping	Reducing computational complexity, improving the forecast accuracy
Settlements with suppliers - settlements with customers (a composition of mappings between a set of input and output data)	Forward-only counterpropagating neural network, which is a nonrecurrent static two-layer ANN [2], assumed that the audit indicators are noisy with Gaussian noise (learner model)	Construction of generalized associative relationships for generalized analysis tasks (in the forward direction)	Automating the formation of generalized features of audit sets and their mapping by means of a forward-only counterpropagating neural network the number of pairs (neurons in the hidden layer N1) is set manually
Release of raw materials - posting of finished products (a composition of mappings between a set of input and output data)	Bidirectional counterpropagating neural network, which is a nonrecurrent static two-layer ANN BCPNN (learner model)	Construction of generalized associative relationships for generalized analysis tasks (in the forward and reverse direction)	Automating the formation of generalized features of audit sets and their mapping by means of a bidirectional counterpropagating neural network the number of pairs (neurons in the hidden layer N1) is set manually

The choice of model in the audit DSS depends on:

1) characteristics of the audit data type (time series data, spatial data (as mappings));

2) audit level (upper middle, lower),

3) audit tasks (internal, external);

4) the type of analysis tasks (detection of anomalies, structural analysis, assessment of indicators);

5) the characteristics of the enterprise (large, medium, small) and the type of activity (industry) at the top level;

6) characteristics of sets and subsets of operations at lower levels (numerological, quantitative, semantic, logical).

This choice is schematically formalized in the form of a binary decision tree for choosing a neural network data audit model (Fig.1).

At the first level, the choice of a model is carried out depending on data type. If map data is analyzed, therefore ANN with associative memory is used, otherwise ANN for forecasting.

When choosing models based on associative memory at the next stage, the choice depends on the type of production: with or without waste. If the production is waste-free, depending on the size of the enterprise and the specified accuracy of the decision maker, a model is selected that is the best in terms of the ratio of learning rate and accuracy. At the next levels, the choice depends on the type of analysis. In the case of structural analysis, models are selected in which the layers correspond to the stages of data transformation, in particular, production data. Also, the choice depends on the direction of analysis: direct or direct and reverse.

The proposed logical-neural network method makes it possible to automate the process of data analysis in the audit DSS and optimize it depending on the characteristics of the audit process and the audit object. One of the main tasks of data analysis of the audit subject area is the identification of anomalies. Let's consider the existing types of anomalies and methods of their operation.

Types of anomalies [7-9]:

- point (are provided by points in character space);
- contextual (usually a point of a time series or the rarefied data which depends on the environment);
- collective (the section of a time series or the rarefied data).

Methods of detection of anomalies [7-9]:

1. Approach on the basis of rules (logical approach):

- methods on the basis of associative rules with classification and without classification (for example, the Apriori method);
- methods on the basis of a decision tree with classification (for example, a method of the isolated wood).

2. Approach on the basis of ANN:

- ANN without classification (for example, the one-class machine of reference vectors, ANN an associative memory (for example, the autoencoder, the self-organizing card of signs, a neural network of Hopfield, Boltzmann's), ANN of the forecast of a time series (for example, NARNN (nonlinear autoregressive neural network), NARMANN (nonlinear autoregressive-moving average neural network), SRN (simple recurrent network), BRNN (bidirectional recurrent neural network), LSTM (long short-term memory), BiLSTM, GRU (gated recurrent unit), BiGRU));
- ANN with classification (for example, MLP, RBFNN).

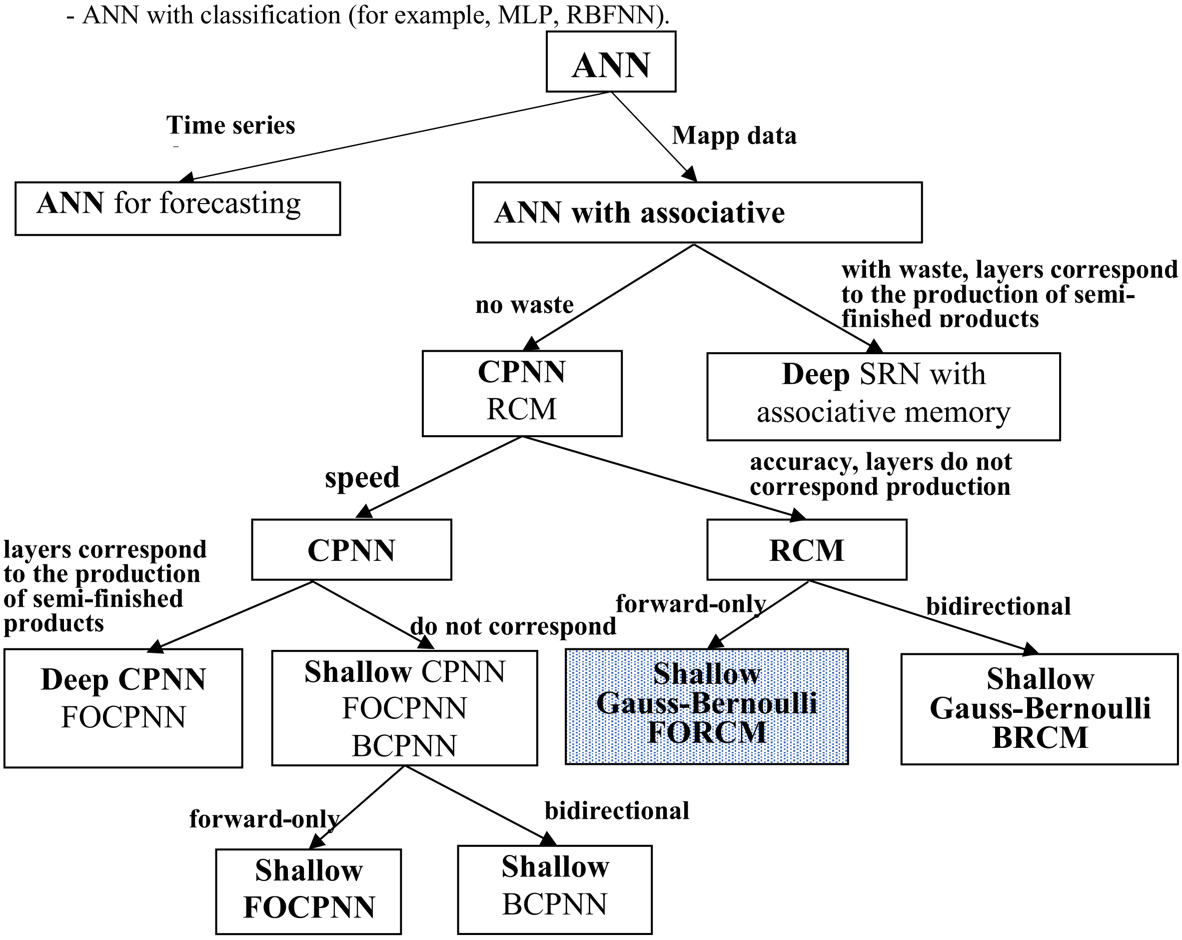

Fig.1 Binary decision tree of neural network model selection for data analysis

3. Approach on the basis of Bayes' networks with classification

4. Approach on the basis of a clustering:

- clustering on the basis of centroid (for example, a method of k-means) or distributions (for example, the EM method);
- clustering on the basis of medoid (for example, the PAM methods (partitioning around medoids), a subtractive clustering);
- density clustering (for example, DBSCAN methods (density-based spatial clustering of applications with noise), OPTICS (ordering points to identify the clustering structure)).

5. Approach on the basis of the neighborhood (metric approach) (for example, methods of the k-nearest neighbors, a local outlier factor (LOF))

6. Approaches on the basis of distributions:

6.1. Parametrical approach on a basis:

- Gaussian distributions (for example, method of the minimum covariance determinant (MCD));
- mixtures of distributions (for example, hidden Markov models (HMM), Gaussian mixture models (GMM)).

6.2. Nonparametric approach on a basis:
- histograms;
- functions of a kernel (for example, Parzen window method).

7. Approach on the basis of regression model (for example, the Box-Jenkins ANN method)

8. Approach on the basis of the spectral theory (matrixes decomposition) (for example, a method of the principal component analysis (PCA))

9. Approach on the basis of information theory (entropy).

In this work, neural networks were chosen to detect anomalies for the following reasons:
- the possibility of their training and adaptation;
- the ability to identify patterns in the data, their generalization, i.e. extracting knowledge from data, so knowledge about the object (for example, its mathematical model) is not required;
- parallel processing of information, which increases computing power.

Now the most popular is approach of detection of anomalies on the basis of neural networks.

Disadvantages of the one-class machine of reference vectors is restriction for quantity of reference vectors. Disadvantage of ANN of the forecast of a time series is that they require existence of a time series. The disadvantage of ANN with classification is the requirement to classify anomalies, which is not always possible due to the laboriousness of obtaining labeled data for each type of anomaly. Therefore, ANNs with associative memory were chosen in this work.

Traditional neural networks with an associative memory are:

1. Neural networks only with heteroassociative memory (for example, the unidirectional neural network of counter distribution [11], a neural network of the principal component analysis [12], a neural networks of the analysis independent a component [13].

2. Neural networks only with an autoassociative memory (for example, the autoencoder [10], sigmoidal network of trust [14], the self-organizing feature map [15], a neural networks of Hopfield [16], Gauss [17], Hamming [18],

3. ANN with a heteroassociative and autoassociative memory (for example, a full (bidirectional) counterpropagation neural network [13], a bidirectional associative memory [19], Boltzmann [20]).

The majority of neural networks with an associative memory possess some or more shortcomings: do not possess a heteroassociative memory; do not work with real data; have no high capacity of an associative memory; have no high accuracy; have high computational complexity.

In this regard, creation of a neural networks which will allow to eliminate the specified disadvantages is actual problem.

The purpose of work is increase an efficiency of audit data analysis of waste-free production on the basis of neural network model of anomalies detection without use of the marked data that simplifies audit.

For achievement of the goal it is necessary to solve the following problems:
– offer model of generalized multiple transformation of audit data;
– offer neural network model of detection of anomalies;
– select criterion for evaluation of efficiency of neural network model of detection of anomalies;
– offer a method of parametrical identification of neural network model of detection of anomalies;
– perform numerical researches.

In this paper, the structure of the data transformation model is determined based on the production structure. It is assumed that the transformation of raw materials into finished products in one step without waste without intermediate products. Each type of raw material is used in the production of one or more types of finished products. The production structure for each planning period (month, quarter, year) is determined on the basis of long-term contracts and short-term (in particular urgent) orders. The production plan is decomposed into quantization periods of the planning period, taking into account the production capacity for different types of products.

In this case, the transformation of these raw materials into finished products for the planning period can be represented in the form of a two-layer neural network. The number of neurons in the input layer is equal to the number of raw materials used in production. The number of neurons in the output layer is equal to the number of types of finished products. The input values are the amount of raw materials by type, the output of the network is the finished product values for the planning period or the quantization period. To train the neural network, the "correct" data are used (the formation of which has been verified). Data that are subject to verification are used as control data.

The formal statement of the learning problem of the neural network is formulated as follows. Let for model of anomalies detection the training set be set $S=\{(\mathbf{x}_m^{in},\mathbf{d}_m^{out})\}$, $m\in\overline{1,M}$, where $\mathbf{x}_m^{in}$ – m-th raw materials vector, $\mathbf{d}_m^{out}$ – m-th the expected reference vector of finished goods.

Then a problem of increase in accuracy of detection of anomalies on neural network model of the $g(\mathbf{x}^{in},\mathbf{w})$

, where $\mathbf{x}^{in}$ – the vector of raw materials, $\mathbf{w}$ – a vector of parameters, is represented as a problem of finding such vector of parameters $\mathbf{w}^*$, which satisfies criterion

$$F = \frac{1}{M}\sum_{m=1}^{M}(g(\mathbf{x}_m^{in}, \mathbf{w}^*) - \mathbf{d}_m^{out})^2 \to \min . \tag{1}$$

Block diagram of neural network model of detection of anomalies

The block diagram of Gauss-Bernoulli model of the forward only restricted Cauchy machine (FORCM) [8] which is recurrent ANN and consisting one visible layer and one hidden layer (fig.2).

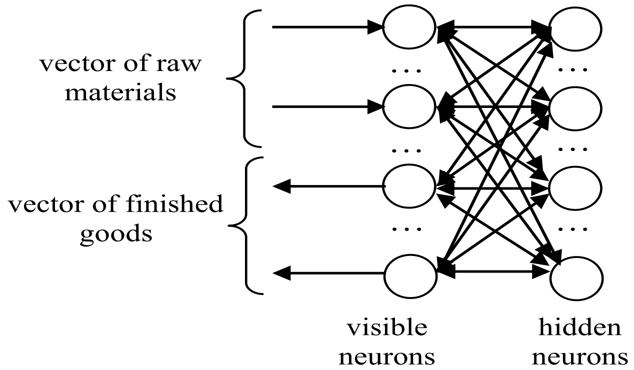

Fig.2. Block diagram of model of Gauss-Bernoulli of the forward only restricted Cauchy machine (FORCM)

Gauss-Bernoulli's components of FORCM are:
stochastic visible neurons which state is described on the basis of Gaussian distribution in the form

$$x_j = \mu_j + \sigma_j N(0,1) , \tag{2}$$

where μ_j – mathematical expectation,

σ_j – a mean square deviation (if the training a vector are normalized and centered, then σ_j =1),

$N(0,1)$ – the function returning standardly normally distributed random number.

Transition probability j -th a stochastic neuron in a state α is defined in the form

$$P_j = \frac{1}{\sigma_j\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{\alpha-\mu_j}{\sigma_j}\right)\right).$$

– the stochastic hidden neurons which state is described on the basis of Bernoulli's distribution in the form

$$x_j = \begin{cases} 1, & \text{with probability } P_j \\ 0, & \text{with probability } 1-P_j \end{cases} . \tag{3}$$

Transition probability j -th a stochastic neuron in a state 1 is defined in the form

$$P_j = \frac{1}{2} + \frac{1}{\pi}\arctan\left(\Delta E_j\right), \tag{4}$$

where ΔE_j – increment of energy of ANN at state change j -th a stochastic neuron with 0 on 1.

Gauss-Bernoulli's advantages FORCM:
1. Unlike the majority ANN possesses a heteroassociative memory.
2. Unlike Boltzmann's machine works real data.
3. Unlike Boltzmann's machine has no restrictions for storage capacity.

4. Unlike Boltzmann's machine has smaller computational complexity.

Neural network model of detection of anomalies

Positive phase include steps 1-2.

1. Initialization of a state of the visible neurons corresponding to raw materials $\mathbf{x1}^{in} = \mathbf{x}^{in}$.

2. Calculation of a state of the hidden neurons ($j \in \overline{1, N^h}$)

$$P_j = \frac{1}{2} + \frac{1}{\pi} \arctan\left(b_j^h + \sum_{i=1}^{N^{in}} w_{ij}^{in-h} \frac{x1_i^{in}}{\sigma_i^{in}} \right), \quad x1_j^h = \begin{cases} 1, & P_j \geq U(0,1) \\ 0, & P_j < U(0,1) \end{cases}. \tag{5}$$

where $U(0,1)$ – the function returning uniform distributed random number in the range $[0,1]$.

Negative phase include step 3.

3. Calculation of a state of the visible neurons corresponding to finished goods ($j \in \overline{1, N^{out}}$)

$$\mu_j^{out} = b_j^{out} + \sigma_j^{out} \sum_{i=1}^{N^h} w_{ij}^{out-h} x1_i^h, \quad x2_j^{out} = \mu_j^{out} + \sigma_j^{out} N(0,1), \tag{6}$$

where b_j^h – bias for j-th of a neuron of the hidden layer,

b_j^{out} – a bias for j-th of a neuron of the visible layer corresponding to finished goods

w_{ij}^{in-h} – communication weight from the neuron i-th in a visible layer corresponding to raw materials to j-th to a neuron of the hidden layer,

w_{ij}^{out-h} – communication weight from the neuron i-th in a visible layer corresponding to finished goods to j - th to a neuron of the hidden layer,

N^h – number of neurons in the hidden layer,

N^{in} – the number of the neurons in a visible layer corresponding to raw materials

N^{out} – the number of the neurons in a visible layer corresponding to finished goods.

Choice of criterion for evaluation of efficiency of neural network model of detection of anomalies

In this work for training of the FORCM model the function of the purpose which means the choice of such values of a vector of parameters is selected $\mathbf{w} = (w_{11}^{in-h}, ..., w_{N^{in}N^h}^{in-h}, w_{11}^{out-h}, ..., w_{N^{out}N^h}^{out-h})$, which deliver a minimum of a root mean square error (the differences of a sample on model and a test sample)

$$F = \frac{1}{M(N^{in} + N^{out})} \sum_{m=1}^{M} \left\| \mathbf{x2}_m^{out} - \mathbf{d}_m^{out} \right\|^2 \to \min_{\mathbf{w}}, \tag{7}$$

where $\mathbf{x2}_m^{out}$ – m-th an evaluation vector of finished goods on model,

$\mathbf{d}_m^{out}$ – m-th vector of finished goods.

Method of parametrical identification of neural network model of detection of anomalies on the basis of algorithm CD-1 (one-step contrastive divergence)

The method of parametrical identification of neural network model of detection of anomalies on the basis of algorithm CD-1 consists of the following blocks (fig. 3).

1. Initialization

Number of iteration of training $n = 1$, initialization by means of uniform distribution on an interval (0.1) or [-0.5, 0.5] bias $b_i^{out}(n)$, $i \in \overline{1, N^{out}}$, $b_j^h(n)$, $j \in \overline{1, N^h}$, and weights $w_{ij}^{in-h}(n)$, $i \in \overline{1, N^{in}}$, $j \in \overline{1, N^h}$,

$w_{ij}^{out-h}(n)$, $i \in \overline{1, N^{out}}$, $j \in \overline{1, N^h}$, $w_{ii}^{in-h}(n) = 0$, $w_{ii}^{out-h}(n) = 0$, $w_{ij}^{in-h}(n) = w_{ji}^{in-h}(n)$, $w_{ij}^{out-h}(n) = w_{ji}^{out-h}(n)$.

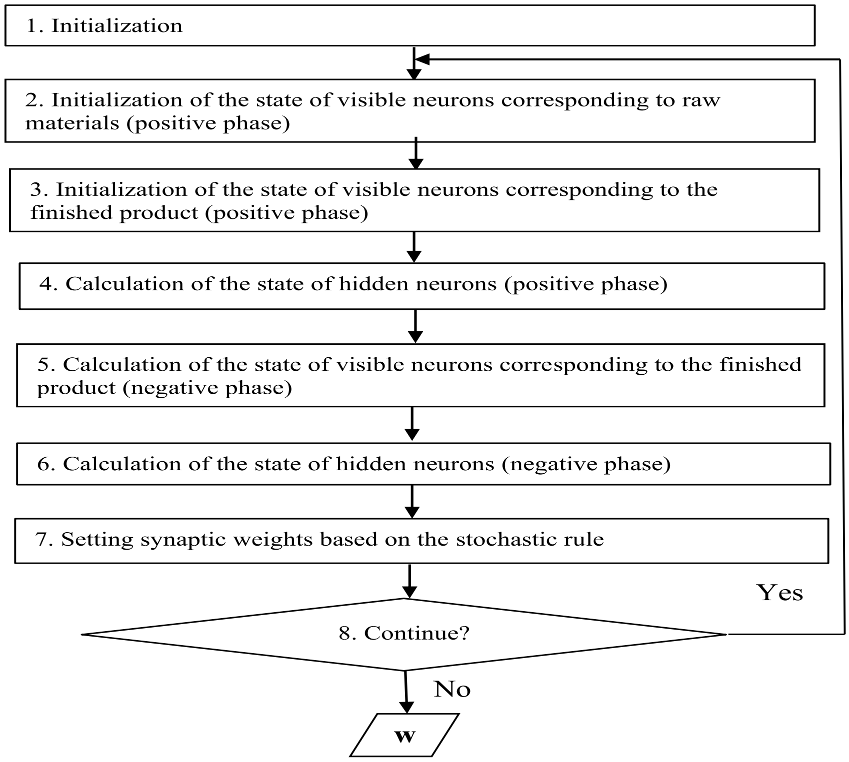

Fig.3. The sequence of procedures of parametrical identification method of neural network model of anomalies detection on the basis of CD-1

The training set is set $\{(\mathbf{x}_m^{in}, \mathbf{x}_m^{out}) \mid \mathbf{x}_m^{in} \in (0,1)^{N^{in}}, \mathbf{x}_m^{out} \in (0,1)^{N^{out}}\}$, $m \in \overline{1, M}$, where $\mathbf{x}_m^{in}$ – m-th raw materials vector, $\mathbf{x}_m^{out}$ – m-th vector of finished goods, M – power of a training set, vector of mean square deviations for a raw materials vector $\boldsymbol{\sigma}^{in} = (\sigma_j^{in}, ..., \sigma_{N^{in}}^{in})$; vector of mean square deviations for a vector of finished goods $\boldsymbol{\sigma}^{out} = (\sigma_j^{out}, ..., \sigma_{N^{out}}^{out})$.

Positive phase include steps 2-4.

2. Initialization of a state of the visible neurons corresponding to raw materials $\mathbf{x1}_m^{in} = \mathbf{x}_m^{in}$, $m \in \overline{1, M}$.

3. Initialization of a state of the visible neurons corresponding to finished goods $\mathbf{x1}_m^{out} = \mathbf{x}_m^{out}$, $m \in \overline{1, M}$.

4. Calculation of the hidden neurons state ($j \in \overline{1, N^h}$)

$$P_{mj} = \frac{1}{2} + \frac{1}{\pi} \arctan\left(b_j^h(n) + \sum_{i=1}^{N^{in}} w_{ij}^{in-h}(n) \frac{x1_{mi}^{in}}{\sigma_i^{in}} + \sum_{i=1}^{N^{out}} w_{ij}^{out-h}(n) \frac{x1_{mi}^{out}}{\sigma_i^{out}} \right), \quad m \in \overline{1, M}, \tag{8}$$

where

$$x1_{mj}^h = \begin{cases} 1, & P_{mj} \geq U(0,1) \\ 0, & P_{mj} < U(0,1) \end{cases}, \quad m \in \overline{1, M}.$$

Negative phase include steps 5-6.

5. Calculation of a state of the visible neurons corresponding to finished goods ($j \in \overline{1, N^{out}}$)

$$\mu_{mj}^{out} = b_j^{out}(n) + \sigma_j^{out} \sum_{i=1}^{N^h} w_{ij}^{out-h}(n) x1_{mi}^h, \quad m \in \overline{1, M}, \quad x2_{mj}^{in} = \mu_{mj}^{in} + \sigma_j^{in} N(0,1), \quad m \in \overline{1, M}. \tag{9}$$

6. Calculation of a state of the hidden neurons ($j \in \overline{1, N^h}$)

$$P_{mj} = \frac{1}{2} + \frac{1}{\pi} \arctan\left(b_j^h(n) + \sum_{i=1}^{N^{in}} w_{ij}^{in-h}(n) \frac{x1_{mi}^{in}}{\sigma_i^{in}} + \sum_{i=1}^{N^{out}} w_{ij}^{out-h}(n) \frac{x2_{mi}^{out}}{\sigma_i^{out}} \right), \ m \in \overline{1, M}, \tag{10}$$

where

$$x2_{mj}^h = \begin{cases} 1, & P_{mj} \geq U(0,1) \\ 0, & P_{mj} < U(0,1) \end{cases}, \ m \in \overline{1, M}.$$

7. Setup of bias and synoptic weights on the basis of the stochastic rule

$$b_i^{out}(n+1) = b_i^{out}(n) + \eta \left(\frac{1}{M} \sum_{m=1}^{M} \frac{x1_{mi}^{out}}{\left(\sigma_i^{out}\right)^2} - \frac{1}{M} \sum_{m=1}^{M} \frac{x2_{mi}^{out}}{\left(\sigma_i^{out}\right)^2} \right), \ i \in \overline{1, N^{out}}, \tag{11}$$

$$b_i^h(n+1) = b_i^h(n) + \eta \left(\frac{1}{M} \sum_{m=1}^{M} x1_{mi}^h - \frac{1}{M} \sum_{m=1}^{M} x2_{mi}^h \right), \ i \in \overline{1, N^h}, \tag{12}$$

$$\rho_{ij}^+ = \frac{1}{M} \sum_{m=1}^{M} \frac{x1_{mi}^{in} x1_{mj}^h}{\sigma_i^{in}}, \ i \in \overline{1, N^{in}}, \ j \in \overline{1, N^h}, \tag{13}$$

$$\rho_{ij}^- = \frac{1}{M} \sum_{m=1}^{M} \frac{x1_{mi}^{in} x2_{mj}^h}{\sigma_i^{in}}, \ i \in \overline{1, N^{in}}, \ j \in \overline{1, N^h}, \tag{14}$$

$$w_{ij}^{in-h}(n+1) = w_{ij}^{in-h}(n) + \eta(\rho_{ij}^+ - \rho_{ij}^-), \ i \in \overline{1, N^{in}}, \ j \in \overline{1, N^h}, \tag{15}$$

$$\rho_{ij}^+ = \frac{1}{M} \sum_{m=1}^{M} \frac{x1_{mi}^{out} x1_{mj}^h}{\sigma_i^{out}}, \ i \in \overline{1, N^{out}}, \ j \in \overline{1, N^h}, \tag{16}$$

$$\rho_{ij}^- = \frac{1}{M} \sum_{m=1}^{M} \frac{x2_{mi}^{out} x2_{mj}^h}{\sigma_i^{out}}, \ i \in \overline{1, N^{out}}, \ j \in \overline{1, N^h}, \tag{17}$$

$$w_{ij}^{out-h}(n+1) = w_{ij}^{out-h}(n) + \eta(\rho_{ij}^+ - \rho_{ij}^-), \ i \in \overline{1, N^{out}}, \ j \in \overline{1, N^h}. \tag{18}$$

8. Check of a termination condition

If $\frac{1}{M \cdot N^{out}} \sum_{m=1}^{M} \sum_{i=1}^{N^{out}} | x1_{mi}^{out} - x2_{mi}^{out} | > \varepsilon$ then $n = n+1$, transition to 2.

Algorithm of parametrical identification of neural network model of anomalies detection on the basis of algorithm CD-1 (one-step contrastive divergence) for implementation on **GPU**

For the offered method (5) – (18) of parametrical identification of neural network model of detection of anomalies on the basis of algorithm CD-1 the algorithm intended for implementation on GPU by means of technology of parallel processing of information of CUDA is considered. This algorithm functions as follows.

Step 1. Input of length a raw materials vector $\mathbf{x}^{in}$ N^{in}, lengths vector of finished goods $\mathbf{x}^{out}$ N^{out}, the number of neurons in the hidden layer N^h, capacities of a training set M, training set $\{(\mathbf{x}_m^{in}, \mathbf{x}_m^{out}) \mid \mathbf{x}_m^{in} \in (0,1)^{N^{in}}, \mathbf{x}_m^{out} \in (0,1)^{N^{out}}\}$, $m \in \overline{1, M}$, a vector of mean square deviations for a raw materials vector $\boldsymbol{\sigma}^{in} = (\sigma_j^{in}, ..., \sigma_{N^{in}}^{in})$; vector of mean square deviations for a vector of finished goods $\boldsymbol{\sigma}^{out} = (\sigma_j^{out}, ..., \sigma_{N^{out}}^{out})$

.

Step 2. Set training iteration number $n=1$. Initialization by means of uniform distribution on an interval (0.1) or [-0.5, 0.5] bias $b_i^{out}(n)$, $i \in \overline{1, N^{out}}$, $b_j^h(n)$, $j \in \overline{1, N^h}$ and weights $w_{ij}^{in-h}(n)$, $i \in \overline{1, N^{in}}$, $j \in \overline{1, N^h}$, $w_{ij}^{out-h}(n)$, $i \in \overline{1, N^{out}}$, $j \in \overline{1, N^h}$, $w_{ii}^{in-h}(n)=0$, $w_{ii}^{out-h}(n)=0$, $w_{ij}^{in-h}(n)=w_{ji}^{in-h}(n)$, $w_{ij}^{out-h}(n)=w_{ji}^{out-h}(n)$.

Step 3. Initialization of a state of the visible neurons corresponding to raw materials $\mathbf{x1}_m^{in}$, $m \in \overline{1, M}$.

Step 4. Initialization of a state of the visible neurons corresponding to finished goods $\mathbf{x1}_m^{out}$, $m \in \overline{1, M}$.

Step 5. Calculation of the sums $\sum_{i=1}^{N^{in}} w_{ij}^{in-h}(n) \frac{x1_{mi}^{in}}{\sigma_i^{in}}$, $m \in \overline{1, M}$, $j \in \overline{1, N^h}$, on the basis of a reduction, using $N^h \cdot M \cdot N^{in}$ GPU threads which are grouped in $N^h \cdot M$ blocks. The sum is result of operation of each block $s_{mj}^{in-h}(n)$.

Step 6. Calculation of the sums $\sum_{i=1}^{N^{out}} w_{ij}^{out-h}(n) \frac{x1_{mi}^{out}}{\sigma_i^{out}}$, $m \in \overline{1, M}$, $j \in \overline{1, N^h}$, on the basis of a reduction, using $N^h \cdot M \cdot N^{out}$ GPU threads which are grouped in $N^h \cdot M$ blocks. The sum is result of operation of each block $s_{mj}^{out-h}(n)$.

Step 7. Calculation of transition probabilities of stochastic neurons in a state 1 in the form $\frac{1}{2}+\frac{1}{\pi}\arctan\left(b_j^h(n)+s_{mj}^{in-h}(n)+s_{mj}^{out-h}(n)\right)$, using $N^h \cdot M$ GPU threads which are grouped in N^h blocks. Probability is result of work of each thread P_{mj}.

Step 8. Calculation of a state of the hidden neurons $x1_{mj}^h$. $m \in \overline{1, M}$, $j \in \overline{1, N^h}$, using $N^h \cdot M$ GPU threads which are grouped in N^h blocks. The state of the hidden neuron is result of work of each thread $x1_{mj}^h$.

Step 9. Calculation of the sums $\sum_{i=1}^{N^h} w_{ij}^{out-h}(n) x1_{mi}^h$, $m \in \overline{1, M}$, $j \in \overline{1, N^{out}}$, on the basis of a reduction, using $N^{out} \cdot M \cdot N^h$ GPU threads which are grouped in $N^{out} \cdot M$ blocks. The sum is result of operation of each block $s_{mj}^{h-out}(n)$.

Step 10. Calculation of a state of the visible neurons corresponding to finished goods $b_j^{out}(n)+\sigma_j^{out} s_{mj}^{h-out}(n)+\sigma_j^{out} N(0,1)$. $m \in \overline{1, M}$, $j \in \overline{1, N^{out}}$, using $N^{out} \cdot M$ GPU threads which are grouped in N^{out} blocks. The state of a visible neuron is result of work of each thread $x2_{mj}^{out}$.

Step 11. Calculation of the sums $\sum_{i=1}^{N^{in}} w_{ij}^{in-h}(n) \frac{x1_{mi}^{in}}{\sigma_i^{in}}$, $m \in \overline{1, M}$, $j \in \overline{1, N^h}$, on the basis of a reduction, using $N^h \cdot M \cdot N^{in}$ GPU threads which are grouped in $N^h \cdot M$ blocks. The sum is result of operation of each block $s_{mj}^{in-h}(n)$.

Step 12. Calculation of the sums $\sum_{i=1}^{N^{out}} w_{ij}^{out-h}(n) \frac{x2_{mi}^{out}}{\sigma_i^{out}}$, $m \in \overline{1, M}$, $j \in \overline{1, N^h}$, on the basis of a reduction, using $N^h \cdot M \cdot N^{out}$ GPU threads which are grouped in $N^h \cdot M$ blocks. The sum is result of operation of each block $s_{mj}^{out-h}(n)$.

Step 13. Calculation of transition probabilities of stochastic neurons in a state 1 in the form

$\frac{1}{2}+\frac{1}{\pi}\arctan\left(b_j^h(n)+s_{mj}^{in-h}(n)+s_{mj}^{out-h}(n)\right)$, using $N^h \cdot M$ GPU threads which are grouped in N^h blocks. Probability is result of work of each thread P_{mj}.

Step 14. Calculation of a state of the hidden neurons $x2_{mj}^h$, $m \in \overline{1,M}$, $j \in \overline{1,N^h}$, using $N^h \cdot M$ GPU threads which are grouped in N^h blocks. The state of the hidden neuron is result of work of each thread $x2_{mj}^h$.

Step 15. Calculation of the sums $\sum_{m=1}^{M} \frac{x1_{mi}^{out}}{\left(\sigma_i^{out}\right)^2}$, $i \in \overline{1,N^{out}}$, on the basis of a reduction, using $N^{out} \cdot M$ GPU threads which are grouped in N^{out} blocks. The sum is result of operation of each block $s1_i^{out}(n)$.

Step 16. Calculation of the sums $\sum_{m=1}^{M} \frac{x2_{mi}^{out}}{\left(\sigma_i^{out}\right)^2}$, $i \in \overline{1,N^{out}}$, on the basis of a reduction, using $N^{out} \cdot M$ GPU threads which are grouped in N^{out} blocks. The sum is result of operation of each block $s2_i^{out}(n)$.

Step 17. Calculation of bias in the form $b_i^{out}(n)+\eta\frac{s1_i^{out}(n)-s2_i^{out}(n)}{M}$. $i \in \overline{1,N^{out}}$, using N^{out} GPU threads which are grouped in one block. The bias is result of work of each thread $b_i^{out}(n+1)$.

Step 18. Calculation of the sums $\sum_{m=1}^{M} x1_{mi}^h$, $i \in \overline{1,N^h}$, on the basis of a reduction, using $N^h \cdot M$ GPU threads which are grouped in N^h blocks. The sum is result of operation of each block $s1_i^h(n)$.

Step 19. Calculation of the sums $\sum_{m=1}^{M} x2_{mi}^h$, $i \in \overline{1,N^h}$, on the basis of a reduction, using $N^h \cdot M$ GPU threads which are grouped in N^h blocks. The sum is result of operation of each block $s2_i^h(n)$.

Step 20. Calculation of bias in the form $b_i^h(n)+\eta\frac{s1_i^h(n)-s2_i^h(n)}{M}$. $i \in \overline{1,N^h}$, using N^h GPU threads which are grouped in one block. The bias is result of work of each thread $b_i^h(n+1)$.

Step 21. Calculation of the sums $\sum_{m=1}^{M} \frac{x1_{mi}^{in} x1_{mj}^h}{\sigma_i^{in}}$, $i \in \overline{1,N^h}$, $j \in \overline{1,N^h}$, on the basis of a reduction, using $N^{in} \cdot N^h \cdot M$ GPU threads which are grouped in $N^{in} \cdot N^h$ blocks. The sum is result of operation of each block ρ_{ij}^+.

Step 22. Calculation of the sums $\sum_{m=1}^{M} \frac{x1_{mi}^{in} x2_{mj}^h}{\sigma_i^{in}}$, $i \in \overline{1,N^h}$, $j \in \overline{1,N^h}$, on the basis of a reduction, using $N^{in} \cdot N^h \cdot M$ GPU threads which are grouped in $N^{in} \cdot N^h$ blocks. The sum is result of operation of each block ρ_{ij}^-.

Step 23. Calculation of weights in the form $w_{ij}^{in-h}(n)+\eta\frac{\rho_{ij}^+ - \rho_{ij}^-}{M}$. $i \in \overline{1,N^h}$, $j \in \overline{1,N^h}$, using $N^{in} \cdot N^h$ GPU threads which are grouped in N^{in} blocks. The bias is result of work of each thread $w_{ij}^{in-h}(n+1)$.

Step 24. Calculation of the sums $\sum_{m=1}^{M} \frac{x1_{mi}^{out} x1_{mj}^h}{\sigma_i^{out}}$, $i \in \overline{1,N^{out}}$, $j \in \overline{1,N^h}$, on the basis of a reduction, using

$N^{out} \cdot N^{h} \cdot M$ GPU threads which are grouped in $N^{in} \cdot N^{h}$ blocks. The sum is result of operation of each block ρ_{ij}^{+}.

Step 25. Calculation of the sums $\sum_{m=1}^{M} \frac{x2_{mi}^{out} x2_{mj}^{h}}{\sigma_{i}^{out}}$, $i \in \overline{1, N^{out}}$, $j \in \overline{1, N^{h}}$, on the basis of a reduction, using $N^{out} \cdot N^{h} \cdot M$ GPU threads which are grouped in $N^{out} \cdot N^{h}$ blocks. The sum is result of operation of each block ρ_{ij}^{-}.

Step 26. Calculation of weights in the form $w_{ij}^{out-h}(n) + \eta \frac{\rho_{ij}^{+} - \rho_{ij}^{-}}{M}$. $i \in \overline{1, N^{out}}$, $j \in \overline{1, N^{h}}$, using $N^{out} \cdot N^{h}$ GPU threads which are grouped in N^{out} blocks. The bias is result of work of each thread $w_{ij}^{out-h}(n+1)$.

Step 27. Calculation of the sums $\sum_{i=1}^{N^{out}} | x1_{mi}^{out} - x2_{mi}^{out} |$, $m \in \overline{1, M}$, on the basis of a reduction, using $M \cdot N^{out}$ GPU threads which are grouped in M blocks. The sum is result of operation of each block s_{m}^{out}.

Step 28. Calculation of the sums $\sum_{m=1}^{M} s_{m}^{out}$, $m \in \overline{1, M}$, on the basis of a reduction, using M GPU threads which are grouped in one block. The sum is result of operation of each block S.

Step 29. If $\frac{s}{M \cdot N^{out}} \le \varepsilon$, transition to a step 30, else increase number of iteration, i.e. $n = n+1$, transition to a step 3.

Step 30. Record of the received bias $b_{i}^{out}(n+1)$, $b_{i}^{h}(n+1)$, weights $w_{ij}^{in-h}(n+1)$ and $w_{ij}^{out-h}(n+1)$ the database.

Experiments

The proposed method was investigated by the indicators of the release of raw materials into production and the posting of finished products of a manufacturing enterprise with a two-year depth of sampling with daily time intervals. Comparison results of the offered neural network model (FORCM) with neural network model are presented by the forward only counterpropagation neural network (FOCPNN) on the basis of criterion of a root mean square error (tab. 2).

Table 2

Comparison of the offered neural network FORCM model with traditional FOCPNN

Root mean square error of neural network model	
FORCM	FOCPNN
0.02	0.05

According to tab.2, use of FORCM reduces a root mean square error and by that increases the accuracy of detection of anomalies.

Results of comparison of the offered algorithm of parametrical identification with use and without use of GPU and technology of parallel processing information of CUDA are provided in tab.3.

According to tab. 3, use of GPU reduces computational complexity approximately in $\frac{2N^{h}M}{\log_2(\sqrt{N^{h}}M)}$ time and by that increases the speed of parametrical identification.

Table 3

Comparison of computational complexity of parametrical identification method with use and without use of GPU

Indicator	Method	
	use of GPU	without use of GPU
Computational complexity	$O(2\log_2(N^{in} N^{out} \sqrt{N^{h}} M))$	$O(4(0.75N^{in} + N^{out})N^{h}M)$

Compared to other neural networks, the proposed neural network has the following advantages:

1. Possesses hetero-associative memory (as opposed to auto-associative stochastic neural networks).
2. Works with real data (as opposed to associative neural networks with binary / bipolar input neurons).
3. Has a high capacity of associative memory (in contrast to associative recurrent neural networks without a hidden layer with binary / bipolar input neurons).
4. Possesses high accuracy (due to stochasticity).
5. Has low computational complexity (due to the possibility of using the CD-1 method in comparison with full stochastic neural networks and the possibility of training based on CUDA technology).

Unlike non-connectionist methods, the proposed neural network does not have a limitation on the number of support vectors, a time series is not required, and preliminary data labeling for each type of anomaly is not required.

Conclusion

1. The actual problem of increase in efficiency of detection of anomalies in data of audit of waste-free production was solved by means of neural network model of Gauss-Bernoulli of the forward only restricted Cauchy machine (FORCM).

2. Cauchy offered neural network model of Gauss-Bernoulli of the forward only restricted Cauchy machine possesses a heteroassociative memory; works real data; has no restrictions for storage capacity; provide high accuracy of detection of anomalies; uses Cauchy's distribution that increases the speed of convergence of a method of parametrical identification.

3. To increase the speed of Gauss-Bernoulli parametric identification of a forward only restricted Cauchy machine, a parametric identification algorithm was developed to be implemented on a GPU using CUDA technology. The offered algorithm allows to increase training speed approximately in $\frac{2N^h M}{\log_2(N^h M)}$ time, where N^h – number of neurons in the hidden layer , M – power of a training set.

To increase the speed of Gauss-Bernoulli parametric identification of a forward only restricted Cauchy machine, a parametric identification algorithm was developed to be implemented on a GPU using CUDA technology. The proposed algorithm allows you to increase the learning rate by about a factor of $\frac{2N^h M}{\log_2(N^h M)}$, where N^h is the number of neurons in the hidden layer, M is the power of the training set.

4. The made experiments confirmed operability of the developed software and allow to recommend it for use in practice in a subsystem of the automated analysis of DSS of audit for detection of anomalies. Prospects of further researches are in checking the offered methods on broader set of test databases.

REFERENCES

1. The World Bank: World Development Report 2016: Digital Dividends, 2016. URL: https://www.worldbank.org/en/publication/wdr2016. (дата звернення: 10.07.2020).
2. Schultz M., Tropmann-Frick M. Autoencoder neural networks versus external auditors: detecting unusual journal entries in financial statement audits: proceedings of the 53rd Hawaii intern. conf. on System Sciences (HICSS-2020), Maui, Hawaii, USA, 2021. P. 5421-5430.
3. Nonnenmacher J., Kruse F., Schumann G., Marx G. Using autoencoders for data-driven analysis in internal auditing: proceedings of the 54th Hawaii intern. conf. on system sciences, Maui, Hawaii, USA, 2021. P. 5748-5757.
4. Neskorodieva T., Fedorov E., Izonin I. Forecast Method for Audit Data Analysis by Modified Liquid State Machine. Intelligent Information Technologies & Systems of Information Security: proceedings of the 1st Intern. Workshop (Khmelnytskyi, Ukraine, 10-12 June. 2020). proceedings CEUR-WS vol. 2623, 2020. P. 25-35.
5. Neskorodieva T., Fedorov E. Automatic analysis method of audit data based on neural networks mapping. Information technology and interaction: proceedings of the VII intern. conf. IT&I, Kyiv, 02-03 decem. 2020. CEUR-WS. vol. 2833, 2021, P.60-70.
6. Neskorodieva T., Fedorov E. Method for automatic analysis of compliance of expenses data and the enterprise income by neural network model of forecast. Modern machine learning technologies and data science: proceedings of the 2nd intern. workshop on (MoMLeT&DS-2020), Lviv-Shatsk, Ukraine, 2-3 June. 2020. CEUR-WS, vol. I: Main conference. vol. 2631, 2020. P. 145-158.
7. Chandola V., Banerjee A., Kumar V. Anomaly detection: a survey. ACM Computing Surveys, 2009. Vol.41, No 3.
8. Aggarwal Ch.C. Outlier Analysis. 2nd edn. Cham : Springer, 2017. 465 p.
9. Aggarwal Ch.C., Sathe S. Outlier Ensembles. An Introduction. Cham : Springer, 2017. 276 p.
10. Langr J., Bok V. GANs in Action Deep learning with Generative Adversarial Networks. Shelter Island : Manning Publications Co., 2019. 214 p.
11. Singh U.P., Jain S., Tiwari A., Singh R.K. Gradient evolution-based counter propagation network for approximation of noncanonical system. Soft Computing. 2019. Vol. 23, P. 4955–4967.
12. Tharwat A. Principal component analysis. Pattern Recognit. 2016. Vol. 3(3). P. 197–240.
13. Chai R., Naik G.R., Nguyen T.N., Ling S.H., Tran Y., Craig A., Nguyen H.T. Driver fatigue classification with independent component by entropy rate bound minimization analysis in an eeg-based system. Biomed. Health Inf. 2017. Vol. 21(3.). P. 715–724.
14. Baggenstoss P.M. Applications of projected belief networks. Signal Processing: proceedings 27th European Conference, 2019. P.1-5.
15. Kohonen T. Self-organizing and associative memory: textbook. 3rd edn. New York : Springer, 2012. 157 p.
16. de Vries H., Memisevic R., Courville A. Deep learning vector quantization. Artificial Neural Networks, Computational Intelligence and Machine Learning: proceedings European Symposium, on. Bruges (Belgium), 27-29 April 2016. ESANN, 2016. P. 503–508.
17. Du K.L., Swamy M.N.S. Neural Networks and Statistical Learning: textbook Springer-Verlag London, 2014. 824 p.
18. Khristodulo O.I., Makhmutova A.A., Sazonova T.V. Use algorithm based at Hamming neural network method for natural objects classification. Procedia Computer Science, 2017. № 103. P. 388-395.

19. Javidmanesh E. Global stability and bifurcation in delayed bidirectional associative memory neural networks with an arbitrary number of neurons. J. Dyn. Sys., Meas., Control. 2017, vol. 139(8).
20. Wang Q., Gao X., Wan K., Li F., Hu Z. A novel restricted boltzmann machine training algorithm with fast gibbs sampling policy. URL: https://www.hindawi.com/journals/mpe/2020/4206457/.

Tetiana Neskorodieva **Тетяна Нескородєва**	Ph. D., Docent, Head of Information Technology Department, Vasyl' Stus Donetsk National University, Vinnytsia, Ukraine e-mail: t.neskorodieva@donnu.edu.ua orcid.org/0000-0003-2474-7697, Scopus Author ID: 57218242548 (Neskorodieva), 25123353600 (Zemlyak), ReseacherID: S-5190-2017 https://scholar.google.com/citations?user=XmShM6EAAAAJ&hl=uk	кандидат технічних наук, доцент, зав. кафедри інформаційних технологій, Донецький національний університет імені Василя Стуса, Вінниця, Україна.
Eugene Fedorov **Євген Федоров**	DrS., Professor of Information Technology Department, Vasyl' Stus Donetsk National University, Vinnytsia, Ukraine e-mail: fedorovee75@ukr.net orcid.org/0000-0003-3841-7373, Scopus Author ID: 47161087200, 47161155900, 57205185819, ReseacherID: AAO-6744-2020 https://scholar.google.com/citations?authuser=1&user=hbzfMWcAAAAJ	доктор технічних наук, професор кафедри інформаційних технологій, Донецький національний університет імені Василя Стуса, Вінниця, Україна.
Anastasiia Neskorodieva **Анастасія Нескородєва**	Laboratory Assistant of Interdepartmental Laboratory of Machine Learning and Data Mining, Vasyl' Stus Donetsk National University, Vinnytsia, Ukraine e-mail: neskorodieva.a@gmail orcid.org/0000-0002-8591-085X, https://scholar.google.com/citations?hl=uk&user=4HS_gHkAAAAJ&view_op=list_works	лаборант міжкафедральної лабораторії машинного навчання та інтелектуального аналізу даних, Донецький національний університет імені Василя Стуса, Вінниця, Україна.
Tetiana Sichko **Тетяна Січко**	Ph. D., Docent, Associate Professor of Information Technology Department, Vasyl' Stus Donetsk National University, Vinnytsia, Ukraine e-mail: t.sichko@donnu.edu.ua https://orcid.org/0000-0003-1766-4981, Scopus Author ID: 57204592112, ReseacherID: L-9019-2018 https://scholar.google.com/citations?user=nlTrwOoAAAAJ&view_op=list_works	кандидат технічних наук, доцент кафедри інформаційних технологій, Донецький національний університет імені Василя Стуса, Вінниця, Україна.
Pavlo Rymar **Павло Римар**	Senior Lecturer of Information Technology Department, Vasyl' Stus Donetsk National University, Vinnytsia, Ukraine e-mail: p.rymar@donnu.edu.ua orcid.org/0000-0002-0647-2020 https://scholar.google.com/citations?user=Sw4C5JMAAAAJ&view_op=list_works	старший викладач кафедри інформаційних технологій, Донецький національний університет імені Василя Стуса, Вінниця, Україна.