

UDC 616.6-071:004.032.26

doi: 10.32620/reks.2019.4.13

M. I. FEDORENKO*National Aerospace University "Kharkiv Aviation Institute", Ukraine***MODELS OF GENERATION OF INPUT DATA OF TRAINING OF NEURAL NETWORK MODULES FOR DIAGNOSTIC OF DISEASES IN UROLOGY**

The **subject** of the research presented in the article is neural network modules (NNMs), which are used to solve problems in the practice of diagnosing diseases in urology. This work aims to develop a mathematical model for generating a multitude of uroflowmetric parameters, in particular, graphs of uroflowgrams of the required volume, used as input data for NNM training. **Objective:** to develop a mathematical model for the formation of uroflowmetric parameters using a probabilistic approach based on a uniform "white noise". To develop an effective algorithm for the procedure for generating new parameter values and tools for its implementation. **Methods used:** NNM training methods, mathematical modeling methods, digital signal processing methods, tools for generating and processing random numerical sequences, digital data filtering methods. The following **results** were obtained: when creating and implementing a mathematical model for generating a large amount of training data, the requirements of randomness are taken into account when obtaining new values of uroflowmetric parameters. And at the same time, the obtained noise values are filtered to values of a given range, which are percentage-wise comparable to the amplitude value of the uroflowmetric parameter. **Conclusions.** The scientific novelty of the results is as follows: the NNM training method for recognizing diseases in urology has been improved by developing a mathematical model to generate uroflowmetric parameters for NNM training. The presented model allows you to create the necessary amount of data for training neural network modules in the course of experimental research on the recognition of diseases. The generation of uroflowmetric parameters is based on adding noise to the parameter values. This allows you to change the input data of the NNM training in a given range. This ensures the creation of the required input volume of the NNM training procedure. In the future, this contributes to the testing process of trained neural network modules with reliable information on the diagnosis of diseases in urology.

Keywords: uroflowmetric parameters; graph of the uroflowrogram; normal noise; neural network modules; training of neural network modules; diagnosis of diseases.

Introduction

Modern medicine has been long using computer diagnosing of diseases. The procedures for collecting and processing diagnostic information (in particular, the results of radiography, electrocardiogram, ultrasound, mammography, tomography, uroflowmetry, etc.) have been automated, in real time as well by means of telemonitoring [1-3]. Another factor influencing the systems of diagnosing diseases is the implementation of intelligent information technologies for solving medical issues. In particular, the neuronet approaches to diagnosing diseases have been widespread in various fields of medicine, such as cardiology, urology, diseases of the lungs, infectious and other diseases [4-6].

A current task with such an approach is to obtain the necessary amount of data sufficient for training, and further correct recognition of diagnosed diseases. In addition, neuronet modules should have a high level of generalizing abilities to identify diseases, taking into account the individual characteristics of patients and the possible influence of regional factors on the diagnostic

parameters associated with the state of the environment or natural changes. In this regard, the current task is to provide a sufficient amount of training data to assure the reliability of the diagnosis of diseases in urology.

Objective of the paper is to develop a model for the formation of uroflowmetric parameters using a probabilistic approach based on a uniform "white noise".

Network training issues

Network generic capabilities can be improved by increasing the amount and updating of training data. The study of the neuronet is performed with the help of additional learning pairs. The purpose of this approach is to provide a network that allows one to identify the incoming data that is slightly modified and has variations in values within the given permissible limits. In other words, when processing incoming signals the network should produce adequate output signals, so it should be insensitive to the variations in input data within a given range. Even when it was tested, the input signals were not included in the training set. There is a mathematical

substantiation for such a requirement. Let the vector of test set be xv . The task of network training is to minimize the target function [7-9]:

$$L = \frac{1}{2} \sum_{j=1}^N \|p_j - f(xl_j)\|^2, \quad (1)$$

where p_j is the type of disease, $f(xl_j)$ is the output value of the net, N is the volume of training sample. However, this does not guarantee that the trained network will accurately recognize the input vector xv_k from the test set $XV = \{(xv_k)\}_{k=1}^Z$, where Z is volume of test set. Further research is based on the assumption that the test vector is xv_k slightly different from the vector xl . Define this difference in the form:

$$xv_k = xl_j + r, \quad (2)$$

where $r = [r_1, r_2, \dots, r_M]^T$ is a noise vector which is a random variable with small values of the amplitude, M – is the number of input neurons.

Influence of noise on uroflowmetric parameters

When forming the training data for learning the neuronet, the author has used the values of the uroflowmetric parameters and their respective diseases. An increase in the amount of training pairs has been carried out by generating a quantity of values of the uroflowmetric parameters.

This is done by adding the noise as indicated in formula (2) to the uroflowmetric parameters, which with 100 % confidence correspond to the disease.

The studies were conducted to analyze the network's response to changes in training data over several stages. In particular, in the first phase of training there were 100 training pairs.

After completing the training, an analysis of quantitative and qualitative indicators of the network operation was carried out.

The network has been trained to diagnose of two classes of diseases: a slight bladder outlet obstruction (class 1), detrusor-sphincter dyssynergia (class 2) [10]. The analysis of the influence of noise on changes of the uroflowmetric parameters was investigated in the graphs of uroflowrograms. Figure 1 shows a typical graph of the uroflowrogram for disease in class 1. Noise formation is based on the following condition white noise is a random value ranging from 0 to $a \times xl_j$, where a is a constant (threshold value of change in amplitude of noise)), $xl_j, \{xl_j\}_{j=1}^M$ is a vector of the values of the uroflowmetric parameter (in particular, of the uroflowrogram) involved in the training, M is a number of vector values of the uroflowmetric parameter (number of input neurons). It should be noted that the value of the constant a is determined experimentally, depending on the value of the reliability of the results obtained by the identifying the uroflowmetric parameter.

Subsequently, an analysis was made of the effect of noise on changes in the values of uroflowmetric parameters. In particular on the graphs of uroflowrograms for a disease of class 1. Figure 1 shows a typical graph of uroflowrograms for this disease.

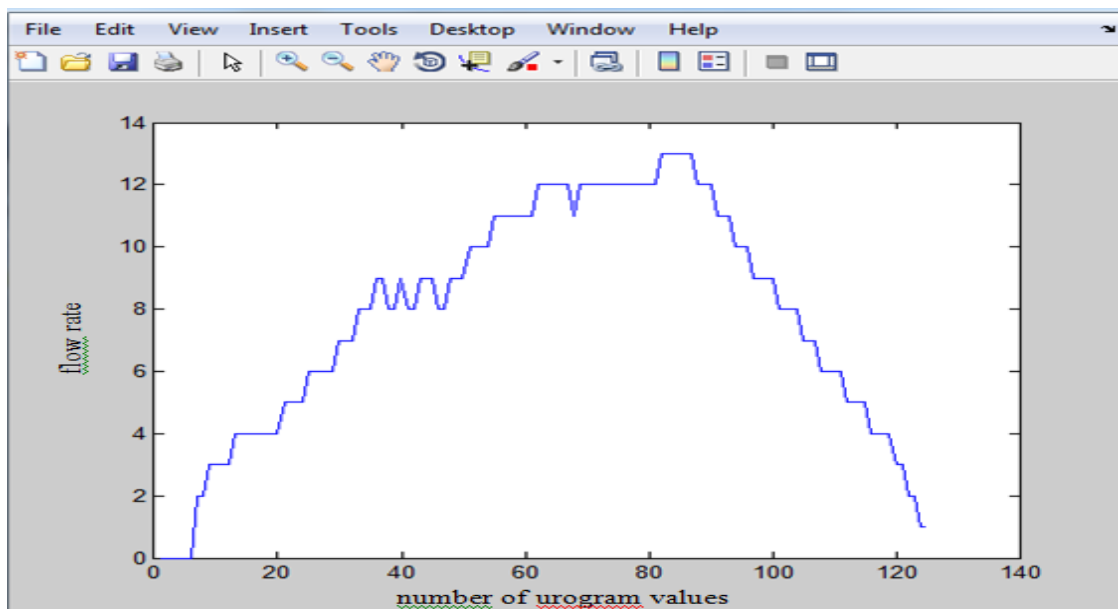


Fig. 1. Uroflowrogram for disease in class 1

To prepare the parameters for learning the network their values are normalized by the formula

$$x_{lj} = \left(\frac{x_{lj}}{x_{l_{\max}}} \right), \quad (3)$$

where $x_{l_{\max}}$ is the maximum value quantity of uroflowmetric parameters, x_{lj} is the current vector value. Figure 2 shows the graphs of the uroflowgram after normalization for two classes.

Then the normalized values of the uroflowgram change under the influence of noise in the same way as in expression 2:

$$x_{lj} = x_{lj} + r_j, \quad (4)$$

where r_j is the noise value corresponding to each value of the vector of the uroflowmetric parameter. MATLAB generates noise sequentially; one random number is generated in the range from 0 to 1 using the function $\text{rand}(1,1)$ which is called even white noise.

Figure 3 shows a graph of generated noise for a set of 125 values of a uroflowgram.

At the initial stage of the constant a has value 0.15. In other words, the noise threshold is 15 % of each current value of the vector of uroflowmetric parameter. After that, the noise value is filtered in the range from 0 to $a = 0.15$. Figure 4 shows the filtered noise graph for class 1 uroflowgram. Further network learning included 100 training pairs for two diseases of class 1

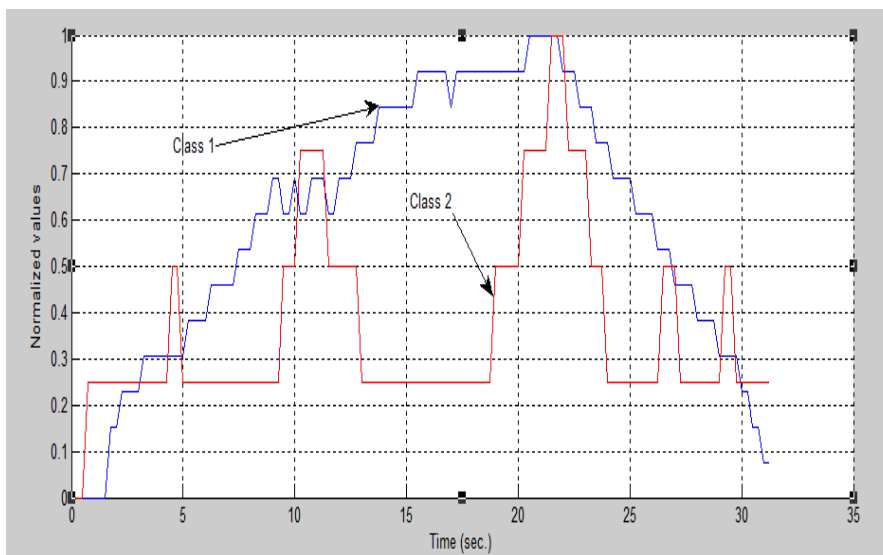


Fig. 2. Graphs of normalized uroflowgrams of two classes

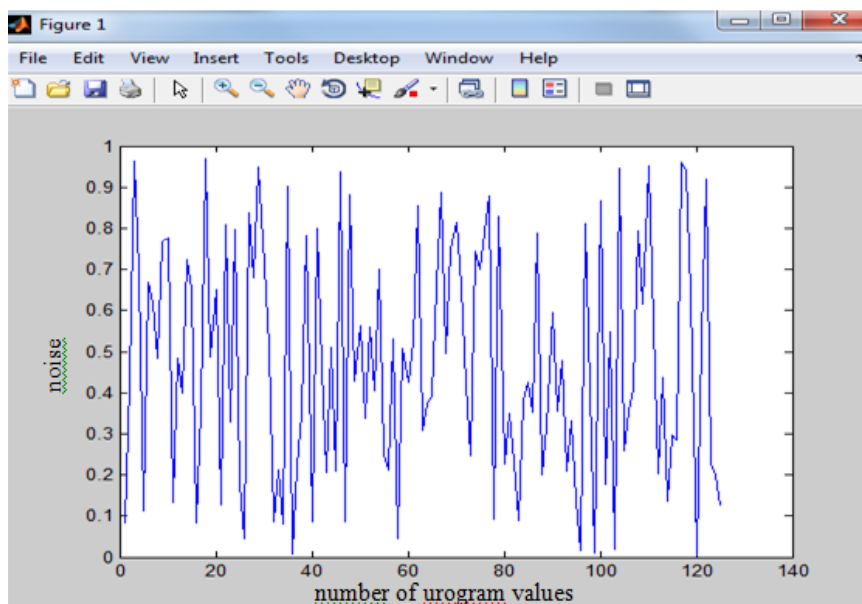


Fig. 3. Generated noise graph for a set of 125 value

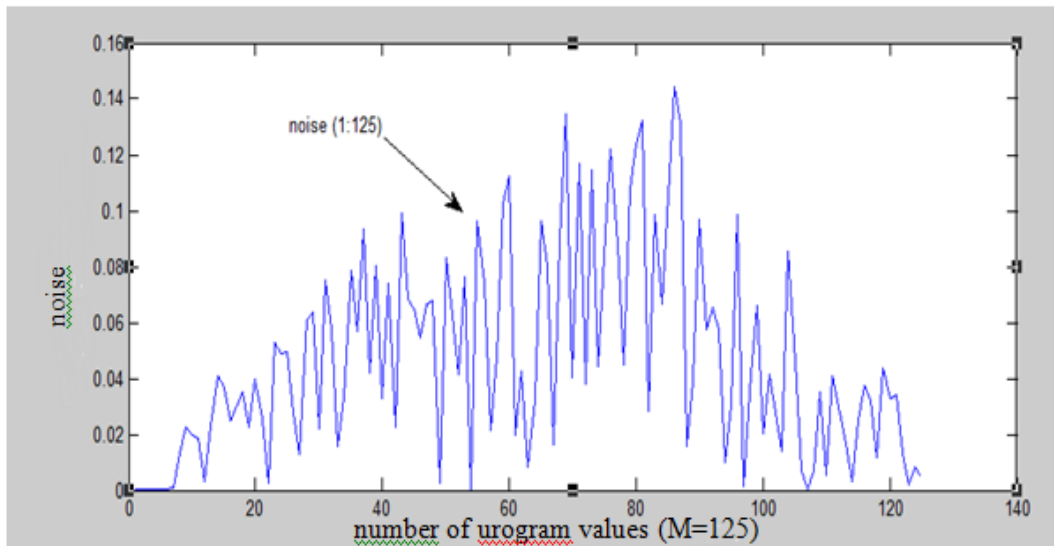


Fig. 4. Filtered noise pattern for class 1 uroflowrogram

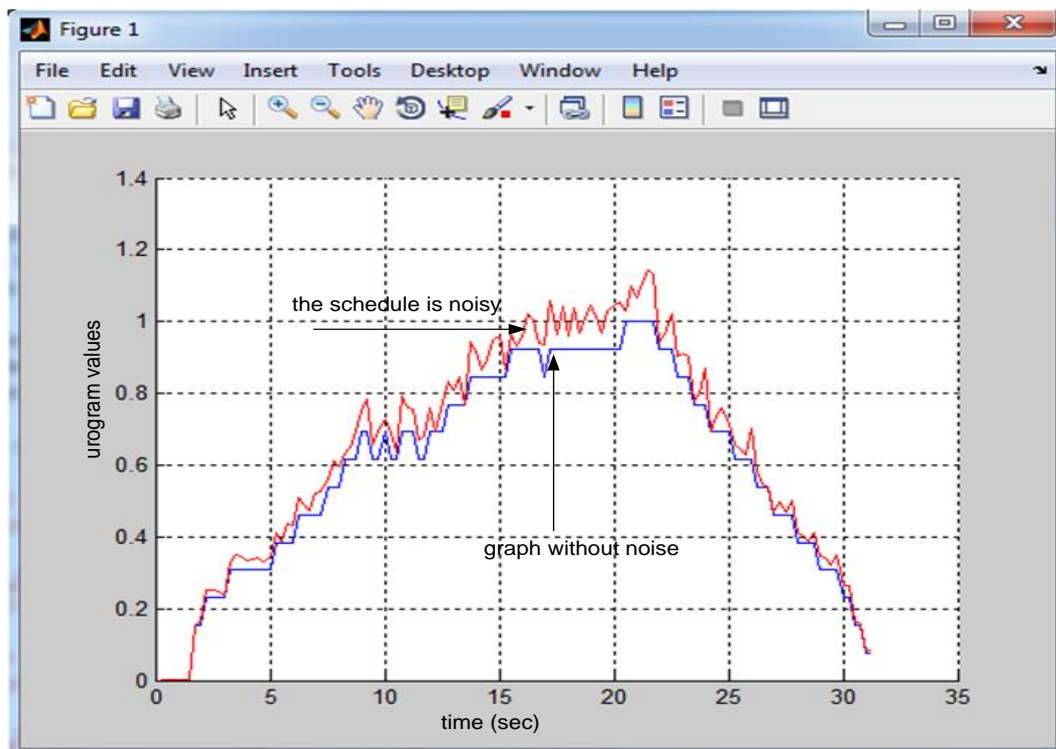


Fig. 5. Graphs of distorted and undistorted uroflowrograms of class 1

and 2. In Figure 5 shows one of the variants of the uroflowrogram of the first class of the disease (distorted and undistorted), which took part in the learning process.

Figure 6 presents a variant of distorted uroflowrograms of two classes of diseases involved in the training.

Figure 7 shows the graphical results of training a neural network module in the volume of 100 training pairs for two classes of diseases with distorted input vectors.

Conclusion

The article presents a mathematical model for generating a sufficient volume of input data for training neural network modules for diagnosing diseases in urology. A solution to this problem is necessary for experiments to analyze the reliability of disease recognition results. The experiments were carried out on uroflowrograms, which are one of the most difficult to recognize uroflowmetric parameters. In the work, we

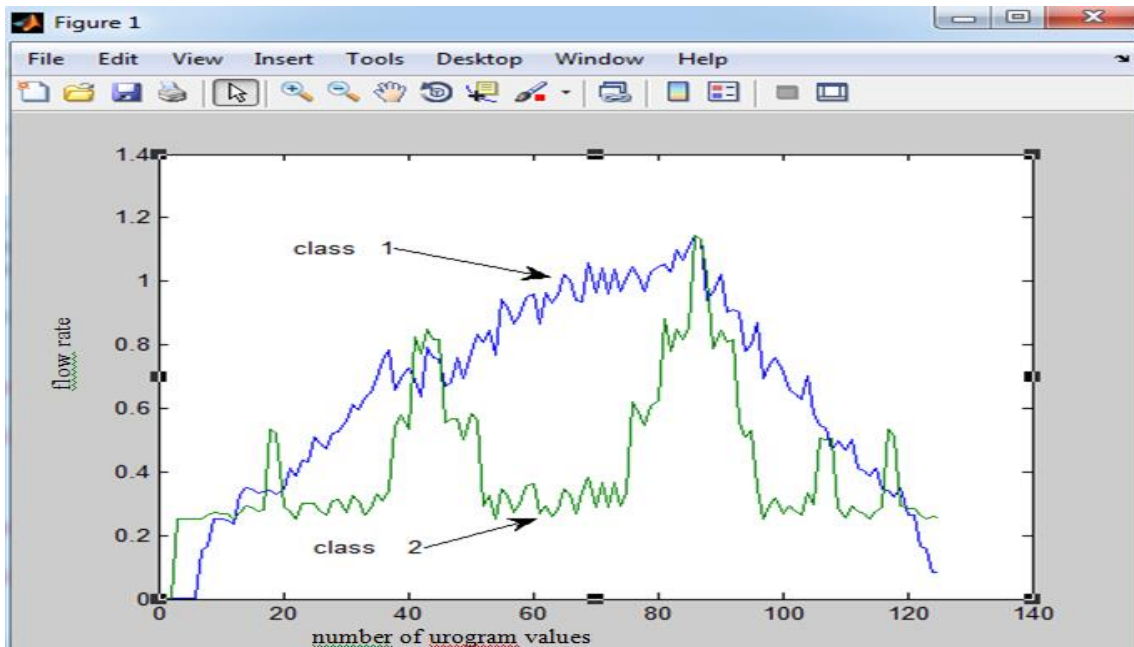


Fig. 6. Graphs of two distorted uroflowgrams that participated in the training network

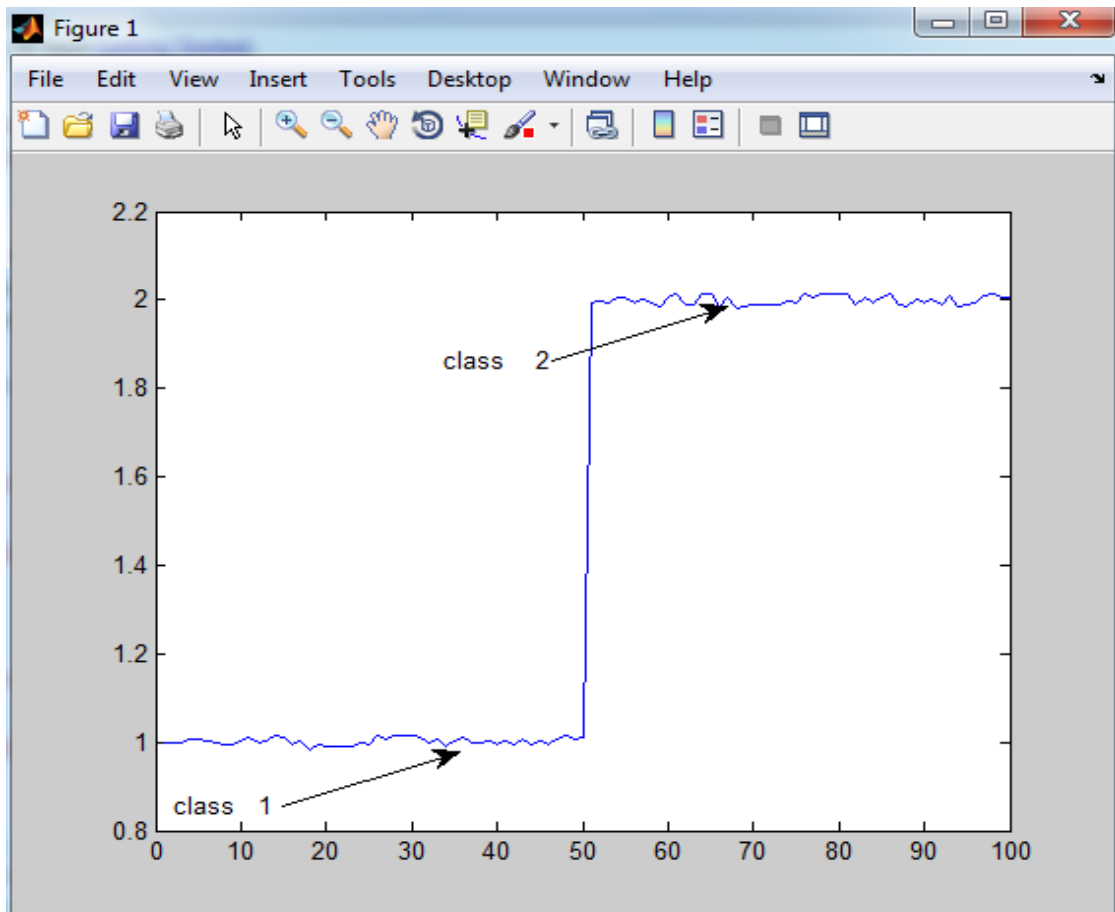


Fig. 7. Results of recognition by the module of distorted input data in training cycles

used graphs of uroflowgrams with 100 % certainty identifying the disease. The required number of graphs is generated by adding uniform “white noise” to the main uroflowgram taking into account the limitation

of its amplitude in a given range. This approach allows you to simulate the presence of patients with the same diseases.

References (GOST 7.1:2006)

1. Хорозов, О. А. Телемоніторинг життєво важливих показників пацієнтів [Текст] / О. А. Хорозов // Управляющие системы и машины. – 2015. – № 5. – С. 37-41.
2. Belciug, S. Error-correction learning for artificial neural networks using the Bayesian paradigm. Application to automated medical diagnosis [Text] / S. Belciug, F. Corunescu // Journal of Biomedical Informatics. – 2014. – Vol. 52. – P. 329-337.
3. Ghorbani, A. Interpretation of neural networks is fragile [Electronic resource] / A. Ghorbani, A. Abid, J. Zou. – Available at: <https://arxiv.org/pdf/1710.10547.pdf>. (accessed 29.10.2017).
4. Artificial Neural Network and Mobile Applications in Medical diagnosis [Text] / G. Pearce, L. Mirtskhulava, K. Bakuria, J. Wong, S. Al-Majeed, N. Gulua // International Conference on Modelling and Simulation 17th UKSIM-AMSS, 2015. – P. 61-65.
5. Banchev, B. Architecture of automated communication and data processing system for forensic medicine [Text] / B. Banchev, T. Georgiev // CompSysTech. – 2015. – P. 160-167.
6. Deutsch, L. Generating Neural Networks with Neural Networks [Electronic resource] / L. Deutsch. – Available at: <https://arxiv.org/pdf/1801.01952.pdf> (accessed 17.02.2018).
7. Осовский, С. Нейронные сети для обработки информации [Текст] / С. Осовский. – М.: Финансы и статистика, 2002. – 344 с.
8. Меркушева, А. В. Нейросетевые методы обработки сигналов в информационных системах. Элементы структуры, принципы обучения и мера многообразия отображений, реализуемых нейронной сетью [Текст] / А. В. Меркушева // Информационные технологии. – 2005. – № 3. – С. 9-20.
9. Хайкин, С. Нейронные сети: полный курс [Текст] / С. Хайкин. – 2-е изд. – М.: Издательский дом «Вильямс», 2006. – 1104 с.
10. The processing of the diagnostic data in a medical information-analytical system using a network of neuro modules with relearning [Text] / N. Fedorenko, V. Kharchenko, Y. Yehorova, L. Lutai // IEEE EAST-WEST DESIGN & TEST SYMPOSIUM, 14- 17 October 2016. – Yerevan, Armenia. – P. 381-383.

References (BSI)

1. Khorozov, O. A. Telemonitoring zhly'ttyevo vazhly'vy'x pokazny'kiv patsiyentiv [Telemonitoring vital parameters of patients]. *Upravlyayushchie sistemy i mashiny*, 2015, no. 5, pp. 37-41.
2. Belciug, S., Gorunescu, F. Error-correction learning for artificial neural networks using the Bayesian paradigm. Application to automated medical diagnosis. *Journal of Biomedical Informatics*, 2014, vol. 52, pp. 329-337.
3. Ghorbani, A., Abid, A., Zou, J. *Interpretation of neural networks is fragile*. Available at: <https://arxiv.org/pdf/1710.10547.pdf>. (accessed 29.10.2017).
4. Pearce, G., Mirtskhulava, L., Bakuria, K., Wong, J., Al-Majeed, S., Gulua, N. Artificial Neural Network and Mobile Applications in Medical diagnosis. *International Conference on Modelling and Simulation 17th UKSIM-AMSS*, 2015, pp. 61-65.
5. Banchev, B., Georgiev, T. Architecture of automated communication and data processing system for forensic medicine. *CompSysTech*, 2015, pp. 160-167.
6. Deutsch, L. *Generating Neural Networks with Neural Networks*. Available at: <https://arxiv.org/pdf/1801.01952.pdf> (accessed 17.02.2018).
7. Osovskii, S. *Neironnye seti dlya obrabotki informatsii* [Neural networks for information processing]. Moscow, Finansy i statistika Publ., 2002. 344 p.
8. Merkusheva, A. V. *Neirosetevye metody obrabotki signalov v informatsionnykh sistemakh*. Elementy struktury, printsipy obucheniya i mera mnogoobraziya otobrazhenii, realizuemykh neuronnoi set'yu [Neural network signal processing methods in information systems. Elements of the structure, principles of training and a measure of the variety of mappings implemented by the neural network]. *Informatsionnye tekhnologii*, 2005, no. 3, pp. 9-20.
9. Khaikin, S. *Neironnye seti: polnyi kurs* [Neural networks: a full course]. Moscow, Izdatel'skii dom «Vil'yams» Publ., 2006. 1104 p.
10. Fedorenko, N., Kharchenko, V., Yehorova, Y., Lutai, L. The processing of the diagnostic data in a medical information-analytical system using a network of neuro modules with relearning. *IEEE EAST-WEST DESIGN & TEST SYMPOSIUM*, 14- 17 October 2016, Yerevan, Armenia, pp. 381-383.

Поступила в редакцию 9.10.2019, рассмотрена на редколлегии 10.12.2019.

МОДЕЛИ ГЕНЕРАЦИИ ВХОДНЫХ ДАННЫХ ОБУЧЕНИЯ НЕЙРОСЕТЕВЫХ МОДУЛЕЙ ДЛЯ ДИАГНОСТИКИ ЗАБОЛЕВАНИЙ В УРОЛОГИИ

Н. И. Федоренко

Предметом исследования, представленного в статье, являются нейросетевые модули (НСМ), которые используются для решения задач в практике диагностики заболеваний в урологии. **Целью** данной работы является разработка математической модели для генерации множества урофлоуметрических параметров, в частности графиков урофлоурограмм необходимого объема, используемых в качестве входных данных для

обучения НСМ. **Задача:** разработать математическую модель формирования множества урофлоуметрических параметров с использованием вероятностного подхода, основанного на равномерном «белом шуме»; разработать эффективный алгоритм для процедуры генерации новых значений параметров и инструменты для его реализации. Используемые **методы:** методы обучения НСМ, методы математического моделирования, методы цифровой обработки сигналов, инструменты для генерации и обработки случайных числовых последовательностей, методы цифровой фильтрации данных. Были получены следующие **результаты:** при создании и реализации математической модели для генерации большого количества обучающих данных требования случайности учитываются при получении новых значений урофлоуметрических параметров. И в то же время полученные значения шума фильтруются до значений заданного диапазона, которые в процентном отношении сопоставимы со значением амплитуды урофлоуметрического параметра. **Выводы.** Научная новизна полученных результатов заключается в следующем: метод обучения НСМ для распознавания заболеваний в урологии был усовершенствован путем разработки математической модели для генерации множества урофлоуметрических параметров для обучения НСМ. Представленная модель позволяет создать необходимый объем данных для обучения нейросетевых модулей в ходе экспериментальных исследований по распознаванию заболеваний. Генерация множества урофлоуметрических параметров основана на добавлении шума к значениям параметров. Это позволяет изменить входные данные обучения НСМ в заданном диапазоне, что обеспечивает создание необходимого объема данных для процедуры обучения НСМ. В дальнейшем это способствует процессу тестирования обученных нейросетевых модулей с достоверной информацией по диагностике заболеваний в урологии.

Ключевые слова: урофлоуметрические параметры; график урофлуорограммы; нормальный шум; нейросетевые модули; обучение нейросетевых модулей; диагностика заболеваний.

МОДЕЛІ ГЕНЕРАЦІЇ ВХІДНИХ ДАНИХ НАВЧАННЯ НЕЙРОМЕРЕЖЕВИХ МОДУЛІВ ДЛЯ ДІАГНОСТИКИ ХВОРОБ В УРОЛОГІЇ

М. І. Федоренко

Предметом дослідження, представленим в статті, є нейромережеві модулі (НММ), які використовуються для вирішення проблем у практиці діагностики захворювань в урології. **Метою** даної роботи є розробка математичної моделі для формування множини урофлоуметричних параметрів, зокрема, графіків урофлуорограм необхідного обсягу, що використовуються як вхідні дані для навчання НММ. **Задача:** розробити математичну модель формування множини урофлоуметричних параметрів, використовуючи ймовірнісний підхід, заснований на рівномірному «білому шумі»; розробити ефективний алгоритм процедури генерації нових значень параметрів та інструментів для його реалізації. Використовувані **методи:** методи навчання НММ, методи математичного моделювання, методи цифрової обробки сигналів, інструменти для генерації та обробки випадкових числових послідовностей, методи цифрової фільтрації даних. Отримані такі **результати:** при створенні та реалізації математичної моделі для формування великого обсягу навчальних даних, вимоги випадковості враховуються при отриманні нових значень урофлоуметричних параметрів. І в той же час отримані значення шуму фільтруються до значень заданого діапазону, які в відсотках порівняні зі значенням амплітуди урофлоуметричного параметра. **Висновки.** Наукова новизна результатів полягає в наступному: метод навчання НММ для розпізнавання захворювань в урології був удосконалений шляхом розробки математичної моделі для створення множини урофлоуметричних параметрів для навчання НММ. Представлена модель дозволяє створити необхідний обсяг даних для навчання нейромережевих модулів в ході експериментальних досліджень з розпізнавання захворювань. Формування множини урофлоуметричних параметрів базується на додаванні шуму, що забезпечує створення необхідного вхідного обсягу даних процедури навчання НММ. Надалі це сприяє процесу тестування навчених нейромережевих модулів з достовірною інформацією по діагностиці захворювань в урології.

Ключові слова: урофлоуметричні параметри; графік урофлуорограми; нормальний шум; нейромережеві модулі; навчання нейромережевих модулів; діагностика захворювань.

Федоренко Микола Іванович – завідувач лабораторії комп'ютерних технологій та мультимедійних засобів навчання кафедри прикладної лінгвістики, Національний аерокосмічний університет ім. М. С. Жуковського “Харківський авіаційний інститут”, Харків, Україна.

Fedorenko Mykola Ivanovych – Head of the Laboratory for Computer Technology and Multimedia Training at the Department of Applied Linguistics, National Aerospace University "Kharkiv Aviation Institute", Kharkiv, Ukraine, e-mail: n.fedorenko@csn.khai.edu,
ORCID Author ID: 0000-0001-8651-7892, Scopus Author ID: 57203146499.