

**NEWS**

OF THE NATIONAL ACADEMY OF SCIENCES OF THE REPUBLIC OF KAZAKHSTAN

**PHYSICO-MATHEMATICAL SERIES**

ISSN 1991-346X

<https://doi.org/10.32014/2020.2518-1726.62>

Volume 4, Number 332 (2020), 28 – 34

UDC 004.032.26

IRSTI 28.23.37

**Marat Nurtas<sup>1</sup>, Zh. Baishemirov<sup>2,3</sup>,  
V. Tsay<sup>1</sup>, M. Tastanov<sup>1</sup>, Zh. Zhanabekov<sup>1</sup>**

<sup>1</sup>International Information Technologies University, Almaty, Kazakhstan;

<sup>2</sup>Abai Kazakh National Pedagogical University, Almaty, Kazakhstan;

<sup>3</sup>RSE Institute of Information and Computational Technology CS MES RK, Almaty, Kazakhstan.

E-mail: [maratnurtas@gmail.com](mailto:maratnurtas@gmail.com); [zbai.kz@gmail.com](mailto:zbai.kz@gmail.com);

[tsay.victor96@gmail.com](mailto:tsay.victor96@gmail.com); [tychty4@gmail.com](mailto:tychty4@gmail.com); [zzhanabekov@gmail.com](mailto:zzhanabekov@gmail.com)

## **APPLYING NEURAL NETWORK FOR PREDICTING CARDIOVASCULAR DISEASE RISK**

**Abstract.** This article concerns the problem of the prevalence of cardiovascular disease in economically developed countries. The purpose of this article is to create a neural network to determine the risk of cardiovascular disease based on the individual characteristics of the patient. In order to predict the risk of cardiovascular disease a neural network has been developed. The model was built in the Python programming language using the open-source library for building neural networks Keras. Data containing patient information for model building were taken from Kaggle.com. The accuracy of the neural network is 82%. With the help of neural network it will be possible to analyze the changes and the development of diseases in the future by changing the patient's input parameters, for instance, age, increase in blood pressure and etc. Also it would be possible to change the predictive diagnosis for the better if follow the parameters such as refusal from addictions, regular sleep, a healthy lifestyle and proper nutrition.

**Key words:** Neural network, cardiovascular system diseases, predicting models, supervised learning, activation function, Keras.

**Introduction.** Cardiovascular system diseases, in particular chronic heart failure, cardiac ischemia and arterial hypertension are a public health problem in economically developed countries. One of the obvious and main reasons is the complexity of the initial diagnosis and the reluctance of people to visit clinics and hospitals when the first symptoms appear. A fast pace of life, poor nutrition, a sedentary lifestyle, addictions, and a lack of proper sleep all negatively affect the cardiovascular system. According to World Health Organization in 2016, more than 17.9 million people died from coronary heart disease, which amounted to 31% of all deaths<sup>[1]</sup>. Given the prevalence of cardiovascular diseases in developed countries, there is an urgent need to create a tool for the diagnosis of cardiovascular disease in the early stages of development. A neural network can be one of solutions to the problem due to its accessibility and ease of use.

The first attempts to create diagnostic neural networks date back to the 40-50s of the XX century, after the publication of the fundamental works of W. McCulloch, W. Pitts and F. Rosenblatt, who laid the theoretical foundations of neural networks<sup>[2]</sup>.

The list of medical spheres where neural networks have begun to be applied for diagnosing diseases are very long, and it is increasing from year to year. The accessibility of the neural networks courses, electronic patient records, a large number of articles about implementing of neural networks, the ability to create neural networks on ordinary laptops contributes to the development and application of neural networks not only in medicine, but also in other areas of human activities<sup>[3]</sup>.

Neural networks largely borrow their structure and principles of work from the cerebral cortex. Like the brain, they learn from life examples, extract and encode this knowledge in the form of synoptic neural connections. After training and testing, neural networks become mathematical models. This means that you can experiment with them, solve practical problems, predict behavior in the subject area (predicting the development of the disease), and diagnose the state of the subject area<sup>[4]</sup>.

The selection of input parameters with high information value, in other words parameters that affect the result of the model, is of great importance. The choice of parameters that have significant impact on the diagnosis of cardiovascular disease is not obvious. Therefore, as the input parameters for the neural network, all provided parameters should be selected that characterize the patient and are able to influence the result. On the other hand, in order to create a neural network accessible for mass use, the most accessible data were used, which did not require complex analyzes in hospitals.

**Neural network.** A neural network is a mathematical model and its software embodiment, which operates on the principle of a biological neural network, where each neuron is connected to another neuron from the previous layer. Each neuron receives and sends signals to other neurons, which provides a connection between them by synapses.

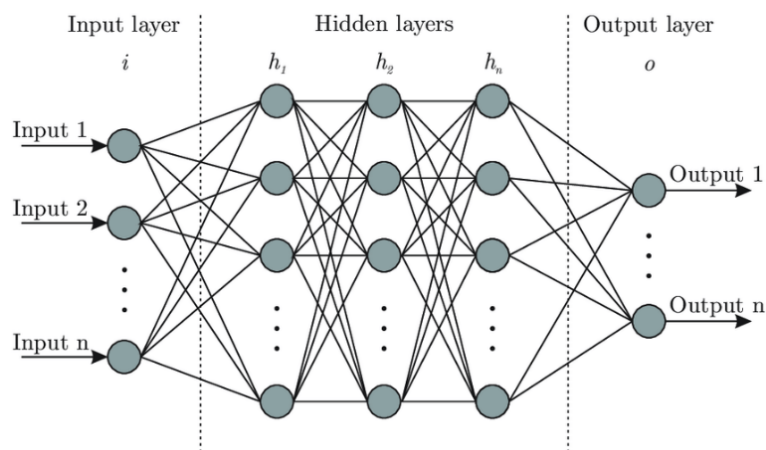


Figure 1 – General structure of neural network

A neuron is a computing unit that receives information, performs simple calculations on it, and passes the changed information on. Neurons are mainly divided into 3 groups: input, hidden and output as shown in Figure 1. There is also a displacement neuron and a contextual neuron. In the case when the number of neurons is large, the term layer is used. Each neuron has input and output data. In the case of an input neuron, income data is equivalent to output. In other cases, the income data is a summary of data from all previous layers. Further, this input information is normalized using the activation function, usually written  $f(x)$ . Neurons operate with numbers in the range of  $[0,1]$  or  $[-1,1]$ . Therefore, it is necessary to carry out the normalization process.

A synapse is a connection between two neurons that has a weight, usually denoted by  $w$ . Due to this weight, the input information changes during transmission from one neuron to another. The higher the weight of one neuron, the more significant the information transmitted to the next neuron will be. Thanks to these scales, the input information is processed and turned into a result. During the initialization of the neural network, weights are randomly distributed.

The activation function is a way to normalize input data, in other words, if you have a large number at the input, going through the activation function, this number will be reduced in the required range. The most commonly used activation functions are: linear, sigmoid (logistic) and hyperbolic tangent. The linear function (1) is used in cases where it is necessary to transmit information without changes or to test a neural network. The sigmoid (2), also called the logistic function, takes values in the range  $[0,1]$  and is the most common activation function. It makes sense to use hyperbolic tangent (3) when the values can be negative and positive. It has a range of  $[-1,1]$ . Using hyperbolic tangent, when the values are only positive, will adversely affect the accuracy of the neural network.

$$f(x) = x \tag{1}$$

$$f(x) = \frac{1}{(1+e^{-x})} \tag{2}$$

$$f(x) = \frac{(e^{2x}-1)}{(e^{2x}+1)} \tag{3}$$

Neural network training is divided into two large clusters: supervised learning and unsupervised learning. Supervised learning is a method of building a neural network, in which the results of observations from dataset, which is used for training the model, are already known initially and the neural network learns to correctly determine the future results using examples. Supervised learning is used for problems such as regression and classification. In unsupervised learning, the result of observation in the dataset is not initially known. Most often, this approach is used when it is necessary to group data by certain parameters or to create a recommendation system.

**Model building.** For building any predicting model. As the dataset on which the neural network was built, the file heart.csv<sup>[5]</sup> was used. The file contains information about 303 patients and 14 significant variables, among which:

- 1) age;
- 2) sex;
- 3) cp: chest pain type (value 1: typical angina, value 2: atypical angina, value 3: non-anginal pain, value 4: asymptomatic);
- 4) trestbps: resting blood pressure (in mm Hg on admission to the hospital);
- 5) chol: serum cholestorol (in mg/dl);
- 6) fbs: fasting blood sugar > 120 mg/dl (1 = true; 0 = false);
- 7) restecg: resting electrocardiographic result (value 0: normal, value 1: having ST-T wave abnormality [T wave inversions and/or ST elevation or depression of > 0.05 mV], value 2: showing probable or definite left ventricular hypertrophy by Estes' criteria);
- 8) thalach: maximum heart rate achieved;
- 9) exang: exercise induced angina (1 = yes; 0 = no);
- 10) oldpeak = ST depression induced by exercise relative to rest;
- 11) slope: the slope of the peak exercise ST segment (value 1: upsloping, value 2: flat, value 3: downsloping);
- 12) ca: number of major vessels (0-3) colored by fluoroscopy;
- 13) thal: 3 = normal; 6 = fixed defect; 7 = reversible defect;
- 14) num: diagnosis of heart disease or angiographic disease status (value 0: < 50% diameter narrowing, value 1: > 50% diameter narrowing).

In [3]: `data.head(10)`

Out[3]:

	age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	ca	thal	target
0	63	1	3	145	233	1	0	150	0	2.3	0	0	1	1
1	37	1	2	130	250	0	1	187	0	3.5	0	0	2	1
2	41	0	1	130	204	0	0	172	0	1.4	2	0	2	1
3	56	1	1	120	236	0	1	178	0	0.8	2	0	2	1
4	57	0	0	120	354	0	1	163	1	0.6	2	0	2	1
5	57	1	0	140	192	0	1	148	0	0.4	1	0	1	1
6	56	0	1	140	294	0	0	153	0	1.3	1	0	2	1
7	44	1	1	120	263	0	1	173	0	0.0	2	0	3	1
8	52	1	2	172	199	1	1	162	0	0.5	2	0	3	1
9	57	1	2	150	168	0	1	174	0	1.6	2	0	2	1

Figure 2 – Dataset for predicting cardiovascular disease risk

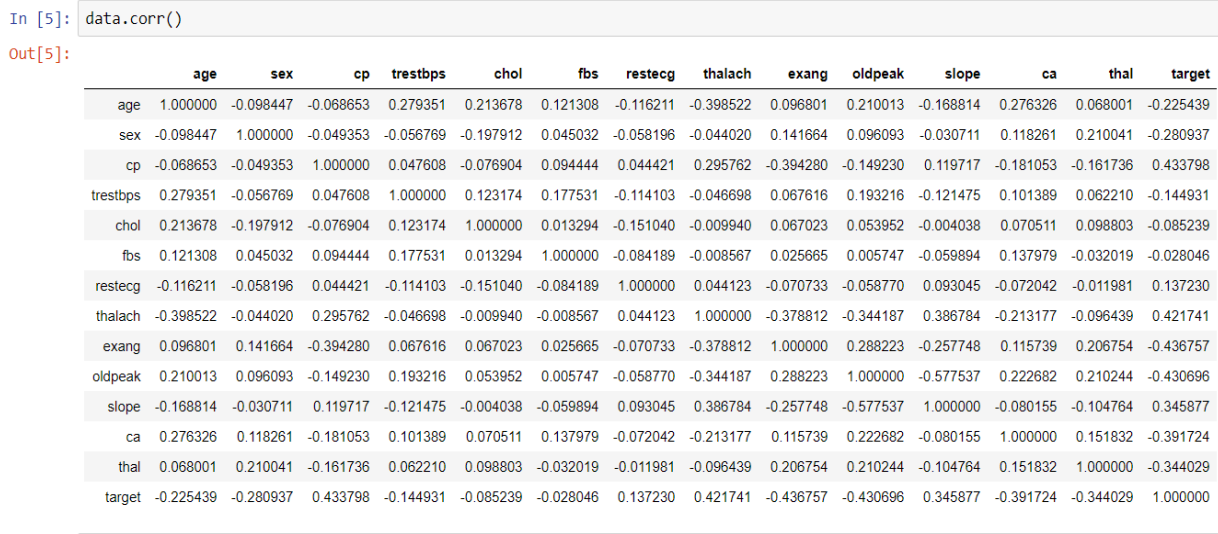


Figure 3 – Correlation matrix of the dataset

The model was built with open-source neural network library Keras written in Python<sup>[6]</sup>. Keras is designed for building deep neural networks. It is user-friendly and modular library, which includes blocks such as layers, optimizers, activation functions, objectives and host of tools to work with image and text data. Keras supports standard, convolutional and recurrent neural networks.

**Results.** The neural network, after training and optimization, was tested on data that was not involved in the learning process of the model. As a result, the neural network has an accuracy of 82% and a standard deviation of 0.026. The most valuable variables were chest pain type, maximum heart rate achieved, exercise induced angina, ST depression induced by exercise relative to rest and number of major vessels (0-3) colored by fluoroscopy.

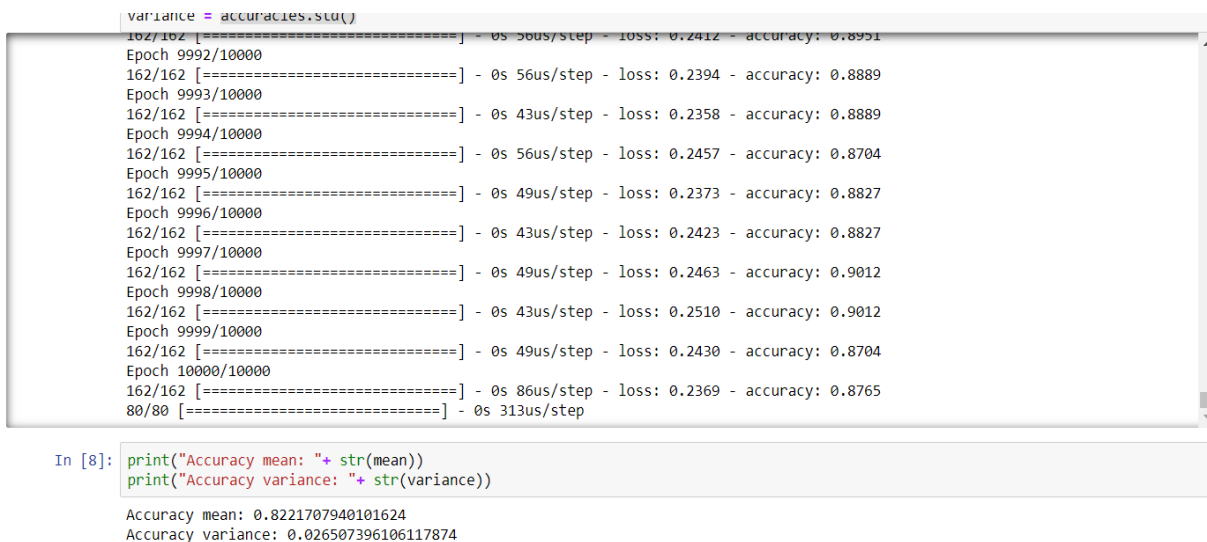


Figure 4 – Results of the neural network model and part of training algorithm

**Conclusion.** A neural network has been developed to predict the risk of cardiovascular disease. The standard deviation was 0.026, which allows us to state that the model is stable and can be used to identify patterns in this area of medicine, including those pattern that are not explicit. Such patterns can be revealed by experiments. By changing the patient's input parameters, for example, age, increase in blood pressure, it will be possible to analyze changes and the development of diseases in the future. You can try to improve the predictive diagnosis of the patient by changing the input parameters. Refusal from

addictions, a healthy lifestyle, proper nutrition, physical activity - all these parameters can change the predictive diagnosis for the better. By observing changes in input parameters and predicted results, it is possible to identify the dependence, on the basis of which you can develop an individual patient treatment plan.

Although the neural network has high predictive power, it is still the primary diagnosis of cardiovascular disease. And more accurate and complete analysis requires visiting the hospital.

**Марат Нуртас<sup>1</sup>, Ж.Д. Байшемиров<sup>2,3</sup>, В. Цай<sup>1</sup>,  
М. Тастанов<sup>1</sup>, Ж. Жанабеков<sup>1</sup>**

<sup>1</sup>Халықаралық ақпараттық технологиялар университеті, Алматы, Қазақстан,

<sup>2</sup>Абай атындағы Қазақ ұлттық педагогикалық университеті, Алматы, Қазақстан,

<sup>3</sup>Ақпараттық және есептеуіш технологиялар институты, Алматы, Қазақстан

### **ЖҮРЕК-ҚАН ТАМЫРЛАРЫ АУРУЛАРЫНЫҢ ҚАУПІН БОЛЖАУ ҮШІН НЕЙРОНДЫҚ ЖЕЛІНІ ПАЙДАЛАНУ**

**Аннотация.** Бұл мақалада экономикалық дамыған елдерде жүрек-қан тамырлары ауруларының таралуы туралы айтылады. Осы мақаланың мақсаты пациенттің жеке ерекшеліктеріне негізделген, жүрек-қан тамырлары ауруларының қаупін анықтау үшін нейрондық желі құру болып табылады. Модель Python-да жазылған Keras нейрондық желілік кітапхананың көмегімен жасалды. Керас терең нейрондық желілерді құруға арналған. Бұл пайдаланушыға ыңғайлы және модульді кітапхана, оның құрамына қабаттар, оптимизаторлар, активтендіру функциялары, мақсаттар, кескіндермен және мәтіндік деректермен жұмыс істеуге арналған құралдар жиынтығы кіреді. Керас стандартты, үйірткілі және рекурентті нейрондық желілерді қолдайды. Үлгіні құру үшін қажет пациенттер туралы мәліметтер Kaggle.com сайтынан алынды. Нейрондық желінің дәлдігі 82% құрайды.

Жүрек-қан тамыр жүйесінің аурулары, атап айтқанда созылмалы жүрек жетіспеушілігі, жүрек ишемиясы және артериялық гипертензия, экономикалық дамыған елдерде қоғамдық денсаулық сақтау мәселесі болып табылады. Бұл жағдайдың басты және анықталған себептерінің бірі - алғашқы диагноздың күрделілігі және де алғашқы белгілер пайда болған кезде адамдардың емханалар мен ауруханаларға барғысы келмеуі. Өмірдің жылдам қарқыны, дұрыс тамақтанбау, отырықшы өмір салты, жаман әдеттер және дұрыс ұйқының болмауы жүрек-қан тамыр жүйесіне кері әсерін тигізеді.

Дүниежүзілік денсаулық сақтау ұйымының мәліметтері бойынша, 2016 жылы жүректің ишемиялық ауруынан 17,9 миллионнан астам адам қайтыс болды, бұл барлық өлімнің 31% құрады [1]. Дамыған елдерде жүрек-қан тамыр ауруларының кең таралуын ескере отырып, дамудың алғашқы кезеңдерінде жүрек-қан тамыр ауруларының диагнозын анықтайтын құрал қажет. Өзінің қол жетімділігі мен қолданылуының қарапайымдылығына байланысты нейрондық желі осы мәселені шешудің бір жолы болуы мүмкін. Жүрек-қан тамыр ауруларының қаупін болжауға арналған нейрондық желі әзірленді. Желінің стандартты ауытқуы 0.026 құрады, бұл модельдің тұрақтылығын және медицинаның осы саласындағы, оның ішінде анықталмайтын модельдерді анықтау үшін пайдаланылуы мүмкін екендігін пайымдауға мүмкіндік береді. Мұндай заңдылықтарды тәжірибелер арқылы анықтауға болады. Оқыту мен оңтайландырудан кейін нейрондық желі модельдік оқыту процесіне қатыспайтын мәліметтер негізінде тексерілді. Ең құнды ауыспалыларға кеудедегі ауырсыну түрлері, максималды жүрек соғуы, физикалық белсенділіктен туындаған стенокардия, демалыспен салыстырғанда физикалық белсенділіктің нәтижесінде пайда болған ST депрессиясы және флюорокопиямен боялған ірі тамырлардың саны (0-3) кірді.

Пациенттің енгізу параметрлерін, мысалы, жасын, қан қысымының жоғарылауын өзгерту, болашақта аурулардың өзгеруі мен дамуын талдауға мүмкіндік береді. Пациенттің енгізу параметрлерін өзгерту арқылы оның болжамды диагнозын жақсартуға мүмкін болады.

Жаман әдеттерден бас тарту, салауатты өмір салты, дұрыс тамақтану, физикалық белсенділік сияқты параметрлер болжамды диагнозды жақсы жаққа өзгерте алады. Енгізу параметрлерінің өзгеруін және болжамды нәтижелерді байқау арқылы тәуелділікті анықтауға болады, соның негізінде пациенттің жеке емдеу жоспарын жасауға болады.

**Түйін сөздер:** нейрондық желі, жүрек-тамыр жүйесі аурулары, болжау модельдері, бақыланатын оқыту, активация функциясы, Keras.

Марат Нуртас<sup>1</sup>, Ж.Д. Байшемиров<sup>2,3</sup>, В. Цай<sup>1</sup>,  
М. Тастанов<sup>1</sup>, Ж. Жанабеков<sup>1</sup>

<sup>1</sup>Международный университет информационных технологий, Алматы, Казахстан;

<sup>2</sup>Казахский национальный педагогический университет имени Абая, Алматы, Казахстан;

<sup>3</sup>Институт информационных и вычислительных технологий КН МОН РК, Алматы, Казахстан

## ПРИМЕНЕНИЕ НЕЙРОННОЙ СЕТИ ДЛЯ ПРОГНОЗИРОВАНИЯ РИСКА СЕРДЕЧНО-СОСУДИСТЫХ ЗАБОЛЕВАНИЙ

**Аннотация.** Данная статья касается проблемы распространения сердечно-сосудистых заболеваний в экономически развитых странах. Цель этой статьи заключается в создании нейронной сети для определения риска сердечно-сосудистых заболеваний на основе индивидуальных характеристик пациента. Модель была построена с открытым исходным кодом нейронной сетевой библиотеки Keras, написанной на Python. Keras разработан для построения глубоких нейронных сетей. Это удобная для пользователя и модульная библиотека, которая включает такие блоки, как слои, оптимизаторы, функции активации, цели и набор инструментов для работы с изображением и текстовыми данными. Keras поддерживает стандартные, сверточные и рекуррентные нейронные сети. Данные, содержащие информацию о пациенте для построения модели, были взяты с сайта Kaggle.com. Точность нейронной сети составляет 82%.

Заболевания сердечно-сосудистой системы, в частности хроническая сердечная недостаточность, ишемия сердца и артериальная гипертензия являются проблемой общественного здравоохранения в экономически развитых странах. Одной из очевидных и основных причин является сложность первоначального диагноза и нежелание людей посещать клиники и больницы, когда появляются первые симптомы. Быстрый темп жизни, плохое питание, сидячий образ жизни, вредные привычки и отсутствие надлежащего сна негативно влияют на сердечно-сосудистую систему.

По данным Всемирной организации здравоохранения в 2016 году, от ишемической болезни сердца умерло более 17,9 миллиона человек, что составило 31% от всех смертей [1]. Учитывая распространенность сердечно-сосудистых заболеваний в развитых странах, существует острая необходимость в создании инструмента для диагностики сердечно-сосудистых заболеваний на ранних этапах развития. Нейронная сеть может быть одним из решений проблемы из-за ее доступности и простоты использования. Нейронная сеть была разработана для прогнозирования риска сердечно-сосудистых заболеваний. Стандартное отклонение составило 0,026, что позволяет нам утверждать, что модель стабильна и может быть использована для идентификации закономерностей в этой области медицины, включая те, которые не являются явными. Такие закономерности могут быть выявлены в ходе экспериментов. Нейронная сеть после обучения и оптимизации была протестирована на основе данных, которые не были задействованы в процессе обучения модели. Наиболее ценными переменными были типы боли в груди, максимальная частота сердечных сокращений, стенокардия, вызванная физической нагрузкой, депрессия ST, вызванная физической нагрузкой, по сравнению с отдыхом, и количество крупных сосудов (0-3), окрашенных с помощью флюороскопии.

Изменение входных параметров пациента, например, возраст, повышение кровяного давления, позволит анализировать изменения и развитие заболеваний в будущем. Можно попытаться улучшить прогнозную диагностику пациента, изменив входные параметры.

Отказ от вредных привычек, здоровый образ жизни, правильное питание, физическая активность – все эти параметры могут изменить прогнозирующий диагноз в лучшую сторону. Наблюдая изменения входных параметров и прогнозируемых результатов, можно определить зависимость, на основе которой можно разработать индивидуальный план лечения пациента.

**Ключевые слова:** нейронная сеть, заболевания сердечно-сосудистой системы, модели прогнозирования, контролируемое обучение, функция активации, Keras.

**Information about authors:**

Nurtas Marat, PhD, Associate Professor of Department of Mathematical and Computer Modeling, International Information Technologies University; [maratnurtas@gmail.com](mailto:maratnurtas@gmail.com), <https://orcid.org/0000-0003-4351-0185>;

Baishemirov Zh.D., PhD, Associate Professor of Abai Kazakh National Pedagogical University, Senior Research Officer of RSE Institute of Information and Computational Technology CS MES RK; [zbai.kz@gmail.com](mailto:zbai.kz@gmail.com), <https://orcid.org/0000-0002-4812-4104>;

Tsay V., master student, Department of Mathematical and Computer Modeling, <sup>1</sup>International Information Technologies University; [tsay.victor96@gmail.com](mailto:tsay.victor96@gmail.com), <https://orcid.org/0000-0002-6073-2030>;

Tastanov M., Master Student of Department of Mathematical and Computer Modeling, International Information Technologies University; [tychty4@gmail.com](mailto:tychty4@gmail.com), <https://orcid.org/0000-0002-1802-6394>;

Zhanabekov Zh., Master Student of Department of Mathematical and Computer Modeling, International Information Technologies University; [zzhanabekov@gmail.com](mailto:zzhanabekov@gmail.com), <https://orcid.org/0000-0001-5984-7132>;

**REFERENCES**

[1] World Health Organization, “Cardiovascular diseases (CVDs)”, available at [https://www.who.int/en/news-room/fact-sheets/detail/cardiovascular-diseases-\(cvds\)](https://www.who.int/en/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)) (accessed 5 March 2020).

[2] Yasnitsky, L.N., Dumler, A.A., Poleschuk, A.N., Bogdanov, K.V., Cherepanov, F.M. (2011), “Neuronetwork system of cardiovascular diseases express-diagnosis”, *Perm Medical Journal*, Vol. XXVIII № 4, pp. 77-86.

[3] Kim, J.K., Sanggil K. (2017), “Neural Network-Based Coronary Heart Disease Risk Prediction Using Feature Correlation Analysis”, *Journal of healthcare engineering*, vol. 2017: 2780501. doi:10.1155/2017/2780501

[4] Mazakov T. Zh., Jomartova Sh. A., Wojcik Waldemar, Ziyatbekova G. Z., Amirkhanov B. S., Zholmagambetova B. R. “Universal complex of psychophysiological testing”, *News of the National Academy of sciences of the Republic of Kazakhstan*, Volume 2, Number 330 (2020), 14 – 22. <https://doi.org/10.32014/2020.2518-1726.10>

[5] Kaggle, Heart Disease UCI, available at <https://www.kaggle.com/ronitf/heart-disease-uci> (accessed 25 March 2020)

[6] Wikipedia, “Keras”, available at <https://en.wikipedia.org/wiki/Keras> (accessed 4 April 2020).