DOI: https://doi.org/10.15276/hait.07.2024.26
UDC 004. 93

Heterogeneous ensemble classifier in computer systems for medical diagnostics

Anatoly I. Povoroznyuk[1)]
ORCID: http://orcid.org/0000-0003-2499-2350; ai.povoroznjuk@ gmail.com. Scopus Author ID: 55225664000
Oksana A. Povoroznyuk[1)]
ORCID: http://orcid.org/0000-0001-7524-5641; povoks76@gmail.com. Scopus Author ID: 55817007400
Anna E. Filatova[1)]
ORCID: https://orcid.org/0000-0003-1982-2322; filatova@gmail.com/ Scopus Author ID: 56448583600
[1)] National Technical University “Kharkiv Polytechnic Institute”, 2, Kyrpychova Str. Kharkiv, 61002, Ukraine

ABSTRACT

The work is dedicated to the solution of an important scientific and technical problem: building a diagnostic decision-support system in medicine. The foundation of this system is a model developed as a heterogeneous ensemble classifier, which implements two primary approaches to formulating a diagnostic conclusion through basic models. The first of these approaches is probabilistic. It is based on the analysis of a training sample of patients with a confirmed diagnosis, which enables estimation of the probability of the presence of a particular disease based on available data. The second approach is expert-based, relying on expert information about the structure of symptom complexes that characterize each individual disease. It is important to note that both of these approaches address the same problem from different perspectives, and their combined use holds great promise for developing effective diagnostic systems. The purpose of this study is to synthesize a heterogeneous ensemble classifier that integrates both expert and probabilistic components into the diagnostic process. An analysis of various diagnostic methods used by doctors in alignment with the current requirements of evidence-based medicine was carried out as part of the study. Methods of constructing diagnostic decision rules in medical decision-support systems were also considered. Based on these studies, a mathematical model of a heterogeneous ensemble classifier was developed, with the choice of its constituent parts being justified. Widely used classification methods were selected as the probabilistic component in this system, particularly the standard comparison method, the *k*-nearest neighbors method, and the potential functions method. Expert knowledge concerning the structure of symptom complexes is formalized by expressing the symptom complexes of each disease in the form of numerical intervals. In this framework, linguistic variables are used, which can indicate “below the norm”, “norm”, or “above the norm”. Various strategies for aggregating different types of basic models within the heterogeneous ensemble classifier are reviewed. This approach preserves the advantages of each method and enhances the overall classification accuracy. Requirements for the developed system's functionality were formulated, design tools and the main development platform (Java) were defined, and the database management system (MySQL). The decision-support system was designed, and a comprehensive evaluation of the developed system was conducted on real medical data. The results of these tests confirmed the effectiveness of the system.

Keywords: Medical diagnostics; ensemble classifier; base model; probabilistic classifier; symptom complex; expert information; model aggregations; decision support system

***For citation*: Povoroznyuk A. I., Povoroznyuk O. A., Filatova A. E. “Heterogeneous ensemble classifier in computer systems for medical diagnostics”. *Herald of Advanced Information Technology.* 2024; Vol. 7 No.4: 371–383. DOI: https://doi.org/10.15276/hait.07.2024.26**

INTRODUCTION

The development of a high-tech society contributes to implementing scientific and technological advancements in the important yet insufficiently formalized field of medical diagnostics. Today, there are seven levels of medical information systems (MIS) [1, 2], [3], ranging from electronic medical records to intelligent decision-support systems [4, 5] that integrate synergistic databases [6], artificial intelligence methods [5, 7], and telemedicine capabilities [8, 9], [10].

Most of these systems are aimed at automating the examination of patients in various areas of health care [11, 12], [13], focusing on the detection and registration of diagnostic signs (disease markers) [14, 15] and the processing of biomedical signals [16, 17], [18] and images [19, 20], [21].

However, despite this progress, the problem of synthesizing diagnostic decision rules remains relevant, as the reliability and accuracy of diagnostic conclusions largely depend on the quality and structure of these rules.

One approach to solving this problem is the integration of medical information from various sources and the development of complex diagnostic models to provide a deeper understanding of the patient's condition and, as a result, improve the quality of medical decisions [22, 23].

The ethical aspects of implementing such systems are also important, particularly in protecting patients' personal data and ensuring transparency in decision-making.

GENERAL PROBLEM STATEMENT

Diagnosis, derived from the Greek word διάγνωσις, means recognition or identification. It is a medical conclusion that characterizes functional health, the patient's morphological condition, as well as existing diseases or injuries. A diagnosis may include observed changes in the body and explanations of the causes that led to a fatal outcome. Accordingly, a diagnosis is a structured conclusion based on established classifications of diseases.

Traditional primary diagnosis is based on a comprehensive examination of the patient, an analysis of medical history (anamnesis), patient complaints, and objective symptoms revealed during a physical examination, as well as instrumental and laboratory testing.

The practicing doctor uses the following logic for diagnosis, which corresponds to the principles of evidence-based medicine [24]:

– detection of deviations from normal indicators of the body's condition;

– interpretation of deviations and their significance (whether they are of physiological or pathological nature, and their possible connection to a specific nosological form);

– the detected signs (symptoms) of the disease are combined by the doctor into symptom complexes (a set of symptoms that share a common pathogenesis and manifest in a specific disease), forming the basis for a diagnostic conclusion.

Important factors in this process are the professional skills of the doctor, as well as the experience of clinicians in this medical specialty, which is documented in medical handbooks and can be considered an expert assessment of each disease.

In modern MIS, diagnostics is formalized as a classic classification problem, where the model of the diagnostic object is perceived as a “black box” and the relationship between the formalized states of the diagnostic object Y and the vector of input features X, i.e., $Y = f(X)$ [1, 25], [26]. This probabilistic relationship is established during the training stage of the classifier model by analyzing a training sample of patients with confirmed diagnoses. However, there are significant challenges. First, it is not always possible to form a representative sample, especially for rare diseases. Second, practicing doctors do not always trust the results of computer diagnostics and may experience difficulties in interpreting these diagnostic results.

Since these approaches consider the same problem from different points of view, a combination of traditional and formalized diagnostic methods is a promising direction. The combination of various classifier models and the aggregation of their results is possible by developing a heterogeneous ensemble classifier, which increases the accuracy and reliability of individual classifiers and allows doctors to use computer system results more confidently in their practice.

ANALYSIS OF RECENT RESEARCH AND PUBLICATIONS

Today, there is a wide range of computer decision support systems (CDS), which are actively used in various fields, particularly in medicine, where their role is extremely important [4, 5], [6]. These systems typically use a variety of mathematical techniques that significantly contribute to the decision-making process. Among them, several main categories can be distinguished, which differ in their nature and approaches to data analysis.

First, deterministic methods [1, 27] include those based on clear and defined rules and algorithms. These methods generally provide accuracy and reproducibility of results, which is an important factor in critical situations such as medical diagnoses.

Secondly, probabilistic methods [28, 29] use statistical approaches to estimate the probability of events and possible outcomes. They are especially useful in conditions of uncertainty, where traditional deterministic models may not be sufficient for adequate decision-making.

Methods based on pattern recognition theory are also an important area [1, 30]. These methods allow for the automatic classification of data and the detection of patterns, which has great potential in medicine, where, for example, medical images can be analyzed to detect pathologies.

Equally significant is fuzzy logic [31, 32], which accounts for uncertainty and vagueness in data. This approach opens up new opportunities for modeling complex systems where traditional binary solutions may not be effective.

Logical-linguistic methods [33] also deserve attention because they combine logic and language, allowing for the consideration of human factors and subjective evaluations during decision-making.

In addition, methods based on artificial neural networks [32, 34] are becoming increasingly popular due to their self-learning and adaptive capabilities. They allow for the modeling of complex relationships between data, which makes them extremely useful in medicine for diagnosis, prediction of treatment outcomes, and personalization of therapy.

Each of these methods has its own unique features and advantages that make them useful in different contexts. In this article, we will consider them in more detail, analyzing how they can be implemented in the practice of supporting decision-making in medical systems and what prospects are open to them in the context of rapid technological development.

Let's take a closer look at the most commonly used methods.

Deterministic methods are used in cases with clear, deterministic connections between features and formalized states of objects. These methods are generally applied at the preliminary classification stage, where exact algorithms based on rules of the type

IF <*condition*> THEN *Res1*; ELSE *Res* 2

can provide a high level of accuracy. Deterministic approaches are also often combined with fuzzy logic techniques to construct fuzzy rules. In this work, fuzzy logic methods are used to formalize expert assessments of the structure of symptom complexes. This approach also allows for considering subjective factors that affect the diagnosis.

Probabilistic or statistical methods are based on mathematical statistics [35]. They are usually applied in cases where the probabilistic characteristics of the classes are known or can be determined from the available training sample. However, it is worth noting that such methods have certain limitations, as they require a sufficient amount of data to obtain adequate estimates, which narrows their scope of application.

Methods based on pattern recognition theory. This group of methods includes various techniques that allow for presenting the results of measuring object characteristics as points in the space of diagnostic features. In this space, different classes of objects should form compact sets. Diagnosis of a new object is then reduced to calculating its degree of proximity to each of the classes. Among the most common methods in this group are the prototype (standard) comparison method, the *k*-nearest neighbor’s method, and the potential functions method. An important part of implementing these methods is the model training stage using the training sample, as well as the classification stage for new objects, which ensures system flexibility and adaptability.

These methods are used in this work as a probabilistic component of the diagnostic decision rule. The combination of different approaches enables the creation of a more stable and accurate system that can operate effectively under conditions of incomplete or inaccurate information, which is a common occurrence in medical practice. This opens new horizons for the development of decision support systems that can significantly improve the quality of medical services and assist doctors in their work.

PURPOSE AND OBJECTIVES OF THE RESEARCH

The purpose of the work is to develop a heterogeneous ensemble classifier that takes into account both the expert and probabilistic components of the diagnosis process.

To achieve the goal, the following tasks are solved:

– to develop a mathematical model of a heterogeneous ensemble classifier, in which the basic models are classifiers using methods based on pattern recognition theory and expert information on the structure of the symptom complex;

– to carry out the formalization of expert information for quantitative and binary diagnostic signs and, on its basis, to develop a standard model for each class of the disease, based on the conclusions of experts regarding the structure of symptom complexes;

– to develop scenarios for the joint use of methods based on the analysis of the training sample and the formalization of expert information on the structure of the symptom complex;

– to develop system software architecture and perform testing based on real medical data.

DEVELOPMENT OF A HETEROGENEOUS ENSEMBLE CLASSIFIER MODEL

The combination of individual classifiers through ensemble methods of machine learning is one of the most effective ways to reduce the impact of random errors or shortcomings of individual classifiers, as well as to increase the accuracy of diagnostics in various fields of application [36, 37]. Today, there are various methods of aggregation, among which bagging (Bootstrap Aggregating), boosting, and stacking stand out [38, 39]. These methods achieve significantly better results than using base models individually, thanks to their ability to combine the advantages of several models.

Bagging, in particular, involves creating multiple random subsamples of data through repeated sampling. Different models are trained on each of these subsamples, and the results of their predictions are combined to form a final conclusion. This method not only reduces variability but also makes the model more resistant to random errors. Thanks to this strategy, even if one of the models does not work quite correctly, the others can compensate for its shortcomings.

Boosting, in turn, is based on the idea of sequential training of models. Each new model focuses on the mistakes made by previous models. This approach greatly reinforces the importance of correct forecasting, especially in cases where the data contain extreme or anomalous values. Thanks to this, boosting ensures high diagnostic accuracy, which is critically important in fields such as medicine, where even minor errors can have serious consequences.

Stacking is another powerful strategy that combines the predictions of different models using meta-models that are trained on the outputs of the underlying models. This strategy allows you to efficiently combine the results of different classifiers, taking into account the diversity of their predictions. However, it should be noted that it requires significant computing resources and complex parameter tuning, which may become a certain obstacle to its wide implementation.

Homogeneous ensembles consist of classifiers of the same type that are trained independently on different training samples. The results of these classifiers are combined to obtain a final prediction. This approach provides some simplicity in implementation and understanding but can limit the variety of models, which can affect overall accuracy.

Heterogeneous ensembles, on the contrary, combine different methods of machine learning.

This allows various aspects of the data to be considered, as different algorithms may have different strengths and weaknesses. This approach opens up possibilities for more detailed analysis and improvement of the overall performance of the model. For example, one algorithm may be more effective at detecting certain patterns, while another may be better at handling noise in the data.

Thus, it is advisable to develop a heterogeneous ensemble classifier for the implementation of the above-mentioned approaches when making a computer diagnosis in decision support systems. Such a classifier can become an effective diagnostic decision rule that will improve the accuracy and reliability of medical predictions. Thanks to the combination of different machine learning methods, it is possible to leverage their strengths and reduce the risk of errors, which, in turn, will contribute to improving the quality of medical services and enhancing treatment outcomes.

To develop a heterogeneous ensemble classifier as part of a computer system for supporting diagnostic decision-making in medicine, let's consider several basic models that can be used as probabilistic components.

The method of comparison with the prototype is an effective approach for the analysis of quantitative characteristics when the classes Ω_m $(m=\overline{1,M})$ form compact sets of objects having a spherical shape in the functional space of features. This method is particularly useful in situations where the data can be adequately described by geometric characteristics.

In this method, each class Ω_m $(m=\overline{1,M})$ is represented by the model of its prototype ω^{mr} which is chosen as the geometric center (centroid) of the class. At the training stage, the coordinates of each prototype are calculated based on the training sample of objects belonging to the corresponding class, according to the formula:

$$x_i^{mr} = \frac{1}{n_m}\sum_{j=1}^{n_m} x_i^j \,, (i=\overline{1,p}), \quad (m=\overline{1,M})\,, \tag{1}$$

where x_i^{mr} is the *i*-th coordinate of the prototype of the class *m*; n_m is the number of objects of the class *m* in the training sample; x_i^j is the *i*-th coordinate of the *j*-th object of the class *m* in the training sample; *p* is the size of the coordinate space (the number of diagnostic features); *M* is the number of classes into which unknown objects are classified (the number of diseases diagnosed in this field of medicine).

This formula allows you to find the average value of each feature for all objects in the class, which determines the position of the prototypes of each class in the feature space.

At the stage of classification, the new object ω is compared with standards ω^{mr} of all classes Ω_m. For this, various metrics can be used, such as Euclidean or Manhattan distance, cosine similarity, etc. For the e nested distance between the object ω and each of the standards ω^{mr} the measure of closeness $R(\omega,\omega^{mr})$ is calculated by the formula

$$R(\omega,\omega^{mr})=\sqrt{\sum_{i=1}^{p}(x_i^{\omega}-x_i^{\omega^{mr}})^2},\quad (m=\overline{1,M}),\tag{2}$$

where $x_i^{\omega^{mr}}$ is the *i*-th coordinate of the reference object of the class *m*; x_i^{ω} is *i*-th coordinates of the object to be classified; *p* is the size of the coordinate space (the number of diagnostic features); *M* is the number of classes into which classification is performed.

Based on the selected metric, the closest standard is determined, and the new object is classified according to the class Ω_t to which this standard belongs

$$R(\omega,\omega^{tr})=\min_{m=\overline{1,M}} R(\omega,\omega^{mr}),\tag{3}$$

where $R(\omega,\omega^{mr})$ is the distance between the object ω and the prototype ω^{mr} of the class Ω_t which is calculated according to expression (2).

The Fix–Hodges method, also known as the *k*-nearest neighbor's method, is a popular classification algorithm that is actively used in cases where the class structure is complex and far from spherical. This method is based on the hypothesis of the continuity of the multidimensional density of the distribution of classes in each local region of the feature space. The main essence of this method is to determine a certain predetermined number *k* of objects of the training sample closest to the unknown object (in the selected metric), which are called "nearest neighbors". The unknown object ω belongs to the class whose number of representatives prevails among these *k* neighbors. The value of *k* is a positive integer, which is usually small. For example, if $k=1$, then the object is simply assigned to the class of this single nearest neighbor, which simplifies the classification process.

Implementation of the method. The method implementation process consists of several main stages. First, the distances between the unknown object ω and all the objects of the training sample, which represent the set of all classes Ω_m $(m=\overline{1,M})$, are determined:

$$R(\omega,\omega_i)=\sqrt{\sum_{j=1}^{p}(x_j^{\omega}-x_j^{\omega_i})^2}\quad (i=\overline{1,N}),\tag{4}$$

where *N* is the size of the training sample.

After that, the found distances are ranked in ascending order, and the first *k* elements are selected as the “nearest neighbors”. The class Ω_t is determined among these neighbors, the number of representatives of which predominates among the selected *k* “nearest neighbors”. Thus, the unknown object ω belongs to class Ω_t, and this classification process is both intuitive and effective.

Since the method uses a distance function for classification, if the features represent different physical units or have very different scales, it is necessary to normalize the training data. This prevents situations where some features may disproportionately influence the outcome due to their magnitude.

Additionally, there is the possibility of assigning weight values to the contributions of neighbors, allowing the influence of closer neighbors to be greater than that of those farther away. For example, each neighbor can be assigned a weight $\alpha_i=1/R(\omega,\omega_i)$, where $R(\omega,\omega_i)$ is the distance to the neighbor, which is calculated according to (4). This approach improves classification accuracy, because the nearest neighbors, which are more relevant for classification, will have a greater influence on the final result than those farther away.

Therefore, the *k*-nearest neighbor's method demonstrates its versatility and effectiveness in various fields of application, including medical diagnostics. Its ease of implementation and high performance at the same time make it a popular choice among researchers and practitioners in the field of machine learning.

The potential function method is an effective approach to object classification, based on using a potential function to determine an object's membership in a particular class. In this method, a positive, monotonically decreasing distance function serves as the potential function, characterizing the likelihood of an object belonging to a particular class. This function has a form that is analogous to the electric potential φ. Examples of such functions can be various mathematical expressions that allow describing the potential as a function of distance.

Formally, the potential can be described as follows:

$$\varphi_m(R(\omega,\omega_{m_i})) = \sum_{i=1}^{N_m} \left| \frac{\sin(\alpha R(\omega,\omega_{m_i})^2)}{\alpha R(\omega,\omega_{m_i})^2} \right|,$$

$$\varphi_m(R(\omega,\omega_{m_i})) = \sum_{i=1}^{N_m} e^{-\alpha R(\omega,\omega_{m_i})^2}, \qquad (5)$$

$$\text{or } \varphi_m(R(\omega,\omega_{m_i})) = \sum_{i=1}^{N_m} \frac{1}{1+\alpha R(\omega,\omega_{m_i})^2},$$

where $R(\omega,\omega_{m_i})$ is the distance between the source point, which represents the object ω_{m_i} $i=\overline{1,N_m}$ of the class Ω_m, $m=\overline{1,M}$ and the sink point ω, where the potential is calculated; N_m is the volume of the training sample of class Ω_m; M is the number of classes into which classification is performed.

An important element of this expression is the weight coefficient α which satisfies the condition $\alpha > 0$. This coefficient characterizes the rate of decrease of the potential $\varphi(R)$ with distance determining how quickly the potential decreases with distance from the source.

In this case, the points-sources of the potential are the objects that belong to the classes Ω_m, and the receiver point is the object ω that is subject to classification. It is important to note that the object ω will be assigned to the class Ω_t for which the total potential calculated according to (5) will be maximum. This means that the potential function method enables accurate classification of an object based on the distribution of potentials generated by other objects within the training sample. Using potential functions allows for modeling various aspects of distance and neighbor influence, which can lead to more accurate and reliable classification.

The Ficks–Hodges method and the potential function method are powerful tools for object classification, especially in scenarios where the data structure is complex and multidimensional. These methods allow taking into account the local structure of the data and provide an efficient approach to classification in a multidimensional feature space. They not only complement classical methods of classification but also open up new opportunities for research and application in various fields where classification is an important component of the decision-making process.

CONSIDERATION OF EXPERT INFORMATION ON THE STRUCTURE OF SYMPTOM COMPLEXES

The structure of the symptom complex is a reflection of the opinion of experts about a specific diagnosis developed by several generations of doctors. Information about the symptoms of diseases in a non-formalized form is given in various medical handbooks [40].

The symptom complex consists of different types of symptoms, which can be categorized as follows:

– Pathognomonic symptoms: These are symptoms that clearly indicate the presence of a specific disease and serve as definitive markers.

– Specific symptoms: These symptoms occur only in certain diseases but are not unequivocal indicators on their own.

– Non-specific symptoms: These symptoms may be characteristic of multiple diseases and do not point to a specific illness.

The structure of the symptom complex reflects the opinions of experts regarding a specific diagnosis. The impact of each symptom is calculated based on the linguistic variables that describe the corresponding symptoms (e.g. “high temperature” or “high blood pressure”).

Each symptom x_j is assigned an expert rating which indicates its significance in the symptom complex.

These ratings take the following values:

e_0 – for pathognomonic symptoms;

e_1 – for specific symptoms;

e_2 – for non-specific symptoms;

e_3 – for indicators that do not belong to the syndrome of this disease.

At the same time, inequality is valid

$$e_0 \ge e_1 \ge e_2 \ge e_3, \quad \sum_{i=0}^{3} e_i = 1. \qquad (6)$$

Let's formalize the experts' opinions regarding the structure of the symptom complex in the creation of an ensemble classification method. When diagnosing, doctors often rely on the concept of normality for each symptom.

The dynamic range of symptom values is divided into three categories: “below normal”, “normal”, and “above normal”. Similarly, binary features are defined, where linguistic variables are denoted by the terms “symptom is present” or "symptom is absent".

Therefore, each pathological syndrome (the prototype for each class) is determined by diagnostic symptoms x_j, each of which takes on the value of the corresponding linguistic variables. Descriptions of typical sets of symptoms, together with variants of linguistic variables, form a formalized expert evaluation of the prototypes for each disease class.

If the values of the dynamic range for all diagnostic features and the threshold values of "norms" are known, the centers of these ranges ("below normal", "norm", "above normal") serve as a quantitative representation of the prototype ω^{ml} of the class Ω_m, based on the conclusions of experts about the structure symptom complex, and not on training data. These prototypes can be used as base models in a heterogeneous ensemble classifier.

Let's consider two scenarios of the joint use of methods based on the analysis of the training sample and the formalization of expert information on the structure of the symptom complex.

Scenario 1: Aggregation of base models' results. In the classic stacking ensemble classifier diagram, the classification results of different models are combined through a meta-model. For each class Ω_m ($m=\overline{1,M}$), basic classification models are created based on the analysis of the training sample and expert assessments of the symptom complex structures. Classification of new objects is carried out by each base model according to the algorithms described above. If the classification results of different models differ, the results are aggregated and the final diagnosis D_k is formed by weighted voting.

Scenario 2: Aggregation of models in the prototype comparison method when determining the coordinates of class prototypes. The results of calculating the coordinates of class prototypes by different models are aggregated. For each class Ω_m ($m=\overline{1,M}$) separate prototypes are created (ω^{mr} based on the training sample and ω^{ml} based on expert assessments).

The coordinates of the prototype of the meta-model are calculated according to the formula:

$$x_i^m = k_1 x_i^{mr} + k_2 x_i^{ml},\ (i=\overline{1,p}),\ (m=\overline{1,M}), \qquad (7)$$

where $k_i > 0,\ \sum k_i = 1,\ \ i=\overline{1,2}$ are the weighting coefficients corresponding to the confidence in the training sample k_1 and in the expert assessment k_2 respectively.

The end-user expert should be able to adjust these values at their discretion and under their responsibility.

After training the meta-model, new objects are classified using the standard algorithms of the prototype comparison method by (2) and (3). The final diagnosis provided by the ensemble classifier serves as a decision support tool, ultimately aiding the physician in making the final diagnostic decision. Considering expert information on the symptom complex structure and the use of various classification methods allows for the creation of a more reliable and accurate heterogeneous ensemble classifier. This combination of expert opinion and statistical methods opens up new opportunities for improving the diagnostic process in medical practice.

IMPLEMENTATION OF THE DECISION SUPPORT SYSTEM AND TEST VERIFICATION

Based on the analysis of modern information technologies and tools, the methods of implementing the decision-making support system have been substantiated. The most attractive tools for software development today are those that enable to building of applications with cross-platform capabilities at the runtime level. This is especially important given the wide variety of platforms and operating systems in use today. Leaders in this field are powerful platforms such as .NET from Microsoft and Java from Oracle Corporation (formerly Sun Microsystems). These technologies offer developers a wide range of opportunities to create effective solutions.

The Windows operating system (OS) occupies nearly 90 % of the market for personal computer operating systems, emphasizing its popularity and importance in the field of information technology. Therefore, the Windows operating system was chosen as the foundation for our system. In addition, its APIs (application programming interfaces) provide the necessary capabilities to build an intuitive and functional application interface. Considering the economic component and the availability of the latest versions of runtime environments for most modern platforms such as Windows, UNIX, and Linux, Java was chosen as the main development platform.

This programming language allows for direct use of the Windows system API and provides flexible memory management capabilities, which are critical for the efficient functioning of the system. Java also has a large number of libraries that simplify working with graphical user interfaces, which significantly enhances developers' productivity. MySQL, a popular and powerful solution for data storage and processing, was chosen as the database design tool for our system.

Based on the analysis of the subject area, the requirements for the functional capabilities of the system are formulated, which should ensure the following tasks:

– working with the database;

– formation of basic models of classifiers and meta-models of a heterogeneous ensemble classifier;

– implementation of the diagnosis procedure.

The system assumes the role of a user physician who will use the system to perform a diagnostic procedure, an expert physician who forms expert information on the structure of symptom complexes, and an administrator who performs system debugging and prepares all necessary data structures.

The architecture of the decision support system software has been developed. In the development of the software, the main templates for constructing the project were identified. Each design pattern is used to solve specific tasks. In recent years, the MVC (Model-View-Controller) design pattern has gained popularity, which divides the system into three parts: the data model, the data view, and the controller. It is used to separate data (the model) from the interface (the view) so that changes to the interface minimally impact data handling, while changes to the data model can be made without altering the interface. The purpose of the pattern is to enable flexible software design, which should facilitate further changes or extensions of programs and provide the possibility of reusing individual components of the program. In addition, the use of this pattern in large systems leads to a certain order in their structure, making them more understandable by reducing complexity.

In the decision-making support system based on the developed heterogeneous ensemble classifier, mandatory components can be distinguished: the system kernel (which must ensure data handling, knowledge generation, and diagnostics), the database, and the user interface (following the MVC pattern in architectural terms). The simplified structure of the system is presented in Fig. 1.

The database stores data about patients, medical institution personnel, the structure and content of examinations, and the system's knowledge base.

The kernel consists of Java classes grouped into packages based on their functions, forming the main subsystems:

– the data management subsystem represents data from the database as Java-class objects of the system;

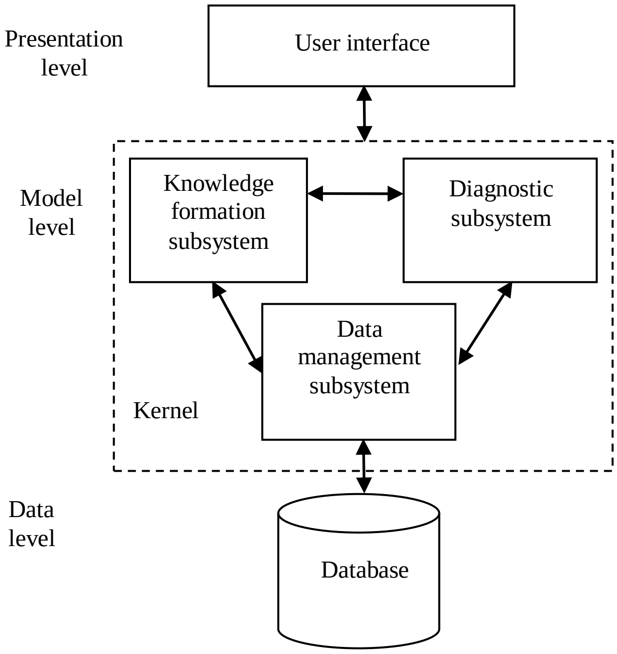

Fig. 1. **System component interaction diagram**

***Source*: compiled by the authors**

– the knowledge formation subsystem constructs sequences of sets of diagnostic features and diagnosable states, along with a formalized structure of symptom complexes;

– the diagnostic subsystem includes classes implementing descriptions of base models and the meta-model of the heterogeneous ensemble classifier and providing a diagnosis for the proposed diagnostic object.

The kernel interacts with the user interface, which allows system debugging and operation. The training sample data are stored in the MySQL database.

A training sample consisting of 200 patients was created to test the developed classifier. The diagnostic features consisted of 9 indicators from clinical blood and urine tests (Table 1).

These signs are used to diagnose breast diseases (Table 2).

Table 1. **Diagnostic indicators of clinical tests**

Sign	Unit of measurement	Norm	
		Lower limit	Upper limit
Urine acidity	pH	5	7
Density	g/l	1.010	1.025
Protein	g/l	0	0.033
Glucose	μmol/l	0	1
Ketone bodies	μmol/l	0	0.5
Erythrocytes	μmol/l	0	2
Leukocytes	μmol/l	0	5
Blood acidity	pH	7	7.35
Erythrocyte Sedimentation Rate (ESR)	mm/hour	2	15

***Source*: compiled by the authors**

Table 2. **Structure of the training sample**

Code	ICD-10	Name	Number of female patients
D1	N60.0	Breast cyst	20
D2	N60.1	Mastopathy	40
D3	D17	Lipoma	60
D4	N60.2	Fibroadenoma	30
D0	0	Practically healthy	50

***Source*: compiled by the authors**

The results of the diagnostics are presented in Table 3.

Table 3. **Diagnostic results**

Code	Number of female patients	Diagnostic results					
		$P(x)$		$\mu(x)$		Ω	
		N	**%**	***N***	**%**	***N***	**%**
D1	40	35	87	34	85	38	95
D2	60	54	90	53	88	56	93
D3	20	17	85	17	85	19	95
D4	30	26	87	24	80	26	87
D0	50	45	90	47	94	49	98
Total	200	177	88	175	87.5	188	94

***Source*: compiled by the authors**

In Table 3, the following notations are used:

$P(x)$ – the results of applying the method of comparison with the prototype;

$\mu(x)$ – the results of using deterministic logic by interpreting the symptom complex in terms of fuzzy logic;

Ω – the results of the work of the developed classifier;

N – the number of correctly identified objects;

% – percentage of the total number of correctly identified objects.

In expression (7), the weighting coefficients have the same weight, i.e., $k_1 = k_2 = 0.5$.

According to the data in Table 1, the share of correctly classified objects according to the results of clinical analysis in the diagnosis of diseases of the mammary gland by the method of comparison with the prototype is 88 %, according to deterministic logic is 87.5 %, and the use of the developed classifier is 94 %. These results confirm the functionality and effectiveness of the developed diagnostic method.

CONCLUSIONS AND PROSPECTS FOR FURTHER RESEARCH

This work substantiates the need for the synthesis of a heterogeneous ensemble classifier for medical diagnosis, which combines two methods of forming diagnostic conclusions: a probabilistic approach based on the analysis of training samples, and a method based on the formalization of expert knowledge about the structure of symptom complexes.

A mathematical model of a heterogeneous ensemble classifier has been developed, in which the basic models are classifiers using methods based on pattern recognition theory, namely: the method of comparison with the prototype, the method of *k*-nearest neighbors, and the method of potential functions.

As a basic model, which uses expert information on the structure of the symptom complex, its formalization was carried out in the case when the dynamic range of values of diagnostic signs can be represented by a linguistic change that takes the following values: “below normal”, “norm” and “above normal” and the threshold values of these ranges are known, or “feature present” or “feature absent” for binary features. The description of typical sets of diagnostic features together with variants of linguistic variables forms a formalized expert evaluation of standards for each disease class by calculating the average values of these ranges, which serve as a quantitative representation of the standard of each disease class, based on the experts' conclusions about the structure of the symptom complex, and not on training data.

Two scenarios of the joint use of methods based on the analysis of the training sample and the formalization of expert information regarding the structure of the symptom complex are considered:

1) According to the classic diagram of the stacking ensemble classifier, the classification

results from different base models are combined through the meta-model using weighted voting.

2) Aggregation of models in the method of comparison with the prototype when the coordinates of the class prototypes are determined. Separate prototypes are created for each class based on the educational sample and based on the formalization of expert evaluations. The coordinates of the prototypes of the meta-model are calculated by their weighted sum, in which the weighting coefficients correspond to the confidence of the training sample and the expert assessment, respectively.

The architecture of the system software has been developed, and test verification has been carried out. The results of testing conducted based on real medical data confirmed the functionality of the developed approach and demonstrated the effectiveness of the diagnostics.

Further research will aim to apply the developed decision rules not only to quantitative parameters but also to a variety of diagnostic data, including biomedical signals and images.

REFERENCES

1. Avrunin, O. G., Bodyanskyi, E. V., Kalashnik, M. V., Semenets, V. V. & Filatov, V. O. “Modern intellectual technologies of functional medical diagnostics: monograph” (in Ukrainian). Kharkiv, Ukraine. 2018. – Available from: https://openarchive.nure.ua/entities/publication/696aa207-7248-4eff-9283-87f9f51e263d. – [Accessed: Oct. 2023]

2. Tymchyk, S. V., Zlepko, S. M. & Kostyshyn, S. V. “Classification of medical information systems and technologies according to the integral aggregate criterion” (in Ukrainian). *Information Processing Systems.* 2016; 3 (140): 194–198.

3. Povoroznyuk, A. I. “Decision support systems in medical diagnostics. Synthesis of structured models and decision rules” (in Russian). *Saarbrücken Germany: LAP LAMBERT Academic Publishing GmbH & Co. KG.* 2011. – Available from: https://my.lap-publishing.com/catalogue/details/gb/978-3-8465-1341-5/*системы-поддержки-принятия-решений-в-медицинской-диагностике.* – [Accessed: Oct. 2023]

4. Eguia, H., Sánchez-Bocanegra, C. L., Vinciarelli1, F., Alvarez-Lopez, F. & Saigí-Rubió, F. “Clinical decision support and natural language processing in medicine: Systematic literature review”. *Journal of Medical Internet Research*. 2024; 26. DOI: https://doi.org/10.2196/55315.

5. Nyiramana Mukamurera P. “The role of artificial intelligence in clinical decision support systems”. *Research Invention Journal of Public Health and Pharmacy.* 2024; 3 (2): 14–17. DOI: https://doi.org/10.59298/RIJPP/2024/321417.

6. Chen, J., Wu, L., Liu, K. et al. “EDST: a decision stump based ensemble algorithm forsynergistic drug combination prediction”. *BMC Bioinformatics.* 2023; 24: 325–346. DOI: https://doi.org/10.1186/s12859-023-05453-3.

7. Rajpoot, R., Gour, M., Jain, S. et al. “Integrated ensemble CNN and explainable AI for COVID-19 diagnosis from CT scan and X-ray images”. *Scientific Reports.* 2024. 14: 24985–25005. DOI: https://doi.org/10.1038/s41598-024-75915-y.

8. Voronenko, Yu. V., Orabina, T. M., Moiseienko, R. I., et al. “Methodical recommendations fordiagnosis and treatment of some diseases in the provision of telemedicine services (for general practitioners –family medicine)” (in Ukrainian). *Publ. Shupyk National Medical Academy of Postgraduate Education.* Kyiv, Ukraine. 2019. – Available from: https://moz.gov.ua/uploads/2/14076-a5_2019_moz_shupik_forweb.pdf. – [Accessed: Oct. 2023].

9. Vladzimirsky, A. V. “Telemedicine: monograph” (in Russian). *Digital Printing House LLC. Donetsk.* Ukraine. 2011. – Available from: http://editor.inhost.com.ua/storage/Телемедицина/Vladzymyrskyy%20A.V.%20Telemedicine%20%20ISBN%20978-617-579-142-4.pdf. – [Accessed: Oct. 2022].

10. Yang, Y. T., Iqbal, U., Ching, H.-Y. J. et al. “Trends in the growth of literature of telemedicine: A bibliometric analysis”. *Computer Methods and Programs in Biomedicine*. 2015; 122 (3): 471–479. DOI: https://doi.org/10.1016/j.cmpb.2015.09.008.

11. Motamed Shariati, M. & Darvish, A. “Clinical decision support systems in ophthalmology: A systematic search and a narrative review”. *Appl Med Inform [Internet].* 2023; 46 (3): 70–82. – Available from: https://ami.info.umfcluj.ro/index.php/AMI/article/view/1065. – [Accessed: Oct. 2023].

12. Ju, R.-Y., Chien, C.-T., Xieerke, E. & Chiang, J.-S. “FCE-YOLOv8: YOLOv8 with feature context excitation modules for fracture detection in pediatric wrist X-ray images”. *arXiv:2410.01031.*

DOI: https://doi.org/10.48550/arXiv.2410.01031.

13. Das, S. & Sanyal, M. K. “Machine intelligent diagnostic system (MIDs): An instance of medical diagnosis of tuberculosis”. *Neural Comput & Applic.* 2020; 32: 15585–15595. DOI: https://doi.org/10.1007/s00521-020-04894-8.

14. Dullabh, P., Dhopeshwarkar, R., Cope, E., Gauthreaux, N., et al. “Advancing patient-centered clinical decision support in today’s health care ecosystem: Key themes from the clinical decision support innovation collaborative’s 2023 annual meeting”. *JAMIA Open*, 2024; 7 (4): ooae109. DOI: https://doi.org/10.1093/jamiaopen/ooae109.

15. Povoroznyuk, A. I., Povoroznyuk, O. A. & Shehna, Kh. “Formalizing the stages of mammographic examinations in designing a medical decision support system”. *Herald of Advanced Information Technology*, 2020; 3 (4): 279–291. DOI: https://doi.org/10.15276/hait.04.2020.6.

16. Biloborodova, T., Scislo, L., Skarga-Bandurova, I., Sachenko, A., Molgad, A., Povoroznjuk, O & Yevsieiva, Y. “Fetal ECG signal processing and identification of hypoxic pregnancy conditions in-utero”. *Mathematical Biosciences and Engineering.* 2021; 18 (4): 4919–4942. https://www.scopus.com/authid/detail.uri?authorId=55817007400. DOI: https://doi.org/10.3934/mbe.2021250.

17. Fainzilberg, L. S. “New approaches to the analysis and interpretation of the shape of cyclic signals”. *Cybernetics and Systems Analysis.* 2020; 56 (4): 665–674. DOI: https://doi.org/10.1007/s10559-020-00283-0.

18. Filatova, A. Ye., Povoroznyuk, A. I., Nosachenko, B. P. & Fahs, M. “Synthesis of an integral signal for solving the problem of morphological analysis of electrocardiograms”. *Herald of Advanced Information Technology.* 2022; 5 (4): 263–274. DOI: https://doi.org/10.15276/hait.05.2022.19.

19. Hein, D., Chen, Z., Ostmeier, S., Xu, J., et al. “Preference fine-tuning for factuality in Chest X-Ray interpretation models without human Feedback”. *arXiv:2410.07025.* DOI: https://doi.org/10.48550/arXiv.2410.07025.

20. Jain, K., Rangarajan, K. & Arora, C. “Follow the radiologist: Clinically relevant multi-view cues for breast cancer detection from mammograms”. *In Linguraru, M.G., et al. Medical Image Computing and Computer Assisted Intervention. MICCAI. Lecture Notes in Computer Science.* 2024; 15001: 102–112. DOI: https://doi.org/10.1007/978-3-031-72378-0_10.

21. Musa, A., Ibrahim Adamu, M., Kakudi, H. A., Hernandez, M. & Lawal, Y. “Analyzing cross-population domain shift in Chest X-Ray image classification and mitigating the gap with deep supervised domain adaptation”. *In Linguraru, M.G., et al. Medical Image Computing and Computer Assisted Intervention – MICCAI 2024. Lecture Notes in Computer Science*, 2024; 15003: 585–595. DOI: https://doi.org/10.1007/978-3-031-72384-1_55.

22. Moiseyenko, M. I., Kuzyshyn, M. M., Turovska, L. V., et al. “Large language models of artificial intelligence in medicine” (in Ukrainian). *Modern Information Technologies and Innovation Methodologies of Education in Professional Training Methodology Theory Experience Problems.* 2024. 72: 73–88. DOI: https://doi.org/10.31652/2412-1142-2024-72-73-88.

23. Povoroznyuk, A. I., Povoroznyuk, O. A. & Shekhna, H. “Synthesis of combined diagnostic decision rule in medical decision support systems” (in Ukrainian). *Management, Navigation and Communication Systems.* 2021; 1 (63): 103–106. DOI: https://doi.org/ 10.26906/SUNZ.2021.1.103.

24. Syvolap, V. D., Kiselyov, S. M. & Lashkul, D. A. “Basic principles of evidence-based medicine: a study guide” (in Ukrainian). *ZDMU.* Zaporizhzhia: Ukraine. 2020. – Available from: http://dspace.zsmu.edu.ua/bitstream/123456789/10909/3/%D0%94%D0%BE%D0%BA%D0%B0%D0%B7%20%D0%BC%D0%B5%D0%B4%20PhD_%D0%B7%D0%B0%D1%82%D0%B2%D0%B5%D1%80%D0%B4.pdf – [Accessed: Oct. 2023].

25. Povoroznyuk, A. I., Filatova, A. E., Surtelb, W., et al. "Design of decision support system when undertaking medical-diagnostic action". *Optical Fibers and Their Applications.* 2015; 9816: 98161O1–98161O17, https://www.scopus.com/authid/detail.uri?authorId=55225664000.
DOI: https://doi.org/10.1117/12.2229295.

26. Harish, V., Morgado, F., Stern, A. D. & Das, S. “Artificial intelligence and clinical decision making: The new nature of medical uncertainty”. *Academic Medicine.* 2021; 96 (1): 31–36. DOI: https://doi.org/10.1097/ACM.0000000000003707.

27. Nascimento, P. M., Medeiros, I. G., Falcão, R. M. et al. “A decision tree to improve identification of pathogenic mutations in clinical practice”. *BMC Medical Informatics and Decision Making.* 2020; 20 (52). DOI: https://doi.org/10.1186/s12911-020-1060-0.

28. Zhukovskaya, O. A. & Fainzilberg, L S. “Bayesian strategy for group decision making and its

interval generalization". *Journal of Automation and Information Sciences.* 2019; 51 (1): 1–14. DOI: 10.1615/JAutomatInfScien.v51.i1.10.

29. Nour, M., Cömert, Z. & Polat, K. "A novel medical diagnosis model for COVID-19 infection detection based on deep features and bayesian optimization". *Applied Soft Computing*. 2020; 97 (Part A): 106580. DOI: https://doi.org/10.1016/j.asoc.2020.106580.

30. Kumar, Y., Koul, A., Singla, R. et al. "Artificial intelligence in disease diagnosis: a systematic literature review, synthesizing framework and future research agenda". *Journal Ambient Intel Human Comput*. 2023; 14: 8459–8486. DOI: https://doi.org/10.1007/s12652-021-03612-z.

31. Innocent, P. R., John, R. I. & Garibaldi J .M. "Fuzzy methods for medical diagnosis". *Applied Artificial Intelligence.* 2004; 19 (1): 69–98. DOI: https://doi.org/10.1080/08839510590887414.

32. Kour, H., Manhas, J. & Sharma, V. "Usage and implementation of neuro-fuzzy systems for classification and prediction in the diagnosis of different types of medical disorders: a decade review". *Artificial Intelligence Review.* 2020; 53: 4651–4706. DOI: https://doi.org/10.1007/s10462-020-09804-x.

33. Korobova, L. A. & Matytsina, I. A. "The use of logical-linguistic apparatus to describe the functioning of medical systems". *Journal of Physics: Conference Series*. 2020; 1515: 052020. DOI: https://doi.org/10.1088/1742-6596/1515/5/052020.

34. Sarvamangala, D. R. & Kulkarni, R. V. "Convolutional neural networks in medical image understanding: a survey". *Evolutionary Intelligence*. 2022; 15: 1–22. DOI: https://doi.org/10.1007/s12065-020-00540-3.

35. Schlesinger, M. & Hlavac, V. "Ten lectures on statistical and structural recognition". – *Kyiv: Naukova Dumka*, 2004. – DOI: https://doi.org/10.1007/978-94-017-3217-8.

36. Kiflay, A. Z., Tsokanos, A. & Kirner R. "A network intrusion detection system using ensemble machine learning". *International Carnahan Conference on Security Technology*. 2021. p. 1–6.

37. Sarkar, A., Sharma, H. S. & Singh M. M. "A supervised machine learningbased solution for efficient network intrusion detection using ensemble learning based on hyperparameter optimization". *International Journal of Information Technology*. 2022; 15 (1): 423–434.

38. Hornostal, O., Chelak, V. & Gavrylenko S. "Research of intelligent data analysis methods for identification of computer system state". *Proceedings of the 30th International Scientific Symposium Metrology and Metrology Assurance (MMA).* 2020. p. 1–5.

39. Povoroznyuk, A. I., Povoroznyuk, O. A. & Filatova, A. Ye. "Heterogeneous ensemble classifier in computer systems of medical diagnostics" (in Ukrainian). *Informatics. Culture. Technology*. 2024; 1 (1): 221–226. DOI: https://doi.org.10.15276/ict.01.2024.33.

40. Grigoriev, K. I. "Pediatrics: A handbook for a practitioner" (in Russian). *Medpres-inform.* Kyiv, Ukraine. 2014. – Available from: https://bigl.ua/p1138777038-grigorev-pediatriya-spravochnik?srsltid=AfmBOoqZatnnVzMH7mXPq2CMY1XQLdDuuyYLNnDBhCAftifMysa1G0Gm. – [Accessed: Oct. 2023].

Conflicts of Interest: The authors declare that they have no conflict of interest regarding this study, including financial, personal, authorship or other, which could influence the research and its results presented in this article

Received 21.08.2024
Received after revision 30.10.2024
Accepted 15.11.2024

DOI: https://doi.org/10.15276/hait.07.2024.26
УДК 004. 93

Гетерогенний ансамблевий класифікатор у комп'ютерних системах медичної діагностики

Поворознюк Анатолій Іванович[1)]
ORCID: http://orcid.org/0000-0003-2499-2350 ai.povoroznjuk@gmail.com. Scopus Author ID: 55225664000
Поворознюк Оксана Анатоліївна[1)]
ORCID: http://orcid.org/0000-0001-7524-5641 povoks76@gmail.com. Scopus Author ID: 55817007400
Філатова Ганна Євгенівна[1)]
ORCID: https://orcid.org/0000-0003-1982-2322 filatova@gmail.com. Scopus Author ID: 56448583600
[1)] Національний технічний університет «Харківський політехнічний інститут». вул. Кирпичова, 2. Харків, 61002, Україна

АНОТАЦІЯ

Робота присвячена вирішенню актуальної науково-технічної задачі побудови діагностичної системи підтримки прийняття рішень в медицині. . Основою даної системи є модель розробленого гетерогенного ансамблевого класифікатора, який в якості базових моделей реалізує два підходи до формулювання діагностичного висновку. Перший з цих підходів є ймовірнісним. Він ґрунтується на аналізі навчальної вибірки з пацієнтів з підтвердженим діагнозом, що дозволяє оцінити імовірність наявності того чи іншого захворювання на основі наявних даних. Другий підхід – експертний, який базується на наявній експертній інформації про структуру симптомокомплексів, які характеризують кожне окреме захворювання. Важливо зазначити, що обидва ці підходи розглядають одну і ту ж проблему з різних точок зору, і їх спільне використання є надзвичайно перспективним для розвитку ефективних діагностичних систем. Мета цього дослідження полягає в синтезі гетерогенного ансамблевого класифікатора, який інтегрує як експертну, так і ймовірнісну складову процесу постановки діагнозу. У рамках дослідження було проведено аналіз різних методів діагностики, що використовуються лікарями відповідно до сучасних вимог доказової медицини. Також було розглянуто методи побудови діагностичних вирішальних правил у системах підтримки прийняття медичних рішень. На основі цих досліджень була розроблена математична модель гетерогенного ансамблевого класифікатора, і було обґрунтовано вибір його складових частин. У якості ймовірнісної складової в даній системі були обрані широко вживані методи класифікації, зокрема метод порівняння з еталоном, метод K-найближчих сусідів та метод потенційних функцій. Експертні знання, які стосуються структури симптомокомплексів, формалізуються шляхом вираження симптомокомплексів кожного захворювання у вигляді числових інтервалів. При цьому використовуються лінгвістичні змінні, що можуть мати значення “нижче норми”, “норма” або “вище норми”. Розглянуті різноманітні варіанти агрегації різнотипних базових моделей в межах гетерогенного ансамблевого класифікатора. Це дозволяє зберігати переваги кожного з методів та підвищувати загальну точність класифікації. Були сформульовані вимоги до функціональних можливостей розробленої системи, визначено засоби проектування, а також основну платформу для розробки – Java, і систему управління базами даних – MySQL. Виконано проектування системи підтримки прийняття рішень та виконано комплексну перевірку розробленої системи на реальних медичних даних. Результати цих перевірок підтвердили ефективність роботи системи.

Ключові слова: медична діагностика; ансамблевий класифікатор; базова модель; ймовірнісний класифікатор; симптомокомплекс; експертна інформація; агрегації моделей; система підтримки прийняття рішень

ABOUT THE AUTHORS

Anatoliy I. Povoroznyuk - Doctor of Engineering Sciences, Professor, Department of “Computer Engineering and Programming” National Technical University “Kharkiv Polytechnic Institute”, 2, Kyrpychova Str. Kharkiv, 61002, Ukraine
ORCID: https://orcid.org/0000-0003-2499-2350; ai.povoroznjuk@gmail.com. Scopus Author ID: 55225664000

***Research field*:** Methods and algorithms for processing experimental data in medicine; processing of medical signals and images; synthesis of decisive rules based on structured models; design of decision support systems in medicine; computer architecture, design of specialized computer systems

Поворознюк Анатолій Іванович - доктор технічних наук, професор, професор кафедри Комп’ютерної інженерії та програмування. Національний технічний університет «Харківський політехнічний інститут» вул. Кирпичова, 2. Харків, 61002, Україна

Oksana A. Povoroznyuk - PhD (Eng), Associate Professor, Computer Engineering and Programming Department, National Technical University “Kharkiv Polytechnic Institute”, 2,Kyrpychova Str. Kharkiv, 61002, Ukraine
ORCID: http://orcid.org/0000-0001-7524-5641 povoks76@gmail.com. Scopus Author ID: 55817007400

Research field: Methods and algorithms of multi-criteria selection of alternatives in medicine; decision-making support systems for diagnostic and therapeutic measures

Поворознюк Оксана Анатоліївна - кандидат технічних наук, доцент, доцент кафедри Комп’ютерної інженерії та програмування. Національний технічний університет «Харківський політехнічний інститут» вул. Кирпичова, 2. Харків, 61002, Україна

Anna E. Filatova - Doctor of Engineering Sciences, Professor, Department of “Computer Engineering and Programming” National Technical University “Kharkiv Polytechnic Institute”, 2, Kyrpychova Str. Kharkiv, 61002, Ukraine
ORCID: https://orcid.org/0000-0003-1982-2322; filatova@gmail.com. Scopus Author ID: 56448583600

Research field: Morphological analysis of medical signals and images; methods and algorithms for processing experimental data in medicine; pattern recognition theory; design of decision support systems in medicine; probability theory and mathematical statistics

Філатова Ганна Євгенівна – доктор технічних наук, професор, професор кафедри Комп’ютерної інженерії та програмування. Національний технічний університет «Харківський політехнічний інститут» вул. Кирпичова, 2. Харків, 61002, Україна