

Development of mathematical methods for diagnosing kidney diseases using fuzzy set tools

Alua Myrzakerimova, Kateryna Kolesnikova

Department of Information Systems, Faculty of Business, Media and Management, International IT University, Almaty, Kazakhstan

Article Info

Article history:

Received Feb 9, 2024

Revised Mar 11, 2024

Accepted Mar 20, 2024

Keywords:

Dataset of pyelonephritis
Decision support system
Diagnosing kidney diseases
Fuzzy logic method
Morbidity analysis

ABSTRACT

An approach based on fuzzy set theory is presented in the scientific article to enhance the efficiency of diagnosing kidney diseases by decreasing the time required for making medical decisions. The suggested approach employs fuzzy models and algorithms that consider the uncertainty and variability of clinical data to optimize the assessment of the functional state of the kidneys, taking into account various risk factors and individual characteristics of patients. The paper suggests a technique to develop a system of fuzzy decision rules. This technique combines E. Shortliff's iterative rules with functions from the studied classes of kidney diseases. Mathematical modeling and experimental studies have indicated relatively high effectiveness in classifying different forms of kidney diseases. The results can be used to formulate intelligent decision support systems in clinical practice and improve diagnostic and monitoring processes. Moreover, the findings may aid in shaping more targeted and effective health policies at the national and regional levels, enhancing access to healthcare, and promoting the population's overall health.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Alua Myrzakerimova

Department of Information Systems, Faculty of Business, Media and Management

International IT University

Almaty-050047, Kazakhstan

Email: a.ospan@iitu.edu.kz

1. INTRODUCTION

Currently, extensive research is focused on developing new mathematical methods for processing and analyzing medical indicators. Also, the use of IT technologies in medicine significantly expands the professional capabilities of doctors through various automated diagnostic systems. The use of information systems based on computer technologies for conducting various studies will lead to a reduction in various types of material, energy, and financial costs [1]. The goal of research paper is to optimize the assessment of the functional state of the kidneys using non-invasive technologies. Although computer data analysis allows solving medical problems and determining treatment strategy, methods of intelligent support for medical decision-making have not yet become widespread. The incorporation of decision-making principles based on fuzzy sets theory has emerged as a corner stone and pertinent issue in medicine. Current research is dedicated to developing mathematical models for diagnosing kidney diseases, aiming to create systems that support physicians with appropriate visualizations throughout the entire diagnostic and treatment process. By leveraging information technology and fuzzy sets theory, these systems can aid medical professionals in making well-informed decisions despite uncertainties in the available data. Proposed algorithm method improves the efficiency and effectiveness of doctors diagnosing process by using algorithm of preliminary disease prediction in a relatively short period of time, in comparison of conventional methods [2].

Healthcare decision support systems have been developed to support medical decision-making, improve patient outcomes, and enhance the efficiency of healthcare systems [3]. They typically use a combination of algorithms, statistical models, and medical knowledge to provide guidance and recommendations to healthcare providers. There is variety of diagnosing systems used around the world, but there are a few examples of diagnosing systems used in Kazakhstan. Thus, preliminary diagnosis algorithm method is highly demanded in such a developing country such as a Kazakhstan, where one third of populations lives in a rural area that takes more travelling time to clinics and hospitals. The information mainly related to the management systems of medical organizations. The subject of developing and utilizing diagnostic systems is largely underexplored. To effectively advance such systems, digital platform could play a crucial role in such automated system [4]. Assistance in medical diagnostic systems holds the potential to not only expedite the diagnostic process for healthcare professionals but also mitigate potential subjectivity in their decision-making. Furthermore, these diagnostic systems can enhance educational and study process, also enhancing the training of medical students and bolstering their competitiveness in the field [5].

Effective diagnostic systems require accurate mathematical models to identify the correct diagnosis. Fuzzy set theory is particularly useful for creating these models because it can convert linguistic variables, such as patient symptoms, into computable data [6]. Fuzzy logic, with its ability to manage vague and uncertain information, is a critical tool for making decisions in areas that involve complex data interpretation, including medical diagnosis. It helps in understanding and analyzing varied medical information sources—such as patient histories, lab results, and microscopic studies—where uncertainty and imprecision are common. Fuzzy logic's framework facilitates the modeling and analysis of this complex data, offering a more nuanced and flexible approach to decision-making. By using fuzzy sets and rules, it allows for the consideration of uncertainty and membership degrees, helping medical professionals interpret and navigate through the complexities of medical databases more effectively.

This research primarily aims to develop and analyze a mathematical model and algorithm to improve diagnostic processes using an expert system, with a specific focus on kidney diseases. The paper discusses the using of a mathematical model based on fuzzy logic that facilitates the diagnosis by leveraging expert system methodologies. It outlines the algorithmic approach utilized in the diagnostic process, emphasizing the integration of expert knowledge and computational efficiency. The selection of kidney diseases as a case study is justified within the problem statement, highlighting the complexity and critical nature of accurate diagnosis in nephrology. The results section presents the outcomes of applying the developed model and algorithm to real-world diagnostic scenarios, demonstrating the potential improvements in accuracy, speed, and reliability of medical diagnoses facilitated by this research.

The structure of the paper is outlined as follows: first section provides a comprehensive literature review on the use of diagnostic systems, focusing on data mining and analysis in the medical field, with a specific emphasis on kidney diseases. It includes an empirical analysis of the morbidity of kidney diseases in the Republic of Kazakhstan. Section 2 introduces method for the non-invasive diagnosis of kidney diseases through the application of fuzzy set theory. This approach is designed to adeptly handle the uncertainties in clinical data. Final section concludes with a discussion that covers the benefits and limitations of the proposed diagnostic systems, including a comparison with existing systems and a critique of the mathematical models used. This is followed by concluding remarks that summarize the significant contributions of the study and propose avenues for future research.

Data mining and analysis of medical data are widely utilized in identifying the population's reproductive, physical development, and health status, as well as the prevalence and duration of various diseases. These methods enable identifying and establishing links between the general level of morbidity and mortality, numerical data collection and analysis of medical institution activities, monitoring of the implementation of health development plans, and evaluation of physicians' work quality. For this study, statistical data on kidney diseases in Kazakhstan were selected. According to statistical data compiled by the Ministry of Health of the Republic of Kazakhstan from 2009 to 2022, specifically the "Health of the population of the Republic of Kazakhstan and healthcare activities" statistical compilation, the population is estimated at 18 million people [7]. This compilation presents statistical materials on healthcare organizations' activities and health indicators of the population of the Republic of Kazakhstan, presented per 100,000 population per year.

The first expert systems were developed in the field of medical diagnosis, including the MYCIN system for diagnosing bacterial infections and the PUFF system for diagnosing lung diseases [8]. With the advent of the internet, research in expert systems for medical diagnosis expanded to include telemedicine and online diagnostic systems. One notable example is the WebMD symptom checker, which uses an expert system to assist in diagnosing symptoms and conditions [9]. The development of more advanced expert systems for medical diagnosis, such as IBM's Watson for Health and Google's DeepMind Health, can be attributed to the progress made in artificial intelligence (AI) and machine learning (ML) [10]. Machine

learning algorithms and natural language processing are utilized in these systems to assess substantial volumes of medical data and aid in the process of diagnosis. Bohr and Memarzadeh [10] in 2020 research in expert systems for medical diagnosis continues to advance, with a focus on developing systems that are more accurate, reliable, and user-friendly. There is also growing interest in using expert systems to assist in the diagnosis of rare and complex diseases. The evolution of expert systems for medical diagnosis has seen remarkable progress, from their inception in the 1970s, through the integration of advanced AI and machine learning in the 2010s, and the ongoing pursuit of greater accuracy and versatility in the 2020s, indicating a promising future for their role in healthcare.

Notable contributions include Mirmozaffari proposed a fuzzy expert system for diagnosing liver diseases based on clinical and laboratory data [11]. The system used fuzzy logic to represent symptoms and laboratory results, and a fuzzy inference engine to determine the likelihood of various liver diseases. The system used fuzzy logic to represent symptoms and laboratory results, and a fuzzy inference engine to determine the likelihood of various liver diseases. Similarly, Nahato *et al.* [12] examines the utilization of data mining methodologies to extract valuable insights from clinical datasets, with the aim of aiding doctors in their decision-making process. The objective is to create a classifier that can accurately predict the presence or absence of diseases by utilizing a small number of features derived from these datasets.

Dankwa-Mullan *et al.* [13] aimed to investigate the capacity of AI and cognitive computing to transform the management and treatment of diabetes, a disorder that impacts around 425 million individuals globally. In the context of substantial healthcare spending and a worrisome number of undiagnosed and untreated persons, the authors sought to find and study clinically significant breakthroughs in AI that could be advantageous for individuals with diabetes. The authors found that AI has the potential to greatly improve medical diagnosis, particularly in areas such as radiology, dermatology, and ophthalmology. AI can analyze large amounts of data quickly and accurately, potentially reducing diagnostic errors and improving patient outcomes. However, the authors also note that there are several challenges to implementing AI in medical diagnosis, including concerns about data privacy, the need for large amounts of high-quality data to train AI models, and the potential for AI to perpetuate existing biases in healthcare. One potential limitation of this review is that it only focuses on studies published in the last five years, which may not provide a complete picture of the use of AI for medical diagnosis.

Literature analysis shows that the same kidney diseases of different types are often diagnosed in diametrically opposite ways [14], [15]. For example, primary non-obstructive pyelonephritis (PL) usually initially manifests with general symptoms, followed by local symptoms after 2-3 days. Conversely, obstructive PLs first presents with local symptoms. The severity of obstructive PLs is influenced by various factors, such as the duration and intensity of the obstruction, and the localization and nature of the inflammation in the renal parenchyma [14]. Differential diagnosis of different forms of PLs presents significant diagnostic challenges. The literature describes numerous clinical, laboratory, and radiological signs that allow differentiating types of kidney diseases. Among these signs are weakness, headache, joint and muscle pain, nausea, vomiting, tachycardia, dehydration, hyperthermia, blurring of muscle contours observed on radiographs, azotemia, dysproteinemia, and hyperleukocytosis with a shift to the left. In diagnosing kidney diseases, an important tool is the analysis of functional kidney tests, which provide information about the functional state of the kidneys [16]. Significant risk factors for kidney dysfunction include clinical signs indicating a high or moderate risk of developing kidney disease [17]. Identifying these risk factors helps determine the appropriate approach to management and treatment for individuals at risk.

Furthermore, the process of analysis assumes a pivotal role in assessing the efficacy of interventions aimed at preventing and treating diseases, hence facilitating more accurate and logical decision-making within the realm of medical practice [18]. Figure 1 illustrates an algorithm utilized for the analysis of medical statistics data. This method was employed to examine the incidence of kidney disease and then map the distribution of illness incidence across different regions [19]. The data analysis consists of stages, which is illustrated in Figure 1.

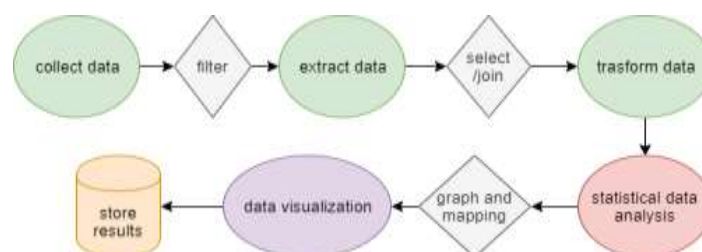


Figure 1. Data analysis phases

In the initial examination of patients in medical practice, several key challenges arise. Firstly, there's often a pressing lack of time to thoroughly evaluate a patient's health condition. This necessitates rapid decision-making to prioritize patients based on the severity of their condition, determining who requires immediate care, urgent consultation, or can afford a more delayed evaluation. Complicating matters further, clinicians' judgments are inherently subjective and can be influenced by personal biases or external factors like fatigue, potentially leading to inconsistencies in diagnosis and treatment [17].

Secondly, the current diagnostic process frequently struggles to efficiently manage patient flow, resulting in prolonged wait times for critically ill individuals and inefficient resource allocation. This not only compromises patient outcomes but also exacerbates the strain on medical resources. However, the integration of a computerized system offers a promising solution. By leveraging algorithms and predefined criteria, such a system could automate patient prioritization, providing rapid assessments to distinguish between urgent and non-urgent cases. Additionally, it could offer preliminary diagnoses based on symptom presentation, equipping medical staff with valuable insights to expedite treatment decisions and optimize patient care pathways. Ultimately, the integration of such a system holds the capacity to greatly augment the efficiency and efficacy of healthcare provision.

Furthermore, ongoing research encompasses an empirical examination of kidney disease morbidity in the Republic of Kazakhstan [19]. This geographic analysis plays a crucial role in pinpointing regions with heightened risks of kidney diseases, providing invaluable insights for the development of targeted and effective healthcare interventions at both national and regional levels. Ultimately, the objective is to enhance accessibility to medical services and elevate the overall health standards of the population by implementing informed strategies tailored to specific geographic areas.

Disease mapping is a tool that provides the ability to identify common spatial trends and areas of increased risk. This methodological approach is great significance for improving epidemiological studies focused on geographical aspects [20]. The advantages of disease mapping include detecting disease clusters in different regions, facilitating ecological studies. Moreover, allows the find the relationship between indicators and the frequency of risk factors, as well as assessing the location and degree of pollution impact on health, and determining the population exposed to a specific influence.

The risk map illustrated in Figure 2, for kidney disease estimation in Kazakhstan reveals notable regional disparities in disease prevalence, with Turkistan, Zhambyl, and Almaty regions identified as high-risk areas, while Qostanay, Astana, and Atyrau regions are characterized as low-risk zones. This geographical distribution of risk provides valuable insights for analyzing factors influencing morbidity rates associated with kidney diseases.



Figure 2. Risk map for kidney disease estimation in Kazakhstan

The literature review highlights the limited availability of sources on medical diagnostic systems used in Kazakhstan and Central Asian countries. The information provided by the Ministry of Health of the Republic of Kazakhstan is mainly related to the management systems of medical organizations. The subject of developing and utilizing diagnostic systems is largely underexplored. The structure of proposed diagnosing system is explained [21], which includes mathematical part. Therefore, proposed research is aimed for explain mathematical method using theory of fuzzy sets.

The contributions and novelty of this research work can be summarized as follows: i) research is aimed for formulation of accurate mathematical models capable of diagnosing and predicting diseases. These models will serve as the foundation for developing diverse automated systems that aid in making well-informed decisions; ii) the application of fuzzy logic methods allows for the optimization of assessing the functional state of kidneys, considering various risk factors and individual patient characteristics. This optimization is a practical contribution that can directly impact clinical practice; iii) application of fuzzy logic improves not only the accuracy of diagnostic results but also their interpretability. This enhancement is crucial for developing intelligent decision support systems that can be used by clinicians in diagnosing and monitoring kidney diseases; iv) the paper includes an empirical analysis of morbidity related to kidney diseases in the Republic of Kazakhstan, identifying high-risk regions. This contribution is valuable for public health planning and management, offering data-driven insights into regional disease prevalence; and v) the findings from this research can inform more targeted and effective healthcare strategies at both national and regional levels. By improving diagnosis and monitoring processes for kidney diseases, the research contributes to better healthcare outcomes. This strategic contribution is significant for policymakers and healthcare providers aiming to improve access to medical care and elevate the overall health level of the population.

2. ANALYSIS AND DATA FORMATION OF INFORMATIVE FEATURES BY EXPERT METHOD

Exploring techniques for decision-making using vague initial data holds significant potential. In the field of medicine, the challenge lies in addressing diagnostic concerns pertaining to a range of illnesses. This involves processing intricate and frequently ambiguous information, encompassing diverse sets of clinical, historical, laboratory, and other details regarding the patient's condition [22]. Frequently, information originates from individuals and might stand as the sole data source, as seen in cases of clinical and historical data. However, this information tends to be incomplete, subjective, and prone to inaccuracy. Hence, in the realm of medicine, the application of fuzzy set theory and the corresponding principles of decision-making, particularly tailored to handle vague initial data, emerge as a critical and pressing challenge.

Fuzzy set theory methods enable the utilization of linguistic variables, which often constitute the medical input data utilized in the diagnostic process. This approach offers the benefit of considering the intensity of symptoms. Given the dynamic nature of a patient's condition in real-world scenarios, where improvements and deteriorations occur, this versatile mathematical method holds the potential to predict the progression of the illness.

Figure 3 illustrates the complete workflow of the proposed system, outlining each stage from the initial patient check-in to the ultimate approval decision made by medical experts. This figure provides visualization of the computational procedure used to evaluate initial causes of diseases. Initially, when the patient enters the healthcare system, it signifies the start of the diagnostic process, during which data and medical background are gathered. To create a database, we have developed a system of criteria designed to provide an objective decision on determining the diagnosis of kidney disease based on self-diagnosis data, laboratory, and other data. The values of the input data are formed based on the examination of patients, according to the standards of medical care approved by the Ministry of Health of the Republic of Kazakhstan [23].

There is a list of 52 parameters x_i , and the contribution of each parameter to determining the object's belonging to the identified classes has been defined. There are data that has been taken into account: i) patient's questionnaire data, ii) laboratory analyses of patients, iii) data from instrumental examination: ultrasound examination, iv) pathology (if exists). All information from the preoperative examination was transformed into a multidimensional vector, and a database was formed, where each patient corresponds to information on 52 features. Each entry ends with a specific conclusion (output parameter) the presence or absence of kidney disease [24], [25]. A retrospective analysis of the medical records of more than 400 patients with various forms of PLs, chronic kidney disease (CKD) and acute kidney injury (AKI) who were treated in a private clinical hospital in Almaty from 2017 to 2022 was carried out. The average age of the patients was 49.3 years, of which 57% were men and 43% were women. During the study, corresponding symptoms were assessed, and clinical-laboratory diagnostics of inflammatory processes in the kidneys were conducted. The data presented in Table 1 is taken from sources and illustrated in short form [26]. Table 1 presents the initial set of informative features related to risk factors for kidney disease, collected during surveys, examinations, and routine studies. Group of factors include: nutritional factors, medical biological factors, laboratory research methods, socio-economic factors, occupational factors, and behavioral factors [26].

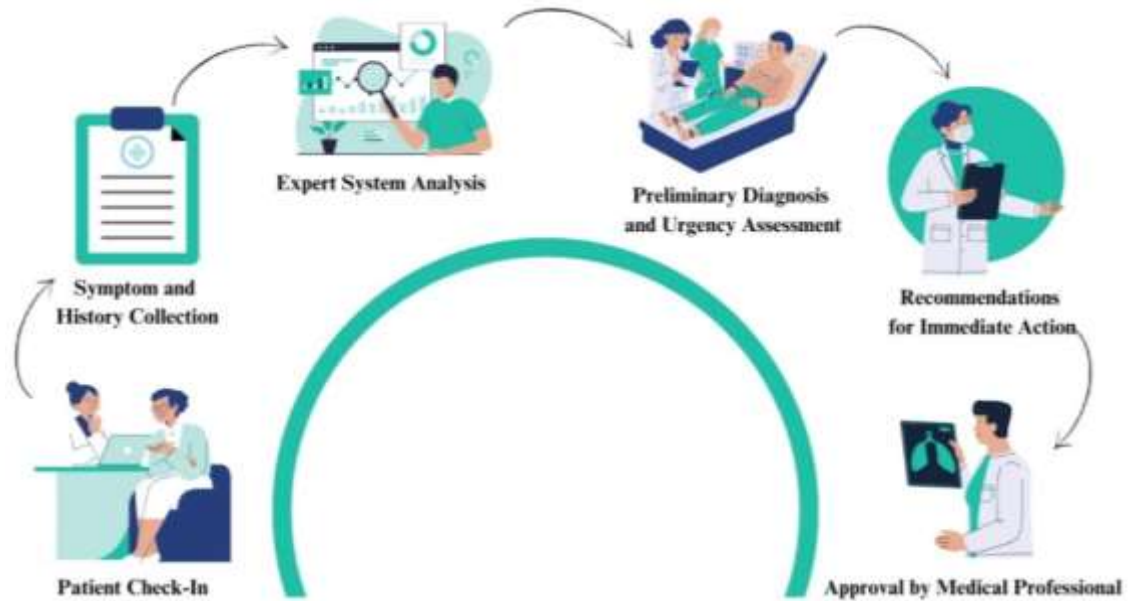


Figure 3. Proposed diagnose system methodology algorithm

Table 1. Informative features of risk factors for kidney disease

Group of factors	Indicators (risk factors)	Gradation
Medical biological factors	Age; chronic urinary tract infection; history of urethral instrumental invasion; history of sexually transmitted infection; chronic disease of the colon, rectum.	
Socio-economic factors	Level of education; Unsatisfactory housing conditions, lack of hot water, inadequate temperature regime; Low-income level; Social status	higher, secondary, below secondary
Behavioral factors	Employment; Work in harmful production conditions; Workers detached from home, shift work, long-haul truck drivers; Outdoor work in unfavorable temperature conditions; Labor associated with physical or psychological stress;	

The analysis of factors causing kidney diseases and distinctive features of different forms of this condition reveals significant differences in the attribute space used in traditional medicine. The boundaries are unclear, leading to dynamic and complex medical interpretations. Therefore, it is recommended to address these issues using decision theory based on fuzzy logic. While using a statistical approach, it is important to consider that the data used for research analysis must be statistically precise enough [27]. However, in practical scenarios, the conditions for statistical precision are often not met, especially in medical applications. Problems include the limitation of information for accurately describing an object, the complexity of verifying data reliability, the presence of important but unconfirmed information, overlapping classes without precise formal models, different structures within the same classes, and the fundamentally fuzzy nature of data. To address such problems, Zadeh [28] introduced fuzzy sets, an extension of classical set theory, as a tool for fuzzy logic decision-making.

In fuzzy representation, the concept of belonging to a set A becomes ambiguous, allowing for partial membership, defined as “degree of membership to set A”. Fuzzy set theory, in addition to digital variables, uses linguistic variables to describe semantic concepts. For example, the linguistic variable AD=[*blood pressure*] can have values such as BPn [*low blood pressure*], BPnd=[*normal pressure*], and BP=[*high blood pressure*]. Another example includes linguistic expressions describing conditions such as “absence of function”, “normal function state”, and “excessive function”, represented by membership functions $\mu_H(x)$, $\mu_{norm}(x)$ and $\mu_I(x)$.

Zadeh [29] developed a method that includes logical operations over membership functions to create fuzzy logic rules of inference. These rules, resembling an “if-then” structure, are successfully used in medical applications for solving forecasting and medical diagnosis tasks. These rules determine how the system will interpret and process input data to produce relevant output data. Each rule consists of an antecedent (the “if” part) and a consequent (the “then” part). Fuzzy “if-then” rules, also known as fuzzy implications, are formally described as follows:

$$\text{IF } x = A, \text{ THEN } y = B \tag{1}$$

Where A and B are linguistic fuzzy values, determined by the corresponding membership functions for variables x and y. Briefly this implication looks like $A \rightarrow B$. For variables x_1, x_2 this expression is ultimately transformed by aggregating a set of “if-then” rules showed in expression (2):

$$\text{IF } x_1 = A_1 \text{ and } x_2 = A_2 \text{ and } \dots \text{ and } x_n = A_n, \text{ THEN } y = B \tag{2}$$

Variables x_1, x_2, \dots represent feature x of N-dimensional vector. Parameters A_1, A_2, \dots, A_n and B indicate the corresponding membership coefficients, $\mu_A(x_i)$ and $\mu_B(y)$. Another approach to decision making with incomplete certainty is E. Shortliff's iterative formulas, derived from extensive observations of the logic of medical decisions [29]. The central element of his reasoning is the confidence coefficient of the ω hypothesis confidence coefficient $HCC(\omega_i/X)$ which in general is determined by the discrepancy between the measures of trust and distrust regarding the hypothesis being tested, illustrated by (3).

$$HCC(\omega_i/X) = MC(\omega_i/X) - MM(\omega_i/X) \tag{3}$$

Where, $HCC(\omega_i/X)$ is confidence in the diagnostic hypothesis ω_i with characteristic(s) X, $MC(\omega_i/X)$ is measure of trust to ω_i with characteristic (s) X, $MM(\omega_i/X)$ is measure of distrust in a hypothesis ω_i with characteristic (s) X.

Indicators of trust and distrust are assessed on a scale from 0 to 1, which reflects the weight of evidence in favor of or against the hypotheses being studied. Confidence coefficient ($HCC(\omega_i/X)$) on a scale from -1 to +1, where «-1» means false, and «+1» means true. Intermediate values indicate varying degrees of confidence in decisions made. When new evidence (x) is received, the measures of trust and distrust (MC and MM) are updated according to the (4) and (5).

$$MC(\omega_i/X, x) = MC(\omega_i/X) + MC(\omega_i/x)(1 - MC(\omega_i/X)) \tag{4}$$

$$MM(\omega_i/X, x) = MM(\omega_i/X) + MM(\omega_i/x)(1 - MM(\omega_i/X)) \tag{5}$$

Sometimes, when solving practical problems, doctors focus only on signs that increase confidence in the hypothesis ω_i , and then expressions (4) and (5) transformed into a decisive rule of the form:

$$HCC_{\omega_i}(p + 1) = HCC_{\omega_i}(p) + HCC_{\omega_i}(Z_s)[1 - HCC_{\omega_i}(p)] \tag{6}$$

where p is number of iterations of $HCC_{\omega_i}(p)$, and Z_s - the basic variable, which is based on conclusion. In frequent cases $Z_s = x_i$.

Research works [30], [31] proved the effectiveness of joint use of rule of type (6) and membership functions. Then model (6) transforms into an expression (7).

$$HCC_{\omega_i}(p + 1) = HCC_{\omega_i}(p) + \mu_{\omega_i}(x_i)[1 - HCC_{\omega_i}(p)] \tag{7}$$

Using these sets of fuzzy decision rules and the synthesis rules of hybrid fuzzy decision rules proposed in research [31], a set of fuzzy decision rules was developed to solve the problem of diagnosing kidney disease. This process is done using a combined fuzzy model:

$$HCC_{OPC}(j + 1) = HCC_{OPC}(j) + HCC_{OPC}(j + 1)[1 - HCC_{OPC}(j)] \tag{8}$$

where, HCC_{OPC} - overall predictive confidence, $HCC_{OPC}(1) = HCC_{lab}$ for signs characterizing laboratory predisposing factors (x_1, x_2, x_3, x_4) functions defined $\mu_1(x_1), \mu_1(x_2), \mu_1(x_3), \mu_1(x_4)$. The membership and partial confidence coefficient are given by:

$$HCC_{lab}(i + 1) = HCC_{lab}(i) + \mu_{\omega_i}(x_i + 1)[1 - HCC_{lab}(i)] \tag{9}$$

where $HCC_{lab}(1) = \mu_{\omega_i}(x_1)$. Similarly for other factors features are x_5, x_6, x_7, x_8 -52. Defuzzification of the output is carried out on the basis of the following implications illustrated with expressions (10) and (11).

$$\text{IF } (HCC_{\omega S1} > HCC_{\omega S2}), \text{ THEN } [pyelonephritis] \text{ OTHERWISE } [Chronic kidney disease] \tag{10}$$

IF ($HCC_{\omega K1} > HCC_{\omega K2}$), *THEN* [kidney stone] *OTHERWISE* [Acute kidney injury] (11)

For the selected classes of diseases, using Delphi technology, belonging function graphs were constructed, shown in Figures 4-8.

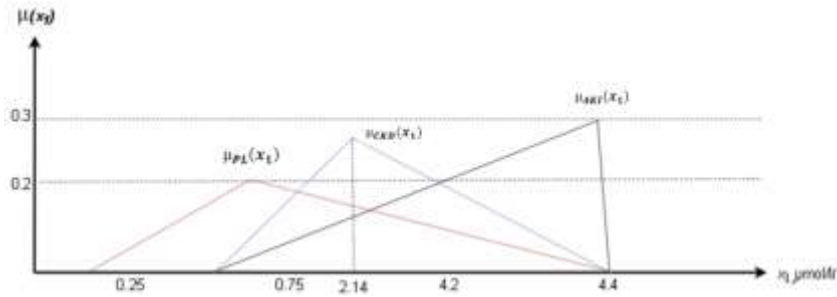


Figure 4. Belonging function graph for feature x_1

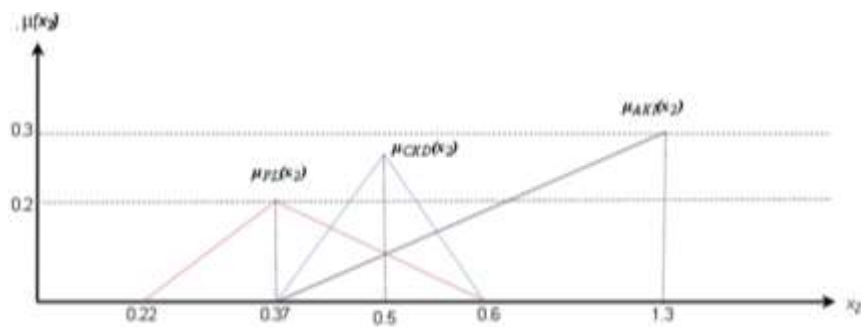


Figure 5. Belonging function graph for feature x_2

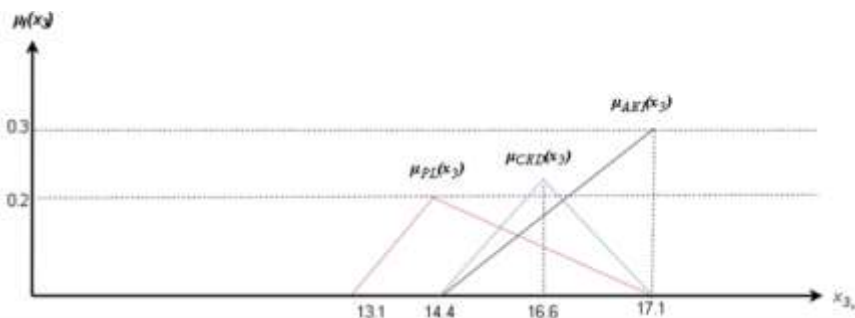


Figure 6. Belonging function graph for feature x_3

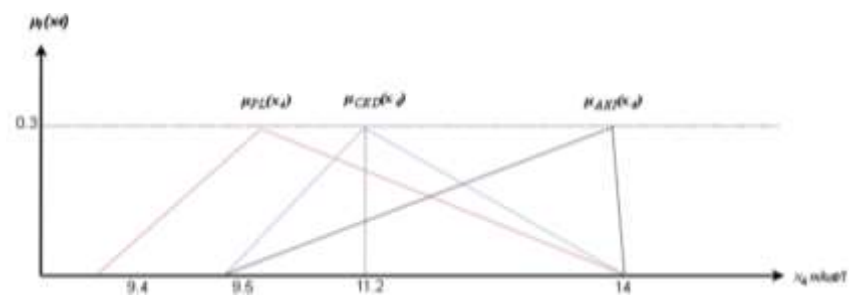


Figure 7. Belonging function graph for feature x_4

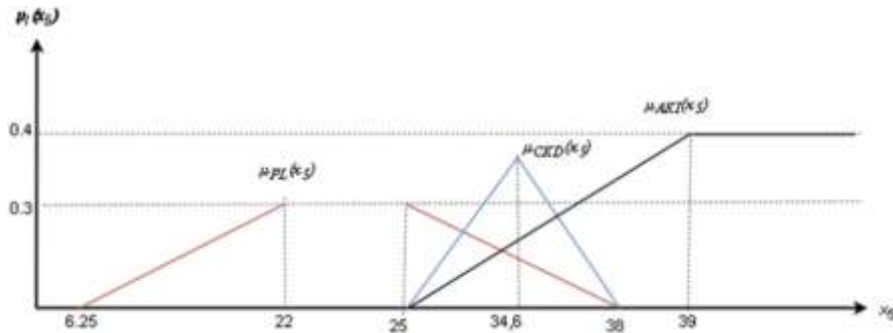


Figure 8. Belonging function graph for feature x_5

Figures 4-8 are illustrated function graphs belong to the kidney disease classes (ω_{PL} , ω_{CKD} , and ω_{AKI}) with basic variables: x_1 ; x_2 ; $x_3 \dots x_5$. In the next stage, we obtain the corresponding analytical expressions using Figures 4-8. For example, we provide analytical calculations of membership functions in selected class states based on characteristics x_1 and x_2 , given by expressions (12)-(17). Below are the analytical expressions for the membership functions of the classes ω_{PL} , ω_{CKD} , and ω_{AKI} based on the levels of bilirubin (factor x_1) and urea (factor x_2):

$$\mu_{PL}(x_1) = \begin{cases} 0, & \text{if } x_1 < 0,19 \\ 0,4x_1 - 0,1, & \text{if } 0,19 \leq x_1 < 0,78 \\ -0,146x_1 + 0,34 & \text{if } 0,78 \leq x_1 < 4,3 \\ 0, & \text{if } x_1 > 4,3 \end{cases} \quad (12)$$

$$\mu_{CKD}(x_1) = \begin{cases} 0, & \text{if } x_1 < 0,32 \\ 0,29x_1 - 0,1, & \text{if } 0,32 \leq x_1 < 5,1 \\ 0,3 & \text{if } x_1 > 5,1 \end{cases} \quad (13)$$

$$\mu_{AKI}(x_1) = \begin{cases} 0, & \text{if } x_1 < 0,17 \\ 0,667x_1 - 0,096, & \text{if } 0,17 \leq x_1 < 0,36 \\ -0,435x_1 + 0,134 & \text{if } 0,36 \leq x_1 < 1,5 \\ 0, & \text{if } x_1 > 1,5 \end{cases} \quad (14)$$

$$\mu_{PL}(x_2) = \begin{cases} 0, & \text{if } x_2 < 16,1 \\ 0,213x_2 - 0,34, & \text{if } 16,1 \leq x_2 < 16,4 \\ -2,15x_2 + 1,4 & \text{if } 16,4 \leq x_2 < 17,1 \\ 0, & \text{if } x_2 > 17,1 \end{cases} \quad (15)$$

$$\mu_{CKD}(x_2) = \begin{cases} 0, & \text{if } x_2 < 11,3 \\ 0,4x_2 - 0,58, & \text{if } 11,3 \leq x_2 < 12,8 \\ -0,065x_2 + 3,4 & \text{if } 12,8 \leq x_2 < 13,6 \\ 0,4, & \text{if } x_2 > 13,6 \end{cases} \quad (16)$$

$$\mu_{AKI}(x_2) = \begin{cases} 0, & \text{if } x_2 < 0,22 \\ 0,322x_2 - 0,12, & \text{if } 0,22 \leq x_2 < 0,37 \\ -0,082x_2 + 0,4, & \text{if } 0,37 \leq x_2 < 0,43 \\ 0,3, & \text{if } x_2 > 0,43 \end{cases} \quad (17)$$

Functions are described in a similar way for each factor from Table 1. Studies indicate that professionals can develop membership functions for individual characteristics but require assistance with fuzzy structures involving multiple dimensions [32]. The E. Shortliffe methodology can be an effective tool for making informed decisions. It employs a fuzzy, iterative, and cumulative approach to tackle complex problems. This methodology substitutes the confidence coefficient with the patient’s symptoms that belong to specific classes, namely ω_{PL} , ω_{CKD} , and ω_{AKI} , as demonstrated in (18)-(20). This approach helps to replicate the reasoning behind decision-making and enhances accuracy.

$$UFP1PL(q + 1) = UFP1PL(q) + \mu_{PL}(x_i + 1)[1 - UFP1PL(q)] \quad (18)$$

$$UFP1CKD(q + 1) = UFP1CKD(q) + \mu_{CKD}(x_i + 1)[1 - UFP1CKD(q)] \quad (19)$$

$$UFP1AKI(q + 1) = UFP1AKI(q) + \mu_{AKI}(x_i + 1)[1 - UFP1AKI(q)] \quad (20)$$

Where $UFP1PL(1)=\mu_{PL}(x_1); UFP1CKD(1)=\mu_{CKD}(x_1); i=1,2,3; q=1,2,3$. Final decisive rules for determining confidence in classes ω_{PL} , ω_{CKD} , and ω_{AKI} aggregated by models of the form presented in expression (21)-(23).

$$UFFPL = UFP1PL + UFP2PL + UFP3PL - UFP1PL \times UFP2PL - UFP1PL \times UFP3PL - UFP2PL \times UFP3PL + UFP1PL \times UFP2PL \times UFP3PL \quad (21)$$

$$UFFCKD = UFP1CKD + UFP2CKD + UFP3CKD - UFP1CKD \times UFP2CKD - UFP1CKD \times UFP3CKD - UFP2CKD \times UFP3CKD + UFP1CKD \times UFP2CKD \times UFP3CKD \quad (22)$$

$$UFFAKI = UFP1AKI + UFP2AKI + UFP3AKI - UFP1AKI \times UFP2AKI - UFP1AKI \times UFP3AKI - UFP2AKI \times UFP3AKI + UFP1AKI \times UFP2AKI \times UFP3AKI \quad (23)$$

Decision on classification according to class diagnosis ω_{PL} , ω_{CKD} , and ω_{AKI} is taken based on the maximum confidence value for the selected classes of patient diseases when it exceeds the threshold, by using expression (24).

$$FUPl = \max(UFF_{PL}, UFF_{CKD}, UFF_{AKI}) \quad (24)$$

Where $l=PL, CKD, AKI, UP^n$ was determined by experts during the work of the expert group, as a result of which the value was determined $UP^n=0,7$. If the confidence scores are equal, the diagnosis is determined in favor of a more severe form of the disease class. With $FUPl \leq UP^n$ a decision is made about the absence of PLs and acute kidney damage. For all tasks, quality indicators generally accepted in medical practice were calculated: diagnostic sensitivity (DS) illustrated with (25), diagnostic specificity (DSp) showed by (26), and diagnostic efficiency (DE) illustrated with (27).

$$DS = \frac{n_l^+}{n} \quad (25)$$

Where n_l^+ is number of correct “match” of the rule ω_l for diseases class $l=PL, CKD, AKI$, n is number of class objects ω_l .

$$DSp = \frac{n_l^-}{n} \quad (26)$$

Where n_l^- number of “does not match” of the rule ω_l for diseases class $l=PL, CKD, AKI$, n is number of class objects ω_l .

$$DE = \frac{n_l^+ + n_l^-}{n_l^+ + n_l^-} \quad (27)$$

3. RESULTS AND DISCUSSION

3.1. Results

The observation results from applying the decision rules based on the recommendations are displayed in the table. The Table 2 presents the results of activation of the decisive rule for the diagnosis of PLs, AKI, and CKD using informative indicators for calculation using formula 24. The presented approach’s accuracy and precision were evaluated by analyzing the medical records of 150 patients diagnosed with AKI, CKD, and PL. Table 2 illustrates the distribution of precise and imprecise results. The number of patients in categories PL, CKD, and AKI were listed, along with the control group’s n_0 patient count.

Table 2. Results of the adequacy of decision rules

Diseases class	Patients	Results	
		Match	Does not match
ω_{PL}	$n_{PL}=150$	125	25
	$n_0=150$	13	137
ω_{CKD}	$n_{CKD}=150$	134	16
	$n_0=150$	9	141
ω_{AKI}	$n_{AKI}=150$	140	10
	$n_0=150$	0	150

n_0 is number of patients

Table 3 presents the outcomes of assessing the suggested instruments for DS, DSp, and DE. Analysis of Table 3 indicates that all metrics are suitable for practical application. Nonetheless, it is important to consider that the quantity of informative attributes within the model has a direct proportional impact on the outcome. This, in turn, influences the time and expense associated with gathering data to address the objectives. The objective focused on crafting hybrid NRPs tailored for distinct problem categories linked to kidney diseases [33]. This involved generating a collection of informative characteristics and intricate markers to accurately depict the health status of patients across different levels of severity. From our perspective, we aim to automate this procedure by implementing a fuzzy hybrid classifier, followed by it is subsequent optimization. The derived fuzzy rules will then serve as a training set for the neural network, facilitating a more efficient and precise diagnostic process.

Table 3. Quality of classification of decision rules

Class	Quality indicators		
	DS	DS _p	DE
FUP_{PL}	0,89	0,87	0,88
FUP_{CKD}	0,91	0,9	0,91
FUP_{AKI}	0,81	0,84	0,83

3.2. Discussions

Diagnostic expert systems have several advantages over traditional diagnostic methods, including increased accuracy, reduced time to diagnosis, and the ability to provide information about diagnostic tests and treatments. They can also be used in low-resource settings where access to specialized medical care may be limited. It is essential to understand that the method proposed in the study cannot serve as a general approach to detecting pyelonephritis or any other kidney disease. However, medical professionals can use it as an additional tool along with basic laboratory and instrumental studies. Also, diagnostic expert systems have certain limitations that must be considered. For example, they rely on the accuracy and completeness of the information they receive and may be unable to help explain rare or complex diseases. Additionally, they may only sometimes make a definitive diagnosis; additional tests and evaluations may be required to confirm the diagnosis. The limitation of the proposed algorithm is that it does not work in case of an emergency.

Despite these limitations, diagnostic expert systems have proven to be valuable tools for healthcare professionals in diagnosing a wide range of diseases. By reducing the time and resources required for accurate diagnosis, they can potentially improve patient outcomes and reduce the overall cost of healthcare. The study’s results lay the foundation for the further use of the developed models and methods to create intelligent medical diagnostic systems that can help diagnose and determine further treatment methods for patients.

4. CONCLUSION

Finally, this research paper has successfully demonstrated the efficacy of employing fuzzy set theory in the non-invasive diagnosis of kidney diseases. Through the innovative application of fuzzy models and algorithms, we have addressed the complex challenge of dealing with the uncertainty and variability inherent in clinical data. Our approach has significantly enhanced the precision and interpretability of diagnostic results, enabling a more nuanced assessment of the functional state of the kidneys that takes into account individual patient characteristics and various risk factors. Throughout the development phase, models for internal medicine were crafted using unconventional instruments. These models demonstrate adequate precision in medical assessments by establishing supplementary functions and enabling the assignment of specific symptoms to the categorization within the group. By utilizing expert evaluations of intricate data, the introduced model has the potential to greatly streamline the diagnostic process. The fuzzy rules derived will then be utilized as a training dataset for the neural network. This entire process is notably

demanding in terms of labor and resources. Moving forward, there is potential to automate this procedure through the application of an optimized fuzzy hybrid classifier.

Furthermore, the morbidity analysis conducted within the Republic of Kazakhstan showed on regional disparities in kidney disease risks, providing a crucial foundation for the formulation of targeted healthcare strategies aimed at mitigating these risks. Our findings have far-reaching implications for the future of medical diagnostics and healthcare delivery. By facilitating the development of more precise and personalized diagnostic tools, we contribute to the advancement of personalized medicine. Moreover, the insights gained from our research can aid in the enhancement of healthcare policies, ensuring better access to medical care and improving the overall health of populations. The use of decision support systems in healthcare is becoming increasingly widespread, as they provide a data-driven and systematic approach to decision-making, and help to overcome some of the limitations of traditional methods.





REFERENCES

- [1] D. M. Wells and J. Niederer, "A medical expert system approach using artificial neural networks for standardized treatment planning," *International Journal of Radiation Oncology Biology Physics*, vol. 41, no. 1, pp. 173–182, Apr. 1998, doi: 10.1016/S0360-3016(98)00035-2.
- [2] W. Zhang, C. Zhang, J. Su, P. Huang, and J. Zhou, "Design of automated medicine-liquid proportioning system based on spectral analysis," in *2020 6th International Conference on Control, Automation and Robotics, ICCAR 2020*, Apr. 2020, pp. 563–567, doi: 10.1109/ICCAR49639.2020.9108091.
- [3] R. R. R. Ikram, M. K. A. Ghani, and N. Abdullah, "An analysis of application of health informatics in traditional medicine: a review of four traditional medicine systems," *International Journal of Medical Informatics*, vol. 84, no. 11, pp. 988–996, Nov. 2015, doi: 10.1016/j.ijmedinf.2015.05.007.
- [4] W. H. Wiese, *Public health informatics and information systems*, vol. 25, no. 1. Springer International Publishing, 2003.
- [5] A. Sheikhtaheri, F. Sadoughi, and Z. H. Dehaghi, "Developing and using expert systems and neural networks in medicine: a review on benefits and challenges," *Journal of Medical Systems*, vol. 38, no. 9, Jul. 2014, doi: 10.1007/s10916-014-0110-5.
- [6] A. Charleonnan, T. Fufaung, T. Niyomwong, W. Chokchueypattanakit, S. Suwannawach, and N. Ninchawee, "Predictive analytics for chronic kidney disease using machine learning techniques," in *2016 Management and Innovation Technology International Conference, MITiCON 2016*, Oct. 2017, pp. 80–83, doi: 10.1109/MITiCON.2016.8025242.
- [7] "Statistical digest Health of the population of the Republic of Kazakhstan," 2022. https://nrchd.kz/index.php/ru/?option=com_content&view=article&id=973. Accessed December 2023.
- [8] W. R. Swartout, "Rule-based expert systems: The mycin experiments of the stanford heuristic programming project. B.G. Buchanan and E.H. Shortliffe. (Addison-Wesley, Reading, MA, 1984); 702 pages, \$40.50," *Artificial Intelligence*, vol. 26, no. 3, pp. 364–366, 1985, doi: 10.1016/0004-3702(85)90067-0.
- [9] M. Furmankiewicz, J. Furmankiewicz, and P. Ziuziański, "Evaluation of the expert system as a stage of the life cycle model ESDLC on the example of WIKex," *Computer Science and Mathematical Modelling*, vol. 2, no. 2, pp. 23–32, Dec. 2015, doi: 10.5604/15084183.1197448.
- [10] A. Bohr and K. Memarzadeh, "The rise of artificial intelligence in healthcare applications," in *Artificial Intelligence in Healthcare*, Elsevier, 2020, pp. 25–60.
- [11] M. Mirmozaffari, "Developing an expert system for diagnosing liver diseases," *European Journal of Engineering and Technology Research*, vol. 4, no. 3, pp. 1–5, Mar. 2019, doi: 10.24018/ejeng.2019.4.3.1168.
- [12] K. B. Nahato, K. N. Harichandran, and K. Arputharaj, "Knowledge mining from clinical datasets using rough sets and backpropagation neural network," *Computational and Mathematical Methods in Medicine*, pp. 1–13, 2015, doi: 10.1155/2015/460189.
- [13] I. Dankwa-Mullan, M. Rivo, M. Sepulveda, Y. Park, J. Snowdon, and K. Rhee, "Transforming diabetes care through artificial intelligence: the future is here," *Population Health Management*, vol. 22, no. 3, pp. 229–242, Jun. 2019, doi: 10.1089/pop.2018.0129.
- [14] F. Sanmarchi, C. Fanconi, D. Golinelli, D. Gori, T. Hernandez-Boussard, and A. Capodici, "Predict, diagnose, and treat chronic kidney disease with machine learning: a systematic literature review," *Journal of Nephrology*, vol. 36, no. 4, pp. 1101–1117, Feb. 2023, doi: 10.1007/s40620-023-01573-4.
- [15] A. V. Novikov, "Fuzzy decision-making technique and algorithm for managing the provision of medical care to patients with pyelonephritis," 2009, <https://www.disserscat.com/content/metodologiya-nechetkogo-prinyatiya-reshenii-i-algoritmizatsiya-upravleniya-okazaniem-meditsi>
- [16] N. A. Korenevskiy *et al.*, "Mathematical models of differential diagnostics of serious and purulent pyelonephritis in pregnant women based on fuzzy logic for decision-making," *Critical Reviews in Biomedical Engineering*, 2021.
- [17] M. M. Hassan *et al.*, "A comparative study, prediction and development of chronic kidney disease using machine learning on patients clinical records," *Human-Centric Intelligent Systems*, vol. 3, no. 2, pp. 92–104, Feb. 2023, doi: 10.1007/s44230-023-00017-3.
- [18] I. A. Huqqani, L. T. Tay, and J. M. Saleh, "Analysis of landslide hazard mapping of Penang island Malaysia using bivariate statistical methods," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 16, no. 2, pp. 781–786, Nov. 2019, doi: 10.11591/ijeecs.v16.i2.pp781-786.
- [19] A. B. Myrzakerimova, M. Shaizat, K. S. Duisebekova, and M. O. Nurmaganbetova, "Forecasting risk of diseases in Kazakhstan with using mapping technique based on 9 years statistics," *Procedia Computer Science*, vol. 170, pp. 75–81, 2020, doi: 10.1016/j.procs.2020.03.010.
- [20] E. I. Papageorgiou, "Fuzzy cognitive map software tool for treatment management of uncomplicated urinary tract infection," *Computer Methods and Programs in Biomedicine*, vol. 105, no. 3, pp. 233–245, Mar. 2012, doi: 10.1016/j.cmpb.2011.09.006.
- [21] A. Myrzakerimova, K. Kolesnikova, and M. Nurmaganbetova, "Development of the structure of an automated system for diagnosing diseases," *Scientific Journal of Astana IT University*, pp. 101–112, Dec. 2022, doi: 10.37943/12avge4585.
- [22] A. Mardani *et al.*, "Application of decision making and fuzzy sets theory to evaluate the healthcare and medical problems: a review of three decades of research with recent developments," *Expert Systems with Applications*, vol. 137, pp. 202–231, Dec. 2019, doi: 10.1016/j.eswa.2019.07.002.





- [23] N. Alpysbay, K. Kolesnikova, T. Chinibaeva, and T. Olekh, "Using project management tools in the process of modernizing the healthcare system of the Republic of Kazakhstan," *CEUR Workshop Proceedings*, vol. 3382, 2022.
- [24] V. Kunwar, K. Chandel, A. S. Sabitha, and A. Bansal, "Chronic Kidney disease analysis using data mining classification techniques," in *Proceedings of the 2016 6th International Conference-Cloud System and Big Data Engineering, Confluence 2016*, Jan. 2016, pp. 300–305, doi: 10.1109/CONFLUENCE.2016.7508132.
- [25] N. Chetty, K. S. Vaisla, and S. D. Sudarsan, "Role of attributes selection in classification of chronic kidney disease patients," in *Proceedings of the 2015 International Conference on Computing, Communication and Security (ICCCS), IEEE, Pamplemousses, Mauritius*, 2015. <https://ieeexplore.ieee.org/document/7374193>
- [26] N. A. Korenevskiy, "Use of fuzzy logic decision making in medical expert systems". *Med Tekh.* 2015 Jan-Feb;(1):33-5. Russian. PMID: 25980218. <https://pubmed.ncbi.nlm.nih.gov/25980218/>
- [27] M. M. El Sherbiny, E. Abdelhalim, H. E.-D. Mostafa, and M. M. El-Seddik, "Classification of chronic kidney disease based on machine learning techniques," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 32, no. 2, pp. 945–955, Nov. 2023, doi: 10.11591/ijeecs.v32.i2.pp945-955.
- [28] L. A. Zadeh, "Advances in fuzzy mathematics and engineering: fuzzy sets and fuzzy information-granulation theory," *Beijing Normal University*, 2005.
- [29] L. A. Zadeh, "Fuzzy sets," *Information and Control*, vol. 8, no. 3, pp. 338–353, Jun. 1965, doi: 10.1016/S0019-9958(65)90241-X.
- [30] P. Shah and T. Shah, "Adaptive neuro fuzzy inference system based classifier in diagnosis of breast cancer," *Results in Control and Optimization*, vol. 14, Mar. 2024, doi: 10.1016/j.rico.2023.100358.
- [31] J. B. Awotunde, O. E. Matiluko, and O. W. Fatai, "Medical diagnosis system using fuzzy logic," *African Journal of Computing & ICT*, vol. 7, no. 2, pp. 99–106, 2014, [Online]. Available: www.ajocict.net.
- [32] H. Bensenane, D. Aksa, W. F. Omari, and A. Rahmoun, "A deep learning-based cardio-vascular disease diagnosis system," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 25, no. 2, pp. 963–971, Feb. 2022, doi: 10.11591/ijeecs.v25.i2.pp963-971.
- [33] T. Saroja and Y. Kalpana, "Hybrid missing data imputation and novel weight convolution neural network classifier for chronic kidney disease diagnosis," *Measurement: Sensors*, vol. 27, Jun. 2023, doi: 10.1016/j.measen.2023.100715.

BIOGRAPHIES OF AUTHORS



Alua Myrzakerimova     received the M.Sc. degree at Nottingham University, UK. She is senior lecturer of Information Systems department, International IT University, Almaty, Kazakhstan. She is Ph.D. candidate at International IT University. Her research interests include using IT technologies in medicine, diagnosing diseases, intelligent systems, mathematics. She can be contacted by email: a.ospan@iitu.edu.kz.



Kateryna Kolesnikova     Doctor of technical science, professor of Information Systems department, Vice-Rector for Research Activities at International IT University, Almaty, Kazakhstan. She has supervised and co-supervised more than 30 masters and 15 Ph.D. students. She has authored or coauthored more than 100 publications, where more than 31 scientific papers, which are indexed on Scopus and Web of Science, with 9 H-index. Her research interests include information technology used in project management, modeling complex systems, manage IT projects in various fields. She can be contacted by email: kkolesnikova@iitu.edu.kz.