Article

Development of method to analyze factors of kidney disease by the use of fuzzy logic

Assel Yembergenova* [1], Azamat Serek [2], and Bauyrzhan Berlikozha [3]

[1]Department of Computer Science, SDU University, Kaskelen, Kazakhstan
[2]School of Information Technologies and Engineering, Kazakh-British Technical University (KBTU), Almaty, Kazakhstan
[3]School of digital engineering technology Narxoz University, Almaty, Kazakhstan

DOI: 10.47344/9pcv0t26

Abstract

The study introduces a new strategy for the analysis of kidney disease parameters based on fuzzy logic. Fuzzy logic is a more accurate way to categorize clinical parameters than statistical analysis because there is uncertainty and variability in medical data. The data is comprised of an extensive amount of clinical parameters including age, blood pressure, specific gravity, albumin, sugar, random blood glucose, blood urea, serum creatinine, sodium, potassium, hemoglobin, packed cell volume, white blood cell count, and red blood cell count.

The methodology utilizes fuzzy logic centroid computation to categorize these parameters into low, medium, and high levels to provide a more dynamic and interpretable assessment of renal health. Fuzzy memberships give the current work the capability to discover intricate interrelationships between clinical variables, which may have been otherwise unattainable by conventional mean, median, and standard deviation-based analyses.

The findings confirm that fuzzy logic and conventional statistical methods enhance the comprehension of kidney disease by incorporating intricate interactions between clinical variables. The method is employed to achieve more accurate prediction and diagnostic models, offering insight to be used in kidney disease assessment and medical decisions.

Keywords: fuzzy logic, kidney disease, clinical parameters, diagnostic modeling, centroid analysis.

*Corresponding author: assel.yembergenova@sdu.edu.kz ORCID: 0009-0005-9051-2070
Email (2): a.serek@kbtu.kz ORCID:0009-0003-8644-7427
Email (3): bauyrzhan.berlikozha@gmail.com ORCID:0009-0003-9096-8719

Received: February 17, 2025. Reviewed: March 19, 2025. Accepted: March 19, 2025. *© 2025 Assel Yembergenova, Azamat Serek and Bauyrzhan Berlikozha. All rights reserved.*

I. Introduction

Kidney disease is a global health problem, and it is found in millions of patients all over the globe, causing morbidity and death in populations all over the world [1]–[3]. The blood pressure, age, and biochemical markers' correlation is most important to establish the kidney's health status. Statistical measures have their applications, but they cannot reflect the natural uncertainty and subtlety in medical data.

The current work introduces a new methodology based on fuzzy logic analysis to separate the complex interaction between disease-related parameters. Fuzzy logic is particularly appropriate to work with imprecise and incomplete data and therefore is a perfect tool to unveil hidden trends, which may not be visible to traditional statistical analysis. With the increasing prevalence of chronic kidney disease, an early and accurate diagnosis is critical to successful management. In response, exploring new analysis methods is relevant to increase diagnostic accuracy and direct clinical decision.

The final goal of the study is to obtain fuzzy membership functions and centroids to quantify and describe trends in the variables of kidney disease. With the application of fuzzy logic to traditional descriptive statistics, we will have a more meaningful understanding of the shape and distribution of the variables, and a better understanding of kidney health.

Fuzzy logic is applied across various fields because of its ability to deal with uncertainty and imprecise values. In control systems, it is applied on a large scale to increase accuracy and flexibility and provide more powerful solutions compared to traditional binary logic [4]–[7]. Fuzzy logic is applied in the automotive industry to improve vehicle stability, fuel efficiency, and performance. Fuzzy logic is applied in medicine to simulate diagnostic neurological and cardiovascular diseases [8]–[11]. Fuzzy logic is applied to assist in financial modeling by dealing with uncertainty in the market, improving risk analysis, and improving investment decisions [12]–[14].

The paper examines a 2019 Chronic Kidney Disease data set [15] with key kidney-related variables. The study attempts to unravel complex relations and trends by performing fuzzy logic analysis, and it opens a door to explore kidney disease development and their major clinical implications.

The subsequent sections encapsulate the research strategy employed, key findings, and implications thereof for the development of diagnostic tests and treatments exactly aimed at attacking kidney disease.

II. Literature Review

One of these fields that have been subjected to intense research by multiple computational approaches is kidney disease detection. Authors in the article [16] have analyzed the serious issue of kidney failure, with an emphasis on slowing CKD progression and minimizing economic burdens to treatments. They have based their work on risk factors causing kidney degeneration, with an emphasis on conducting regular checks to determine risks before severe health decline.

Unlike other research that employed conventional statistical techniques, researchers [17] employed fuzzy and adaptive neural fuzzy inference systems to diagnose CKD more precisely. They attempted to increase the reliability of medical tests to diagnose disease. They considered important parameters such as nephron function, glucose, blood pressure, age, body mass index (BMI), and smoking status in minute detail while creating a fuzzy inference system. Their system distinguished between CKD stages 1 to 5 and provided real-time results on the extent of the disease. Their fuzzy system was simulated with the assistance of MATLAB to demonstrate how it can be employed in real-world healthcare applications.

One contribution to the study was by authors in [18], where they found natural imprecision in CKD detection in clinics. They used a fuzzy inference system in MATLAB to correct the issue, based on the capability of fuzzy logic to handle uncertainty. Their work, however, did not have a robust validation system, and there is a chance to enhance prediction accuracy.

In contrast to other studies, the current work is centered on the development of a more accurate fuzzy expert system (FES) from a large dataset, clinical guidelines, and expert opinion to enhance CKD diagnosis.

This contribution is original in its application of systematic tests for normality to establish the input parameter effect and system performance assessment by exhaustive testing. Surface analysis identified nephron function, blood sugar, and body mass index (BMI) as the most accurate parameters to predict CKD. The FES was validated with 80 test cases, and the accuracy level was 93.75%, confirming its feasibility in real-world applications.

Follow-up studies also reiterated the significance of CKD early detection. The authors in [19] and [20] employed deep learning models to predict CKD, with satisfactory performance but requiring large training data and computational resources. Our study bridges the gap by presenting an interpretable and computationally light alternative based on fuzzy logic.

The current work introduces a new prediction model based on the fuzzy logic toolbox in MATLAB to screen CKD. There are five key steps in the process: (1) selection of key input parameters such as blood urea nitrogen, eGFR, and serum creatinine, (2)

Received: February 17, 2025. Reviewed: March 19, 2025. Accepted: March 19, 2025. *© 2025 Assel Yembergenova, Azamat Serek and Bauyrzhan Berlikozha. All rights reserved.*

input-output relation normalization by carrying out min-max normalization, (3) construction of a fuzzy inference engine, (4) fuzzy rule aggregation, and (5) defuzzification to obtain clear diagnostic outputs. 70 test cases from patients have been analyzed, and 47 have been diagnosed as CKD positive, which validated the reliability of the proposed model.

With the inclusion of results from past studies and enhancement of fuzzy logic-based techniques, the work is meant to enhance diagnostic accuracy and achieve early detection of CKD. The study results demonstrate the potential of fuzzy expert systems to aid clinical decision and enhance healthcare treatments.

III. Methods

The study makes use of a range of data collected from (Chronic kidney disease data set, 2019), with differing parameters such as age, blood pressure, specific gravity, albumin, sugar, random blood 5 glucose, blood urea, serum creatinine, sodium, potassium, hemoglobin, packed cell volume, white cell count, and red cell count, and so on.

The dataset, formulated by merging two consecutive hospital reports, is an abstraction of one of the most important prediction modeling indicators of chronic kidney disease (CKD). Since its multivariate set is heterogeneous, it is appropriate to use in classification applications in health care and is important to use in order to gain insight into causality of chronic kidney disease. The dataset comprises 25 attributes and 400 instances, all being of utmost relevance that will be used to contribute to prediction modeling.

The set of data is composed of physiologic tests, clinical features, and lab values that describe a diverse set of conditions of the patients. Key features include blood pressure (bp), specific gravity (sg), albumin (al), sugar (su), red cells (rbc), pus cells (pc), pus cell clumps (pcc), bacteria (ba), random blood glucose (bgr), blood urea (bu), serum creatinine (sc), sodium (sod), potassium (pot), hemoglobin (hemo), packed cell volume (pcv), white cell count (wc), red cell count (rc), and target feature (class) that describe absence or presence of chronic kidney disease (CKD).

The wide range of features enables effective classification modeling and contributes to establishing causes of CKD. The compound nature of the dataset makes it suitable to use within processes of machine learning of CKD or its initiation forecasting, enabling possible better-timed effective treatments.

The following are processes that fall under fuzzy logic analysis:

- *Variable selection.* The most important identified variables were age, blood pressure, and biochemical markers. Fuzzy logic is developed on top of these variables.
- *Membership function generation.* Membership functions corresponding to low, medium, and high levels of membership were created using skfuzzy library for every identified variable. The functions give a pictorial representation of every range of variables of the dataset.
- *Centroid Calculation.* For determination of numerical values of each level of each variable range, each of centroid, or centre of gravity, of the membership function were computed
- *Descriptive Statistics.* For better understanding of the dataset, mean, median, and standard deviation were also calculated on each of the variables.

Fig. 1 summarizes briefly the key steps followed to use fuzzy logic on parameters of renal disease, from preprocessed data through fuzzification, inference system designing, defuzzification, to classification verification.

All processes are organized to adhere to a process of fuzzy logic that is implemented using clinical data guidelines.

The aim of merging fuzzy logic analysis into classical descriptive statistics is to have a complete understanding of complex relationships among factors of renal disease. The outcomes of this process will provide valuable insight to practice, informing practice and research of kidney health in the future.

IV. Results and Discussion

Table 1 presents the fuzzy analysis of clinical attributes pertaining to kidney disease in low, medium, and high levels. This enhances understanding by applying mean, median, and standard deviation to determine patterns of distribution, range, and variability in the dataset. Fuzzy logic centroids provide more accurate classification of clinical attributes.

The fuzzy logic analysis of the centroids (low, medium, and high) provides a mean of around 51.48 years and a median of 55.00 years, showing a reasonably symmetric distribution. The standard deviation of 17.15 indicates a moderate spread in the distribution of age, with a broad range of values.

Blood pressure classification shows that the mean (76.47) and median (80.00) are almost equal, which implies a symmetric distribution. The standard deviation of 13.67 reflects moderate variation in blood pressure level between patients.

Received: February 17, 2025. Reviewed: March 19, 2025. Accepted: March 19, 2025. *© 2025 Assel Yembergenova, Azamat Serek and Bauyrzhan Berlikozha. All rights reserved.*

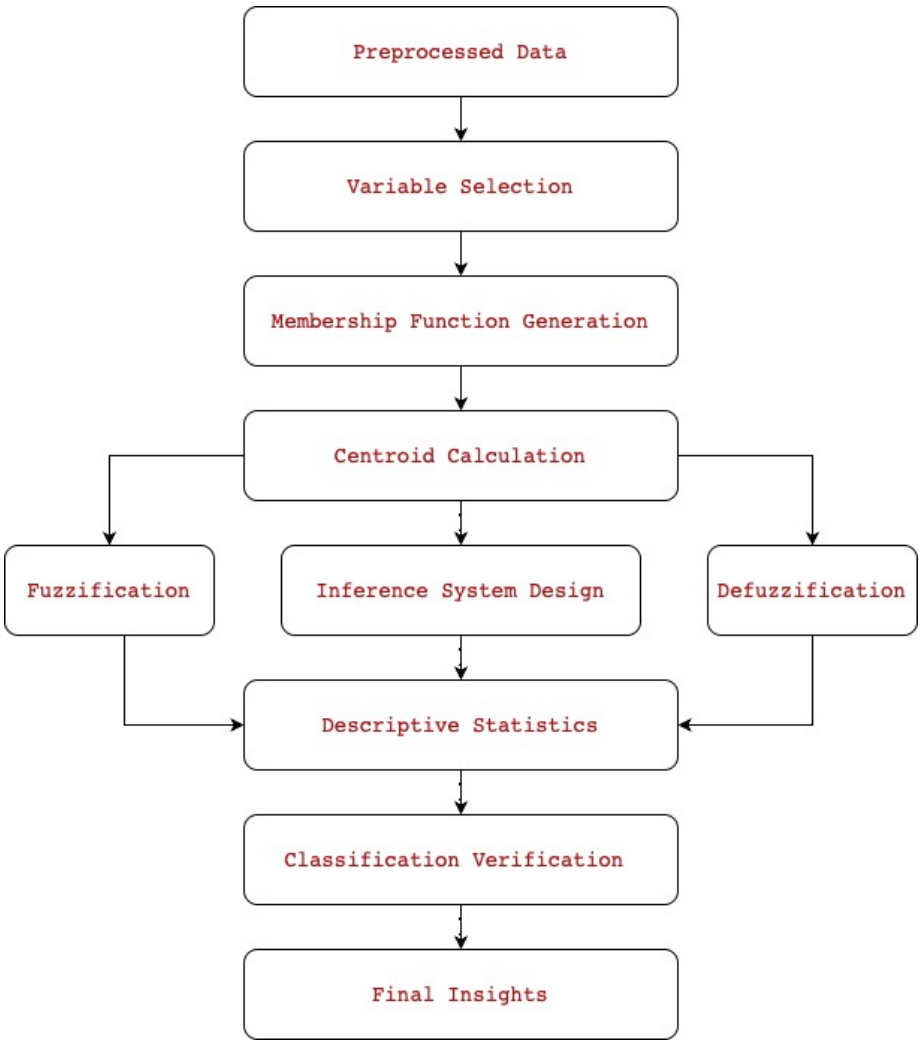

Fig. 1: Flowchart summarizing the fuzzy logic-based methodology for kidney disease analysis

Specific gravity shows a clear distinction between centroids, with the low level being 1.01 and medium and high levels being 0.00. The close proximity between the mean and median and a standard deviation of 0.01 indicate homogeneity, which suggests that values of specific gravity are comparatively consistent across levels of kidney disease.

The albumin values reflect a striking difference between the median (0) and the mean (1.02), indicating a skewed distribution. The standard deviation of 1.35 reflects moderate variation in albumin values, an indication of variable degrees of kidney dysfunction in the patients.

Red blood cell count reveals a closely similar mean (4.71) and median (4.8), indicating a near-symmetrical distribution. The standard deviation of 1.02 indicates moderate variation, and it could reflect progression in anemia in patients with renal disease.

Other parameters, including blood glucose, blood urea, serum creatinine, sodium, potassium, hemoglobin, packed cell volume, and white blood cell count, have varying degrees of skewness and spread. An example is blood glucose random levels, which have a high standard deviation (79.17), indicating the great range in blood sugar levels between patients, an important aspect in the

Received: February 17, 2025. Reviewed: March 19, 2025. Accepted: March 19, 2025. *© 2025 Assel Yembergenova, Azamat Serek and Bauyrzhan Berlikozha. All rights reserved.*

Algorithm 1 Fuzzy Logic-based Analysis of Kidney Disease Factors

Input: Dataset D with features $F = \{f_1, f_2, \ldots, f_n\}$, where $f_1, f_2, \ldots, f_n$ are clinical parameters related to kidney disease
Output: Fuzzy logic centroids for classification of kidney disease factors
Step 1: Data Preprocessing
Clean dataset D by handling missing values and outliers
Normalize or scale features in F if required for uniformity
Step 2: Fuzzification
for each feature $f_i \in F$ **do**
    Define membership functions $\mu_{\text{low}}(f_i), \mu_{\text{medium}}(f_i), \mu_{\text{high}}(f_i)$
    Apply fuzzification to convert crisp values of f_i into fuzzy sets
end for
Step 3: Fuzzy Inference System (FIS) Design
Define fuzzy rules based on medical knowledge and clinical thresholds
for each feature $f_i \in F$ **do**
    Use predefined rules to establish relationships between input features and kidney disease diagnosis
    Apply fuzzy operators (e.g., AND, OR) to combine rules
end for
Step 4: Defuzzification
for each fuzzy output **do**
    Calculate the centroid of the fuzzy output using the centroid method:

$$\text{Centroid} = \frac{\sum x \cdot \mu(x)}{\sum \mu(x)}$$

    Convert fuzzy outputs into crisp values for interpretation
end for
Step 5: Classification and Evaluation
Use defuzzified values to classify the clinical parameters into risk categories (low, medium, high)
Evaluate the classification accuracy using standard performance metrics (e.g., accuracy, precision, recall)
End

Fig. 2: Pseudocode of methodology

development of the disease. Similarly, serum creatinine is very variable (standard deviation 5.73), as is to be expected given that it is a key indicator of renal function.

Clinical Implications. The statistical trends identified in the present study have important clinical implications. By classifying clinical features into low, medium, and high levels based on fuzzy logic, medical professionals are able to:

- *Improve Early Diagnosis.* lassification of parameters such as blood pressure, albumin, and serum creatinine into levels enables early-stage kidney disease to be diagnosed so that action can be taken in a timely fashion.
- *Improve Patient Monitoring.* Regular monitoring of blood urea, serum creatinine, and hemoglobin levels will help clinicians monitor disease progression and adjust treatment regimens.
- *Facilitate Personalized Treatment.* With an understanding of the distribution and variation in important parameters, physicians can individualize medication and diet recommendations to fit patient profiles.

Received: February 17, 2025. Reviewed: March 19, 2025. Accepted: March 19, 2025. *© 2025 Assel Yembergenova, Azamat Serek and Bauyrzhan Berlikozha. All rights reserved.*

TABLE I: Summary of Experimental Results

	Centroid Low	Centroid Medium	Centroid High	Mean	Median	Std Dev
Age	18.5	47.82	76.49	51.48	55.00	17.15
Blood Pressure (bp)	158.83	102.15	144.82	76.47	80.00	13.67
Specific Gravity (sg)	1.01	0	0	1.02	1.02	0.01
Albumin (al)	10.35	1.88	3	1.02	0	1.35
Sugar (su)	0.33	1.88	2.77	0.45	0	1.1
Blood Glucose Random (bgr)	64.01	220.01	375.35	148.04	121.00	79.17
Blood Urea (bu)	120.14	149.98	279.48	57.43	42.00	50.44
Serum Creatinine (sc)	1.34	26.51	51.29	3.07	1.3	5.73
Sodium (sod)	48.84	101.67	154.18	137.53	138.00	10.39
Potassium (pot)	3.25	18.05	32.54	4.63	4.4	3.19
Hemoglobin (hemo)	6.26	11.1	15.54	12.53	12.65	2.91
Packed Cell Volume (pcv)	118.96	33.93	48.29	38.88	40.00	8.98
White Blood Cell Count (wc)	4268.71	12335.37	20401.37	8406.12	8000.00	2939.46
Red Blood Cell Count (rc)	3.03	4.83	6.24	4.71	4.8	1.02

A. Discussion

The fuzzy logic analysis provides a complete examination of the membership function and distribution characteristics of the variables of kidney disease. Some of the observations and their potential implications are as follows:

Centroids for low, medium, and high membership levels in age indicate an even distribution across the range of the ages. This is an indication that the dataset represents a range of ages, as is fitting to explore kidney health throughout all stages in one's life.

The diversity enhances the capability of the dataset to give meaningful insights into risk factors based on kidney disease by age. The distribution of blood pressure indicates a prevalent cluster in medium membership levels. This result indicates that the subjects in the dataset have mostly moderate blood pressure, which may have an effect on kidney health. Additional studies are needed to explore moderate blood pressure and outcomes in kidney disease. Specific gravity (sg) Anomalies: Specific gravity indicates an abnormality, with centroids measured as 1.01 and 0.00, respectively, for low and medium and high membership levels. In spite of these abnormalities, the median and the mean values are within expected ranges. This abnormality needs to be explored to establish if there is an issue with the quality of data or to explore special features in the dataset.

Centroids for sugar and albumin indicate clear progression from low to high as would be clinically expected. Higher glucose and albumin concentrations are typical to indicate renal impairment, and the distinctly clear fuzzy logic centroids indicate strong ability to classify. The results provide a strong basis to determine risk levels for kidney disease. The blood glucose random values have large variance, and their fuzzy logic centroids are dispersed across low, medium, and high membership values. This indicates strong variance of blood glucose levels, underlining how important it is to account for renal function. The strong variance of distribution underlines the necessity to carefully determine how renal function is influenced by blood glucose levels. The white cell count centroids indicate a strong shift to higher counts, indicating varied immunological responses in sets. More intense observation of correlation between immunological and renal functions could indicate important correlation with practice implications on managing and predicting kidney disease. 9 The centroids of the red cell counts indicate symmetric count distribution across counts of low, medium, and high values. The resulting homogeneity indicates balanced coverage of the counts of the red cells in sets, hence eliminating sources of bias and providing reliability to corresponding analyses of other parameters of interest of a clinical nature.

The findings justify the use of fuzzy logic to establish complex relations in kidney disease data. The process ensures complete classification of parameters of interest of a clinical type, hence better comprehension of complex determinants of renal function. Future studies should continue to evolve based on findings to establish variable interactions and related practice implications. Finally, correction of aberrations, as that of specific gravity, requires complete determination of related data to give reliability to subsequent assessments.

V. Conclusion and Future Work

Fuzzy modeling of clinical parameters adds a new dimension to kidney disease analysis, with more refined grouping of variables as low, medium, and high. This grouping enables us to understand better how parameters influence kidney function. The addition of

Received: February 17, 2025. Reviewed: March 19, 2025. Accepted: March 19, 2025. *© 2025 Assel Yembergenova, Azamat Serek and Bauyrzhan Berlikozha. All rights reserved.*

fuzzy logic centroids and traditional measures, such as mean, median, and standard deviation, provides a complete perspective on variance and central tendency. The key findings are that parameters such as random blood glucose, blood urea, and serum creatinine have high variance, and they can be potential markers in the progression of kidney disease.

Despite these advances, there are a few restrictions to this work. The dataset used may not capture all the variation present in patient populations with diversity, and the fuzzy logic-based approach cannot account for potential non-linear interactions between parameters. Additionally, the lack of direct comparison with conventional machine learning methods prevents assessment regarding the relative performance of fuzzy logic in clinical decision support.

Future studies should integrate this fuzzy logic-based categorization into clinical decision support systems to assist medical personnel in risk assessment and early diagnosis. Comparison to machine learning techniques, including decision trees, support vector machines, and deep learning techniques, would provide more proof of efficacy. Larger dataset size and real-time patient monitor data could also enhance the predictability of the model, adding to its utility in the clinical environment.

REFERENCES

[1] Cockwell, P., Fisher, L.-A. (2020). The global burden of chronic kidney disease. The Lancet, 395(10225), 662–664.

[2] Jha, V., Garcia-Garcia, G., Iseki, K., Li, Z., Naicker, S., Plattner, B., Saran, R., Wang, A. Y.-M., Yang, C.-W. (2013). Chronic kidney disease: Global dimension and perspectives. The Lancet, 382(9888), 260–272.

[3] Mills, K. T., Xu, Y., Zhang, W., Bundy, J. D., Chen, C.-S., Kelly, T. N., Chen, J., He, J. (2015). A systematic analysis of worldwide population-based data on the global burden of chronic kidney disease in 2010. Kidney International, 88(5), 950–957.

[4] Cherry, A., Jones, R. (1995). Fuzzy logic control of an automotive suspension system. IEE Proceedings-Control Theory and Applications, 142(2), 149–160.

[5] Ivanov, V. (2015). A review of fuzzy methods in automotive engineering applications. European Transport Research Review, 7(3), 1–10.

[6] Von Altrock, C., Krause, B., Zimmermann, H.-J. (1992). Advanced fuzzy logic control technologies in automotive applications. In [1992 Proceedings] IEEE International Conference on Fuzzy Systems (pp. 835–842). IEEE.

[7] Vachtsevanos, G., Farinwata, S. S., Pirovolou, D. K. (1993). Fuzzy logic control of an automotive engine. IEEE Control Systems Magazine, 13(3), 62–68.

[8] Baig, F., Khan, M. S., Noor, Y., Imran, M., Baig, F., others. (2011). Design model of fuzzy logic medical diagnosis control system. International Journal on Computer Science and Engineering, 3(5), 2093–2108.

[9] Awotunde, J. B., Matiluko, O. E., Fatai, O. W. (2014). Medical diagnosis system using fuzzy logic. African Journal of Computing ICT, 7(2), 99–106.

[10] Das, S., Guha, D., Dutta, B. (2016). Medical diagnosis with the aid of using fuzzy logic and intuitionistic fuzzy logic. Applied Intelligence, 45, 850–867.

[11] Phuong, N. H., Kreinovich, V. (2001). Fuzzy logic and its applications in medicine. International Journal of Medical Informatics, 62(2–3), 165–173.

[12] Gil-Lafuente, A. M., others. (2005). Fuzzy logic in financial analysis. Springer.

[13] Korol, T. (2012). Fuzzy logic in financial management. INTECH Open Access Publisher.

[14] Mohammadian, M., Kingham, M. (1997). Hierarchical and feed-forward fuzzy logic for financial modelling and prediction. In Australian Joint Conference on Artificial Intelligence (pp. 147–156). Springer.

[15] Chronic kidney disease data set. (2019). UCI Machine Learning Repository.https://archive.ics.uci.edu/dataset/336/chronic+kidney+disease (Accessed: 1 September 2023).

[16] Ahmed, T. I., Bhola, J., Shabaz, M., Singla, J., Rakhra, M., More, S., Samori, I. A., others. (2022). Fuzzy logic-based systems for the diagnosis of chronic kidney disease. BioMed Research International, 2022.

[17] Bai, Y., Wang, D. (2006). Fundamentals of fuzzy logic control—Fuzzy sets, fuzzy rules and defuzzifications. In Advanced fuzzy logic technologies in industrial applications (pp. 17–36).

[18] Celikyilmaz, A., Turksen, I. B. (2009). Modeling uncertainty with fuzzy logic. Studies in Fuzziness and Soft Computing, 240(1), 149–215.

[19] John, R. I., Innocent, P. R. (2005). Modeling uncertainty in clinical diagnosis using fuzzy logic. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 35(6), 1340–1350.

[20] Larsen, P. M. (1980). Industrial applications of fuzzy logic control. International Journal of Man-Machine Studies, 12(1), 3–10.

Received: February 17, 2025. Reviewed: March 19, 2025. Accepted: March 19, 2025. *© 2025 Assel Yembergenova, Azamat Serek and Bauyrzhan Berlikozha. All rights reserved.*