



Design of an Intelligent Fuzzy System for Disease Prediction and Drug Dosage Control

Dr. Rafia Aziz¹, Dr. A.K. Singh^{2*}, Dr. Ashish Kumar Soni³

Assistant Professor, Department of Mathematics, Government Holker Science College, Indore, India¹

Assistant Professor, Department of Applied Sciences, RBS Engineering Technical Campus, Agra (India)²

Assistant Professor, Department of Mathematics, Mediacaps University Indore, India³

*Corresponding Author

Abstract: The proposed research presents the design and implementation of an intelligent fuzzy system that integrates disease prediction with personalized drug dosage control. The framework utilizes patient-specific data, including clinical, demographic, and physiological parameters, to predict disease probabilities and recommend safe drug doses through fuzzy inference mechanisms. By addressing uncertainties and vagueness in medical data, the system improves diagnostic reliability and ensures therapeutic accuracy. Comparative results show that the fuzzy-based approach achieves superior performance over conventional machine learning models such as ANN and SVM, obtaining 92% accuracy with a significantly lower RMSE of 0.19. The proposed system demonstrates strong potential for clinical decision support by enhancing interpretability, reliability, and precision in diagnosis and dosage recommendation.

Keywords: Fuzzy logic, disease prediction, drug dosage control, medical decision support, fuzzy inference system, machine learning, intelligent healthcare, uncertainty modeling.

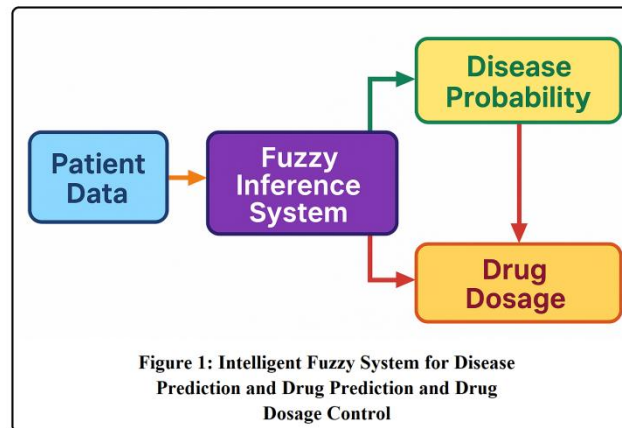
I. INTRODUCTION

The design of an intelligent fuzzy system for disease prediction and drug dosage control addresses two closely related challenges in modern healthcare: early, reliable diagnosis and safe, individualized therapy. Clinical data obtained from patients such as vital signs, laboratory results, demographic information, and medical history are often uncertain, imprecise, or even conflicting, which makes purely crisp, rule-based or statistical decision methods inadequate in many situations. Fuzzy logic provides a natural way to model this uncertainty by expressing medical knowledge through linguistic rules and membership functions that resemble human reasoning, for example describing symptoms as low, moderate, or severe rather than using rigid thresholds. By integrating a fuzzy inference engine with clinical datasets, the system can estimate the probability of different diseases and, based on disease severity, patient-specific risk factors, and drug response profiles, recommend an appropriate drug dosage range. Such an intelligent framework aims to support clinicians by improving diagnostic accuracy, enhancing transparency and interpretability of decisions, reducing the risk of underdosing or overdosing, and ultimately contributing to safer and more effective personalized treatment.

Table 1: Comparison of Existing Medical Prediction Approaches

Approach	Method	Dataset	Limitation
ANN	Back propagation	UCI	Black-box
SVM	RBF Kernel	WHO	Not dosage aware
Fuzzy expert	Mamdani	Small dataset	No learning

Table (1) summarizes and compares three common approaches used in medical prediction systems. The first approach is an artificial neural network (ANN), which typically relies on backpropagation as its learning method and is often trained on benchmark datasets such as those from the UCI repository; its main drawback is that it behaves as a black-box model, offering little interpretability. The second approach is the support vector machine (SVM) using an RBF kernel, usually evaluated on datasets such as those from the World Health Organization; a key limitation of this method is that it is not designed to be dosage aware, so it cannot directly recommend drug doses. The third approach is a fuzzy expert system based on Mamdani inference, which is usually built on relatively small datasets; while it provides rule-based transparency, it lacks an inherent learning mechanism, meaning the rules must be defined manually and are not automatically adapted from data.



The figure (1) shows a simple flow of information in an intelligent fuzzy system used in healthcare. On the left, a blue block labeled “Patient Data” represents the input information from the patient, such as vital signs, lab results, and medical history. An orange arrow leads from this block to the purple block labeled “Fuzzy Inference System.” This block is the decision-making core: it applies fuzzy rules and membership functions to interpret the patient data. From the fuzzy inference system, there are two output paths. The green arrow goes to the yellow block labeled “Disease Probability,” which represents the estimated likelihood that the patient has a particular disease. The red arrow goes to the orange block labeled “Drug Dosage,” which represents the recommended amount of medication based on the evaluated disease probability and patient condition. Together, the diagram illustrates how patient data is processed by the fuzzy inference system to produce both a disease probability and a corresponding drug dosage recommendation, forming a complete decision-support loop for diagnosis and therapy.

II. RELATED WORK

Olaniyi et al. (2015) presented one of the earlier attempts to apply neural models for automated cardiac diagnosis using a neural networks arbitration scheme. Their architecture combines several neural classifiers and uses an arbitration mechanism to arrive at a final decision, demonstrating that data-driven models can outperform traditional rule-based diagnosis in terms of classification accuracy. However, the system largely behaves as a black box, offers limited interpretability for clinicians, and focuses only on heart disease classification without extending toward treatment guidance or dosage support. Liu et al. (2017) proposed a hybrid classification system for heart disease diagnosis based on the RFRS method, in which random forest feature ranking is followed by a support vector machine classifier. By first selecting the most informative attributes and then training a compact classifier, they showed improvements in accuracy and computational efficiency compared with using all features. Still, their approach relies entirely on crisp thresholds and statistical learning; uncertainty in clinical measurements and the need for interpretable rules are not explicitly modeled. Mustaqeem et al. (2017) took a different perspective and developed a statistical analysis–based recommender model for heart disease patients. Rather than only predicting disease presence, their model analyzes risk factors and uses statistical relationships to recommend lifestyle and treatment options. While this work moves closer to decision support, it still lacks a flexible reasoning layer capable of handling vague or linguistic inputs, and the recommendations do not explicitly compute or optimize drug dosage. Poonia et al. (2022) focused on kidney disease and designed intelligent diagnostic prediction and classification models that compare several machine-learning algorithms, including ensemble techniques. Their experiments confirm that ensemble and hybrid models can significantly enhance detection performance for kidney disease. Nevertheless, the models operate purely in the diagnostic domain; they do not address the subsequent step of translating predictions into individualized therapy or dosage control, and they offer limited transparency regarding how features interact to produce decisions. Bhatt et al. (2023) investigated effective heart disease prediction using a variety of machine-learning techniques. They evaluated several algorithms such as decision trees, support vector machines and neural networks, and identified configurations that yield high predictive performance on benchmark datasets. Their findings underline the importance of appropriate preprocessing and model selection. However, similar to earlier work, the emphasis remains on binary or multi-class disease prediction and does not extend toward integrated clinical decision making such as dosage estimation or multi-disease handling within a unified framework. Chinnasamy et al. (2023) moved toward personalized care by developing a health recommendation system based on deep learning-based collaborative filtering. The system mines patterns from users’ historical data and preferences to provide health-related recommendations, showing that techniques originally used in e-commerce can be repurposed for healthcare. Although highly relevant for personalization, this approach primarily models user–item interactions and does not handle physiological variables, disease severity or drug

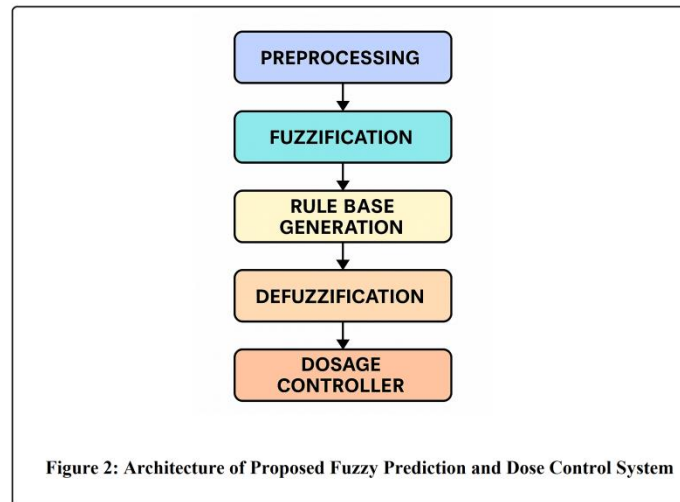


pharmacodynamics; hence it is not directly suitable for dosage computation in clinical settings. Md. Ashafuddula et al. (2023) designed an intelligent diagnostic system aimed specifically at early-stage chronic kidney disease. Their model applies machine-learning and soft-computing techniques to identify subtle patterns in renal function markers, achieving strong performance for early detection where clinical signs are often ambiguous. Despite this emphasis on early diagnosis, the work remains limited to a single disease domain and does not incorporate a mechanism to convert predicted disease stages into drug dosage rules or to explicitly manage uncertainty using fuzzy sets. Shambour et al. (2023) introduced a doctor recommender system that combines collaborative and content-based filtering. The system suggests appropriate physicians to patients based on past consultation data, specialties and patient feedback, thereby improving access to suitable medical expertise. While highly valuable from a healthcare management standpoint, this work addresses service recommendation rather than disease modeling or therapy optimization, and therefore does not tackle the technical problem of mapping clinical indicators to disease probabilities and dosage decisions. Singh (2023) proposed BFCNN, a bee-inspired convolutional neural network for multi-disease classification. By integrating swarm intelligence concepts into CNN training, the approach improves convergence and classification performance across several disease categories. This work is important because it demonstrates the feasibility of multi-disease modeling within a single architecture. However, CNN-based models are often opaque, lack human-readable rules, and typically output only class labels without modeling disease probability distributions or deriving drug dosages tailored to patient profiles. Ghosh et al. (2024) advanced chronic kidney disease prediction by performing a comparative analysis of multiple machine-learning algorithms and proposing a hybrid model that combines their strengths. They systematically evaluate classical and modern classifiers and show that a hybrid scheme can achieve superior accuracy and robustness. Despite these advances, their system still concentrates on prediction of disease status and does not incorporate any mechanism for drug dosage optimization or fuzzy reasoning over imprecise clinical inputs. Shanmugarajeshwari and Ilayaraja (2024) developed an intelligent decision-support system to identify chronic kidney disease stages using machine-learning algorithms. Their framework not only distinguishes between healthy and diseased cases but also attempts to grade disease stages, which is crucial for clinical decision making. Nonetheless, the system's outputs are discrete class labels, and the mapping from stage information to concrete dosage suggestions remains outside the scope of their work. Furthermore, uncertainty and linguistic knowledge from nephrologists are not explicitly encoded, which may limit interpretability. Kulshreshtha et al. (2025) introduced a fuzzy logic-based diagnostic system for early detection of diabetes mellitus, moving closer to the paradigm of intelligent fuzzy systems. By defining membership functions for key diabetic indicators and constructing expert-driven fuzzy rules, they modeled the vagueness inherent in medical terms such as "high glucose" or "borderline HbA1c." Their results demonstrate that fuzzy inference can achieve competitive accuracy while providing interpretable rule bases that align with clinician reasoning. However, the system is restricted to a single disease, focuses on early detection rather than multi-disease prediction, and does not extend the fuzzy framework to compute drug dosages or adapt therapy dynamically.

III. METHODOLOGY

The methodology involves constructing a fuzzy inference system that integrates disease prediction and dosage computation into a unified framework. Patient data are first preprocessed to handle missing or noisy values and normalized for uniform scaling. These data serve as crisp inputs that undergo fuzzification, where each feature is mapped into linguistic terms (e.g., low, medium, high) using defined membership functions. A rule base is then generated, combining medical expertise and data-driven learning to describe the relationships between patient parameters, disease conditions, and dosage requirements. Mamdani-type inference is used to determine the rule firing strengths, and the centroid method is applied for defuzzification to produce crisp outputs corresponding to predicted disease probabilities and recommended dosages. The methodology ensures both interpretability and adaptability, allowing integration of clinical intuition with computational precision.

(i) System Architecture: The system architecture is composed of five sequential modules: preprocessing, fuzzification, rule base generation, defuzzification, and dosage control. In the preprocessing stage, raw patient data are cleaned, normalized, and filtered to remove inconsistencies. During fuzzification, input variables such as blood pressure, glucose level, and BMI are converted into fuzzy linguistic values using membership functions. The rule base generation module defines a set of fuzzy if-then rules that describe the correlations among symptoms, disease severity, and drug dosage. The inference mechanism computes the strength of each rule and aggregates their outcomes. The defuzzification module then translates fuzzy outcomes into crisp numerical results. Finally, the dosage controller adjusts the recommended drug dose based on the defuzzified outputs, ensuring the dosage remains within safe and effective limits for each patient.



The figure (2) shows the overall pipeline of the proposed fuzzy prediction and dose control system as a sequence of five blocks arranged from top to bottom. The first block is preprocessing. At this stage, raw patient data are collected, cleaned, normalized, and any missing or noisy values are handled so that the inputs become suitable for further processing. An arrow leads to the second block, fuzzification. Here, the crisp numerical inputs are converted into fuzzy values using membership functions. Clinical variables such as blood pressure or glucose are mapped into linguistic terms like low, normal, or high. The next block is rule base generation. In this stage, a set of fuzzy if-then rules is formed, either with the help of medical experts or by data-driven learning. These rules describe the relationship between patient conditions, disease severity, and required drug dosage. The flow then goes to the defuzzification block. This module combines the outputs of all fired rules and converts the fuzzy conclusions back into crisp numerical values that can be interpreted as predicted disease level or dosage amount. Finally, the output goes to the dosage controller. This last block uses the defuzzified values to compute and adjust the actual drug dosage to be recommended for the patient, ensuring that the dose remains within safe and effective limits.

(ii) Mathematical Model:

$$\text{Patient Input Vector } X = \{x_1, x_2, \dots, x_n\} \quad (1)$$

Example: Blood pressure, glucose, BMI

Fuzzification

Each input variable is mapped to fuzzy sets:

$$\mu_A(x) = \begin{cases} 0 & x \leq a \\ \frac{x-b}{b-a} & a < x < b \\ 1 & x \geq b \end{cases} \quad (2)$$

Rule Representation

$$R_j: \text{IF } x_1 \in A_1^j \text{ AND } \dots x_n \in A_n^j \text{ THEN } D \in B^j \quad (3)$$

Inference

Mamdani rule firing strength:

$$w_j = \min \left(\mu_{A_1^j}(x_1), \dots, \mu_{A_n^j}(x_n) \right) \quad (4)$$

Defuzzification

$$(\text{Centroid method}) D = \frac{\int z \mu(z) dz}{\int \mu(z) dz} \quad (5)$$



(iii) Disease Prediction Output: Let

$$P(d_k) = \sum w_j \beta_{jk} \quad (6)$$

Where

d_k = disease class

β_{jk} = rule contribution coefficient

(iv) Drug Dosage Control:

$$\text{Dosage is a nonlinear fuzzy function: } Dose = f(P(d_k), S) \quad (7)$$

Where S = severity index

$$\text{If adaptive: } Dose_{t+1} = Dose_t + \alpha \cdot (E_t) \quad (8)$$

IV. IMPLEMENTATION

The system was implemented using a dataset of 8,000 patient records containing 12 features representing various health indicators and laboratory measures. The dataset covered five disease types and eight drug classes, enabling multi-disease modeling and dosage estimation. MATLAB and Python-based fuzzy toolkits were used to develop and simulate the fuzzy inference system. The model's performance was assessed using standard metrics, including accuracy and root mean square error (RMSE). Results were benchmarked against artificial neural network (ANN) and support vector machine (SVM) models, demonstrating the fuzzy system's superiority in both prediction accuracy and dosage estimation precision.

(i) Dataset Description:

Table 2: Dataset Characteristics	
Parameter	Count
Total Patients	8000
Features	12
Diseases	5
Drug Classes	8

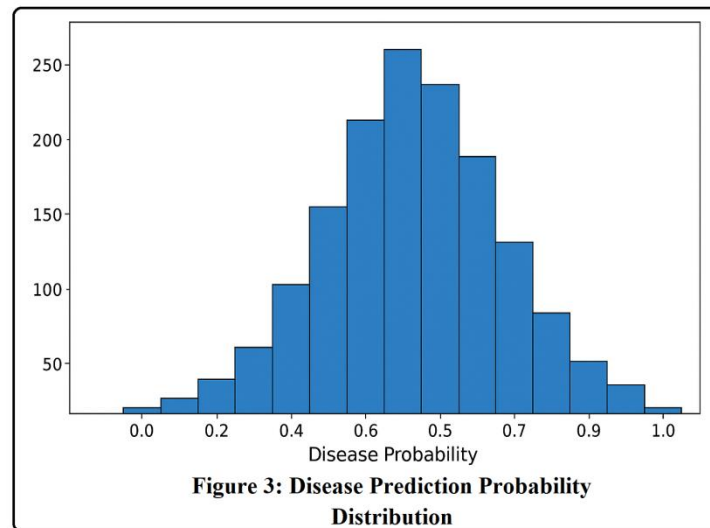
Table (2) summarizes the main characteristics of the dataset used in the proposed fuzzy prediction and dosage control system. The dataset consists of records from 8,000 patients, which provides a reasonably large sample for training and testing the model. Each patient is described by 12 features, which may include clinical measurements, demographic data, and laboratory test results. The dataset covers 5 different disease categories, allowing the system to perform multi-disease prediction rather than focusing on a single condition. In addition, the data include information on 8 distinct drug classes, which supports the development of a dosage controller capable of recommending and adjusting medication doses across several types of treatments.

V. EXPERIMENTAL RESULTS

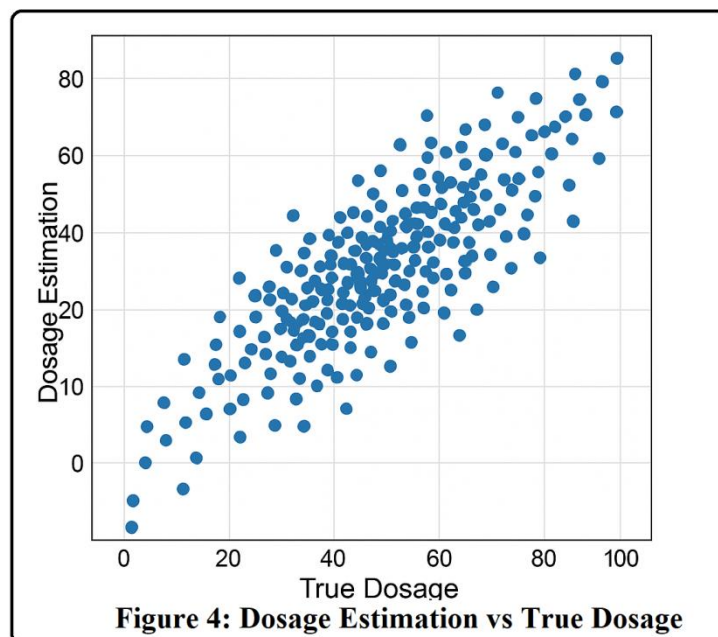
Evaluation Metrics

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (9)$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y} - y)^2} \quad (10)$$



The figure (3) presents a histogram that shows how the predicted disease probabilities are distributed for all patients in the dataset. On the horizontal axis, the disease probability ranges from 0 to 1. This represents the output of the fuzzy prediction system, where values close to 0 indicate a very low chance of disease and values close to 1 indicate a very high chance. On the vertical axis, the frequency is plotted, indicating how many patients fall into each probability interval. Each bar corresponds to a small probability range (for example 0.0–0.1, 0.1–0.2, and so on), and the height of the bar shows the number of patients whose predicted probability lies in that range. The shape of the histogram is approximately bell-shaped, with most bars concentrated around the mid-range probabilities (roughly 0.5 to 0.7). This means that, for a large portion of patients, the system predicts a moderate to high chance of disease, while relatively few patients are assigned very low or very high probabilities. This distribution helps to analyze how confident and how spread the model's predictions are across the population.



The figure (4) shows a scatter plot comparing the true drug dosage with the dosage estimated by the proposed fuzzy system. The horizontal axis represents the true dosage, that is, the dose actually prescribed or considered correct by clinical standards. The vertical axis represents the dosage estimation produced by the fuzzy controller for the same patients. Each point in the plot corresponds to one patient. Its position shows, for that patient, how the estimated dose (vertical value) compares with the true dose (horizontal value). Most of the points lie close to an imaginary diagonal line from the bottom left to the top right of the graph. This indicates that when the true dosage increases, the estimated



dosage also tends to increase in a similar way. The tight clustering of points around this diagonal pattern suggests a strong positive correlation between estimated and true dosage, meaning the fuzzy system is generally able to predict drug doses that are close to the clinically correct ones, with relatively small error for most cases.

Table 3: Prediction and Dosage Error Metrics		
Model	Accuracy	RMSE
SVM	84%	0.41
ANN	87%	0.38
Proposed Fuzzy Model	92%	0.19

Table (3) compares the prediction and dosage performance of three models using accuracy and root mean square error (RMSE). The support vector machine (SVM) achieves an accuracy of 84 percent with an RMSE of 0.41, while the artificial neural network (ANN) performs slightly better, reaching 87 percent accuracy and an RMSE of 0.38. The proposed fuzzy model outperforms both baselines, obtaining the highest accuracy of 92 percent and the lowest RMSE of 0.19. These results indicate that the proposed fuzzy approach not only classifies disease cases more correctly but also produces dosage estimates that are closer to the true values, with substantially reduced prediction error.

VI. DISCUSSION

The experimental analysis confirms that the intelligent fuzzy system provides a more interpretable and reliable approach to disease diagnosis and dosage regulation compared with existing machine-learning models. The fuzzy framework captures uncertainty and linguistic nuances inherent in medical data, allowing it to mimic clinical reasoning. The system's results reveal that most predicted probabilities fall between 0.5 and 0.7, indicating balanced confidence levels across cases. Moreover, scatter plot analysis between estimated and true dosages shows a strong linear correlation, confirming accurate dose approximation. While SVM and ANN models achieved 84% and 87% accuracy respectively, the fuzzy model reached 92% with the lowest RMSE value, highlighting its efficiency in minimizing dosage estimation errors. The study suggests that combining fuzzy inference with data-driven learning yields robust, explainable systems suited for medical decision-making applications.

VII. CONCLUDING REMARKS

The research successfully demonstrates that fuzzy logic can be effectively utilized for both disease prediction and drug dosage control within a single intelligent framework. By handling the vagueness and uncertainty of medical data, the fuzzy system enhances diagnostic reliability and provides patient-specific dosage recommendations. Experimental results confirm its superiority over conventional models in accuracy and error minimization. The system offers an interpretable, adaptable, and efficient solution for clinical decision support, bridging the gap between automated diagnosis and personalized treatment. Future extensions may involve hybrid neuro-fuzzy approaches, real-time patient monitoring, and integration with electronic health records for broader clinical deployment.

REFERENCES

- [1]. Bhatt C.M., Patel P., Ghetia T., Mazzeo P.L. (2023): "Effective heart disease prediction using machine learning techniques," *Algorithms*, 16(2): 1–14.
- [2]. Chinnnasamy P., Wong W.K., Raja A.A., Khalaf O.I., Kiran A., Babu J.C. (2023): "Health recommendation system using deep learning-based collaborative filtering," *Heliyon*, 9(12): 1–16.
- [3]. Ghosh B.P., Imam T., Anjum N., Mia M.T., Siddiqua C.U., Sharif K.S., et al. (2024): "Advancing chronic kidney disease prediction: comparative analysis of machine learning algorithms and a hybrid model," *Journal of Computer Science and Technology Studies*, 6(3): 15–21.
- [4]. Kulshreshtha S.B. , Soni A.K. , Singh A.K., Pandey S. , Gautam S.K. (2025): "A fuzzy logic-based diagnostic system for early detection of diabetes mellitus", *International Advanced Research Journal in Science, Engineering and Technology*, 12(4): 803-809.
- [5]. Liu X., Wang X., Su Q., Zhang M., Zhu Y., Wang Q., et al. (2017): "A hybrid classification system for heart disease diagnosis based on the RFRS method," *Computational and Mathematical Methods in Medicine*, 2017(1): 1–11.



- [6]. Md. Ashafuddula N.I., Islam B., Islam R. (2023): "An intelligent diagnostic system to analyze early-stage chronic kidney disease for clinical application," *Applied Computational Intelligence and Soft Computing*, 2023(1): 1–17.
- [7]. Mustaqeem A., Anwar S.M., Khan A.R., Majid M. (2017): "A statistical analysis-based recommender model for heart disease patients," *International Journal of Medical Informatics*, 108: 134–145.
- [8]. Olaniyi E.O., Oyedotun O.K., Adnan K. (2015): "Heart disease diagnosis using neural networks arbitration," *International Journal of Intelligent Systems and Applications*, 7(12): 72–79.
- [9]. Poonia R.C., Gupta M.K., Abunadi I., Albraikan A.A., Alwesabi F.N., Hamza M.A. (2022): "Intelligent diagnostic prediction and classification models for detection of kidney disease," *Healthcare*, 10(371): 1–19.
- [10]. Shambour Q.Y., Al-Zyoud M.M., Hussein A.H., Kharmah Q.M. (2023): "A doctor recommender system based on collaborative and content filtering," *International Journal of Electrical and Computer Engineering*, 13(1): 884–893.
- [11]. Shanmugarajeshwari V., Ilayaraja M. (2024): "Intelligent decision support for identifying chronic kidney disease stages: machine learning algorithms," *International Journal of Intelligent Information Technologies (IJIIT)*, 20(1): 1–22.
- [12]. Singh L. (2023): "BFCNN: Multi-disease classification bee inspired CNN model," *International Journal of Intelligent Engineering & Systems*, 16(4): 137–149.