Research Article

Hybrid Fuzzy Feature Augmentation and XG Boost Learning for Chronic Kidney Disease Detection

Ashish. I. Anerao[1] Omprakash. S. Jadhav[2]

[1]Research Scholar, Department of Statistics, Dr Babasaheb Ambedkar Marathwada University, Chhatrapati Sambhajinagar, 431004. statistics.ain@bamu.ac.in
[2]Associate Professor, Department of Statistics, Dr Babasaheb Ambedkar Marathwada University, Chhatrapati Sambhajinagar, 431004. osj.statistics@bamu.ac.in

Abstract

Chronic kidney disease (CKD) is a progressive health disorder and timely diagnosis is crucial to prevent serious complications. Machine learning techniques are extensively used for disease prediction. however, traditional models face limitations such as uncertainty in clinical data and class imbalance. To overcome these challenges, this study develops a Hybrid Fuzzy–XGBoost classification model for CKD prediction. In the proposed method numerical clinical features are transformed into fuzzy linguistic levels like "low", "medium" and "high" using triangular membership functions which better represent the uncertainty in clinical measurements. Fuzzy sample weighting method is employed to improve the accuracy of CKD case identification. The expanded dataset is then used to train the XG Boost classification algorithm. The model's performance was assessed using Accuracy, Precision, Recall, F1-Score and ROC–AUC. Experimental results demonstrated that the proposed Hybrid Fuzzy–XG Boost model outperformed traditional classification techniques, achieving 98% Accuracy, 100% Recall and 0.984 AUC. These findings underscore the high predictive capability of this model for CKD diagnosis.

Keywords: *Chronic Kidney Disease, Fuzzy Logic, Ensemble Learning, Hybrid Machine Learning, XG Boost.*

1. Introduction

Chronic kidney disease (CKD) is a progressive disease in which kidneys gradually become unhealthy. The early features of this disease are not readily apparent and may only be diagnosed once there is extensive kidney damage. Due to its association with diabetes, hypertension and other factors including older age, CKD has emerged as a major health problem worldwide. It is important to diagnose early as early treatment will help slow the disease and minimize the possibility of end stage kidney failure. The early signs of CKD are not very distinct and are frequently mistaken for normal physiological changes, making early diagnosis of CKD difficult in standard clinical practice [1] [2]. The ideal CKD prediction system would accurately predict high-risk patients with clinical data routinely collected in the clinic. The system should deal with the uncertainty and incompleteness of health data. Medical datasets in fact have categorical features, variables with repetitions and missing data, making suitable preprocessing steps vital to the process of training the model. The classification of data into training set and testing set is a crucial process in the pre-processing stage and also feature selection and encoding of categorical variables are extremely important in the performance of predictive models for diagnosis of CKD [3].

However, recently machine learning methods have been gaining popularity in predicting CKD since they enable modelling complex, non-linear relationships between clinical features. Some of the promising classification algorithms include support vector machines, random forests and various boosting algorithms. Ensemble boosting methods such as XG Boost have been shown to get better classification accuracy on the CKD datasets than conventional methods in several studies [4] [5] . The traditional machine learning models make an assumption that clinical variables have accurate numerical values. In real clinical practice, there is uncertainty for many measurements and the transition from normal to abnormal is also gradual. For example, laboratory parameters may be in the 'borderline' range, such as blood pressure, serum creatinine or haemoglobin, where they are not easy to classify numerically. In these circumstances, fuzzy logic is a useful mathematical tool that is able to express uncertainty. In this case, values are assigned partial membership in linguistic categories: "low", "medium" and "high" is a simulation of human reasoning in medical decision making [6]. Fuzzy-based systems are not powerful enough to effectively learn high-dimensional patterns in real world health data. The rule-based structure might restrict recognition of complex interaction among clinical variables. So, researchers have resorted to hybridization of fuzzy logic with machine learning techniques to achieve both interpretability and predictive accuracy. Recent research indicates that the use of intelligent algorithms coupled with uncertainty modelling can greatly improve the skill of predicting a disease [7] [8]. Hybrid machine learning models are gradually getting more attention in the field of medical data analysis, but their application in the prediction of CKD is still limited. Most existing research employs old-fashioned machine learning workflows and doesn't explicitly use fuzzy membership representations within feature development or training. The application of fuzzy based weighting method to most of the ensemble algorithms including XG Boost has not been studied in detail. This site-to-site gap emphasizes the need for new hybrid frameworks incorporating uncertainty modelling as a part of the learning process. Given these problems, the current study introduces a Hybrid Fuzzy–XGBoost model to improve the prediction accuracy of chronic kidney disease (CKD) based on clinical data. The data is then pre-processed, and stratified sampling is applied to split the data into training and testing subsets. For numerical values, further fuzzy features are created using fuzzy membership functions like “low”, “medium”, “high” along with the quartile limits based on triangular functions. A Fuzzy Sample Weighting technique is used to focus more on the disease-class related patterns, to solve class imbalance problem. Then the classification model is trained using the extended data set (XG Boost). Metrics used to evaluate the performance of the model include those such as accuracy, precision, recall, F1-score, and ROC–AUC. The results show that there is a significant enhancement in Recall, and no change in prediction accuracy.

The objective of this research is:

- To develop a Hybrid Fuzzy XG Boost model for predicting chronic kidney disease.
- To incorporate fuzzy membership-based feature enhancement to capture uncertainty in clinical attributes.
- To improve the model's sensitivity and overall predictive performance through fuzzy weighting and gradient boosting.

The proposed framework aims to improve predictive accuracy without sacrificing interpretability in CKD diagnosis. These hybrid approaches can aid clinical decision-making and help identify high-risk patients early.

2. Review of Literature

Table 1 provides a comparative review of recent research on predicting chronic kidney disease (CKD) using fuzzy logic and machine learning methods. Most existing studies use fuzzy systems or machine-learning classification techniques independently. There is relatively little research that combines fuzzy logic with powerful ensemble algorithms such as XG Boost. The proposed Hybrid Fuzzy–XGBoost model aims to address this research gap. In this model, fuzzy membership-based feature expansion and fuzzy sample-weighting techniques are integrated with the XG Boost classification algorithm. This combined approach improves sensitivity and overall predictive performance in CKD prediction.

Table 1: Comparative Review of Recent Research

Author	Year	Method Used	Key Idea	Limitation	Contribution
Murugesan et al. [9]	2022	ANFIS (Adaptive Neuro-Fuzzy Inference System)	Uses fuzzy rules for CKD diagnosis	Limited scalability for large datasets	Demonstrated usefulness of fuzzy inference in CKD prediction
Khalid et al. [10]	2023	Hybrid Machine Learning Models	Compared classifiers such as Decision Tree and Gradient Boosting	Limited interpretability and uncertainty handling	Showed that hybrid ML improves CKD prediction accuracy
Rehman et al. [11]	2023	Machine Learning Classification	Applied classification algorithms for CKD detection	Crisp decision boundaries may ignore uncertainty	Highlighted the effectiveness of ML models for CKD detection
Praveen et al. [12]	2024	Neuro-Fuzzy Model	Combines neural networks with fuzzy inference	Computational complexity	Improved classification performance through fuzzy reasoning
Patil et al. [13]	2024	Hybrid Classification Framework	Integrates preprocessing and classification stages	Limited focus on class imbalance	Improved feature representation in CKD prediction
Yan et al. [14]	2024	Fuzzy Clustering + Machine Learning	Uses fuzzy clustering to improve classification	High computational cost	Enhanced detection of uncertain medical patterns
Haider et al. [15]	2025	Ensemble Machine Learning	Random Forest and other ensemble classifiers	Limited interpretability	Improved predictive accuracy

Sharma et al. [16]	2025	Fuzzy and Neuro-Fuzzy Systems	Uses fuzzy systems for CKD staging	Limited integration with advanced ensemble learning	Demonstrates the flexibility of fuzzy logic in medical diagnosis
Liu et al. [17]	2025	Hybrid Deep Learning Model	Combines deep learning with prediction models	Requires large datasets	Improved robustness for CKD prediction
Proposed Work	**2025**	**Hybrid Fuzzy–XG Boost Model**	Combines fuzzy membership-based feature augmentation with XG Boost classifier	—	Improves recall and handles uncertainty and class imbalance in CKD prediction

3. Research Methodology

3.1 Data Description

This work uses the Chronic Kidney Disease (CKD) dataset from UCI Machine Learning Repository. A widely-used and known machine learning resource. The data was initially collected from medical records to help with the early diagnosis of chronic kidney disease. UCI CKD is a dataset of patient records with 400 patients, and 24 clinical features associated with kidney function are recorded. This set of data contains both numerical and categorical data. It contains details of physical and lab measurements age, blood pressure, serum creatinine, haemoglobin, blood glucose level. The classification variable in the data set is the classification of the patient, whether or not they have CKD. This is the variable to be predicted by the classification model. The 24 features, some are numerical like age, blood pressure, serum creatinine, blood urea, haemoglobin while some are categorical such as red blood cell status, hypertension, diabetes mellitus, appetite. The presence of both numeric and categorical data makes the dataset more useful for assessing the performance of hybrid machine learning models capable of processing various types of data. [18]

3.2 Data Preprocessing

Data preprocessing is an essential part of the machine learning process that enhances data quality and ensures it can be used effectively in predictive models. At first, features not needed or redundant features were removed from the data set. The dataset had both numeric and categorical features, therefore the categorical features were encoded; this is known as label encoding. This process can effectively communicate the categorical information to the machine learning algorithms. The data set was divided into training set and test set by stratified sampling after pre-processing. This approach ensures equal number of cases in both the CKD and non-CKD sets. This is often applied in medical classification problems when we need to keep the proper class distribution and want to avoid biasing the evaluation of the model.

3.3 Fuzzy Feature Generation

Medical data is uncertain and slowly changing in between normal and abnormal state. Such uncertainty cannot be accounted for by the conventional machine learning approaches that assume these variables are numerical values set in advance. Hence, fuzzy logic is used to represent numerical features with membership instead of a hard constraint. [19] In this study, triangular membership functions were employed to transform numerical clinical attributes into three linguistic fuzzy sets: Low, Medium and High. The triangular membership function is defined as:

$$\mu(x) = \begin{cases} 0 & x \leq a \\ \dfrac{x-a}{b-a} & a < x \leq b \\ \dfrac{c-x}{c-b} & b < x \leq c \\ 0 & x \geq c \end{cases}$$

Here, a, b and c represent the lower, peak, and upper values of the fuzzy set respectively. In this study, quartile-based thresholds were used to determine the magnitude for each numerical variable. The membership values obtained for the categories “low”, “medium” and “high” served as additional fuzzy features. These fuzzy features were integrated with the original dataset and used as input for the Hybrid Fuzzy–XG Boost classification model. [20]

3.4 Fuzzy Sample Weighting

Medical datasets often exhibit class imbalance, with more healthy patients than diseased ones. This leads the classification model to favour the majority class, increasing the risk of missing CKD cases. To mitigate this a fuzzy sample weighting technique is used during training. It gives higher weights to samples from the disease class, boosting sensitivity and reducing false negatives.

The weight of each instance in the sample can be expressed as:

$$w_i = \begin{cases} \alpha & y_i = 1 \\ 1 & y_i = 0 \end{cases}$$

where y_i represents the class label and $\alpha > 1$ is the weighting factor assigned to CKD samples. This approach enhances the model's capacity to accurately recognise minority class instances in imbalanced datasets [21] [22]

3.5 Hybrid Fuzzy–XG Boost Model

This study presents a Hybrid Fuzzy–XG Boost model designed to predict chronic kidney disease (CKD). It combines fuzzy-based feature representation with the XG Boost classification algorithm to address uncertainty in clinical features, resulting in reliable and consistent predictions. XG Boost is an advanced gradient boosting method that builds multiple decision trees sequentially and merges them into an ensemble, thereby minimizing prediction errors and improving classification performance [23]. Recent studies demonstrate the effectiveness of gradient boosting and hybrid machine learning methods in medical diagnosis tasks [24].

Initially, numerical variables are transformed using triangular fuzzy membership functions, which denote linguistic categories such as "low", "medium" and "high". The fuzzy membership values obtained from this process are integrated as additional features, more effectively capturing gradual and subtle changes in clinical measurements [19] [20].

The objective function of XG Boost is specified as:

$$L(\theta) = \sum_{i=1}^{n} l(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k)$$

where $l(y_i, \hat{y}_i)$ represents the loss function and $\Omega(f_k)$ denotes the regularization term controlling model complexity.

The ensemble model's prediction can be expressed as:

$$\hat{y}_i = \sum_{k=1}^{K} f_k(x_i)$$

where $f_k$represents the k^{th}decision tree in the ensemble.

The Hybrid Fuzzy–XG Boost model integrates fuzzy logic-based feature representation with boosting-based decision-tree learning to better detect nonlinear patterns and uncertainties in medical data. This integration improves the accuracy and overall efficiency of CKD prediction.

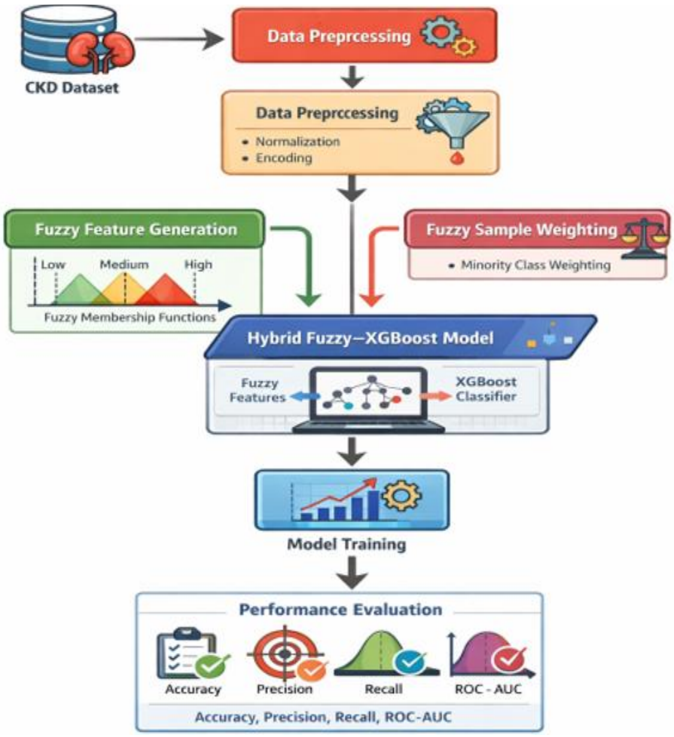

Figure 1: Flowchart of Proposed Model

3.6 Evaluation Matrice

The performance of the proposed Hybrid Fuzzy–XG Boost model was evaluated using multiple standard classification metrics, including Accuracy, Precision, Recall, F1-Score and ROC–AUC. These metrics provide a thorough assessment of the model's effectiveness in accurately distinguishing CKD from non-CKD cases. They are commonly employed in machine learning research and medical diagnostic studies [25].

Accuracy

Accuracy assesses the overall correctness of a classification model by determining the ratio of correctly predicted instances to the total number of samples [26]. It is calculated as follows:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

where TP represents true positives, TN represents true negatives, FP represents false positives, and FN represents false negatives.

Precision

Precision quantifies the ratio of correctly identified positive cases to all cases predicted as positive. positive cases. It reflects how reliable positive predictions are and is expressed as:

$$Precision = \frac{TP}{TP + FP}$$

Recall (Sensitivity)

Recall measures how well the model correctly detects true positive cases, which is especially crucial in medical diagnosis tasks, since missing disease cases can have serious consequences [26].

$$Recall = \frac{TP}{TP + FN}$$

F1-Score

The F1-Score represents the harmonic mean of precision and recall, providing a balanced metric when both false positives and false negatives are important.

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

This metric is especially valuable for assessing classification models on imbalanced datasets because it treats precision and recall equally [27].

ROC Curve and AUC

The Receiver Operating Characteristic (ROC) curve illustrates the trade-off between the True Positive Rate (TPR) and False Positive Rate (FPR) across different classification thresholds.

$$TPR = \frac{TP}{TP + FN}$$

$$FPR = \frac{FP}{FP + TN}$$

The Area Under the ROC Curve (AUC) measures the overall ability of the classifier to differentiate between classes; a higher AUC value indicates improved classification performance across all thresholds [28]. These evaluation metrics collectively provide a comprehensive assessment of the predictive capability of the proposed Hybrid Fuzzy–XGBoost model.

4 Result and Discussion

This section presents the experimental results obtained with the proposed Hybrid Fuzzy–XGBoost model for predicting chronic kidney disease (CKD). The performance of this method was compared with popular machine learning classification models, including Logistic Regression (LR), Support Vector Machine (SVM), Random Forest (RF) and traditional XGBoost. Ensure a fair comparison all models were trained using the same dataset and preprocessing steps. The models' performance was assessed using widely accepted classification metrics such as Accuracy, Precision, Recall and F1-Score.

4.1 Comparative Performance Analysis

Table No.2 presents a comparison of performance results between the baseline models and the newly proposed Hybrid Fuzzy–XGBoost model.

Table 2: Comparison of Performance

Model	Accuracy	Precision	Recall	F1-Score
Logistic Regression	91%	83.72%	94.73%	88.88%
Support Vector Machine	90%	81.81%	94.73%	87.80%
Random Forest	97%	97.22%	92.10%	94.59%
XG- Boost	97%	97.22%	94.73%	95.96%
Hybrid Fuzzy–XGBoost	98%	94.73%	100%	97.29%

The results indicate that the proposed Hybrid Fuzzy–XGBoost model outperformed all other tested classification models in overall performance. It achieved 98% accuracy, exceeding traditional machine learning models. Logistic Regression and Support Vector Machine attained 91% and 90% accuracy respectively. While these models showed relatively high recall, they experienced more false positives due to lower precision. The Random Forest and standard XGBoost models performed slightly better due to their ensemble learning capabilities, both reaching 97% accuracy and demonstrating strong classification performance. Nevertheless, their recall was marginally lower than that of the hybrid model. The Hybrid Fuzzy–XGBoost model achieved 100% recall, accurately identifying all CKD cases without false negatives, which is crucial in medical diagnosis, where missing a diseased case can lead to severe health issues. It achieved an F1 Score of 97.29%, indicating a well-balanced performance between precision and recall.

4.2 ROC Curve Analysis

To evaluate the proposed model's classification performance a Receiver Operating Characteristic (ROC) curve analysis was performed. The ROC curve shows how the True Positive Rate (TPR) relates to the False Positive Rate (FPR) across various classification thresholds.

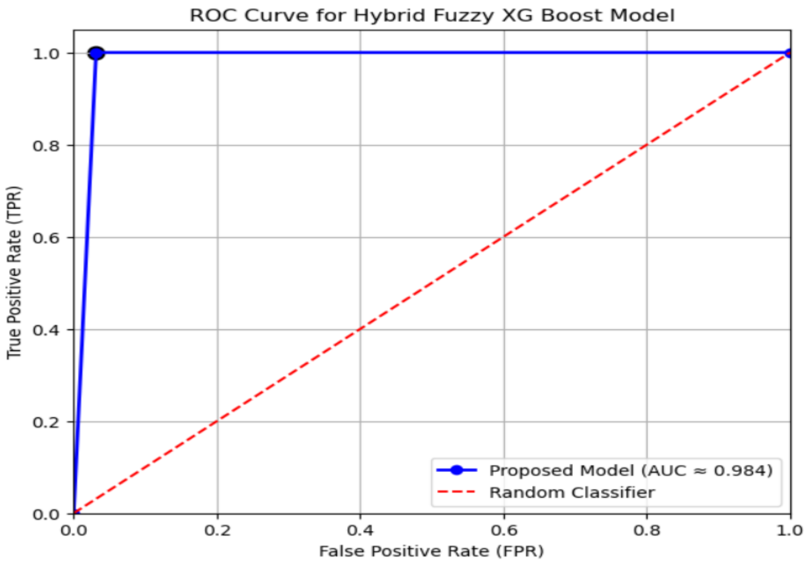

Figure 2: ROC Curve for Hybrid Fuzzy–XGBoost Model

The proposed model achieved an Area Under the Curve (AUC) of about 0.984, reflecting excellent classification performance. An AUC near 1 indicates the model's strong ability to distinguish CKD from non-CKD cases. Additionally, the ROC curve is well above the diagonal, indicating that the classifier performs better than random. This confirms that the Hybrid Fuzzy–XGBoost model offers reliable and effective classification results.

4.3 Discussion

The main advantage of the proposed model is that it has been designed in a way that it has created an improvement in the performance of the system primarily by integrating the fuzzy feature generation with the ensemble learning technique. It translates numerical clinical data to linguistic terms like "low," "medium" and "high" using fuzzy membership functions which is an effective way of modelling the uncertainty in medical data. Class imbalance is handled by fuzzy sample weighting by giving more weight to CKD samples for training, which will help to enhance the accuracy of detecting diseases and minimize the number of false negatives. The Hybrid Fuzzy XG Boost model is proposed to better model the nonlinear relationships between clinical features and handle the measurement uncertainty by integrating fuzzy logic with the XG Boost boosting framework. It therefore results in significantly better predictive results than the classical machine learning classification methods. Overall, the experimental results indicate that the Hybrid Fuzzy–XG Boost framework is an accurate, reliable and efficient approach for the prediction of the patients with CKD, which has high precision, excellent recall and high classification performance.

5. Conclusion

This study presents a novel Hybrid Fuzzy XG Boost classification model for chronic kidney disease (CKD) prediction using clinical data. It combines fuzzy membership-based feature creation, sample weight techniques and the XG Boost ensemble algorithm, enhancing the way uncertainty and class imbalance are addressed in medical datasets. Triangular membership functions, gradual and subtle changes in medical measurements are effectively captured and numerical clinical variables are translated into linguistic categories: "low", "medium" and "high". The proposed Hybrid Fuzzy XG Boost model is better than the popular machine learning classification models such as Logistic Regression, Support Vector Machine, Random Forest and the traditional XG Boost model. The model's predictive ability was good, with it attaining 98% accuracy, 94.73% precision, 100% recall and 97.29% F1-score. The ROC analysis showed an AUC of 0.984, highlighting the model's good capability of differentiating between CKD and non-CKD cases. This finding indicates that the use of fuzzy logic combined with ensemble learning can be improved to represent the features and to improve the percentage of case identification for CKD.

In summary, the proposed Hybrid Fuzzy-XG Boost approach offers a reliable and effective approach to predict the CKD. It may aid early diagnosis by health care providers, and in clinical decision making.

Conflict of Interest

The authors explicitly state that they have no conflict of interest concerning the publication. of this research article.

Data Availability Statement

The dataset utilised in this study is publicly available via the UCI Machine Learning Repository. The Chronic Kidney Disease (CKD) dataset can be retrieved from the following link: https://archive.ics.uci.edu/ml/datasets/chronic_kidney_disease.

References

[1] K. K. Zadeh., H. Jafar, D. Nitsch, L. B. Neuen and V. Perkovic, "Chronic kidney disease," *The Lancet,* vol. 398, no. 10302, pp. 786-802, 2021.

[2] Q. Pang and M. Tong, "Artificial intelligence in predicting chronic kidney disease prognosis. A systematic review and meta-analysis," *Renal Failure,* vol. 46, no. 2, 2024.

[3] R. K. Halder and et.al, "ML-CKDP: Machine learning-based chronic kidney disease prediction with smart web application," *Journal of Pathology Informatics,* vol. 15, no. 100371, 2024.

[4] S. M. Ganie, P. K. D. Pramanik, S. Mallik and Z. Zhao, "Chronic kidney disease prediction using boosting techniques based on clinical parameters," *PLOS One,* vol. 18, no. 12, 2023.

[5] M. J. Raihan, M. A.-M. Khan, S. H. Kee and A.-A. Nahid, "Detection of the chronic kidney disease using XGBoost classifier and explaining the influence of the attributes on the model using SHAP," *Scientific Reports,* 2023.

[6] L. Pickering, K. Cohen and B. Baets, "A Narrative Review on the Interpretability of Fuzzy Rule-Based Models from a Modern Interpretable Machine Learning Perspective," *International Journal of Fuzzy Systems,* 2025.

[7] H. Iftikhar, A. Hashem, M. Qureshi and P. Rodrigues, " Clinical Application of Machine Learning Models for Early-Stage Chronic Kidney Disease Detection," *Diagnostics (Basel),* vol. 15, no. 20, 2025.

[8] M. S. Arif, A. Rehman and D. Asif, "Explainable Machine Learning Model for Chronic Kidney Disease Prediction," *Algorithms.,* vol. 17, no. 10, 2024.

[9] G. Murugesan, T. Ahmed, J. Bhola, M. Shabaz, J. Singla, M. Rakhra, S. More and I. A. Samori, "Fuzzy Logic-Based Systems for the Diagnosis of Chronic Kidney Disease," *Biomed Res Int.,* 2022.

[10] H. Khalid, A. Khan, M. Z. Khan, G. Mehmood and M. S. Qureshi, "Machine Learning Hybrid Model for the Prediction of Chronic Kidney Disease," *Computational Intelligence and Neuroscience,* 2023.

[11] A. REHMAN, T. SABA, H. ALI, N. ELHAKIM and N. AYESHA, "Hybrid machine learning model to predict chronic kidney diseases using handcrafted features for early health rehabilitation.," *Turkish Journal of Electrical Engineering and Computer Sciences,* vol. 31, no. 6, pp. 951-968, 2023.

[12] S. P. Praveen, E. J. Veerapaneni, C. Anuradha, K. Venugopal, V. Shariff and S. Sindhura, "Chronic Kidney Disease Prediction Using ML-Based Neuro-Fuzzy Model," *International Journal of Image and Graphics,* vol. 24, no. 06, 2024.

[13] S. Patil and S. Chaudhary, "Hybrid classification framework for chronic kidney disease prediction model," *Hybrid classification framework for chronic kidney disease prediction model,* vol. 72, no. 3, pp. 367-381, 2024.

[14] F. Yan, H. Huang, W. Pedrycz and K. Hirota, "A disease diagnosis system for smart healthcare based on fuzzy clustering and battle royale optimization," *Applied Soft Computing,* vol. 151, 2024.

[15] H. Haider, M. Hussain and I. L. Kharisma, "Predictive Analysis of Chronic Kidney Disease in Machine Learning," in *The 7th International Global Conference Series on ICT Integration in Technical Education & Smart Society*, Aizu, Japan, 2025.

[16] A. Sharma, D. Singh, A. Singh, M. P. Kantipudi, S. Potharaju and B. Babu, "Comparative Development and Evaluation of Multilayer Fuzzy and Adaptive Neuro-Fuzzy Inference Systems for Early Diagnosis of Chronic Kidney Disease," *Informatics in Medicine Unlocked,* vol. 60, 2025.

[17] J. Liu, B. Sun, J. Ye and X. Zhao, "Hybrid deep-learning prediction model based on kernel multi-granularity fuzzy rough sets and its application in the diagnosis and treatment of chronic kidney disease," *Engineering Applications of Artificial Intelligence ,* vol. 147, no. 2, 2025.

[18] D. Dua and C. Graff, "UCI Machine Learning Repository.," University of California, Irvine, School of Information and Computer Sciences., 2019. [Online]. Available: https://archive.ics.uci.edu/ml/datasets/Chronic_Kidney_Disease.

[19] L. A. Zadeh, "Fuzzy Sets," *Information and Control,* vol. 8, pp. 338-353, 1965.

[20] T. J. Ross, Fuzzy Logic with Engineering Applications. Wiley, John Wiley & Sons, Ltd, 2010.

[21] G. Charizanos, H. Demirhan and D. İçen, "Binary classification with fuzzy logistic regression under class imbalance and complete separation in clinical studies," *BMC Medical Research Methodology,* vol. 24, no. 145, 2024.

[22] Y. R. Alvarez, M. M. G. Lorenzo, Y. C. Mota, Y. F. Cabrera, M.Isabel, G. Hilarión, D. M. M. d. Oca and R. B. Pérez., "Fuzzy prototype selection-based classifiers for imbalanced data," *Pattern Recognition Letters,* vol. 163, pp. 183-190, 2022.

[23] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in *Proceedings of the ACM SIGKDD Conference*, San Francisco, CA, USA, 2016.

[24] W. Chen, K. Yang, Z. Yu, Y. Shi and C. P. Chen, "A survey on imbalanced learning: latest research, applications and future directions," *Artificial Intelligence Review,* vol. 57, no. 137, pp. 1-51, 2024.

[25] O. Rainio, J. Teuho and R. Klén, "Evaluation metrics and statistical tests for machine learning," *Scientific Reports,* vol. 14, 2024.

[26] A. Anerao and D. O. jadhav, "A Hybrid Multlayer Stacked Ensemble Model for Early Prediction of Liver Disease," *Internatinoal journal of Enviromemntal Sciences,* vol. 11, no. 20, pp. 1452-1461, 2025.

[27] K. M. Sujo, R. Hassan, K. Choi and M. A. Samad, "Accuracy, precision, recall, f1-score, or MCC? empirical evidence from advanced statistics, ML, and XAI for evaluating business predictive models," *Journal of Big Data,* vol. 12, 2025.

[28] J. Li, "Area under the ROC Curve has the most consistent evaluation for binary classification," *PLoS One,* vol. 19, no. 12, 2024.