The Performance of Machine Learning for Chronic Kidney Disease Diagnosis

Tsehay Admassu Assegie[1,*], Yenework Belayneh Chekol[2]

[1]Department of Computer Science, Injibara University, Injibara, Ethiopia

[2]Department of Information Technology, Injibara University, Injibara, Ethiopia

Received 16 December 2022; received in revised form 19 January 2023; accepted 04 March 2023

DOI: https://doi.org/10.46604/emsi.2023.11285

Abstract

This paper aims to review the performance of different machine learning (ML) models and develop models for the automated diagnosis of chronic kidney disease. To detect chronic kidney disease with better precision, selecting the right and better-performing ML model is significant as it improves the precision and accuracy of the chronic kidney disease diagnosis. The study uses the Joana Briggs Institute (JBI) scoping review methodology, which involves different steps such as searching relevant literature, conducting the review, and reporting the review result. In the search, the year of publication and the indexing of journals where the studies are published is used as inclusion and exclusion criteria. The review result shows that the current chronic kidney disease detection has focused on the development of ensemble-based and deep-learning methods. The deep learning method can achieve a higher accuracy of 99.75%.

Keywords: chronic kidney disease, machine learning, performance, scoping review

1. Introduction

Chronic kidney disease is one of the deadliest diseases on the globe if not treated early. According to Shanmugarajeshwari and Ilayaraja [1], chronic kidney disease affects 10% of the world's population. In chronic kidney disease detection, the complexity of differentiating symptoms or common individual factors can be used to identify the disease at the early stages during the diagnosis. The researchers have developed different machine learning (ML) models, to aid human experts in automating the detection of chronic kidney disease at the early stages of its occurrence. In the traditional method, a urine test is used to identify whether a patient is suffering from kidney disease or not. In the urine test, medical experts examine urine albumin levels to detect kidney disease. However, the urine test has some weaknesses, such as the albumin level of the urine test can be normal in the early stages, the condition of equipment, and the highly experienced medical experts are required for accurate detection of chronic kidney disease.

Alternative chronic kidney disease examination techniques with the help of an automated ML model have been widely proposed by numerous researchers. The use of automated ML detection ranges from simple linear models such as support vector machines and tree-based models to complex deep neural networks and ensemble learning methods. In addition, the research on automated detection of chronic kidney disease has attracted many researchers from the field of artificial intelligence (ANN) especially, data mining and ML. For instance, Amirgaliyev et al. [2] analyzed the performance of logistic regression (LR), support vector machine (SVM), and ANN for chronic kidney disease detection. The comparative result shows that better accuracy of 94.60% is achieved using the SVM model. The model was developed with the chronic kidney disease data collected from the University of California Irvine (UCI) repository.

* Corresponding author. E-mail address: tsehayadmassu2006@gmail.com

English language proofreader: Chih-Wen Teng

Similarly, Saringat et al. [3] applied the ensemble method to develop an automated kidney disease detection model. In the study, various features of chronic kidney disease are used to train the model to improve its accuracy. The performance of the developed Naive Bayes (NB) model shows that the ensemble method detects the disease with an accuracy of 98.5% using the chronic kidney disease dataset collected from the UCI repository.

This study aims to identify the development of current research on the automated chronic kidney disease detection system. By conducting an in-depth study to examine several aspects such as research trends of the past five years, the models, the dataset, the methods used, the performance produced, and the contribution of each study, which is reviewed in this paper. This study is expected to provide a future automated chronic kidney disease detection development. This paper is organized as follows: Section 2 discusses the review-based research of automated chronic kidney disease detection. The research methodology is explained in Section 3. Section 4 discusses the results and answers to the research questions. Finally, Section 5 concludes the key results of the research.

2. Related Work

Different ML algorithms have been applied to detect chronic kidney disease problems. For instance, the deep learning method is applied to the automated detection of chronic kidney disease [4]. The performance of different machine learning models for chronic kidney diagnosis was investigated [5]. Among them, the ANN model achieves 95% accuracy on the test dataset collected from Imam Khomeini Hospital Iran.

Similarly, Sharma et al. [6] proposed a deep learning-based automated chronic kidney disease detection model that is effective for screening patients for early cases of the disease. The degree of precision of the existing retinal image screening systems is increased with the automated deep learning model. Overall, the proposed system achieved 91.6% accuracy as shown in the simulation result.

Furthermore, Senan et al. [7] analyzed the performance of different machine learning models on an Indian chronic kidney disease dataset obtained from the University of California Irvine (UCI) machine learning data repository. ML algorithms, namely the decision tree (DT) and K-nearest neighbors (KNN) gradient descent are compared on the UCI test dataset. The comparative result shows that DT has achieved 98.6% accuracy outperforming the KNN gradient descent.

In addition, Ifraz et al. [8] conducted a comprehensive analysis of intelligent ML methods such as DT, LR, and KNN. This study compared the performance of these models on the simulation using accuracy as a measure of performance in the evaluation. The simulation result shows that the LR model performed with an accuracy of 97%, the DT model achieved an accuracy of 96.25%, and the KNN model performs with the least accuracy of 71.25% in the simulation dataset.

In the literature surveys [1-3], the researchers' review method does not adopt the standard literature review methodology that governs the review process. Moreover, their review does not show a detailed summary of what has been archived and the future direction in the automated chronic kidney disease diagnosis. For this reason, the Joana Briggs institute scoping review method for literature review research contributes to summarizing a literature review-based research flow. The aim is for the reader or other researchers who read this work easily find the information they are looking for from each point being reviewed.

3. Review Methodology

The JBI methodology is employed to conduct a scoping review on chronic kidney disease detection systems using an automated ML learning model. Furthermore, over the past few years, scoping review has become one of the literature review standards where the reviewer's goal is to determine what kind of evidence (quantitative and/or qualitative) is available on the topic and to represent this evidence by mapping or charting the data. The purpose of this review is therefore to apply the JBI scoping review to summarize merging trends and the use of the ML model for the effective detection of chronic kidney disease. In addition, scoping review is a more and more common technique to synthesize the ever-growing suite of methodological guidance and resources to assist reviewers with the planning, conducting, and reporting of the scientific literature review [9].

The first step of the review of the performance of the ML model is the collection of the chronic kidney disease dataset. Phase two involves processing such as duplicate removal, label encoding, and replacing missing values. After that, the dataset is split into training and testing. Then the model is trained on a training set and the trained model is evaluated in the next step. Finally, the model performance is compared with different performance metrics such as accuracy, f-score, and receiver operating characteristic (ROC) curve. Fig. 1 demonstrates the methodology followed to evaluate the performance of different supervised learning models for chronic kidney disease detection.

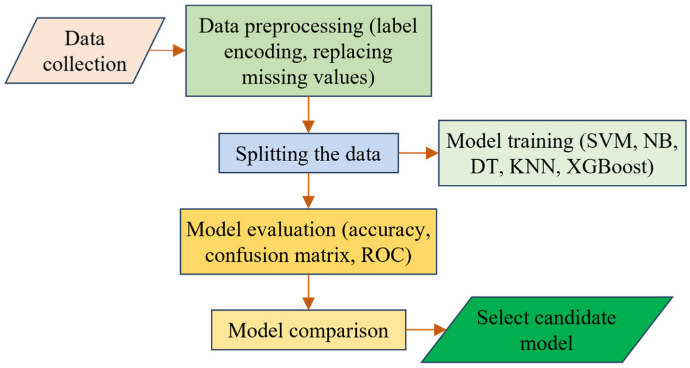

Fig. 1 Flowchart for the chronic kidney disease diagnosis

3.1. Research questions

To maintain the focus of the scoping review, the research question and the objectives of the review are identified. Accordingly, the researchers followed the five criteria known as Population, Intervention, Comparison, Outcomes, and Context (PICOC) and formulated research questions for searching and summarization of reviewed studies related to automated chronic kidney disease detection as shown in Table 1. The population in the field refers to the target group of the research [10-12], intervention in the field refers to issues that attract researchers [13-14], comparison in the field refers to the comparison of interventions [15], and outcomes in the field refer to the result of the research. The research questions formulated based on the PICOC criteria are detailed in Table 2.

Table 1 The PICOC criteria

Criteria	Scope
Population	Automatic detection, ML systems
Intervention	Urine analysis, dataset
Comparison	N/A
Outcomes	The accuracy of chronic kidney disease detection & performance of the ML model employed
Context	Automated medical diagnosis, scholarly articles

Table 2 Research questions

No.	Question	Motivation
1	What is the number of automated chronic kidney disease detection trends	To identify the number of studies conducted each year related to chronic kidney disease detection
2	What ML models are used to detect chronic kidney disease	To identify the type of ML model used to detect chronic kidney disease
3	What model has the better accuracy for kidney disease detection	To identify the model that is better suited for kidney disease detection
4	What types of the dataset are widely used in automated kidney disease detection studies	To identify the types of the dataset used for chronic kidney disease detection

3.2. Database search strategy

The search strategy aims to identify the primary studies considered for reporting the review result. The search strategy is used to collect literature from relevant sources and select appropriate literature for review based on the PICOC shown in Table 1 and research questions summarized in Table 2. In preparation for keywords searching relevant pieces of literature, the researchers followed the steps:

(1) Identify the search terms or keywords based on the PICOC and research question

(2) Identify the search keywords

(3) Identify the alternative search keywords such as synonym words

(4) Construct the search keyword using Boolean such as AND, OR

The keywords are used to collect the related literature depending on the relevance and requirements of the search database. This review only involves studies that were published in the open-access publication model. The year of publication is used as exclusion and inclusion criteria, only studies published in the years between 2017 and 2021 and journals publications are used excluding conference proceedings. The databases used in the literature search are presented as follows:

(1) IEEE- Explore

(2) Science Direct and Springer Link

3.3. Screening strategy

Pieces of literature that are not duplicate search results and similar studies are considered for the review process. The inclusion-exclusion criteria are summarized in Table 3. As demonstrated in Table 3, the studies published in the last five years were included in the literature survey. The studies that are not result-oriented are excluded from the literature review.

Table 3 Inclusion-exclusion criteria

Inclusion criteria	The study is published in the last five years The study is published in conference proceedings The study has a contribution to the literature and an appropriate method is used
Exclusion criteria	The study is not published in English The study is not result-oriented and does not provide an experimental analysis

3.4. Data extraction and eligibility criteria

With the searching strategy discussed in Subsection 3.2, 50 pieces of literature are retrieved from the major databases. However, 26 pieces are duplicated and unnecessary and do not meet the inclusion review criteria. Studies that are not published in the year between 2017 and 2021 were not considered for review as well. Similarly, studies that do not contain detailed results and discussions are also not qualified for the review report.

In addition, the sources of the dataset employed by each study are considered eligibility criteria for the inclusion and exclusion of literature used in the review. After inclusion, all of the literature pieces are analyzed to obtain evidence and provide answers to formulated research questions. Thus, the synthesis of the literature answers the following questions:

(1) What is the research trend of automated ML-based chronic kidney disease detection in the last five years?

(2) What are the contributions of the authors of each research published?

(3) Which ML model is widely used for automated chronic kidney disease diagnosis?

(4) What is the highest accuracy achieved by the ML model?

4. Results and Discussion

This section discusses the research trend, and the number of articles published each year that are related to automated chronic kidney disease detection using ML. The contributions of the different researchers to automated chronic kidney disease detection with ML. The ML method used for automated diagnosis of chronic kidney disease is the highest accuracy achieved so far in the literature.

4.1. Research trends

This review collects twenty-four pieces of literature on an automatic detection system using ML methods. The literature is collected through a selection process based on the inclusion and exclusion criteria and the formulated research questions. The collected literature is then rearranged based on the year of publication to see research trends related to the automation of kidney disease diagnosis using ML methods. The research trend on automated chronic kidney disease detection using ML methods provides the answers to research question 1. The yearly research trend on the automation of chronic kidney disease diagnosis with ML methods is shown in Fig. 2.

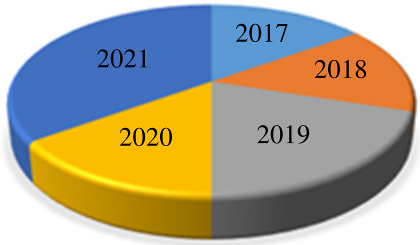

Fig. 2 Research trend for automated chronic kidney disease detection

The number of studies on automation of the chronic kidney disease diagnosis with ML methods has been increasing from year to year with the highest number of studies conducted in the year 2021 and the lowest number of studies conducted in the year 2017. In addition, this scoping review revealed a growing trend in the use of automated chronic kidney disease diagnosis, which is expected to grow and be widely adopted in healthcare centers soon.

4.2. Commonly used ML methods for kidney disease diagnosis

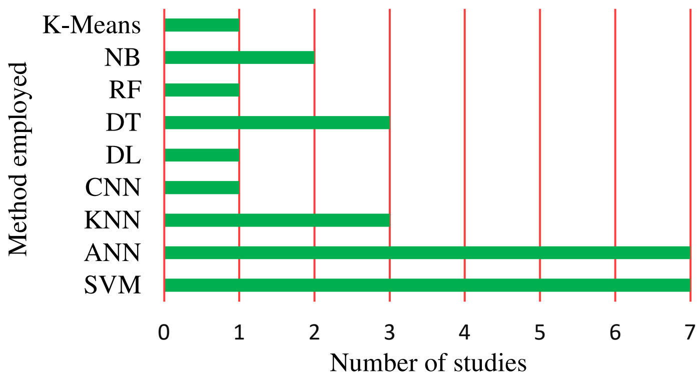

Fig. 3 The ML-model trend for automation of chronic kidney disease diagnosis

The majority of the studies conducted on the automation of chronic kidney disease focus on improving the performance of the ML model for automated chronic kidney disease diagnosis. E.g., different preprocessing methods such as principal component analysis (PCA) for dimensionality reduction of retinal images [16-17]. Different types of ML methods are employed to develop automated chronic kidney disease. The distribution of different algorithms used for chronic kidney disease is shown in Fig. 3. As demonstrated in Fig. 3 ANN and SVM are the most widely used ML method for automated chronic kidney disease detection.

4.3. State-of-the-art on automated chronic kidney disease diagnosis

This is the first scoping review to examine the use of automated chronic kidney disease detection. While research has been conducted on the automation of chronic kidney disease with the ML method [18] and performance evaluation of different ML methods used for chronic kidney disease as separate research topics [19], now is the time to realize the critical relationship between ML method and their performance on chronic kidney disease detection. Automated kidney disease diagnosis systems cannot be implemented effectively without a solid foundation of the state of the art, in both academic and clinical practice settings. The findings of this review will help data scientists and experts from the ANN field to proactively shape the future research direction and focus, ensuring that the ML method with the core decision-making process and personal assistant systems. The automated chronic kidney disease detection methods and the best accuracy achieved as well as the ML method employed are summarized in Table 4.

Table 4 List of selected studies in the literature

No.	Study	Method employed	Dataset	Performance
1	Amirgaliyev et al. [2]	SVM	UCI repository	94.60%
2	Saringat1 et al. [3]	Decision Tree	UCI repository	98.5%
3	M. Bhatt and T. Kasbe [5]	ANN	Imam Khomeini Hospital Iran	95%
4	Jayalakshmi et al. [10]	RF	UCI repository	94%
5	M. Almasoud and Ward [12]	Gradient Boost	UCI repository	99%
6	Tekale et al. [13]	SVM	UCI repository	96.75%
7	Pujianto et al. [14]	SVM	UCI repository	100%
8	Assegie et al. [21]	KNN	UCI repository	99.86%

4.4. The performance of the ML method on kidney disease diagnosis

As demonstrated in Table 4, the highest accuracy achieved on chronic kidney disease detection using the ML method is 99.75% with ANN. To determine the ML method that has the better performance on chronic kidney disease diagnosis, many aspects should be taken into account. For instance, to apply ML and deep learning algorithms optimally to chronic kidney disease detection, the researcher must recognize the characteristics of the problem and the type of dataset used for experimental simulation.

4.5. MLM model on chronic kidney disease detection performance

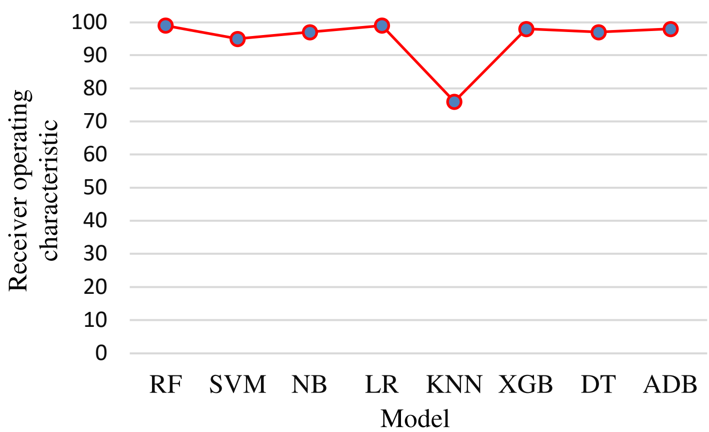

Fig. 4 The performance of MLM on chronic kidney disease diagnosis

This paper developed a different machine learning model (MLM) for the automation of kidney disease diagnosis. The model is developed using the chronic kidney disease dataset collected from the online Kgagle data repository previously employed by [21-22]. The RF, DT, SVM, NB, XGB, ADB, KNN, and LR are considered in this study because these models have higher performance and are commonly employed for chronic kidney disease diagnosis. The MLM is developed by collecting the dataset that contains 25 attributes, 11 numeric and 14 nominal. The dataset contains 400 instances, and 224 of them are used for the training to predict deep learning algorithms after removing instances with missing values. 105 out of the 224 cases have been labeled “chronic kidney disease” and 119 are “non-chronic kidney disease.” After pre-processing the original dataset, removing the missing values, and scaling the feature with a min-max scaler, the MLM is tested on the testing set. The ROC of the MML on chronic kidney disease diagnosis is demonstrated in Fig. 4.

Table 5 indicates the performance of supervised ML models for the diagnosis of chronic kidney disease. The RF and NB models have higher precision and F-score than the other ML models. The XGB model has the lowest precision and F-scored compared to other supervised ML models.

Table 5 The performance of MLM on chronic kidney disease detection

Model	Precision	Recall	F-Score
RF	0.97	1.00	0.98
SVM	0.96	0.96	0.96
NB	0.97	1.00	0.98
KNN	0.93	1.00	0.97
XGB	0.90	1.00	0.95
DT	0.93	1.00	0.97
ADB	0.93	1.00	0.97

Fig. 5 shows the accuracy of implemented models for chronic kidney disease diagnosis. RF and NB models achieved the highest accuracy score. ADB, DT, and KNN are the second highest. Furthermore, XGB and SVM models scored the lowest accuracy compared with the other models for chronic kidney disease diagnosis.

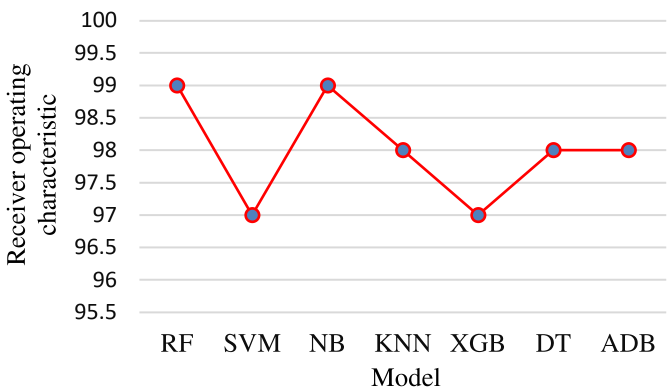

Fig. 5 The accuracy of proposed MML on chronic kidney disease diagnosis

5. Conclusions

In this study, an ML model was developed for chronic kidney disease diagnosis. The existing models for chronic kidney disease diagnosis were also compared using accuracy as the performance metric. The literature survey shows that the application of the ANN and SVM methods and the public UCI chronic kidney disease data repository are common for the studies conducted on the automation of kidney disease detection. The trend analysis shows that studies on automated kidney disease detection have increased in recent years as compared to the previous year. In future work, the use of explainable ML in the diagnosis of chronic kidney disease will be explored.

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] V. Shanmugarajeshwari and M. Ilayaraja, "Chronic Kidney Disease for Collaborative Healthcare Data Analytics Using Random Forest Classification Algorithms," International Conference on Computer Communication and Informatics (ICCCI), pp. 1-14, January 2021.
[2] Y. Amirgaliyev, S. Shamiluulu, and A. Serek, "Analysis of Chronic Kidney Disease Dataset by Applying Machine Learning Methods," IEEE 12th International Conference on Application of Information and Communication Technologies (AICT), pp. 1-4, October 2018.
[3] Z. Saringat, A. Mustapha, R. R. Saedudin, and N. A. Samsudin, "Comparative Analysis of Classification Algorithms for Chronic Kidney Disease Diagnosis," Bulletin of Electrical Engineering and Informatics, vol. 8, no. 4, pp. 1496-1501, December 2019.
[4] K. B. Chimwayi, N. Haris, R. D. Caytiles, and N. C. S. Iyengar, "Risk Level Prediction of Chronic Kidney Disease Using Neuro-Fuzzy and Hierarchical Clustering Algorithm (s)," International Journal of Multimedia and Ubiquitous Engineering, vol. 12, no. 8, pp. 23-36, 2017.
[5] M. Bhatt and T. Kasbe, "A Survey on Chronic Kidney Disease Diagnosis Using Fuzzy Logic," IEEE International Symposium on Smart Electronic Systems (iSES) (Formerly iNiS), pp. 252-256, December 2019.
[6] S. Sharma, V. Sharma, and A. Sharma, "A Two Stage Hybrid Ensemble Classifier Based Diagnostic Tool for Chronic Kidney Disease Diagnosis Using Optimally Selected Reduced Feature Set," International Journal of Intelligent Systems and Applications in Engineering, vol. 6, no. 2, pp. 113-122, April-June 2018.
[7] E. M. Senan, M. H. Al-Adhaileh, F. W. Alsaade, T. H. H. Aldhyani, A. A. Alqarni, N. Alsharif, et al., "Diagnosis of Chronic Kidney Disease Using Effective Classification Algorithms and Recursive Feature Elimination Techniques," Journal of Healthcare Engineering, vol. 2021, article no. 1004767, June 2021.
[8] G. M. Ifraz, M. H. Rashid, T. Tazin, S. Bourouis, and M. M. Khan, "Comparative Analysis for Prediction of Kidney Disease Using Intelligent Machine Learning Methods," Computational and Mathematical Methods in Medicine, vol. 2021, article no. 6141470, December 2021.
[9] M. D. J. Peters, C. Marnie, H. Colquhoun, C. M. Garritty, S. Hempel, T. Horsley, et al., "Scoping Reviews: Reinforcing and Advancing the Methodology and Application," Systematic Reviews, vol. 10, article no. 263, October 2021.
[10] V. Jayalakshmi, L. Nayak, and K. Dharmarajan, "A Survey on Chronic Kidney Disease Detection Using Novel Methods," International Journal of Pure and Applied Mathematics, vol. 119, no. 10, pp. 125-130, 2018.
[11] K. N. Komal, R. L. Tulasi, and D. Vigneswari, "An Ensemble Multi-Model Technique for Predicting Chronic Kidney Disease," International Journal of Electrical and Computer Engineering, vol. 9, no. 2, pp. 1321-1326, April 2019.
[12] M. Almasoud and T. E. Ward, "Detection of Chronic Kidney Disease using Machine Learning Algorithms with Least Number of Predictors," International Journal of Soft Computing and Its Applications, vol. 10, no. 8, pp. 80-96, April 2019.
[13] S. Tekale, P. Shingavi, S. Wandhekar, and A. Chatorikar, "Prediction of Chronic Kidney Disease Using Machine Learning Algorithm," International Journal of Advanced Research in Computer and Communication Engineering, vol. 7, no. 10, pp. 92-96, October 2018.
[14] U. Pujianto, N. A. Ramadhani, and A. P. Wibawa, "Support Vector Machine with Purified K-Means Clusters for Chronic Kidney Disease Detection," 2nd East Indonesia Conference on Computer and Information Technology (EIConCIT), pp. 56-60, November 2018.
[15] S. Gupta, V. Bharti, and A. Kumar, "A Survey on Various Machine Learning Algorithms for Disease Prediction," International Journal of Recent Technology and Engineering, vol. 7, no. 6c, pp. 84-87, 2019.
[16] A.Vijayalakshmi and V. Sumalatha, "Survey on Diagnosis of Chronic Kidney Disease Using Machine Learning Algorithms," 3rd International Conference on Intelligent Sustainable Systems (ICISS), pp. 590-595, December 2020.
[17] B. Deepika, V. K. R. Rao, D. N. Rampure, P. Prajwal, and D. G. Gowda, "Early Prediction of Chronic Kidney Disease by Using Machine Learning Techniques," American Journal of Computer Science and Engineering Survey, vol. 8, no. 2, article no. 7, September 2020.
[18] K. M. Almustafa, "Prediction of Chronic Kidney Disease Using Different Classification Algorithms," Informatics in Medicine Unlocked, vol. 24, article no. 100631, 2021.
[19] O. A. Jongbo, A. O. Adetunmbi, R. B. Ogunrinde, and B. Badeji-Ajisafe, "Development of an Ensemble Approach to Chronic Kidney Disease Diagnosis," Scientific African, vol. 8, article no. e00456, July 2020.
[20] W. Wang, G. Chakraborty, and B. Chakraborty, "Predicting the Risk of Chronic Kidney Disease (CKD) Using Machine Learning Algorithm," Applied Sciences, vol. 11, no. 1, article no. 202, December 2020.
[21] T. A. Assegie, "Automated Chronic Kidney Disease Detection Model with Knearest Neighbor," International Journal of Computer and Information Technology, vol. 10, no. 3, pp. 111-115, May 2021.
[22] C. Mondol, F. M. J. M. Shamrat, Md. R. Hasan, S. Alam, P. Ghosh, Z. Tasnim, et al., "Early Prediction of Chronic Kidney Disease: A Comprehensive Performance Analysis of Deep Learning Models," Algorithms, vol. 15, no. 9, article no. 308, August 2022.