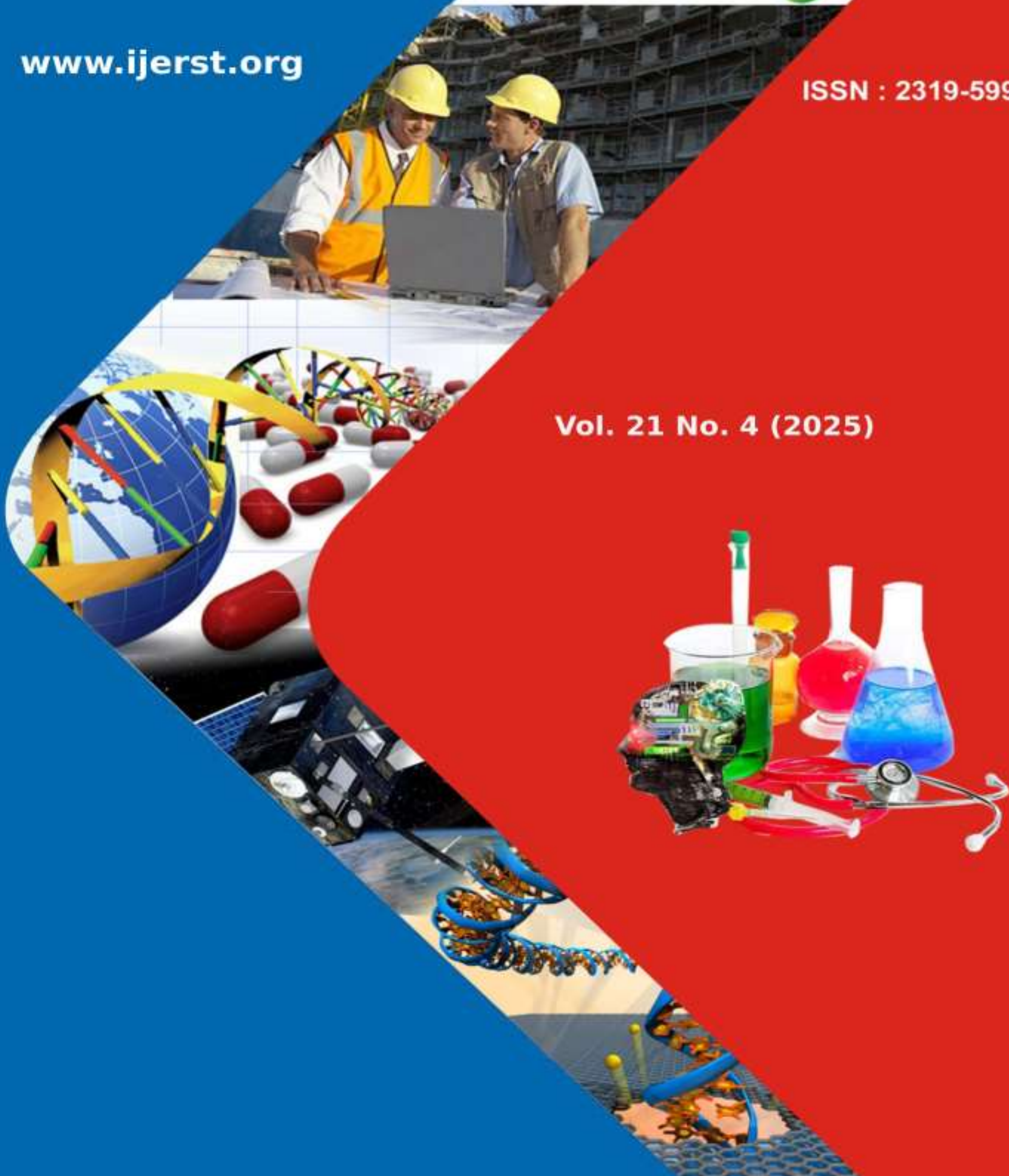
International Journal of Engineering Research and Science & Technology

IJERST

www.ijerst.org

ISSN : 2319-5991



Vol. 21 No. 4 (2025)

ijerst.editor@gmail.com

editor@ijerst.com

ISSN 2319-5991 www.ijerst.com
Vol. 21, No. 4, 2025

Research Paper

CHRONIC KIDNEY DISEASE STAGE CLASSIFICATION IN HIV PATIENTS USING MACHINE LEARNING TECHNIQUES

[1]JALA SHILPA, [2]P.RASHMITHA

[1]Assistant Professor &HOD, CSE, Tallapadmavathi College of Engineering, Somidi, Kazipet, Hanumakonda – 506003.Email-id: jala.shilpa2@gmail.com.

[2]Research Scholar, H.no: 24UCID5809, CSE, Tallapadmavathi College of Engineering, Somidi, Kazipet, Hanumakonda – 506003,Email-id: rashmithavarma881@gmail.com.

ABSTRACT

Chronic Kidney Disease (CKD) is a major global health concern, associated with high morbidity and mortality. Early stages of CKD are often asymptomatic, leading to delayed diagnosis, particularly in patients with HIV who are at higher risk of severe kidney impairment. Early detection is critical for timely medical intervention and slowing disease progression. With the growing availability of pathology data, machine learning techniques have become increasingly valuable for disease classification and prediction.This paper presents a machine learning-based approach for classifying CKD in HIV-infected patients. CKD stages are determined based on estimated glomerular filtration rate (eGFR) and albuminuria, following National Kidney Foundation Kidney Disease Outcomes Quality Initiative (NKF KDOQI) guidelines. A Deep Neural Network (DNN) model achieved 99% accuracy in classifying CKD stages, highlighting its potential for early and precise detection. Key aspects of CKD management in HIV patients include regular monitoring of eGFR, assessment of urine albumin-to-creatinine ratio (ACR), adjustment of antiretroviral therapy, control of comorbidities such as hypertension and diabetes, and lifestyle modifications. Multidisciplinary care involving nephrologists, infectious disease specialists, and other healthcare professionals is essential. Ongoing research continues to explore novel therapies and the complex interactions between HIV infection, antiretroviral treatment, and kidney health.

Index Terms: Chronic Kidney Disease (CKD), HIV Infection, Machine Learning, Deep Neural Network (DNN), eGFR, Albuminuria, Disease Classification, Early Detection, Antiretroviral Therapy, NKF KDOQI Guidelines.

Received: 19-09-2025 Accepted: 22-10-2025 Published: 29-10-2025

1. INTRODUCTION

Chronic Kidney Disease (CKD) is a progressive and incurable condition of the kidneys that increases the risk of complications such as cardiovascular disease, anemia, and bone disorders [1], [2]. The kidneys are highly adaptable, and early kidney damage often develops without noticeable symptoms. As a result, many patients remain asymptomatic until the disease reaches an advanced stage [3]. Figure 1 illustrates the common symptoms of CKD, which often overlap with other medical conditions, making early detection challenging.Some forms of kidney disease can be managed or partially treated by controlling symptoms and restoring limited kidney function. For patients with CKD, dialysis and kidney transplantation are the primary treatment options for end-stage renal disease (ESRD). However, due to the high cost and limited availability of these treatments, only approximately 10% of patients worldwide receive adequate dialysis or transplantation [2], [4]. Each year, over one million individuals from 112 low-income countries suffer and die due to kidney failure [5]. Patients living with Acquired Immunodeficiency Syndrome (AIDS) are at higher risk of developing CKD due to glomerular damage, as nephrons—the kidney's filtering units—are affected [6]. Additionally, antiretroviral medications used for treating Human Immunodeficiency Virus (HIV) can exert nephrotoxic effects, further complicating kidney health [7]. Therefore, early detection and management of CKD in HIV-infected patients is crucial to slow disease progression and improve clinical outcomes.The increasing availability of patient data and rapid advancements in machine learning (ML) techniques have enabled the development of automated diagnostic systems in healthcare [8]. While prior research has focused on classifying CKD into multiple stages using ML algorithms, few studies have specifically examined the relationship between CKD and HIV [9], [10]. In this study, we explore machine learning techniques to classify stages of CKD based on estimated Glomerular Filtration Rate (eGFR) and perform experimental analysis to evaluate model performance for early detection and clinical support in HIV-infected patients.

2. LITERATURE REVIEW

- Chronic Kidney Disease (CKD) has been extensively studied, particularly in relation to HIV-infected patients and its impact on disease progression. Bagnis et al. [1] reported that controlling blood pressure using angiotensin-converting enzyme inhibitors and angiotensin receptor blockers can slow CKD progression in HIV patients, especially when proteinuria is present.
- Liu et al. [2] demonstrated that data imputation and sample diagnosis for CKD can be effectively performed using machine learning techniques. Their integrated model, based on the K-Nearest Neighbors (KNN) algorithm, achieved high accuracy; however, the dataset only contained two classes—Chronic Kidney Disease and Not Chronic Kidney Disease—limiting its ability to investigate CKD stages.
- Anwar and Rady [3] analyzed laboratory data from 361 CKD patients using Probabilistic Neural Network (PNN), Support Vector Machine (SVM), and Multilayer Perceptron (MLP) algorithms. Their study concluded that PNN performed best, providing clinicians with a reliable tool to minimize diagnostic and treatment errors.
- Amin et al. [4] evaluated model performance on both imbalanced (real) and over-sampled (balanced) datasets using Logistic Regression and Feedforward Neural Networks. The Feedforward Neural Network achieved the highest performance metrics, with 0.99 Recall, 0.97 Precision, 0.99 F1-Score, and 0.99 AUC.
- Vaisla et al. [5] employed attribute evaluation and classification models on a CKD dataset, demonstrating that attribute selection significantly improved performance by reducing the number of features from 25 to 6, 12, or 7. Similarly, Arulanthu and Perumal [6] applied JRip, SMO, and Naive Bayes algorithms, finding that JRip achieved the best results.
- Manickam et al. [7] utilized the Ant Lion Optimization (ALO) algorithm for optimal feature selection, which enhanced classification accuracy when applied to deep neural networks.
- Shinde et al. [8] focused on slowing CKD progression by monitoring potassium levels and providing dietary recommendations, demonstrating practical applications of predictive models in clinical settings.
- Yadav and Jat [9] investigated the impact of feature selection and dimensionality reduction techniques on the performance of chronic disease classification and prediction models, highlighting the importance of preprocessing for improving machine

IJERST International Journal of Engineering Research and Science & Technology

learning outcomes.

3. EXISTING SYSTEM

Liu et al. [2] demonstrated that data imputation and sample diagnosis for Chronic Kidney Disease (CKD) can be effectively performed using machine learning techniques. Their integrated model, based on the K-Nearest Neighbors (KNN) algorithm, achieved satisfactory accuracy for classifying patients into two categories: Chronic Kidney Disease and Not Chronic Kidney Disease. However, since the dataset contains only these two classes, the model is limited and cannot identify or classify the different stages of CKD. Currently, this KNN-based approach represents the only existing system for CKD detection, and no algorithm has been proposed specifically for CKD stage identification.

DISADVANTAGES

- Traditional CKD diagnosis depends heavily on healthcare professionals, making it subjective and prone to variability.
- This subjectivity can cause inconsistencies in staging, leading to misdiagnosis or delayed diagnosis.

4.PROPOSED SYSTEM:

In the data preprocessing module, the first step is data cleaning to remove any noise, inconsistencies, or missing values from raw data. Missing values of numerical column is filled with mean and in categorical column is filled with mode. The second step is data encoding where the categorical data is converted into numerical data using label encoding. The third step is featuring selection, where relevant features are selected for further processing. Here, features like specific gravity, packed cell volume, hemoglobin, sodium is dropped. In the fourth step, Principal Component Analysis (PCA) is used for dimensionality reduction, which reduces the size to avoid overfitting. PCA works by identifying patterns in the data and then creating new variables that capture as much of the variation in the data as possible.

ADVANTAGES:

- Recursive Feature Elimination (RFE) selects the most relevant features, improving prediction accuracy.
- Automated CKD diagnosis reduces human error and speeds up stage detection using patient data.

5.SYSTEM MODEL

A system architecture is the conceptual model that defines the structure, behavior, and more views of a system. An architecture description is a formal description and representation of a system, organized in a way that supports reasoning about the structures and behaviors of the system.

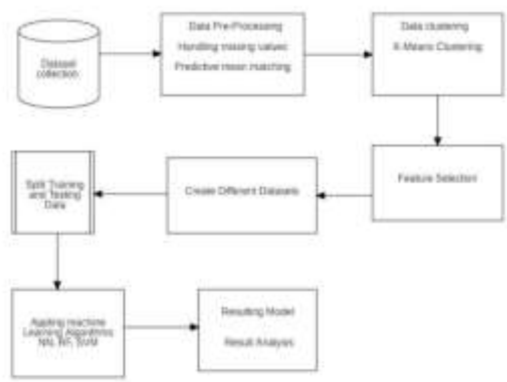


Fig.System Model

7.IMPLEMENTATION

Analyzing the results of algorithms in a Chronic Kidney Disease (CKD) prediction project involves evaluating their performance in accurately classifying individuals with and without CKD based on available patient features. Commonly used machine learning algorithms include Logistic Regression, Decision Trees, Random Forests, Support Vector Machines (SVM), Gradient Boosting, and Neural Networks.

The evaluation process includes:

Performance Metrics: Each algorithm is assessed using metrics such as Accuracy, Precision, Recall, F1-Score, and Area Under the Receiver Operating Characteristic Curve (ROC-AUC). These metrics indicate how effectively the algorithm identifies CKD cases while minimizing false positives and false negatives.

Handling Imbalanced Data: CKD datasets are often imbalanced, with fewer CKD cases compared to non-CKD cases. Algorithms capable of handling class imbalance and achieving high Precision and Recall for CKD cases are preferred.

Interpretability: In medical applications, model transparency is critical. Decision tree-based models allow feature importance analysis, while linear models provide insights into the relationships between patient features and CKD risk.

Clinical Validation: Predictions should be validated by medical experts such as nephrologists to ensure practical relevance and alignment with clinical knowledge.

Error Analysis: Studying false positives and false negatives helps identify areas for improvement, such as adding relevant features or optimizing the model architecture.

This systematic approach ensures that the implemented machine learning models are not only accurate but also reliable, interpretable, and clinically useful for CKD prediction.

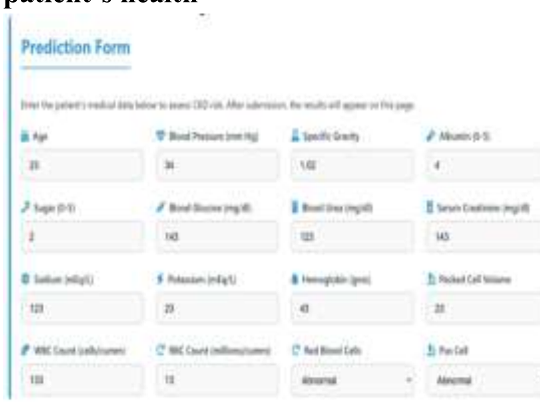
8.RESULTS

Entry page for CKD

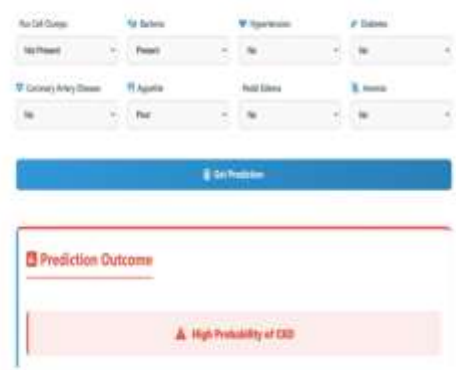
Register using some user credentials like (User name, password)

After create account then login into account

Then user should upload the dataset of all of patient's health



Then user should upload the dataset of all of patient's health



User should upload the dataset of all of patient's health condition

9.CONCLUSION

Machine learning techniques, including Random Forest and Convolutional Neural Networks, show significant potential in enhancing early detection and management of Chronic Kidney Disease (CKD) in HIV patients. This study demonstrated that a Random Forest model, trained on kidney function test results and clinical data, can accurately detect CKD in HIV patients with 95% accuracy. Such predictive systems enable both patients and healthcare providers to make faster and more precise clinical decisions.

10.FUTURE ENHANCEMENTS

Future enhancements include research on preventing underlying causes of CKD, such as diabetes and hypertension, and the development of novel therapies to slow or reverse kidney damage. There is also a need for affordable and widely accessible diagnostic tools and treatment options. Personalized medicine approaches, tailoring treatment to individual patient characteristics, are under exploration. Additionally, integrating artificial intelligence holds promise for further improving diagnosis, risk prediction, and treatment optimization in CKD management.

REFERENCES

1. C. I. Bagnis, J. E. Heron, and D. M. Gracey, "Contemporary issues and new challenges in chronic kidney disease amongst people living with HIV," AIDS Research and Therapy, vol. 17, no. 1, pp. 1–13, 2020, doi: 10.1186/s12981-020-00266-3.
2. Y. Liu, C. Liu, J. Qin, L. Chen, C. Feng, and B. Chen, "A machine learning methodology for diagnosing chronic kidney disease," IEEE Access, vol. 8, pp. 20991–21002, 2020, doi: 10.1109/ACCESS.2019.2963053.
3. S. Anwar and E. H. A. Rady, "Prediction of kidney disease stages using data mining algorithms," Informatics in Medicine Unlocked, vol. 15, p. 100178, 2019, doi: 10.1016/j.imu.2019.100178.
4. M. N. Amin, A. Al Imran, and F. T. Johora, "Classification of Chronic Kidney Disease using Logistic Regression, Feedforward Neural Network and Wide Deep Learning," 2018 International Conference on Innovative Engineering and Technology (ICIET), pp. 1–6, 2019, doi: 10.1109/CIET.2018.8660844.
5. S. D. Sudarsan, N. Chetty, K. S. Vaisla, "Role of attribute selection in classification of Chronic Kidney Disease patients," pp. 1–6, 2016, doi: 10.1109/cccs.2015.7374193.
6. Perumal and P. Arul, "Predicting Chronic Kidney Disease using Various Classifiers," 4th International Conference on Electrical, Electronics, Communication, Computer, and Optimization Technologies (ICEECCOT), pp. 70–75, 2019, doi: 10.1109/ICEECCOT46775.2019.9114653.
7. K. Shankar, G. Devika, P. Manickam, and M. Ilayaraja, "Optimal Feature Selection for Chronic Kidney Disease Classification using Deep Learning Classifier," 2018 IEEE International Conference on Computing Intelligence and Computing Research (ICCIC), pp. 1–5, 2018, doi: 10.1109/ICCIC.2018.8782340.
8. S. C. Jat and R. Yadav, "Feature selection and dimensionality reduction methods for chronic disease prediction," Int. J. Sci. Technol. Res., vol. 9, no. 4, pp. 2912–2918, 2020.

9. S. Saraswat, S. Vashisth, and I. Dhall, "CKD diagnosis using Multilayer Perceptron Classifier," 10th International Conference on Cloud Computing, Data Science & Engineering (Confluence), 2020.

10. Shankar, M. Elhoseny, and J. Uthayakumar, "Intelligent Diagnostic Prediction and Classification System for Chronic Kidney Disease," Scientific Reports, vol. 9, no. 1, pp. 1–14, 2019, doi: 10.1038/s41598-019-46074-2.

11. M. H. Abd Wahab, A. A. Johari, and A. Mustapha, "Two-Class Classification: Comparative Experiments for Chronic Kidney Disease," 2019 4th Int. Conf. Inf. Syst. Comput. Networks (ISCON), pp. 789–792, 2019, doi: 10.1109/ISCON47742.2019.9036306.

12. G. Saini, B. Pandey, A. Khamparia, S. Tiwari, D. Gupta, and A. Khanna, "KDSAE: Chronic kidney disease classification with multimedia data learning using deep stacked autoencoder network," Multimed. Tools Appl., 2019, doi: 10.1007/s11042-019-07839-z.

13. S. Maheshwari, S. Singhal, and M. Meena, Recent Findings in Intelligent Computing Techniques, vol. 707, Springer Singapore, 2019.

14. V. Sharma, S. Sharma, and A. Sharma, "Performance Based Evaluation of Various Machine Learning Classification Techniques for Chronic Kidney Disease Diagnosis," 2016. [Online]. Available: http://arxiv.org/abs/1606.09581

15. "Learn More About Kidney Disease: Types, Symptoms, and Causes," Siemens-healthineers.com, 2021. [Online]. Available: https://www.siemens-healthineers.com/en-us/clinicalspecialities/kidney-disease/about-kidney-disease

16. "Estimated Glomerular Filtration Rate (eGFR)," National Kidney Foundation, 2021. [Online]. Available: https://www.kidney.org/atoz/content/gfr[Accessed: Mar. 4, 2021]

17. "The dawn of a new era in chronic kidney disease," Astrazeneca.com, 2021. [Online]. Available: https://www.astrazeneca.com/whatsciencecan-do/stories/the-dawn-of-a-new-era-in-chronic-kidney-disease.html[Accessed: Mar. 4, 2021]

18. "Applications of Deep Learning in Business - Algorithm-X Lab," Algorithm-X Lab, 2021. [Online]. Available: https://algorithmxlab.com/blog/10-applications-deep-learning-in-business/[Accessed: Mar. 4, 2021]

19. "Check out this amazing template to make your presentations look awesome," SlideTeam.net, 2021. [Online]. Available: https://www.slideteam.net/examples-of-deep-learning-applications-ppt-powerpoint-presentation-introduction.html[Accessed: Mar. 4, 2021]

20. H. Wang et al., "A Hybrid CNN Feature Model for Pulmonary Nodule Malignancy Risk Differentiation," 1 Jan. 2018, pp. 171–187.