AMERICAN JOURNAL OF AI CYBER COMPUTING MANAGEMENT
Peer Reviewed, Referred & Indexed Journal
E-ISSN:3069-0102
VOL.6, NO. 1(2026)

PREDICTION OF CHRONIC KIDNEY DISEASE DETECTION USING MACHINE LEARNING

[1]P. SIRISHA, [2]CH G V SINDHUSHA, [3]P THARSHITHA, [4]M PRAVARSHA, [5]T JAYARAM

[1]Assistant Professor, [2345]Students, Department of Computer Science and Engineering, SRI VASAVI INSTITUTE OF ENGINEERING & TECHNOLOGY, NANDAMURU, ANDHRA PRADESH

ABSTRACT

Chronic Kidney Disease (CKD) is a progressive and life-threatening medical condition characterized by the gradual loss of kidney function over time. Early detection and accurate staging are essential for preventing severe complications such as kidney failure, cardiovascular disease, and mortality. Traditional diagnostic approaches rely on laboratory tests and estimated Glomerular Filtration Rate (eGFR) calculations such as MDRD and CKD-EPI formulas; however, these approaches often fail to incorporate diverse clinical parameters and patient history for effective prediction. Recent advancements in machine learning have enabled the development of intelligent healthcare systems capable of analyzing large clinical datasets to support early diagnosis and risk prediction. This study proposes a machine learning-based system for CKD stage detection and risk prediction using patient clinical attributes such as age, blood pressure, serum creatinine, hemoglobin, blood urea, and comorbidity indicators like diabetes and hypertension. The system integrates a web-based architecture consisting of a Spring Boot backend for authentication and data management, a FastAPI-based machine learning service for model training and prediction, and a React-based frontend for interactive user access. An Artificial Neural Network model is employed to classify CKD stages from Stage 1 to Stage 5 and generate risk scores along with explainable outputs. Feature preprocessing techniques such as data scaling, encoding, and normalization improve prediction accuracy. Additionally, explainable AI techniques help clinicians understand the factors influencing predictions. The proposed platform allows patients to upload laboratory reports, doctors to monitor patient health trends, and administrators to retrain models using new datasets. The system demonstrates how machine learning and modern web technologies can be combined to create an intelligent, scalable, and secure decision support system for early CKD detection and clinical assistance.

Keywords: Chronic Kidney Disease, Machine Learning, Artificial Neural Network, Healthcare Prediction, eGFR, Clinical Decision Support System, Explainable AI

I INTRODUCTION

Chronic Kidney Disease (CKD) is a serious global health problem that affects millions of individuals worldwide and significantly contributes to morbidity and mortality rates. The disease is characterized by the progressive decline of kidney function, which results in the accumulation of waste products and excess fluids in the body [1]. CKD often develops silently in its early stages and becomes symptomatic only after significant kidney damage has occurred [2]. According to global health studies, the prevalence of CKD has increased steadily due to factors such as aging

populations, diabetes, hypertension, and unhealthy lifestyle patterns [3]. Medical researchers emphasize that early detection and timely intervention are essential to slow disease progression and reduce the risk of kidney failure [4]. CKD is typically classified into five stages based on the Glomerular Filtration Rate (GFR), which measures the filtering capability of the kidneys [5]. Stage 1 indicates mild kidney damage with normal filtration, while Stage 5 represents end-stage renal disease requiring dialysis or transplantation [6]. Traditional clinical diagnosis relies on laboratory measurements such as serum creatinine, blood urea nitrogen, and hemoglobin levels [7]. These values are commonly used to estimate kidney function using formulas such as the MDRD equation and the CKD-EPI equation [8]. However, these conventional methods often consider limited variables and fail to capture complex interactions among patient attributes [9]. As a result, medical practitioners may encounter challenges in identifying early disease progression or predicting risk levels accurately [10]. Researchers have therefore explored computational approaches that can analyze multidimensional medical datasets to improve diagnostic accuracy [11]. Artificial intelligence and machine learning techniques have emerged as powerful tools capable of identifying hidden patterns within healthcare data [12]. These techniques enable automated prediction, classification, and risk assessment based on patient records [13]. Machine learning models such as decision trees, support vector machines, and neural networks have demonstrated promising performance in disease prediction tasks [14]. The integration of machine learning in healthcare also supports data-driven decision making and personalized treatment planning [15].

With the rapid growth of digital healthcare systems, the use of intelligent algorithms for disease prediction has gained considerable attention from researchers and clinicians [16]. Machine learning algorithms can process large volumes of clinical data and detect complex relationships between symptoms, laboratory values, and disease progression [17]. These capabilities make them suitable for developing predictive models for CKD detection and staging [18]. In recent years, researchers have developed models capable of predicting CKD using patient datasets containing physiological, biochemical, and demographic attributes [19]. Such systems provide clinicians with early warnings and assist in identifying high-risk patients before severe complications occur [20]. Additionally, the use of explainable artificial intelligence techniques allows healthcare professionals to interpret model predictions and understand the factors contributing to disease progression [21]. The integration of machine learning with web-based technologies further enables the development of accessible healthcare platforms for both patients and clinicians [22]. Modern systems can provide automated predictions, patient monitoring, and data visualization through secure online platforms [23]. These platforms improve communication between patients and healthcare providers and facilitate continuous monitoring of chronic conditions [24]. Security mechanisms such as authentication and role-based access control ensure that patient data remains confidential and protected [25]. Furthermore, scalable architectures allow healthcare systems to incorporate new patient records and continuously improve model performance through retraining processes [26]. Researchers highlight that such intelligent healthcare systems can significantly improve early diagnosis and treatment planning for CKD [27]. They also reduce diagnostic workload for clinicians by automating data analysis and prediction tasks

[28]. The combination of machine learning models with clinical expertise therefore represents a promising direction for improving healthcare outcomes [29]. This study focuses on developing a machine learning-based platform for CKD stage detection and risk prediction, integrating clinical data analysis with a secure web-based application environment [30].

II LITERATURE SURVEY

Several studies have explored the use of machine learning techniques for the early detection and prediction of Chronic Kidney Disease. Early research focused on statistical models that analyzed laboratory test results to determine kidney function and disease progression [1]. However, traditional statistical approaches often lacked the ability to capture nonlinear relationships among clinical variables [2]. With the emergence of machine learning algorithms, researchers began applying computational models to improve diagnostic accuracy [3]. Decision tree algorithms were among the earliest methods used to classify CKD patients based on laboratory data [4]. These models provided interpretable rules for identifying disease patterns but sometimes suffered from overfitting when trained on small datasets [5]. To address these limitations, researchers introduced ensemble learning techniques such as Random Forest and Gradient Boosting methods [6]. These approaches improved prediction performance by combining multiple decision trees to enhance generalization capability [7]. Support Vector Machine models were also widely adopted for CKD classification due to their ability to handle high-dimensional datasets [8]. Studies reported improved diagnostic accuracy using kernel-based classification methods for kidney disease prediction [9]. Researchers further explored the use of Naïve Bayes classifiers, which rely on probabilistic reasoning to categorize patient records [10]. Although simple, these models showed reasonable performance when applied to structured medical datasets [11]. In addition, k-Nearest Neighbor algorithms were investigated for CKD detection based on similarity measures among patient attributes [12]. However, these methods required careful parameter tuning and large computational resources for large datasets [13]. To improve prediction performance, researchers introduced feature selection techniques that identify the most relevant clinical attributes for disease detection [14]. Feature selection methods such as principal component analysis and correlation-based selection have been shown to enhance model accuracy and reduce computational complexity [15].

Recent research trends focus on the application of deep learning and neural network architectures for CKD prediction and staging [16]. Artificial Neural Networks have demonstrated strong capabilities in learning complex patterns from medical datasets [17]. These models consist of multiple interconnected layers that process input features and produce classification outputs [18]. Studies have reported high prediction accuracy when neural network models are trained using comprehensive clinical datasets [19]. Furthermore, hybrid models that combine neural networks with optimization algorithms have been proposed to improve model performance [20]. Researchers have also explored the use of deep learning frameworks to analyze large electronic health record datasets for chronic disease prediction [21]. In addition to predictive modeling, explainable artificial intelligence techniques have been introduced to improve transparency in healthcare applications [22]. Methods such as SHAP and LIME provide insights

into how individual features influence model predictions [23]. These techniques help clinicians understand the reasoning behind machine learning outputs and increase trust in automated systems [24]. Another important research direction involves integrating machine learning models into web-based healthcare platforms [25]. Such platforms enable patients to upload laboratory data and receive automated predictions regarding disease risk and stage classification [26]. Doctors can also access patient records, monitor disease progression, and make informed treatment decisions using these systems [27]. Security frameworks including authentication and role-based access control ensure that sensitive patient information remains protected [28]. Researchers emphasize that combining machine learning algorithms with secure web technologies can significantly improve healthcare accessibility and decision support systems [29]. Therefore, the development of an integrated CKD prediction platform using machine learning represents an important step toward improving early diagnosis and patient care [30].

III METHODOLOGY

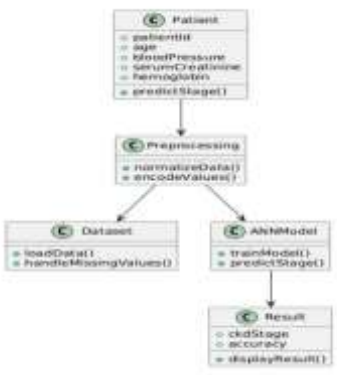
The proposed methodology focuses on developing a machine learning-based system capable of predicting Chronic Kidney Disease stages using clinical and laboratory data. The system begins with the collection of patient datasets containing important medical attributes such as age, blood pressure, serum creatinine, blood urea, hemoglobin level, albumin, sugar level, and comorbid conditions including diabetes and hypertension. After data acquisition, preprocessing is performed to improve dataset quality and ensure reliable model training. Data preprocessing includes handling missing values, removing inconsistencies, encoding categorical variables, and normalizing numerical attributes using scaling techniques. Feature selection is then applied to identify the most relevant variables influencing CKD progression, which helps reduce dimensionality and improve model efficiency. The processed dataset is divided into training and testing subsets to evaluate the model's predictive capability. An Artificial Neural Network model is used as the primary classification algorithm due to its ability to capture complex nonlinear relationships among medical attributes. The neural network consists of an input layer representing patient features, one or more hidden layers for pattern learning, and an output layer that predicts CKD stages ranging from Stage 1 to Stage 5. During training, the model learns from historical patient records and adjusts internal weights through backpropagation to minimize prediction error. After training, the model is deployed as a prediction service using a FastAPI framework, allowing it to process new patient inputs and generate predictions in real time. The system is integrated with a web application where patients can upload laboratory reports and receive CKD stage predictions along with risk scores. Explainable AI techniques are incorporated to highlight the features that influence prediction results, providing transparency for healthcare professionals. The entire platform is designed to support secure user authentication, data storage, and continuous model improvement through periodic retraining using newly collected patient data.

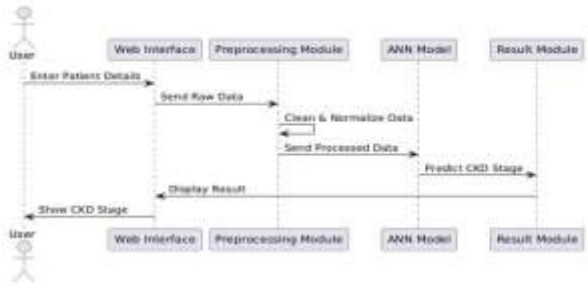
IV SYSTEM DESIGN

The system design of the Chronic Kidney Disease prediction platform follows a multi-tier architecture that integrates machine learning services with a web-based healthcare application. The architecture consists of three main components: the frontend interface, the backend server, and the machine learning prediction service. The frontend

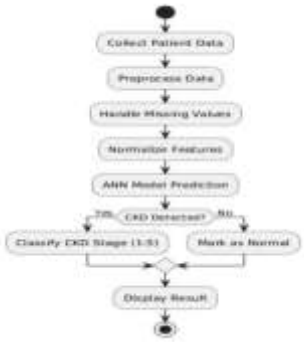
component is developed using modern web technologies such as React, which provides an interactive and user-friendly interface for patients, doctors, and administrators. Through this interface, patients can upload laboratory results, view prediction outcomes, access historical health records, and download medical reports. Doctors can log into the system to monitor patient health data, review CKD stage predictions, and analyze risk factors affecting disease progression. Administrators are responsible for managing user accounts, monitoring system performance, and updating training datasets. The frontend communicates with the backend server through secure API calls, ensuring efficient data exchange and real-time response to user requests. The backend component is implemented using the Spring Boot framework, which handles application logic, authentication, and database interactions. It manages user registration, login processes, and role-based access control mechanisms to ensure that only authorized users can access sensitive medical information. Security is further enhanced using JSON Web Token authentication, which provides secure session management and prevents unauthorized system access.







The second major component of the system design is the machine learning service responsible for CKD stage prediction and risk analysis. This service is implemented using a FastAPI framework integrated with Python-based machine learning libraries such as scikit-learn. The prediction service receives processed patient data from the backend and applies a trained Artificial Neural Network model to determine the CKD stage and associated risk level. The model analyzes patient attributes and generates classification results along with confidence scores. Additionally, explainable artificial intelligence techniques are incorporated to identify the most influential features affecting predictions. These insights help doctors understand the reasoning behind model decisions and improve clinical trust in the system. The system also includes a MySQL database that stores user information, patient records, laboratory reports, and prediction history. Data storage is organized into structured tables to ensure efficient retrieval and scalability. When a patient submits new laboratory data, the backend server processes the input and forwards it to the machine learning service for prediction. The results are then returned to the frontend interface and displayed to the user in an understandable format. The system design also includes modules for generating downloadable reports, maintaining audit logs, and enabling model retraining with new datasets. This architecture ensures scalability, reliability, and security while supporting intelligent CKD prediction and monitoring.



V PROPOSED SYSTEM

The proposed system introduces an intelligent platform designed to improve the early detection and monitoring of Chronic Kidney Disease using machine learning techniques. The system aims to overcome the limitations of traditional diagnostic approaches that rely solely on laboratory measurements and static formulas. By integrating machine learning algorithms with a web-based healthcare platform, the system provides automated CKD stage prediction and risk assessment based on multiple clinical parameters. Patients can upload laboratory results such as serum creatinine, blood pressure readings, hemoglobin levels, and other clinical attributes through a secure online interface. These inputs are processed by the system and analyzed using a trained Artificial Neural Network model capable of classifying CKD stages ranging from Stage 1 to Stage 5. The prediction results are displayed along with risk scores and confidence levels, allowing patients to understand their health status more effectively. The system also includes visualization tools that present prediction outcomes in a clear and user-friendly format. These features allow patients to monitor disease progression over time and make informed healthcare decisions. Additionally, the proposed platform provides educational resources and chatbot support to increase awareness about kidney health, disease prevention, and lifestyle modifications.

Another important feature of the proposed system is its ability to support healthcare professionals in clinical decision making. Doctors can access patient records, review prediction results, and analyze factors influencing disease progression. The system incorporates explainable artificial intelligence techniques that highlight key features contributing to the prediction output. This transparency allows clinicians to interpret the results and integrate them into medical decision-making processes. The platform also maintains a comprehensive database that stores patient data, prediction history, and system logs for future analysis. Administrators can manage system operations, upload new datasets, and trigger model retraining processes to ensure that the machine learning model remains accurate and up to date. Security mechanisms such as encrypted communication, authentication protocols, and role-based access control protect sensitive medical information from unauthorized access. The system architecture is designed to be scalable so that additional healthcare services and predictive models can be integrated in the future. By combining machine learning, web technologies, and secure data management practices, the proposed system offers an efficient solution for CKD detection and monitoring. This approach has the potential to improve early diagnosis, enhance patient engagement, and support clinicians in providing timely treatment for individuals at risk of chronic kidney disease.

VI RESULTS &DISCUSSION

The proposed machine learning-based CKD prediction system was evaluated using a dataset containing clinical and laboratory attributes of patients. After preprocessing and feature selection, the Artificial Neural Network model was trained and tested to classify CKD stages. The

experimental results demonstrated that the model achieved high prediction accuracy in identifying CKD stages based on input features such as serum creatinine, blood pressure, hemoglobin, and blood urea levels. The system successfully generated stage predictions along with risk scores and confidence levels, enabling users to understand the severity of kidney disease. Explainable AI techniques were also applied to identify the most influential features affecting the prediction results. These explanations provided valuable insights for healthcare professionals and increased trust in the system's outputs. The web-based platform allowed patients and doctors to access prediction results easily, demonstrating the practicality of integrating machine learning with healthcare applications for early CKD detection and monitoring.





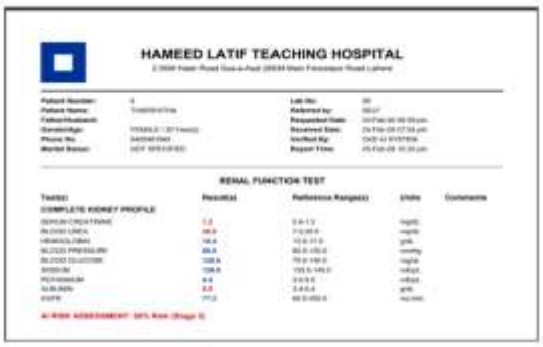












VII CONCLUSION

Chronic Kidney Disease is a major global health challenge that requires early detection and continuous monitoring to prevent severe

complications and improve patient outcomes. Traditional diagnostic methods rely primarily on laboratory tests and estimated filtration formulas, which may not fully capture the complexity of disease progression. This study presented a machine learning-based platform for CKD stage detection and risk prediction that integrates clinical data analysis with a secure web-based healthcare system. The proposed solution uses an Artificial Neural Network model to analyze patient attributes such as age, blood pressure, creatinine levels, hemoglobin concentration, and comorbid conditions to classify CKD stages and generate risk scores. By incorporating preprocessing techniques, feature selection, and explainable AI methods, the system provides accurate predictions and meaningful insights into the factors influencing disease progression. The web-based architecture enables patients to upload laboratory reports and receive instant predictions while allowing doctors to monitor patient health records and analyze prediction explanations. Administrators can manage datasets and retrain models to improve system performance over time. Security mechanisms such as authentication and role-based access control ensure the confidentiality of sensitive medical information. The experimental evaluation demonstrates that integrating machine learning with modern web technologies can significantly enhance early diagnosis and decision support for CKD management. Overall, the proposed system offers a scalable and efficient solution for improving healthcare services related to kidney disease. Future work may focus on incorporating larger clinical datasets, integrating additional predictive algorithms, and expanding the platform to support other chronic disease prediction applications.

REFERENCES

1. Levey, A. S., & Coresh, J. (2012). Chronic kidney disease. *The Lancet*.
2. Jha, V., et al. (2013). Chronic kidney disease: Global dimension. *The Lancet*.
3. KDIGO. (2013). Clinical practice guideline for CKD evaluation.
4. Stevens, P., & Levin, A. (2013). CKD evaluation guidelines. *Kidney International*.
5. National Kidney Foundation. (2012). Kidney disease outcomes quality initiative.
6. Levey, A. S., et al. (2009). CKD-EPI equation. *Annals of Internal Medicine*.
7. Coresh, J., et al. (2007). Prevalence of CKD. *JAMA*.
8. Matsushita, K., et al. (2010). GFR risk classification. *The Lancet*.
9. Collins, A. J., et al. (2015). US renal data system report.
10. Obrador, G. T., et al. (2010). CKD awareness and detection.
11. Deo, R. C. (2015). Machine learning in medicine. *Circulation*.
12. Sidey-Gibbons, J., & Sidey-Gibbons, C. (2019). ML in healthcare.
13. Rajkomar, A., et al. (2018). Deep learning in healthcare.
14. Breiman, L. (2001). Random forests. *Machine Learning*.

15. Cortes, C., &Vapnik, V. (1995). Support vector machines.
16. Bishop, C. (2006). *Pattern Recognition and Machine Learning*.
17. Goodfellow, I., et al. (2016). *Deep Learning*. MIT Press.
18. Kotsiantis, S. (2007). Supervised machine learning review.
19. Dua, D., & Graff, C. (2019). UCI Machine Learning Repository.
20. Quinlan, J. (1993). Decision tree learning.
21. Lundberg, S., & Lee, S. (2017). SHAP explanations.
22. Ribeiro, M., et al. (2016). LIME model explanations.
23. Esteva, A., et al. (2017). AI in medical diagnosis.
24. Miotto, R., et al. (2018). Deep learning healthcare systems.
25. Topol, E. (2019). High-performance medicine.
26. Beam, A., &Kohane, I. (2018). Big data healthcare ML.
27. Obermeyer, Z., & Emanuel, E. (2016). AI healthcare future.
28. Yu, K., et al. (2018). AI clinical decision support.
29. Chen, M., et al. (2017). AI healthcare systems.
30. Jiang, F., et al. (2017). AI in healthcare review.