International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

CHRONIC KIDNEY DISEASE PREDICTION USING CNN,LSTM &ENSEMBLE METHOD

[1]Kindoddi Sandhya Rani, [2]Maram Reddy Ram Mohan Reddy, [3]Gomasa Sidhartha, [4]Gandhari Sai Vardhan, [5]Badavath Veeranna, [6]D.Raj Kumar

[1,2,3,4]Students, Department of computer Science And Engineering, Malla Reddy Engineering College (Autonomous),Hyderabad Telangana, India 500100

[5]Assistant Professor, Department of computer Science And Engineering, Malla Reddy Engineering College (Autonomous),Hyderabad Telangana, India 500100

ABSTRACT

The global prevalence of chronic kidney disease (CKD) is increasing at an alarming rate. Many cases remain asymptomatic, and guideline-based monitoring for early detection is often underutilized. Computer-Aided Diagnosis (CAD) systems, particularly those leveraging deep learning, offer promising solutions for the early prediction of CKD due to their superior classification capabilities. This study explores the use of various clinical features associated with CKD and implements seven advanced deep learning algorithms—Artificial Neural Networks (ANN), Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), Bidirectional LSTM, Bidirectional GRU, Multilayer Perceptron (MLP), and Simple Recurrent Neural Network (RNN)—for effective prediction and classification of CKD. These models, grounded in artificial intelligence, were trained using five distinct feature extraction approaches on preprocessed CKD datasets. Performance was evaluated using metrics such as accuracy, precision, recall, loss, and validation loss. Additionally, the study assessed computation time, prediction ratios, and Area Under the Curve (AUC), alongside statistical significance tests to comprehensively compare model performance.

Keywords: Chronic Kidney Disease (CKD), deep learning, computer-aided diagnosis, artificial neural networks, LSTM, GRU, bidirectional networks, multilayer perceptron, clinical data analysis, feature extraction, medical prediction, classification models, healthcare AI, performance evaluation.

I. INTRODUCTION

Chronic Kidney Disease (CKD) has emerged as a significant public health issue due to its increasing global prevalence and long-term impact on human health. CKD encompasses a group of disorders that cause a gradual loss of kidney function over time. The kidneys play a vital role in filtering waste and excess fluids from the blood, regulating blood pressure, and maintaining essential electrolyte balance. When kidney function declines, it can lead to severe health complications and irreversible damage. The progression of CKD is often asymptomatic in the early stages, making early diagnosis challenging without proper screening. If undetected and untreated, CKD may advance to end-stage renal disease (ESRD), requiring dialysis or kidney transplantation for survival. Additionally, CKD is commonly associated with other health problems such as hypertension, cardiovascular diseases, anemia, and neuropathy, which further

complicate the clinical management of affected individuals. Globally, the burden of CKD is rising rapidly. It is estimated that around 2.2 million people worldwide suffer from renal failure. While the disease affects populations across all countries, its impact is especially pronounced in developing regions. Countries such as Pakistan, India, Nepal, Bangladesh, Bhutan, Sri Lanka, and Afghanistan are facing growing challenges in managing CKD due to limited healthcare resources, lack of awareness, and inadequate early screening programs. In more developed nations, such as the United States, CKD remains a pressing concern. Approximately 750,000 Americans are diagnosed with CKD each year, highlighting the need for proactive measures and better healthcare infrastructure to manage the disease. The increasing prevalence of CKD across both low- and high-income countries underscores the importance of developing effective and scalable diagnostic tools.

One of the critical barriers to timely CKD diagnosis is the underutilization of clinical guidelines and diagnostic protocols. Traditional methods often rely heavily on laboratory tests, which may not be accessible in rural or under-resourced areas. This calls for innovative solutions that can aid healthcare providers in detecting CKD early and accurately, even in challenging environments. Recent advancements in artificial intelligence (AI) and machine learning have opened new avenues for improving disease diagnosis and management. Computer-aided diagnostic (CAD) systems, particularly those utilizing deep learning algorithms, offer high potential for identifying patterns in complex medical data. These systems can enhance the speed and accuracy of CKD detection, ultimately contributing to better patient outcomes and reduced disease progression. Given these concerns and opportunities, this study focuses on the development and evaluation of deep learning-based models for CKD prediction and classification. By leveraging clinical features and modern AI techniques, the goal is to create robust tools that support early diagnosis and assist healthcare professionals in managing CKD more effectively.

II. LITERATURE SURVEY

J. Coresh, "Update on the burden of CKD," J. Amer. Soc. Nephrol., vol. 28, no. 4, pp. 1020–1022, Apr. 2017.

J. Coresh provides a concise yet impactful overview of the global burden of Chronic Kidney Disease (CKD), emphasizing its rising prevalence and underrecognized impact on public health systems. The article highlights how CKD, often asymptomatic in early stages, contributes significantly to cardiovascular morbidity and mortality, making it an urgent health priority. Coresh stresses that despite awareness efforts, CKD remains underdiagnosed and undertreated, especially in lower-income populations. The author also points out disparities in access to treatment and the limited implementation of preventive strategies. This paper is particularly important for setting the stage for predictive diagnostics, as it underscores the need for earlier detection to reduce progression to end-

stage renal disease. Coresh's work supports the idea that health systems must evolve to include better surveillance, patient education, and integration of predictive tools, such as AI and machine learning, to address the growing healthcare burden posed by CKD globally.

R. Kazancioğlu, "Risk factors for chronic kidney disease: An update," Kidney Int. Supplements, vol. 3, no. 4, pp. 368–371, Dec. 2013.

In this paper, R. Kazancioğlu examines both traditional and emerging risk factors contributing to the onset and progression of Chronic Kidney Disease (CKD). The article categorizes key risk factors into modifiable and non-modifiable groups, with special attention to diabetes mellitus, hypertension, cardiovascular conditions, and demographic attributes such as age and ethnicity. Kazancioğlu emphasizes that early intervention in modifiable risk areas could significantly delay CKD progression and improve outcomes. The article also highlights lifestyle-related factors, such as obesity and smoking, which are gaining recognition for their impact on kidney health. This survey is critical for informing machine learning models, as it provides a detailed understanding of input features that may enhance prediction accuracy. By drawing connections between clinical and environmental factors, the article forms a foundational basis for developing diagnostic systems that incorporate multifactorial risk assessments in CKD detection and classification.

G. Abraham et al., "Chronic kidney disease hotspots in developing countries in South Asia," Clin. Kidney J., vol. 9, no. 1, pp. 135–141, Feb. 2016.

G. Abraham and colleagues investigate the geographical and epidemiological distribution of CKD in South Asia, focusing on regions like India, Pakistan, Nepal, and Bangladesh. The paper identifies "hotspots" where CKD prevalence is unusually high, often unrelated to traditional risk factors like diabetes or hypertension. The authors suggest that environmental exposures, such as contaminated water, occupational hazards, and inadequate healthcare access, may play a significant role. This work underscores the need for region-specific public health strategies and diagnostic approaches. Importantly, the article calls for increased surveillance and research to understand local etiologies of CKD, which often differ from global patterns. For AI-driven diagnostic models, this study reinforces the necessity of incorporating regional data and contextual variables. It supports the integration of geographically sensitive datasets into predictive systems to ensure they are effective across diverse populations, particularly in under-resourced settings where CKD remains a silent epidemic.

T. J. Hoerger et al., "The future burden of CKD in the United States: A simulation model for the CDC CKD initiative," Amer. J. Kidney Diseases, vol. 65, no. 3, pp. 403–411, Mar. 2015.

T. J. Hoerger and colleagues present a robust simulation model predicting the future burden of CKD in the U.S. under the CDC's CKD Initiative. Using demographic trends, disease progression rates, and healthcare usage data, the model forecasts a substantial rise in CKD prevalence over the next few decades. The study provides actionable insights into the

potential strain on healthcare infrastructure and the economic impact of untreated CKD. It also emphasizes the importance of early detection and preventive care to mitigate future challenges. For researchers in the field of artificial intelligence and medical diagnostics, this paper is particularly useful. It justifies the urgency of implementing predictive systems that can identify at-risk populations before significant kidney damage occurs. The model's data-driven approach aligns well with machine learning methodologies, offering a framework that complements algorithmic prediction and classification techniques for CKD.

M. Almasoud and T. E. Ward, "Detection of chronic kidney disease using machine learning algorithms with least number of predictors," Int. J. Adv. Comput. Sci. Appl., vol. 10, no. 8, pp. 1–9, 2019.

M. Almasoud and T. E. Ward explore the application of various machine learning algorithms for CKD detection using a minimal set of predictors. The study evaluates models like Decision Trees, Support Vector Machines (SVM), and k-Nearest Neighbors (k-NN), focusing on optimizing accuracy while reducing computational overhead. One of the key contributions of this paper is its emphasis on feature selection to enhance model efficiency without compromising performance. By identifying the most influential clinical parameters, the authors demonstrate how accurate predictions can be achieved using fewer inputs, making the system more practical for real-time healthcare applications. This work is highly relevant for researchers developing CAD systems in resource-constrained environments, as it balances accuracy with simplicity. The paper validates the potential of machine learning in automating early CKD diagnosis and provides a foundational comparison point for integrating more complex deep learning algorithms in future studies.

III.PROPOSED METHODOLOGY

This study employs a range of deep learning algorithms, including Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM) networks, and an ensemble model combining CNN with Random Forest, for the prediction and classification of Chronic Kidney Disease (CKD). The models are trained and evaluated using the publicly available Chronic Kidney Disease dataset sourced from the Kaggle repository.

The dataset consists of multiple clinical and diagnostic attributes, with the final column representing the class label indicating the presence of CKD ("CKD") or absence of the disease ("NO_CKD"). To ensure compatibility with deep learning models, all categorical (non-numeric) features in the dataset are converted into numeric form using the Label Encoder technique from the scikit-learn library. This preprocessing step is essential, as most machine learning algorithms require numerical input.

Following data preprocessing, each model is trained on the prepared dataset. The performance of the models is assessed using standard evaluation metrics such as accuracy, precision, recall, F1-score, and the confusion matrix. These metrics provide insights into the classification ability of each algorithm and help compare their effectiveness in identifying CKD cases accurately.

By implementing and analyzing the performance of individual and ensemble models, this methodology aims to identify the most robust and reliable predictive model for early CKD detection. This approach not only enhances diagnostic accuracy but also contributes to developing automated, data-driven healthcare solutions for CKD risk management.

IV. CONCLUSION

This study presented a comprehensive methodology for the detection of Chronic Kidney Disease (CKD) using seven advanced deep learning algorithms. The primary objective was to enhance early diagnosis and identify key risk factors that contribute to the progression of CKD toward end-stage renal failure. Through systematic evaluation, the study provides valuable insights into the performance and applicability of various deep learning models in a medical diagnostic context.

A significant contribution of this research lies in its rigorous approach to data preprocessing, particularly the handling of missing values—a common challenge in real-world clinical datasets. The study employed linear regression to impute missing values in numerical attributes and logistic regression for categorical features, ensuring a scientifically sound and robust dataset for training and evaluation.

Overall, the proposed methodology advances the use of artificial intelligence in healthcare by demonstrating the potential of deep learning techniques for accurate and early prediction of CKD, thereby supporting timely clinical interventions.

V.REFERENCES

[1] J. Coresh, ‘‘Update on the burden of CKD,’’ J. Amer. Soc. Nephrol., vol. 28, no. 4, pp. 1020–1022, Apr. 2017.

[2] R. Kazancioğlu, ‘‘Risk factors for chronic kidney disease: An update,’’ Kidney Int. Supplements, vol. 3, no. 4, pp. 368–371, Dec. 2013.

[3] G. Abraham, S. Varughese, T. Thandavan, A. Iyengar, E. Fernando, S. A. J. Naqvi, R. Sheriff, H. Ur-Rashid, N. Gopalakrishnan, and R. K. Kafle, ‘‘Chronic kidney disease hotspots in developing countries in South Asia,’’ Clin. Kidney J., vol. 9, no. 1, pp. 135–141, Feb. 2016.

[4] T. J. Hoerger, S. A. Simpson, B. O. Yarnoff, M. E. Pavkov, N. R. Burrows, S. H. Saydah, D. E. Williams, and X. Zhuo, ‘‘The future burden of CKD in the United States: A simulation model for the CDC CKD initiative,’’ Amer. J. Kidney Diseases, vol. 65, no. 3, pp. 403–411, Mar. 2015.

[5] M. Almasoud and T. E. Ward, ‘‘Detection of chronic kidney disease using machine learning algorithms with least number of predictors,’’ Int. J. Adv. Comput. Sci. Appl., vol. 10, no. 8, pp. 1–9, 2019.

[6] F. Ma, T. Sun, L. Liu, and H. Jing, ‘‘Detection and diagnosis of chronic kidney disease using deep learning-based heterogeneous modified artificial neural network,’’ Future Gener. Comput. Syst., vol. 111, pp. 17–26, Oct. 2020.

[7] M. A. Islam, S. Akter, M. S. Hossen, S. A. Keya, S. A. Tisha, and S. Hossain, ‘‘Risk factor prediction of chronic kidney

disease based on machine learning algorithms,'' in Proc. 3rd Int. Conf. Intell. Sustain. Syst. (ICISS), Dec. 2020, pp. 952–957.

[8] I. U. Ekanayake and D. Herath, ''Chronic kidney disease prediction using machine learning methods,''