

“PRECISION DIAGNOSTICS FOR CHRONIC KIDNEY DISEASE USING DEEP LEARNING”

Presented By:

CHETAN

Dept.Of Computer Science And Engineering (Mca)

Vtu Cpgs Kalaburagi,

chetanlaskar@gmail.com

Under The Guidance Of:

Mrs. Shobha S. Biradaar

(Assistant Professor)

Dept. Of Computer Science And Engineering(Mca)

Vtu Cpgs Kalaburagi.,

baichabal@gmail.com

Abstract:— *Chronic kidney disease poses a significant global health challenge exacerbated by modern lifestyles, dietary habits, and medical practices. Its prevalence highlights the urgent need for effective detection and management strategies. Early identification of CKD is crucial as a postponed diagnosis may result in to irreversible renal damage and increased health risks. Traditional diagnostic methods often lack cost-effectiveness and efficiency, delaying timely intervention. However, advancements in technology, particularly deep learning, offer promising prospects for improving CKD prediction. The three deep learning models—Artificial Neural Network (ANN), Recurrent Neural Network (RNN), and Long Short-Term Memory (LSTM)— were trained using the provided dataset to predict Chronic Kidney Disease (CKD) based on selected features. This study explores the application of deep learning techniques in early CKD detection, highlighting their potential to revolutionize diagnostic protocols. By elucidating and comparing these strategies, valuable insights are provided to optimize CKD detection. Harnessing modern technology, particularly through deep learning, promises significant advancements in timely CKD identification and management, ultimately leading to improved patient outcomes and public health benefits. This research advocates for innovation in diagnostic methodologies and underscores the integration of deep learning to propel CKD detection towards greater efficiency and accuracy. Regarding chronic kidney disease (CKD) detection, the ANN algorithm demonstrated superior performance compared to LSTM and RNN models. Achieving an impressive accuracy of 92.1%, the ANN model.*

1. INTRODUCTION:-

Chronic kidney disease is characterized by the impaired role of the kidneys, leading to an inability to effectively filter blood. The kidneys play a crucial part in removing excess water and waste from the bloodstream, which is then excreted as urine. In individuals with CKD, this filtration process is compromised, leading to the build-up of toxins in the body. Symptoms of CKD include back pain, diarrhea, nosebleeds, fever, rash, vomiting, and abdominal discomfort. Progressive damage to the kidneys can have significant impacts over time implications for

overall health, predisposing individuals to various other medical conditions. In its advanced stages, CKD can be life-threatening. CKD imposes a substantial economic burden on patients, healthcare services, and governmental resources. Treatments for end-stage renal disease (ESRD), such as hemodialysis, peritoneal dialysis, and transplantation, are either expensive or complex. Furthermore, the correlation between CKD and other chronic conditions exacerbates the issue. From a public health perspective, it is crucial to forecast CKD occurrence trends to enable timely decision-making by policymakers, insurers, and

healthcare administrators. The lack of early detection methods for various diseases has led to an increment in mortality rates related with these conditions. Consequently, there has been a surge in research efforts aimed at addressing this critical issue.

Advanced computer-based detection methods have become essential assets in this endeavor, aiming to assist healthcare professionals in timely disease identification and management. By utilizing the strength of technology, these methods hold the potential to significantly reduce mortality rates by enabling early intervention and treatment. As such, research plays an essential part in improving healthcare outcomes and saving lives. Possible mitigation tactics could also involve expanded population screening for CKD risks and awareness initiatives. Lifestyle modifications, such as weight management, dietary improvements, increased physical activity, decreased alcohol intake and giving up smoking, early nephrologist referral, appropriate medication use, and control of other risk factors, have proven to be the most effective measures.

2. SYSTEM ANALYSIS EXISTING SYSTEM:

[1] Mohammad Abdul Quayyum et al. [2]. Multiple XG-boost combined with a simple deep neural network (DNN) evaluated KNN, SVM, random forest, decision tree, ada-boost, and other machine learning algorithms to assess if HIV patients had Chronic Kidney Disease. (CKD). They compared the execution of these algorithms in achieving high accuracy, reporting a notable accuracy rate of 98% for CKD classification in HIV-infected patients. This research underscores the significance of accurate classification for guiding timely clinical decisions beneficial to both patients and healthcare providers.

[2] Prof. Asfakahemad Y Darveshwala et al. [3] presented the experimental setup depicted in Figure 2. They explored various state-of-the-art SVM and other

methods for machine learning (Support Vector Machine) and K-nearest neighbors (KNN), for binary classification tasks. These algorithms were assessed according to their efficacy in dividing HIV-positive patients' phases of chronic kidney disease. The research's objective was to assess its functionality in conjunction with a deep neural network (DNN), highlighting a notable accuracy of 93%. The research underscores the importance of accurate CKD classification in HIV patients for facilitating timely and precise clinical decisions beneficial to both patients and healthcare providers.

[3] Naresh et al. [6] explored multiple a variety of machine learning techniques, such as a decision tree, xg-boost, adaboost, random forest, and a simple deep neural network (DNN), to ascertain whether they are experiencing Chronic Kidney Disease (CKD). They compared the performance of these state-of-the-art algorithms, noting that the DNN demonstrated superior performance in CKD classification. The study also introduced the use of the EGFR formula to accurately identify disease stages, achieving an impressive accuracy rate of 99%. These findings underscore the efficacy of DNNs and computational methods in enhancing the precision of CKD diagnosis and staging.

[4] Deepanita Baidya et al. [3] conducted preprocessing on their dataset, dividing it into two sets: a training dataset comprising over Three quarters of the data for predicting attribute values, and a test dataset with 25% allocated for evaluation. They applied machine learning models, including AdaBoost, Random Forest, and Decision Tree Gradient Boost, K-Nearest Neighbors, XGBoost, Gaussian Naive Bayes (GNB), and Extra Tree Classifier on the training data. Subsequently, they evaluated these algorithms on metrics, such as accuracy, precision, recall, and f1-score for each, are measured using the test data. The research focuses on the challenging task of accurately predicting Chronic

Kidney Disease (CKD), showcasing machine learning's efficacy in efficiently detecting CKD using real-time features. Their findings highlight achieving a remarkable level of precision rate of 99%, underscoring the potential of these methodologies in advancing medical science research.

[5] Sara Alshakrani et al. [3] utilized the Pandas library functions "dropna", "backfill", and "ffill" in order to address missing values in the dataset. Columns or rows that contained incomplete data. dropped using "dropna", and missing data was filled in by utilizing "backfill" and "ffill" to replace NAs with attribute mean values. The dataset was then processed by dropping specific columns (Rbc, Wc, and Rc) and transforming nominal labels into binary format (0,1). The research sought to ascertain which machine learning classifier is most effective in forecasting Chronic Kidney Disease (CKD). characterized by traditional methods. They evaluated seven classifiers: KNN, Gaussian Naive Bayes (GNB), Random Forest , Decision Tree, Logistic Regression , Gradient Boosting, and AdaBoost. Notably, GNB and RF achieved exceptional accuracy percentages, together with GNB and RF achieving 100%. The study underscores the need for a mechanism for decision support to enable early CKD prediction, given the limited existing research in this critical area. This suggests that GNB and RF could be suitable for applications involving similar datasets related to medical conditions, emphasizing their potential in advancing predictive healthcare applications.

Traditional statistical In the contemporary clinical decision support systems (CDSS) intended to predict the development Models and machine learning algorithms are crucial components of chronic kidney disease (CKD). These models are often trained using structured clinical data, which includes demographics, medical history, and diagnostic test results. However,

their ability to accurately predict CKD is often limited by challenges in handling intricate relationships among risk factors and evolving patient characteristics. While existing frameworks achieve reasonable accuracy, there remains a pressing need for more effective models.

Future models must integrate diverse data sources, adapt to real-time information updates, and deliver timely and precise predictions autonomously to support clinical decision-making. Advancing beyond current methodologies entails developing models capable of robustly analyzing and interpreting complex datasets. This evolution is crucial for ensuring accurate CKD prediction and improving overall patient care outcomes.

3. PROPOSED SYSTEM:

We developed a chronic kidney disease model focusing on identifying crucial variables using logistic regression, enriched by insights from deep learning techniques and algorithms. Our approach involved Analyzing the dataset with data mining techniques before further processing. and classification techniques implemented through the Scikit-learn library in Python. Additionally, We assessed the efficiency of the Naïve Bayes method to benchmark accuracy results. The data set consists of. key variables such as Age, Albumin, Sugar, Blood Glucose Random, Blood Urea, Serum Creatinine, Hemoglobin, Packed Cell Volume, White Blood Cell Count, Hypertension, and a target variable indicating the existence or lack of kidney disease. After importing and preparing we created a dataset Logistic Regression model. This model utilizes the sigmoid function to visually represent and classify the dataset based on organized features of kidney disease patients. To conclude our study, we compared the prognostic performance of the model utilizing logistic regression against other methodologies using techniques like the Confusion Matrix. This assessment

aimed to gauge the model's accuracy in chronic kidney disease detection and refine our approach for enhanced predictive capability.

4. SYSTEM DESIGN

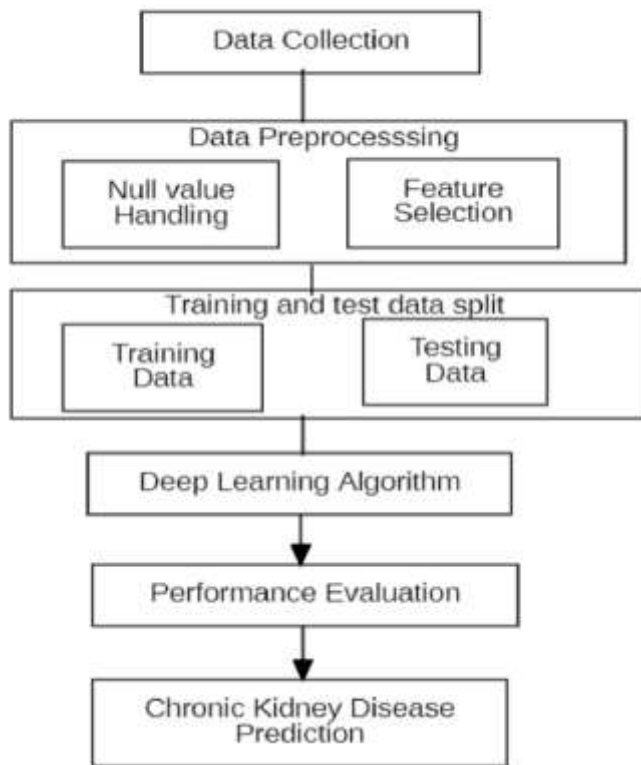


Figure. System Architecture

The flowchart outlines the process of using a deep learning approach to predict chronic kidney disease. Initially, data collection is performed to collect all pertinent data regarding the disease, such as patient records and health indicators. Once the data is collected, it undergoes preprocessing to ensure it is suitable for analysis. This preprocessing step includes handling null values to address any missing data and performing feature selection to identify the key components that will contribute to the prediction model. After preprocessing, the data is split into training and testing sets. Utilizing the training data, one can teach the deep learning algorithm, allowing it to learn patterns and relationships within the dataset. The testing data, on the other hand, is reserved to assess how well it is doing model.

5. METHODOLOGY:

A. Divide the dataset

The dataset was partitioned splitting the information into training and testing subsets to facilitate the artificial intelligence model's development and evaluation. The training set, which constituted a large amount of the information (usually 70-80%), model using selected features. During this phase, the model learned patterns and relationships by processing input-output pairs (features and corresponding classifications). Following training, the model's performance was assessed using the testing set, This took place apart from the training process. The testing set, typically smaller in size compared to the training set, served as an independent dataset to evaluate the degree of effectiveness model generalized to unseen data. By presenting the model featuring new, unseen observations, this step enabled an assessment of its predictive accuracy or other performance metrics, like the region beneath the ROC curve, specifically in classifying CKD status based on the chosen features. This approach ensured a reliable evaluation of the model's capability to make accurate predictions on real-world data.

B. Training the deep learning algorithm

The three deep learning models—Artificial Neural Network (ANN), Recurrent Neural Network (RNN), and Long Short-Term Memory (LSTM)—were trained using the provided dataset to predict Chronic Kidney Disease (CKD) based on selected features. The ANN, serving as the foundational model, iteratively adjusted internal parameters to effectively classify CKD status. RNN, specialized in handling sequential data, processed input features in chronological order to capture temporal patterns relevant to CKD detection. LSTM, a refined form of RNN, focused on learning

long-term dependencies crucial for accurate prediction of CKD progression. Each model was trained to maximize accuracy in CKD classification, leveraging their respective strengths to interpret and learn from the dataset's intricacies.

C. Hyper parameter tuning

Hyperparameter tuning is the procedure for enhancing the parameters of a machine learning model to improve its performance. For advanced learning algorithms such as ANN, RNN, and LSTM, critical hyperparameters include learning rate, quantity of concealed layers, quantity number of neurons in per layer, batch size, activation functions, epochs, dropout rate, optimizer choice, and regularization techniques. Methods such as grid search or random search are typically employed to identify the optimal combination of these hyperparameters. The objective is to enhance the model's accuracy and ability to generalize by refining its settings As stated by the specific traits of the dataset and the requirements of the predictive task.

D. Evaluate the model

Accuracy serves as a direct metric for evaluating the models by measuring how frequently they correctly predict the existence or lack of kidney disease, chronic (CKD) relative to all forecasts. It determines the proportion of accurate predictions equal to the overall quantity of predictions generated by the model. While a higher accuracy score signifies better performance, values approaching 1 indicate more precise predictions.

Result And Discussion

The experimental results demonstrate that the Artificial Neural Network (ANN) algorithm significantly outperforms LSTM and RNN in the context of chronic kidney disease (CKD) detection. The ANN model achieved an impressive accuracy of 92.1%, showcasing its effectiveness in accurately predicting CKD. In contrast, the LSTM model achieved a lower accuracy

of 61.87%, and the RNN model's performance, as depicted in Figure 3, further illustrates its inferiority compared to ANN. These findings underscore the superiority of ANN over LSTM and RNN for CKD detection, emphasizing its potential as a robust and reliable predictive model in healthcare applications.

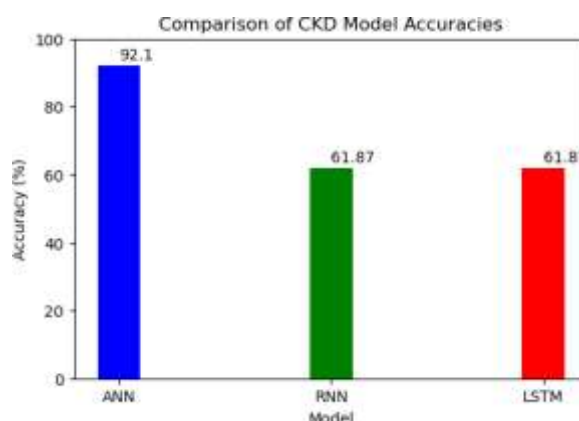


Figure 1. Performance analysis

The accuracies of the ANN model steadily increased with each epoch, as depicted in Figure 1. This upward trend indicates that the model's performance improved over time as it underwent additional training epochs.

The confusion matrix of the ANN model were given in the figure 2. The figures describes that the model correctly classified more number of samples.

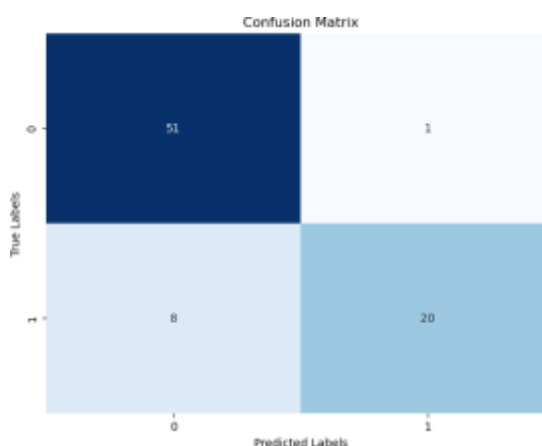


Figure 2. Confusion matrix

The table describes the precision, recall, F1 score and AUC score that are gained by the ANN model.

Merit	Score
Precision	95%
Recall	71%
F1 score	81%
AUC	98%

The figure 3 describe the ROC curve of the ANN model achieved in the training of CKD detection

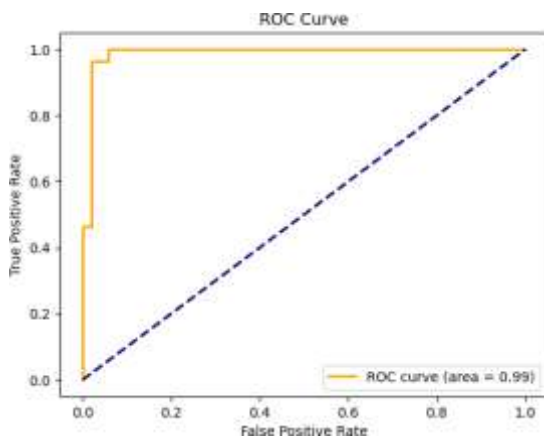


Figure 3. ROC Curve

The incremental growth in accuracy reflects the model's ability to learn and adapt to the training data, capturing more complex patterns and refining its predictive capabilities. This observed phenomenon underscores the effectiveness of iterative training in enhancing the ANN's predictive accuracy and highlights the importance of patience and persistence in achieving optimal model performance.

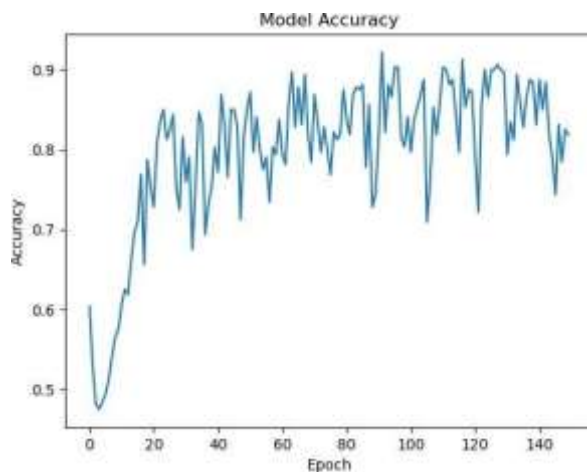


Figure 4. Accuracy of ANN over training

The registration and login system for the chronic kidney disease (CKD) detection platform facilitates user access and security. During registration, users provide necessary details like username, email, and password, which are validated and stored securely. Upon successful registration, users can log in using their credentials, undergoing authentication checks against stored data. Stringent security measures, such as password hashing and session management, safeguard user accounts. Account management features

allow users to update information and recover lost credentials, ensuring a seamless and secure user experience. The registration and login interface of the proposed system given in the figure 5 and 6 wisely.



Figure 5. User registration



Figure 6. User login

Upon the successful login the user will redirect into the CKD prediction page. The CKD detection input form enables users to provide medical data crucial for analysis. It features clearly labeled fields for parameters like age, blood pressure, and serum creatinine levels. Validation ensures accurate input, with error messages for invalid data. Users receive assistance through tooltips and instructions. The form prioritizes user experience, with intuitive design and accessibility. Once submitted, data is securely processed for CKD

analysis. The form for CKD detection is given in the figure 7.



Figure 7. CKD input form



Figure 9. Report Generation



Figure8. CKD prediction

Once after the prediction the report will be generated and the user can be downloading the report for their future reference. The generated report is shown in the figure 9.

6. CONCLUSION :

In summary, our research focused on developing a sturdy model for precise prediction renal illness that is chronic (CKD). Recognizing the shortcomings of conventional machine learning methods, we explored using feature selection and methods for deep learning to streamline complexity and enhance model performance. Our findings demonstrate that when compared to other classifiers, the Artificial Neural Network (ANN) fared better at detecting Chronic Kidney Disease. accuracy. Our efforts culminated in the successful deployment of the CKD detection system as a web application. This implementation effectively demonstrates our capability to determine CKD accurately, offering a vital instrument for early identification and intervention. It underscores the potential of advanced deep learning approaches and underscores the importance of leveraging technology to tackle critical healthcare challenges.

7. REFERENCES

- [1] Saif D et al. "Deep-kidney: an effective deep learning framework for chronic kidney disease prediction" Health Information System. (2023)
- [2] [Kumar, K et al. "Deep Learning Approach for Kidney Disease Recognition and Prediction through Image Processing. Appl." Sci. \(2023\)](#)
- [3] Hira Khalid et al. "Machine Learning Hybrid Model for the Prediction of Chronic Kidney Disease" Multitask Deep Learning and Semantic Knowledge in Intelligent Healthcare, Volume (2023)
- [4] Singh V et al. "A Deep Neural Network for Early Detection and Prediction of Chronic Kidney Disease". Diagnostics (Basel). (2022)
- [5] Sitote, T.M. et al. "Chronic kidney disease prediction using machine learning techniques." J Big Data 9, 109 (2022).
- [6] K. Jhumka et al. "Chronic Kidney Disease Prediction using Deep Neural Network," (2022)
- [7] Khadiime Jhumka et al. "Explainable Chronic Kidney Disease (CKD) Prediction using Deep Learning and Shapley Additive Explanations (SHAP)." (2024)
- [8] Tasiu et al. "Prediction of Chronic Kidney Disease Using Deep Neural Network". (2020).
- [9] Ma, Fuzhe et al. "Detection and diagnosis of chronic kidney disease using deep learning-based heterogeneous modified artificial neural network". (2020).
- [10] Chen, Guozhen et al. "Prediction of Chronic Kidney Disease Using Adaptive Hybridized Deep Convolutional Neural Network on the Internet of Medical Things Platform." (2020).