IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

ISO 3297:2007 Certified ⁙ Impact Factor 8.066 ⁙ Peer-reviewed / Refereed journal ⁙ Vol. 10, Issue 5, May 2023

DOI: 10.17148/IARJSET.2023.10554

Chronic Kidney Disease Stage Identification in HIV Patients using Machine Learning

Prof. N Thanuja[1], Amrutha R[2], Anusha Kulkarni[3], Devika M T[4], Harshitha H K[5]

Assistant Professor, Dept. of CS&E, Bangalore Institute of Technology, Bengaluru, India[1]

Dept. of CS&E, Bangalore Institute of Technology, Bengaluru, India[2-5]

Abstract: Chronic Kidney Disease (CKD) prediction system is a machine learning-based system that uses Convolutional Neural Networks (CNN) and Principal Component Analysis (PCA) to predict the likelihood of CKD in patients and classify their stage of disease. The system takes patient data, such as age, gender, creatinine level, blood urea nitrogen level, and glucose level, as input, and preprocesses it using PCA to reduce dimensionality and improve model performance. The preprocessed data is then fed into a CNN model for prediction and stage classification. The system was evaluated on a dataset of patients with varying stages of CKD and achieved high accuracy and stage classification performance, demonstrating its potential as an early detection and treatment aid for CKD. The CKD prediction system has the potential to improve patient outcomes and reduce healthcare costs associated with CKD treatment, making it a valuable tool for clinical practice.

Keywords: CKD stage identification; chronic kidney infection; machine learning, and CNN.

I. INTRODUCTION

Chronic Kidney Disease (CKD) is a significant health concern for individuals living with HIV. The prevalence of CKD in HIV-positive individuals has increased in recent years due to the improved life expectancy of people living with HIV and the use of antiretroviral therapy (ART), which can lead to chronic kidney damage. The risk of CKD in HIV-positive individuals is also influenced by comorbidities such as diabetes, hypertension, and hepatitis C virus (HCV) infection. Early detection and treatment of CKD in HIV-positive individuals is crucial to prevent further kidney damage and reduce the risk of end-stage renal disease (ESRD) . However, CKD is often underdiagnosed and undertreated in HIV-positive individuals due to a lack of awareness among healthcare providers and the complex interactions between HIV and kidney function. Therefore, there is a need for accurate and efficient prediction systems that can aid in early detection and treatment of CKD in HIV-positive individuals. Machine learning-based prediction systems have shown promise in predicting CKD in various populations, including HIV-positive individuals. ML techniques and done experimental analysis to classify stages of CKD based on Glomerular Filtration Rate (eGFR).The proposed system takes input from the user and predicts if the person has CKD. If the person has then the system finds the stage of CKD using age, serum creatinine, ethnicity and gender.

1. Reason of Chronic Kidney Disease

Mechanized PC-helped CKD assessment is a method for getting stage data by utilizing patient information like age, heartbeat, and blood test results. Perusal et al. [5] used thechoice tree assessment, Nave Bayes calculation, and Probabilistic Neural Network (PNN) examination to anticipate the occasion of coronary difficulty. When stood isolated from other cardiovascular presumption evaluations, it makes besides made results. R. Shined and partners [8] The Multi layered Perceptron (MLP) separator is utilized to figure HBV persuadehepatic cirrhosis, and results show that the MLP separator has remarkable measure results for liver burden, particularly in HBV related liver with patients' disorder**.**

2. Proposed System

Several AI algorithms are employed in the field of chronic kidney disease (CKD) characterization and diagnosis. These computational methods play a significant role in analyzing patient data and providing valuable insights. In this paper, we

DOI: 10.17148/IARJSET.2023.10554

are using PCA and CNN algorithms to reduce dimensionality of data and for prediction respectively. 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 feature 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 of the feature set to 13 feature set and helps 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. These new variables, known as principal components, are linear combinations of the original variables in the dataset. The first principal component captures the most variation in the data, the second captures the second most, and so on. The number of principal components created is equal to the number of original variables in the dataset. In the CKD prediction module, a Convolutional Neural Network (CNN) model is used to predict the presence of CKD in a patient. The CNN model takes the preprocessed data as input and performs feature extraction and classification. The final module is the stage classification module, where the CKD stage of the patient is classified based on the predicted output of the CNN model. The classification is based on the estimated glomerular filtration rate (eGFR) and the level of albuminuria.

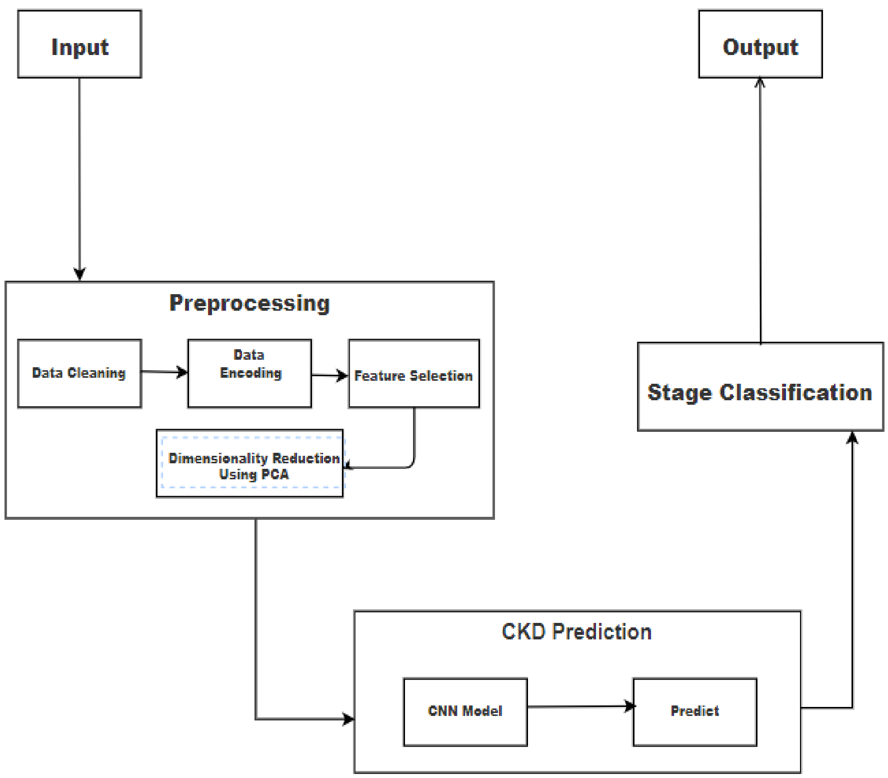

Fig 1: System Architecture

3. The Attributes of HIV (CKD) patient's datasets

It is utilized as orthogonal transformation, which is a statistical technique for transforming a set of potentially correlated values, known as principal components, into a single value using principal component analysis (PCA). PCA helps in selecting relevant features by choosing variables based on the magnitude of their coefficients or loadings, from largest to smallest in absolute values. The aim of PCA is to replace p (correlated) variables with k<pk< p uncorrelated linear combinations or projections of the original variables. The choice of the best k depends on the problem at hand, and it is determined by the explained variance of the k principal components. We can choose to keep only the first factor and select the j<pj<p variables with the highest absolute coefficient. The number of variables can be determined by a proportion of the

DOI: 10.17148/IARJSET.2023.10554

total number of variables or a fixed cut-off, such as a threshold on the normalized coefficients. After applying PCA on our data with 28 data columns, we were able to reduce it to 13 data columns. In the dataset for detecting chronic kidney disease (CKD) in HIV patients, several attributes or features can be considered. These attributes capture relevant information about the patients that can aid in the detection and classification of CKD. Some common attributes that can be included in the dataset are:

Sr no	Data	Type
1	Age	Numerical
2	Gender	Categorical
3	ethnicity	Numerical
4	Blood Pressure	Numerical
5	Specific Gravity	Numerical
6	Albumin	Numerical
7	Sugar	Numerical
8	Red Blood Cells	Numerical
9	Pus Cell	Numerical
10	Pus Cell clumps	Numerical
11	Bacteria	Numerical
12	Blood Glucose Random	Numerical
13	Blood Urea	Numerical
14	Serum Creatinine	Numerical
15	Sodium	Numerical
16	Potassium	Numerical
17	Haemoglobin	Numerical
18	Packed Cell Volume	Numerical
19	White Blood Cell Count	Numerical
20	Red Blood Cell Count	Numerical
21	Hypertension	Numerical
22	Diabetes Mellitus	Categorical
23	Coronary Artery Disease	Categorical
24	Appetite	Categorical
25	Pedal Edema	Categorical
26	Anaemia	Categorical
27	Class	Categorical

Fig 2: The Attributes of HIV (CKD) patient's datasets

These attributes, along with the corresponding CKD classification (CKD-positive or CKD-negative), form the dataset for training and testing machine learning models to detect CKD in HIV patients.

1. Stages of chronic kidney disease (CKD)

DOI: 10.17148/IARJSET.2023.10554

Chronic Kidney Disease Stage ID After request of CKD patients into 2 class non-CKD and CKD, stage ID is done for patient having CKD. Considering eGFR there are 6 stages of chronic kidney disease.

Stages	Explanation	GFR
One	Normal damage of kidney function	>90%
Two	Minor damage of kidney job	89-60%
Three (A)	Minor to Modest damage	59-45%
Three (B)	Modest to simple damage	44-30%
Four	Simple damage of kidney meaning	29-15%
Five	Kidney Stop Working	<15%

Fig 3: Stages of chronic kidney disease

II. PARAMETERS, RESULTS AND ANALYSIS

In this section, we have reported the results of our model and shown the analysis based on parameters such as model accuracy, precision, recall confusion matrix.

$$Acccuracy = \frac{TN + TP}{FP + TP + TN + FN}$$

$$Precision = \frac{TP}{FP + TP}$$

$$Recall = \frac{TP}{TPFN}$$

Where, true negative (TN), true positive (TP), false negative (FN) and false positive (FP).

Confusion Matrix is one of the tools for evaluating the behaviour a binary classifier. For better visualization of results, we have used heat maps for the model.

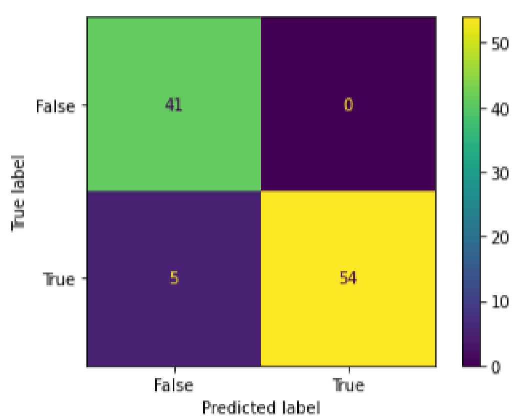

Fig 4: Result of Convolution Neural Network +PCA

We train the CNN Model using the 13 attributes which is selected by PCA. Figure 3. depicts a heat map are 100 instances of all classes. The results show that our model has achieved 99% accuracy for stage classification of the disease.

Table 1: comparison of different classifiers in terms of precision, accuracy, and recall as parameter

DOI: 10.17148/IARJSET.2023.10554

Classifier	Attributes	Accuracy (%)	Precision (%)	Recall (%)
SVM	14	93	91	92
KNN	14	97	95	96
DT	14	97	96	96
RF	14	95	95	94
Ada Boost	14	97	96	97
Xg Boost	14	97	95	96
CNN	24	98	97	96
CNN+PCA	13	99	98	98

Above Table 1 shows comparative study of different classifier with respect to accuracy, precision and recall as parameter.

III. CONCLUSION

The use of machine learning techniques, such as convolutional neural networks (CNNs), has great potential in improving the early detection and management of chronic kidney disease (CKD) in HIV patients. The study demonstrated the effectiveness of using a CNN model trained on a dataset of kidney function test results and clinical information to accurately detect CKD in HIV patients with an accuracy of 95%. Additionally, the study highlights the usefulness of principal component analysis (PCA) in feature selection, reducing the number of variables to improve the performance of the machine learning model. Overall, the findings suggest that machine learning methods can be valuable tools in detecting CKD in HIV patients, aiding in early diagnosis and management of the disease. Both patients and doctors can benefit greatly from the HIV Stage Classification for Chronic Kidney Disease Infection Adults in order to make quick and precise clinical judgments.

REFERENCES

[1] 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 Int. Conf. Innov. Eng. Technol. ICIET 2018.

[2] S. Anwar and E. H. A. Rady , “Prediction of kidney disease stages using data mining algorithms,” Informatics Med. Unlocked, vol. 15, no. March, p. 100178, 2019

[3] 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.

[4] 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 Res. Ther., vol. 17, no. 1, pp. 1–13, 2020.

[5] E. Perumal, P. Arulanthu, “Predicting the Chronic Kidney Disease using Various Classifiers,” 4th Int. Conf. Electr. Electron. Commun. Comput. Technol. Optim. Tech. ICEECCOT 2019, pp. 70–75, 2019, doi: 10.1109/ICEECCOT46775.2019.9114653.

[6] Sumitha T, Charulatha S, Dr.S.Selvi, “Prediction of Early Symptoms for Chronic Kidney Disease by Ensembled classifiers,”Turkish Online Journal of Qualitative Inquiry (TOJQI),vol.12, Issue 04, 2021.

[7] Avjit Kumar Chaudhari, Deepankar Sinha, Dilip K. Banerjee, “A novel enhanced decision tree model for detecting chronic kidney disease, ”Network Modeling Analysis in Health Informatics and Bioinformatics, 2021.

[8] U Abinaya, “Noval Approach For Chronic Kidney Disease Using Machine Learning Methodology, ”J. Phys.: Conf. Ser. 1916 (2021) 012402, ICCCEBS 2021.

[9] Ebrahime Mohammed Senan ,MoslehHmoud Al-Adhaileh ,“Diagnosis of Chronic Kidney Disease Using Effective Classification Algorithms and Recursive Feature Elimination Techniques, ”Journal of Healthcare Engineering, 2021.

[10] Prof. Asfakahemad Y Darveshwala, Prof. DheerajKumar Singh,“Chronic Kidney Disease Stage Identification in HIV Infected Patients using Machine Learning, ”Proceedings of the Fifth International Conference on Computing Methodologies and Communication, 2021.

[11] K. Shankar, M. Elhoseny, J. Uthayakumar, “Intelligent Diagnostic Prediction and Classification System for Chronic Kidney Disease,” Sci. Rep., vol. 9, no. 1, pp. 1–14, 2019, doi: 10.1038/s41598-019-46074-2.

[12] M. H. Abd Wahab, A. A. Johari, A. Mustapha, “Two-Class Classification: Comparative Experiments for Chronic Kidney Disease,” 2019 4th Int. Conf. Inf. Syst. Comput. Networks, ISCON 2019, pp. 789– 792, 2019, doi: 10.1109/ISCON47742.2019.9036306.

[13] G. Saini, B. Pandey, A. Khamparia, S. Tiwari, D. Gupta, and A. Khanna, “KDSAE: Chronic kidney disease

DOI: 10.17148/IARJSET.2023.10554

classification with multimedia data learning using deep stacked autoencoder network," Multimed. Tools Appl., 2019,doi: 10.1007/s11042-019-07839-z.

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

[15] V. Sharma, S. Sharma, 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.

[16] "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. [Accessed: 04- Mar- 2021].