PREDICTION OF CHRONIC KIDNEY DISEASE BY DEEP LEARNING METHOD

S. MADHIVATHANA[1], Dr.B.VANATHI [2], K. SHANMUGAM [3]

[1] *Student,Computer Science and Engineering, SRM Valliammai Engineering College,Tamilnadu*
[2] *Professor, Computer Science and Engineering, SRM Valliammai Engineering College,Tamilnadu*
[3] *Assistant Professor, Computer Science and Engineering, SRM Valliammai Engineering College,Tamilnadu*

ABSTRACT

Chronic kidney disease (CKD) is among the highest twenty causes of death worldwide and affects some 10 percent of the globe adult population. CKD could be a disorder that disrupts traditional urinary organs operation. glomerular Filtration Rate (GFR) and kidney damage markers are employed by researchers round the world to spot CKD as a condition that ends up in reduced urinary organ function over time. An individual with CKD encompasses a higher probability of dying young. Doctors face a tough task in identification the various illnesss connected to CKD at associate early stage so as to stop the disease. This project presents a completely unique deep learning model for the early detection and prediction of CKD. The objectives to make a deep neural network and compare its performance to that of other contemporary machine learning techniques. In tests, the typical of the associated options was accustomed to replace all missing values within the data set. After that, the neural network's optimum parameters were mounted by establishing the parameters and running multiple trials. The proposed Deep neural model outperformed the other 3 classifiers (Support Vector Machine (SVM), call Tree and Random Forest) by achieving smart accuracy. The planned approach can be a great tool for nephrologists in detecting CKD.

Keywords: *Chronic kidney Disease, Deep learning, Deep Neural Network.*

1. INTRODUCTION

Chronic Renal Disease may be a condition during which the excretory organ structure is uncommon or its operation reduces for three months and a lot of with a reduced glomerular filtration rate. Nowadays, (CKD) has been turned into a difficult public pathological condition with increasing prevalence globally, extremely burdening low-income countries, wherever detection, prevention and treatment rates stay low. CKD may be a substantial monetary burden on patients, attention services, and also the government. Treatments of the ESRD with urinary organ Replacement medical aid are either costly (hemodialysis and serosa dialysis) or advanced (transplantation). The association of CKD with other chronic diseases conjointly exacerbates matters. From the general public health perspective, it is thus imperative to be able to predict the trends in terms of CKD prevalence thus that timely selections will be taken by the decision-makers (ministries, insurers, hospital managers, etc.) to mitigate a possible surge within the range of cases. Such mitigation measures will embody increased population screening for CKD-related risks and awareness campaigns, because it has been incontestable that way changes (reducing weight, improving diet, increasing physical activity, reducing alcohol consumption, avoiding smoking, early referral to nephrologists, correct use of medication, and coverings to regulate alternative risk factors) area unit the foremost effective measures to combat the exacerbation of the condition with borderline associated prices.Currently, 850

million folks throughout the globe are doubtless to own excretory organ illness from various factors and therefore the report world kidney day of 2019 indicated that the illness become a significant disease from that a minimum of 2.4 million dies each year and currently it's the sixth fastest-growing reason behind death worldwide. The main objective of this research is to develop a Chronic Kidney Disease Prediction using Deep Learning technique which makes the people to take regular checking to know the status of the kidney and help them to diagnose at early stage. To review related works on chronic kidney disease and its prediction using Deep learning techniques. To collect chronic kidney disease patient history data and prepare the dataset. To identify the best feature selection method for chronic kidney disease prediction. To evaluate the performance of the chronic kidney disease prediction models and detect the chronic kidney disease based on the evaluation result. To implement a chronic kidney disease prediction prototype.

2. PROPOSED SYSTEM

CKD is that the Non-Communicable Disease (NCD) that shows increase in morbidity, mortality associated an admission rate of patients throughout the world. As early detection of this illness isn't potential, it's found solely at the later stages that ends up in an excretory organ Transplantation and Kidney Dialysis. The proposed system uses deep learning techniques to get high predcition when compared to Decision tree, Random Forest and SVM techniques. While using CNN it will generate high accuracy and prediction is good. The main advantage of DNN compared to its predecessors is that it automatically detects the important features without any human supervision. It is easy to understand and fast to implement. It has the highest accuracy among all algorithms that predicts. The proposed Multi-class design is employed to predict Chronic Kidney Disease. The planned technique takes the dataset as input then the dataset is pre-processed to form the planned technique complete. Dataset pre-processing method includes data cleanup, handling missing values, handling categorical data. once pre-processing the dataset is done then the feature selection is done to pick relevant options for the prediction of Chronic Kidney Disease. Before and once the applying of the feature choice ways, the model developed using Deep neural network deep learning algorithms. Evaluation metrices are used. Finally, the epitome which will take new input file to predict chronic kidney disease is developed using the selected prognosticative model.

There are several approaches to develop the disease prediction system. This study focuses on applying the Deep-learning classification algorithms to predict the risk level of disease as NOT-CKD, CKD and MILD that enables the experts to give appropriate treatment recommendations based on a dataset developed from locally collected patient history data from UCI Machine Learning Repository. The Multi class dataset is also used to identify the presence or absence of the disease.

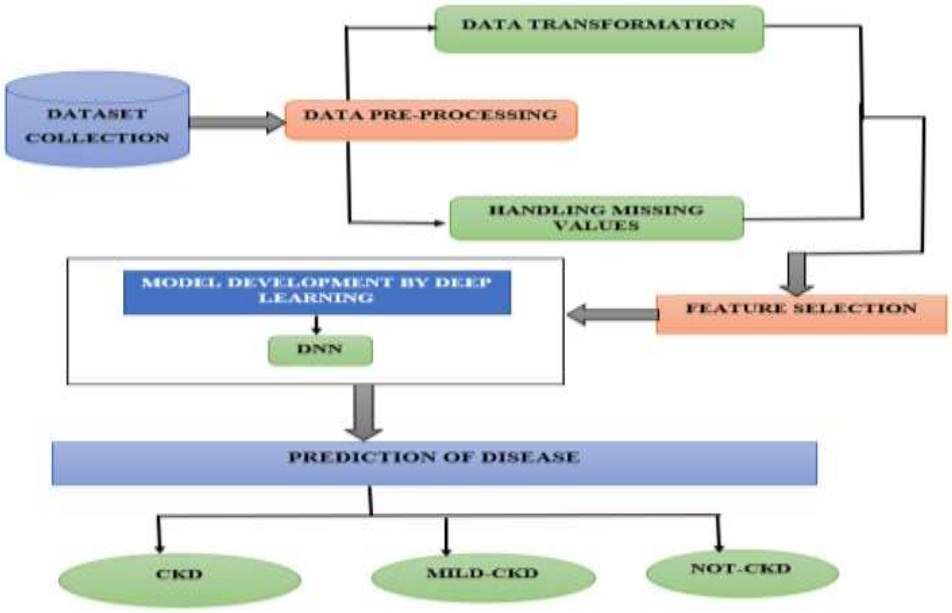

Fig -1: Architecture of Proposed System

3. PROPOSED SYSTEM

3.1 DATASET COLLECTION

The Chronic Kidney Disease dataset was collected from UCI Machine Learning Repository. The dataset contains twentyfive options as features that has been divided into fourteen numeric options and eleven categorical options, additionally to the category options of classification class, like “ckd” and “notckd” for classification. The collected data contains socio-demographic (Age, Gender) features, electrolyte (sodium, potassium, and chloride) features, laboratory test and urine analysis (serum creatinine, hemoglobin, blood urine nitrogen, mean cell volume, platelet count, red blood cell count, white blood cell count, blood pressure) features, risk factors (hypertension, diabetic mellitus, anemia, heart disease,) features, and dependent class (ckd_status). The diagnostic category contains 3 values: ckd ,mild and notckd. All options contained missing values apart from the diagnostic feature. The dataset is unbalanced as a result of it contains 380 cases of “ckd” , 225 cases of mild and 395 cases of “notckd”.

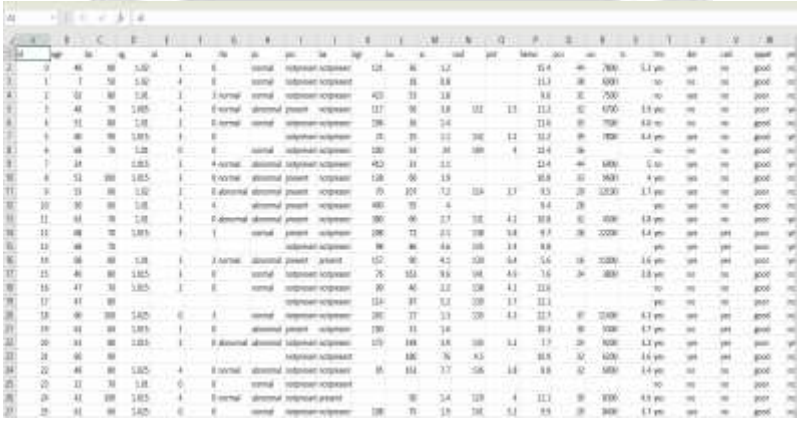

Fig -2: Raw Dataset

3.2 PRE-PROCESSING

The data cleanup step concerned dropping duplicates, missing and incorrect values. additionally, we have a tendency to solely enclosed patients of ages below 100.The estimation of missing values and therefore the removal of noise like outliers, as well as the normalisation and validation of unbalanced information, were all a part of the preprocessing stages. once assessing a patient, some measurements might be missing or incomplete. There are 158 completed cases within the information assortment, with the rest missing. Ignoring records is that the simplest technique to touch upon missing values; but, this is often not practical for tiny information sets. the information set is examined throughout the information preparation method to see whether or not any attribute values are missing. The missing values for numerical options were calculable exploitation the applied math technique of mean imputation. The mode technique was accustomed replace the missing values of nominal options. Because most machine learning algorithms solely settle for numeric values as input, category values should be encoded into numerical values. The binary values "0" and "1" are used to represent the characteristics of classes like "no" and "yes".

3.3 FEATURE SELECTION

Feature selection is that the next step of proposed method. The feature selection methodology performs the selection of necessary features from the initial dataset to enhance the performance of prognostic models. This action takes place when handling missing values and categorical information preprocessing step. Recursive Feature Elimination (RFE) removes options recursively, building a model based on different features. It applies greedy search to search out the foremost economical set of options. Use model accuracy to work out that features are most applicable for predicting a feature. It develops models iteratively, decisive the most effective or worst feature for each iteration. The traits are then classified supported the sequence during which they were removed. If the information set contains N functions, algorithmic feature elimination can importantly search for a mixture of 2N options within the worst-case state of affairs.

3.4 MODELS BUILDING

DEEP NEURAL NETWORK

DNN is a subset of Artificial Neural network which simulate the structure and functionalities of biological neural network consisting of an input, weights and activation function, the structure of DNN has an input, hidden layer and an output.In DNN, a is referred to the output, where Wi and Xi are the weight and input respectively. This is represented as: We used model has ten (10) inputs, two (2) hidden layers and one (1) output in the adopted model. Input layer is set of neurons which take the actual input data for processing. The number of neurons can be decided in accordance with the input data. To train the model, three layers were used. For input and the hidden layer, "Relu" is used as activation function. Relu outputs in 0 or 1. The output layer has only one output result either KD or NKD with Sigmoid as activation function. Stochastic Gradient was used as the optimizer of the model. The experiments are constructed with Python Programming Language.

4. INTEGRATING DEEP LEARNING MODEL WITH SERVER

The proposed system aims to predict chronic kidney disease using three models and recommend the best model to be used for deployment on the server after it is evaluated with the best accuracy. This is done by preparing the HTML form to insert the new input that is connected to the server and selected model that will allow the specialist to diagnose and give the best treatment and advice for the patient fast and correctly.

5. CONCLUSION

Early prediction is very crucial for both the experts and the patients to prevent and slow down the progress of chronic kidney disease to kidney failure. The main goal of this research is to use DNN model for the prediction of kidney disease to high degree of accuracy. Although the existing prediction by algorithms such as Support Vector Machine, Random Forest and Decision Tree, the possibility of identifying its accuracy to diagnose kidney disease, becomes very difficult. Hence, The adopted DNN model proved to be efficient and suitable for the prediction of kidney disease. Therefore, it is necessary to develop more works and research in this area in order to improve the identification of CKD. So that, it will be very useful in assisting the doctors to identify this illness early and providing the patient with possibility of recovering from the abnormal renal function. As future work, it is better to develop android app by using image deep learning as a classifiers and evaluating their performance turn the classification into a multi-class classification. It is better to work with real-time dataset to achieve higher accuracy.

6. REFERENCES

1. Arif-Ui-Islam, Shammin H Ripon, "Rule Induction and Prediction of Chronic Kidney Disease using Boosting Classifiers", IEEE-2019
2. Karthikeyen Harimoorthy and Menakadevi Thangavelu, "Multi-disease Prediction Model using Improved SVM-radial bias Technique in Healthcare Monitoring System", Journal of Ambient Intelligence and Humanized Computing-2021
3. Karna Vishnu Vardhana Reddy, Irraivan Elamvazuthi , Azrina Abd Aziz, Sivajothi Paramasivam , Hui Na Chua and S. Pranavanand, " Heart Disease Risk Prediction using Machine Learning Classifiers with attribute evaluators", MDPI -2021
4. X. Yu and T. Zhang, "Convergence and Runtime of an Ant Colony Optimization", Information Technology Journal 8(3) ISSN 1812-5638, 2009
5. Sahil Sharma, Vinod Sharma, Atul Sharma, "Performance based Evaluation of various Machine Learning Classification Techniques for Chronic Kidney Disease Diagnosis," July18, 2016
6. Amogh Babu A, Priyanka K, Raghavendra Babu T, "Chronic Kidney Disease Prediction based on naive bayes Technique", International Journal of Research in Engineering and Technology- 2019
7. Huseyin Polat, Homay Danaei Mehr , Aydin Cetin, "Diagnosis of Chronic Kidney Disease based on Support Vector Machine by Feature Selection Methods , National Library of Medicine-2019
8. Reshma S, Salma Shaji, S R Ajina , Vishnu Priya S R , Janisha A, "Chronic Kidney Disease Prediction using Machine Learning"- International Journal of Research in Engineering and Technology,2020
9. Hussein Abbass, "Classification Rule Discovery with Ant Colony Optimization", Research Gate Article, 2004
10. Mohammed Deriche, "Feature Selection using Ant Colony Optimization", International Multi-Conference on Systems, Signals and Devices, 2009
11. X. Yu and T. Zhang, "Convergence and Runtime of an Ant Colony Optimization", Information Technology Journal 8(3) ISSN 1812-5638, 2009