# Machine Learning Algorithm Incorporated Chronic Kidney Failure Prediction

## *G. Sandhiya[1], Dr. E. Ranjith[2]*

[1]Master of Computer Application, Krishnasamy College of Engineering &Technology, Cuddalore
[2]MCA., M.Phil.,Ph.D, Assistant Professor, Master of Computer Application, Krishnasamy College of Engineering &Technology, Cuddalore.

**ABSTRACT**

Kidney disease is considered as an important health problem that is currently affecting people of all age groups. To forecast the condition of chronic kidney disease, a large number of researches have been done using machine learning and deep learning techniques. The early identification of chronic kidney disease in HIV infected patients seems to be more fruitful as they will get appropriate treatment at early stage. Machine Learning uses various attributes such as blood pressure, diabetes test, time interval and quantity of urine output, history of kidney injuries, drug consumption and heart problems to predict the kidney failure. The values are stored as dataset and compared with trained dataset whose values are taken from respective hospitals. Thus, the prediction of kidney disease in HIV patients seems more effectively and their treatments can be started more earlier.

Keywords: Deep learning, kidney disease, HIV

## I. INTRODUCTION

Disorders of the kidney sometimes are of such a nature that they prevent the passing of wastes from the blood. The resulting accumulation of wastes, together with toxins that may be formed in the blood, can cause uraemic poisoning. Nephritis, bacteria-caused inflammation of the glomeruli, is often accompanied by the presence of blood in the urine. It results in elevated arterial blood pressure culminating not only in damage to the cardiovascular system but also in kidney and liver failure. When the kidneys do not remove sufficient water from the blood, the water accumulates in the tissues of the body, causing swelling known as oedema or dropsy.CKD is a condition in which the kidneys are damaged and cannot filter blood as well as they should. Because of this, excess fluid and waste from blood remain in the body and may cause other health problems, such as heart disease and stroke. CKD has varying levels of seriousness. It usually gets worse over time though treatment has been shown to slow progression. If left untreated, CKD can progress to kidney failure and early cardiovascular disease. When the kidneys stop working, dialysis or kidney transplant is needed for survival. Kidney failure treated with dialysis or kidney transplant is called end-stage renal disease (ESRD).

## II. RELATED WORKS

Yu et al. [1] closed SVM to see and anticipate diabetic's patients what's more pre diabetic patient .The results show that Support Vector Machine can see patients with average ailments. Perusal et al. [2] used the choice 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 [3] The Multi layered Perceptron (MLP) separator is utilized to figure HBV persuade hepatic cirrhosis, and results show that the MLP separator has remarkable measure results for liver burden, particularly in HBV related liver with patients' disorder.

## III. METHODOLOGY AND IMPLEMENTATION

Methodology of the proposed system includes identifying the various modules like data collection, data preparation, model selection, analyze and prediction etc. Figure 1 shows the block diagram of the proposed system. This paper aims to assist in the prediction of chronic kidney disease (CKD) by utilizing the support vector machine (SVM) classifier in medical domain.In this paper, we have explored ML techniques and done experimental analysis to classify stages of CKD. In this proposed system, we have built ML model using SVM to classify weather a patient has CKD or not. Before applying classification algorithm, we have eliminated few features using feature selection method.
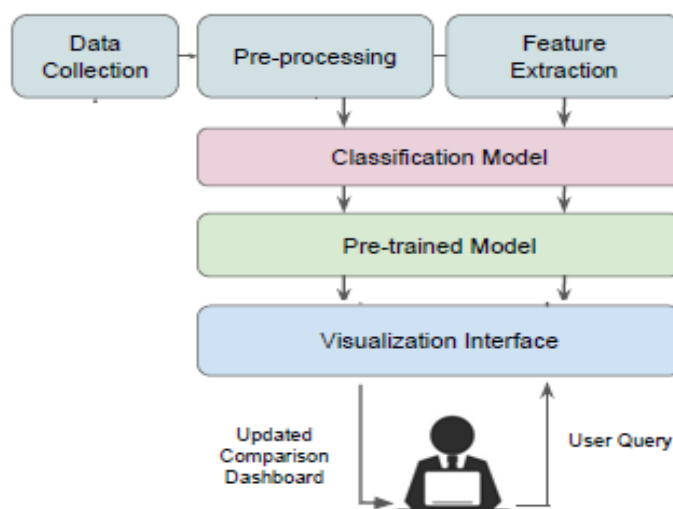
Figure 1 Proposed model

### Data Collection

This is the first real step towards the real development of a machine learning model, collecting data. This is a critical step that will cascade in how good the model will be, the more and better data that we get;the better our model will perform.There are several techniques to collect the data, like web scraping, manual interventions and etc.

The dataset used in this Cleveland Heart Disease dataset taken from the following link:

https://archive.ics.uci.edu/ml/datasets/chronic_kidney_disease

### Dataset

The dataset consists of 401 individual data. There are 25 columns in the dataset, which are described below.

### Data Preparation

Wrangle data and prepare it for training. Clean that which may require it (remove duplicates, correct errors, deal with missing values, normalization, data type conversions, etc.).Randomize data, which erases the effects of the particular order in which we collected and/or otherwise prepared our data Visualize data to help detect relevant relationships between variables or class imbalances (bias alert!), or perform other exploratory analysis Split into training and evaluation sets

### Model Selection

We used support vector machine algorithm, We got a accuracy of 0.9375 on test set so we implemented this algorithm.

Support Vector Machines (SVM) are learning systems that use a hypothesis space of linear functions in a high dimensional feature space, trained with a learning algorithm from optimization theory that implements a learning bias derived from statistical learning theory

Goal of the SVM is to find the optimal hyperplane that divides the two classes. There can be different planes that can divide the two classes, but the main focus is on to finding out such plane that we can achieve maximum margin between the classes. It means pick the hyperplane so that the distance from the hyperplane to the nearest data point is maximized

### Analyze and Prediction

In the actual dataset, we chose only 15 features :

15.dm - diabetes mellitus

### Accurcy on test set

We got a accuracy of 0.9375% on test set.

**Saving the Trained Model**

- Once you're confident enough to take your trained and tested model into the production-ready environment, the first step is to save it into a .h5 or .pkl file using a library like pickle.

- Make sure you have pickle installed in your environment.

- Next, let's import the module and dump the model into .pkl file.

## IV. RESULTS AND DISCUSSION

In this part discuss about different parameters definition and also showcase results. At the end comparative analysis is done withexisting methods.
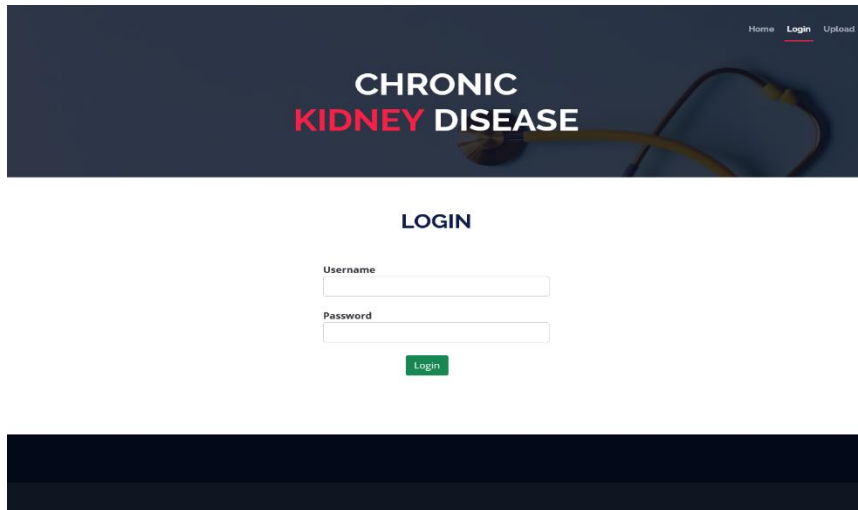


This is login page
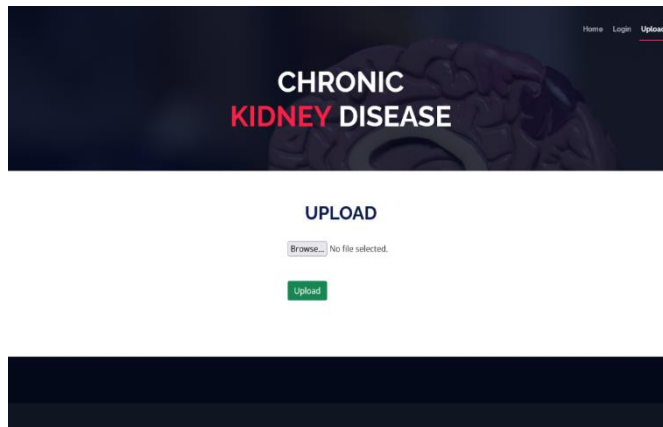


This is upload page where CKD dataset is upload



This is preview page where CKD data set to view

This is page show the prediction result

## V. CONCLUSION

Classification of Chronic Kidney diseases stage in HIV infected patient are extremely useful to patients as well as doctor for timely and accurate clinical decisions. In this paper we have implemented SVM for classification of CKD for patients having HIV. Our study indicates that SVM has outperformed in CKD classification. However, the performance of the CKD model using SVM depends heavily on the quality of the training data, the selection of features, and the tuning of hyper parameters.

## REFERENCES

[1] Yu , R. Naseem, F. Muhammad, G. Abbas, and S. Kim, ''An empirical evaluation of machine learning techniques for chronic kidney disease prophecy,'' IEEE Access, vol. 8, pp. 55012–55022, 2020.

[2] R. Kanduri, T. Bathini, A. Chewcharat, N. Leeaphorn, M. L. Gonzalez-Suarez, and W. Cheungpasitporn,''Promises of big data and artificial intelligence in nephrology and transplantation,'' J. Clin. Med., vol. 9, no. 4, p. 1107, 2020, doi: 10.3390/jcm9041107.

[3] A. Ogunleye and Q.-G. Wang, ''XGBoost model for chronic kidney disease diagnosis,'' IEEE/ACM Trans.Comput. Biol. Bioinf., vol. 17, no. 6, pp. 2131–2140, Nov. 2020

[4] 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, pp. 1–6, 2019, doi: 10.1109/CIET.2018.8660844.

[5] S. D. Sudarsan,N. Chetty, K. S. Vaisla, "Role of attributes selection in classification of Chronic Kidney Disease patients," pp. 1–6, 2016, doi: 10.1109/cccs.2015.7374193.

[6] 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.

[7] K. Shankar, G. Devika, P. Manickam, M. Ilayaraja, "Optimal Feature Selection for Chronic Kidney Disease Classification using Deep Learning Classifier," 2018 IEEE Int. Conf. Comput. Intell. Comput. Res. ICCIC 2018, pp.1–5, 2018, doi: 10.1109/ICCIC.2018.8782340.

[8] R. Shinde, S. John, R. Jadhav, A. Maurya, R. Wable, and R. Dakshayani, "Chronic Kidney Disease Prediction and Recommendation of Suitable Diet Plan by using Machine Learning," 2019 Int. Conf. Nascent Technol. Eng. ICNTE 2019 - Proc., no. Icnte, pp. 1–4, 2019, doi: 10.1109/ICNTE44896.2019.8946029.

[9] S. C. Jat . R. Yadav, "Feature selection and dimensionality reduction methods for chronic disease prediction," Int. J.Sci. Technol. Res., vol. 9, no. 4, pp. 2912–2918, 2020.

[10] S. saraswat, S. Vashisth, I. dhall, CKD diagnosis using Multilayer Perceptron Classifier' 10th International Conference on Cloud Computing, Data Science & Engineering (Confluence)" 2020