**Early Detection of Chronic Kidney Disease using Machine Learning**

**ABSTRACT:**

Chronic kidney disease (CKD) is a widespread disease worldwide that causes the majority of lives. CKD is the 11th leading cause of death worldwide, with 1.2 million deaths each year, and according to the Kidney Foundation in Bangladesh, about 40,000 people with CKD experience kidney failure each year, and several thousand die in the short stage of life due to CKD.

Predictive analytics for healthcare using machine learning is a challenging task to help doctors make accurate treatment decisions to save lives.

Together, the researchers researched chronic kidney disease, with most of their work being on purely statistical models, which created numerous gaps in the development of machine learning models.

In this paper, we discussed the current methods and proposed an improved technology based on XGBoost (Extreme Gradient Boost), which combined the significant characteristics of the F-score and evaluated four pre-processing scenarios.

In addition, we provided machine training methods for predicting chronic kidney disease with clinical information. Four master learning techniques including Support Vector Regressor (SVR), Logistic Regressor (LR), AdaBoost, Gradient Boosting Tree, and Decision Tree Regressor are explored.

Components are constructed from the UCI CKD dataset and the results of these models are compared to determine the best regression model for prediction.

**INTRODUCTION:**

**About Project:**

Chronic Kidney Disease refers to the kidneys' inability to fulfill their normal blood filtration role and other functions (CKD). The term "chronic" refers to progressive deterioration of kidney cells over time.

A kind of artificial intelligence is machine learning (ML) (AI). Its heart is algorithmic procedures, which allow the machine to solve issues without the need for specialist computer programming.

The widespread use of ML in the medical industry promotes medical innovation, lowers medical expenses, and improves medical quality. However, further research on using ML to solve clinical problems in nephrology is needed.

Hence, the prediction and diagnosis of CKD in its early stages is quite essential, it may be able to enable patients to receive timely treatment to ameliorate the progression of the disease.

Machine learning refers to a computer program, which calculates and deduces the information related to the task and obtains the characteristics of the corresponding pattern . This technology can achieve accurate and economical diagnoses of diseases hence, it might be a promising method for diagnosing CKD.

It has become a new kind of medical tool with the development of information technology  and has a broad application prospect because of the rapid development of electronic health record . In the medical field, machine learning has already been used to detect human body status , analyze the relevant factors of the disease  and diagnose various diseases.

For example, the models built by machine learning algorithms were used to diagnose heart disease , diabetes and retinopathy , acute kidney injury , cancer  and other diseases .

**Features**:

The widespread use of ML in the medical industry promotes medical innovation, lowers medical expenses, and improves medical quality. However, further research on using ML to solve clinical problems in nephrology is needed.

**METHODOLOGY ADOPTED:**

The current system of diagnosis is based on urine examination using the serum creatinine level. Many medical methods are used for this purpose, such as screening, ultrasound method.

During the screening, patients with hypertension, cardiovascular disease in the anamnesis, diseases in the past and patients who have relatives with kidney disease are examined.

This technique involves calculating an estimated GFR from the serum creatinine level and measuring the urinary albumin-to-creatinine ratio (ACR) in the first morning urine sample.

This paper focuses on machine learning techniques such as ACO and SVM by minimizing features and selecting the best features to improve prediction accuracy.

**PROPOSED METHOD:**

A proposed framework for developing a prediction enginelearning models and their comparison .The main goal of current research is to design a machinelearning techniques to predict CKD using associative andclassification algorithms.

The proposed technique generatesclassification association rules (CARs) to determine techniqueswith a high percentage of correctly classified cases andidentified classifiers may facilitate early diagnosis of CKD and a comparative analysis of the proposed technique is performed. using other state-of-the-art techniques.It briefly describesdifferent stages:

(i). Data set selection phase:

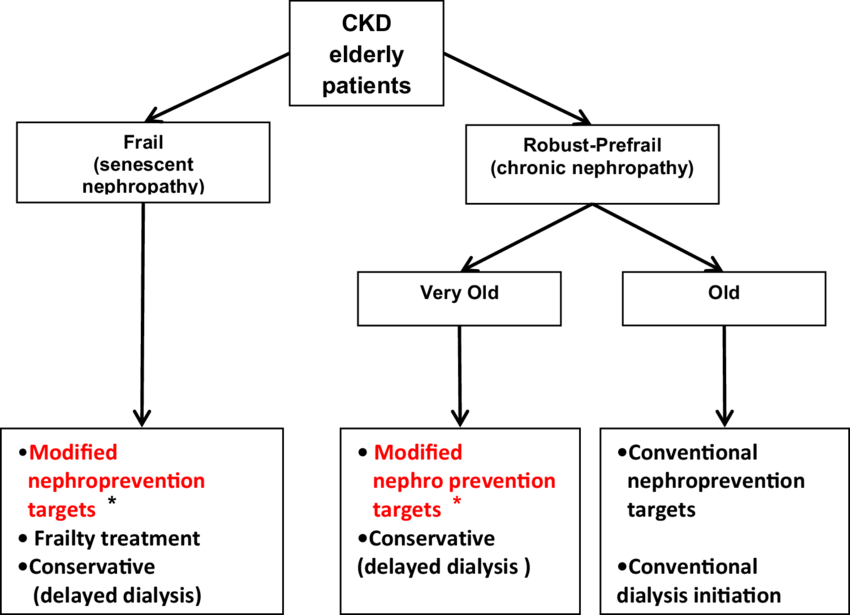
The data set is selectedpredict CKD for data analysis and effective knowledge.Enough data is needed to implement the machinelearning technique for the selected data set. In this setexperiments, CKD data are obtained from UCImachine learning repository.

(ii). Pre-processing and transformation phase:

Data set is prepared in file format with attribute 16attributes. The data set is converted to binomial format implement associative techniques. Moreover, it is missing records, duplicate records and unnecessary fieldsremoved for standard data format.

(iii). Feature Selection Phase:

The most promising featureof the CKD dataset are selected using the WEKA pro toolbetter results. Feature evaluators and search methods areused for this purpose. A function based on correlationthe selection subset evaluator is used as function evaluator,and a greedy stepwise search method is used.



CONCLUSIONS:

This article deals with the early prediction of CKD in humans. The envelope method used here for feature selection is ACO. ACO is a meta-heuristic optimization algorithm. Out of the 24 attributes present, the 12 best attributes are taken for prediction. The prediction is done using a machine learning technique, SVM. In this classification problem, SVM classifies the output into two classes with CKD and without CKD. The main objective of this study was to predict patients with CKD using fewer attributes while maintaining higher accuracy.

REFERENCES:

1. Qin, J., Chen, L., Liu, Y., Liu, C., Feng, C., & Chen, B. (2019). A machine learning methodology for diagnosing chronic kidney disease. *IEEE Access,* *8,* 20991–21002.
2. Alloghani, M., Al-Jumeily, D., Baker, T., Hussain, A., Mustafina, J., & Aljaaf, A. J. (2018, August). Applications of machine learning techniques for software engineering learning and early prediction of students’ performance. In *International Conference on Soft Computing in Data Science* (pp. 246–258). Springer, Singapore.
3. Park, N., Kang, E., Park, M., Lee, H., Kang, H. G., Yoon, H. J., & Kang, U. (2018). Predicting acute kidney injury in cancer patients using heterogeneous and irregular data. *PLoS ONE,* *13*(7), e0199839.
4. Alickovic, E., & Subasi, A. (2016). Medical decision support system for diagnosis of heart arrhythmia using DWT and random forests classifier. *Journal of Medical Systems,* *40*(4), 108.
5. Masetic, Z., & Subasi, A. (2016). Congestive heart failure detection using random forest classifier. *Computer Methods and Programs in Biomedicine,* *130,* 54–64.
6. Arasu, D., & Thirumalaiselvi, R. (2017). Review of chronic kidney disease based on data mining techniques. *International Journal of Applied Engineering Research,* *12*(23), 13498–13505.