NALAIYA THIRAN

(Professional Readiness for Innovation, Employability and Entrepreneurship)

LITERATURE SURVEY 2022-2023

Team Id: PNT2022TMID38667

Team Title: Early Detection of Chronic Kidney Disease using Machine Learning

Team Leader : Sahana R

Members List: Kanimozhi D, Keerthana V, Tamilarasi S, Bhavani G

INTRODUCTION		SURVEY/BODY OF REVIEW					Conclusion		
Year	Title	Keywords	Problem Definition	Methodology (Algorithm, Protocol…Etc)	Input Parameters	Result	Advantages	Disadvantages/Drawbacks	Research Gap/Research Question
Sahana R									
1. 2022	Chronic Kidney Disease diagnosis using Decision Tree algorithm	Decision Tree, Machine learning, CKD, Data Mining	A reduction in the kidney disease burden. Longer lives and improved quality of life for people with CKD.	Collect the datasets like age, sex, race and serum. Then, calculate the GFR rate and categorize the stages based on the result.	Age, BP, Glucose, Albumin, GFR.	Predicts the stage of Chronic Kidney Disease	*Produces 80% accurate result. *Handle both numerical and categorical data.	*Slow and ineffective for real time prediction	J48-Decision tree model is not sufficient for huge data sets. It is in effective for real time prediction
2. 2022	Early Detection of Chronic Kidney Disease Using Advanced Machine Learning Models	Chronic Kidney Disease (CKD), Machine Learning (ML), Support Vector Machine (SVR), Random Forest (LR), Artificial Neural Network (ANN),	To avoid early death cases need to predict early stage disease using data analytics methodologies	*CKD Dataset *Data Preprocessing *Training Models *Prediction	Dataset like age, serum, albumin, creatinine which are obtained from patients	Precision, Accuracy, F1 score, and Recall	*Gives accurate prediction	*More time required for execution	The traditional algorithms were used here. It gets more duration to category and classifies the give data set

		Decision Tree (DT).							
3. 2022	A Deep Neural network for Early detection and Prediction of Chronic kidney disease	Regressive feature, elimination, support vector machine, machine learning	The proposed approach should be a useful tool for nephrologists in detecting CKD	*Data processing *Categorical data encoding *Data transformation *Outlier detection	Serum, Race, Historical data, genetic problems	The study concluded that it is more efficient in detecting CKD	*Execute feature engineering by itself	*It can handle only small datasets	To improve the model performance, significant volumes of increasingly sophisticated and representative CKD data will be collected in the future to detect disease severity
4. 2022	Early identification of CKD- a scoping review of the global publications	Chronic Kidney Disease (CKD), Machine Learning (ML), Support Vector Machine	Decisions on whether to screen for chronic kidney disease (CKD) or not remain contentious in nephrology. This study provides a global overview of early CKD identification efforts.	Data extracted from included studies focused on the following 4 themes: study population measurement methods, interventions used, and available policies.	Gender, bp, cholesterol, pulse rate	We identified 290 CKD screening and detection programs from 83 countries.	*Time saving process	This study has some limitations, including inability to use World Bank income grouping at time of the study	To explore the effect of ethnicity because most studies were not performed in homogeneous populations or did not report the racial composition of participants
5. 2022	Chronic Kidney Disease Predictionusing K-Means algorithm	Chronic kidney disease, k-means, Logistic regression Support vector machine	Elimination of disparities among kidney disease patients.	The best measure of kidney function is the glomerular filtration rate (GFR). To measure GFR is complica- ted in clinical practice, requiring substantial time and resources	*Albuminaria *GFR *Sodium	Stage 1-Normal Stage-2 Slightly damaged kidney Stage-3 Fully damaged kidney	1.Scales to large data sets. 2.Guarantees convergence.	1.Scaling with number of dimensions. 2.Clustering data of varying sizes and density.	To beat the performance of other classifiers using normalized dataset
Tamilarasi S									
6. 2021	Diagnose of Chronic Kidney Disease by using	Naïve Bayes, Random Forest, eGFR, CKD,	The relationship of CKD to chronic kidney	Scan The Dataset, Calculate the	Blood pressure, sugar, bacteria,	Helps medical predicting the	*No need for anticoagula	*Protein loss through dailysate	It has to be enhanced based on the ground truth recommended

	Naïve Bayes Algorithm	Medical Diagnosis	diseases such as diabetes. Among patients with chronic kidney diseases	Probability of each attribute values, Apply the Formulae, Multiply the probabilities value of p,Compare the values.	appetite, red blood cells, hemoglobin	CKD disease based on the CKD parameter.	tion treatment *Reduced risk of hypertensio n *Early diagnosis of infections	*Increased risk of peritonities	by leading medical practitioners to ensure that our application worked as intended.
7. 2021	Diagnosis of Chronic Kidney Disease using effective classification algorithms and recursive feature elimination techniques	CKD, Classification, Recursive, Elimination, GFR	These techniques are supportive for experts and doctors in early diagnosis to avoid developing kidney failure	*Dataset of ckd *Preprocessing *Feature selection *Evaluation of ckd	Age, diabetis, obesity, high blood pressure, stroke	It was followed by the decision tree algorithm, which reached the accuracy, precision, recall and F1-score.	*Easy to configure and use	*Costlier to maintain	Need to select the most strongly representative features of CKD
8. 2021	Detection and evaluation of Chronic Kidney disease using different regeression and classification algorithm in machine learning	Kidney disease, KNN, Logistic regression	We have tried to reduce the clinical effect by automating theprocess of detection	*Dataset attributes were collected *Classification algorithms was applied *Based on that results were predicted	Anemia, RBC count, WBC count, serum, creatinine	The preprocessing of the dataset was the major part which leads to the good accuracy of our model	*Probabilis tic approach *Simple to understand	*Not the best choice for large number of features *Complex	Need to work on huge real life dataset and can develop a healthcare system prototype for CKD
9. 2020	Diagnosis of Chronic Kidney Disease based on Support Vector Machine by feature selection methods	Feature selection . Support vector machine . Chronic kidney disease . Machine learning	Convincing evidence that CKD can be detected using simple laboratory tests.	Wrapper and filter methods based on best first and greedy stepwise were developed to evaluate the feature selection methods and the accuracy of classification. algorithms	Kernel parameters determine the distance between patterns into the new space, dimensions of the new space and the complexity of the classification	High accuracy rate in the diagnosis of chronic kidney disease compared to other methods.	*Effective *Memory efficient	Svm algorithm is not suitable for large data sets. Svm does not perform very well when the data set has more noise target classes are	Need to prepare the dataset with the lowest dimension by feature selection

					model			overlapping.	
10. 2020	Detailed review of chronic kidney disease	Chronic kidney disease, Glomerular filtration rate, Albumin creatinine rate, Acute kidney injury, Cystatin C	With the help of various engineering techniques one can easily design controllers to assess as well as to prevent CKD permanently.	*Collect the dataset *Pre-process it *Apply machine learning algorithm *Predict the stage of CKD	RBC, WBC count, age, bp, sugar content	With the help of engineering techniques we can design predictive control systems for assessing as well as preventing CKD	*Simple to implement *More efficient	*More resources required	More accurate data mining techniques has to be implemented for better prediction
Bhavani G									
11. 2020	The case for early identification and intervention of Chronic Kidney disease	Chronic kidney disease, Glomerular filtration rate, Albumin, creatinine rate	Participants identified strategies for screening, risk stratification, and treatment for early CKD and the key health system and economic factors for implementing these processes	*Input dataset *CKD Screening *Treatment	GFR, blood sample, urine sample	A fundamental justification for the early detection of CKD is the availability of evidence-based interventions to slow the progression of CKD and reduce its complications	*Easy to carry out *More throughput	*Need to work on additional features	Need to support systematic approaches for CKD screening
12. 2020	Risk Level Prediction of Chronic Kidney Disease using Neuro-Fuzzy and Hierarchical Clustering algorithm	Heatmap, clustering multivariant statistical analysis	The treatment can prevent or delay complications of decreased kidney function	A fusion of neural networks with fuzzy logic is generally defined as a system trained using a particular learning algorithm which is derived from neural network	BP, Glucose, Urine protein, urea	The results of the prediction showing the risk of any patient having CKD given the ten features.	*Data mining in healthcare detects fraud and abuse. *Help physicians to identify effective treatments	*Ethical, Legal and Social Issues. *Data Ownership issues.	Implement more bigdata oriented tools and technique which makes the process faster and effective

				foundations			and best practices.		
13. 2019	Machine Learning approach of Chronic Kidney Disease prediction using clustering techniques	Clustering, data mining, machine learning, fuzzy c means	Among patients with chronic diseases there is an increased risk of complication related to those diseases.	*Cluster analysis *K-means algorithm *K-medoids algorithm *Validity indices	Age, GFR, BP, Sugar	The experimental result shows the performance of the fuzzy c means algorithm was improved significantly	*Allows gradual membershi p of data points	*More number of iteration is needed	Must increase the performance of the clustering techniques
14. 2019	Prediction of kidney disease stages using data mining algorithms	Prediction of kidney disease stages Data mining techniques Probabilistic neural networks Multilayer perceptron Support vector machine Radial basis function	Use of efficient data mining techniques is shown to reveal and extract hidden information from clinical and laboratory patient data	*Find the dataset *Compute GFR *Apply data mining algorithms *Compare the result	Albumin, Sodium, Potassium, Chloride, pulse rate	The experimental comparison of the utilized algorithms was done based on the performance measures of classification accuracy and execution time	*Detect risks and fraud *Analyse large quantity of data	*It requires large database *Expensive	Need to implement best algorithm in order to eliminate diagnostic and treatment errors
15. 2019	Early detection of Chronic Kidney disease in Australia	Albuminuria, chronic kidney disease (CKD), glomerular filtration rate (GFR), proteinuria	The principles underlying and justifying the establishment of a screening programme for CKD.	*KNN *Naïve-Bayes *Support vector machine *Logistic regression	Risk factors like GFR, BP, Urine protein etc	The benefit from such a programe lies in reducing deaths from heart attacks and kidney failure and in reducing the number of people needing dialysis ortransplantat ion.	*Includes all the features	*Costly	Because of the cost issues, screening programmes should not be compromised

				Kanimozhi D					
16. 2019	Chronic Kidney Disease Prediction Using Machine Learning Techniques	Chronic kidney disease Decision Tree ·Support Vector Machine Logistic Regression and Bagging Ensemble methods	To develop and validate a predictive model for the prediction of chronic kidney disease.	*CKD dataset *Apply logistic regression and decision tree algorithm *Bagging ensemble technique *Result	Glucose, BP, Albumin, Creatinine	The machine learning classifiers used in two stages, at first the logistic regression classifier is used to predict the results	*Easy to implement *Accurate	*Not efficient	The model can be further turned by applying feature selection methods to increase the performance of the prediction
17. 2019	Chronic Kidney Disease Prediction using Machine Learning Models	Chronic Kidney Disease, Decision Tree, Machine Learning, Random Forest, Support Vactors.	Presents evidences of early identification and care of CKD can improve the quality of the patients life.	*Create dataset *Pre-processing *Apply decision and support vector machine.	Glucose, BP, RBC, WBC, Albumin	Models has been constructed using training data set instances which is 70% of original CKD data set. Constructed models have been validated using test data.	*Efficient functioning *Simple to implement	*Not compared the data at the time of execution	The comparison must be done based on the time of execution and feature set selection
18. 2019	Detection of Chronic Kidney Disease Using Machine Learning Algorithms with Least Number of Predictors	Chronic kidney disease, Random forest, Gradient boosting, Logistic Regression, Support vector machines, Machine learning.	The target is to diagnose the CKD using different intelligent techniques.	*Data preprocessing *Missing values *Data reduction *Modelling *Result	Age, sex, serum, creatinine, albumin	The experiments are conducted using Python 3.3 programming language through the Jupyter Notebook	*Effective method	*Disassociation between classes may occur	It is aimed to validate the results by using bigdata sets or compare the result using another dataset that contains the same features.
19. 2019	Using machine learning models to predict the initiation of renal	A novel approach of screening CKD patients to predict the chances of future RRT based	The objective of our study is to develop a screening tool based on disease	*Study design *Data source *Petient selection *Outcome definition	Age, serum, urea, creatinine, albumin, gfr	They, tested the effect of all the implemented data	*Easy data aquisation	*High error prone *Time consuming	Future scope lies in coming up with a prediction model that would factor in the more clinical data in

	replacement therapy among chronic kidney disease patients	on the clinical data using ML algorithms	history using various ML models.			preprocessing approaches on the results of ML algorithms.			predicting the outcomes.
20. 2018	Neural network and support vector machine for the prediction of chronic kidney disease: A comparative study	Machine learning Artificial Neural Network (ANN) Support Vector Machine (SVM) Chronic Kidney Disease (CKD)	Apply different machine learning classification algorithms to a dataset of 400 patients and 24 attributes related to diagnosis of chronic kidney disease.	Models of the two proposed techniques were developed using the best-obtained parameters and features. The empirical results from the experiments indicated that ANN performed better than SVM.	Age, rbc,wbc count, creatinine, gfr	The empirical results from the experiments indicated that ANN performed better than SVM.	*Good fault tolerance *Distributed memory	*Does not aimed on selecting important features	Need to obtain the results using most important parameter and features
Keerthana V									
21. 2018	Prediction of Chronic Kidney Disease Using Machine Learning Algorithm	CKD, Decision Tree, GFR, SVM, Machine Learning	To build a model with maximum accuracy of predicting whether CKD or not and if yes then its Severity.	*Data preprocessing *Cleanm dataset *Train and test the data *Prediction	Age, Blood Pressure, Albumin, Red Blood cells, Pus cell, Serum creatinine, Hemoglobin.	To build a machine learning model targeting chronic kidney disease with overall accuracy of 99.99%, will need millions of records with zero missing values.	*Prediction process is less time consuming.	*Less security provided for the data	Strength of the data need to be improved
22. 2018	A Deep Neural Network for Early Detection and Prediction of Chronic Kidney Disease	Chronic kidney disease; feature selection; recursive feature elimination	Chronic kidney failure makes to difficulties in removing extra fluids from the body blood.	*Data set description *Data processing *Handling missing value *Data	BP, hypertension, urea, sodium	To improve the model performance, significant volumes of increasingly	*Can work with insufficient knowledge	*Large datasets need to be implemented.	Significant volumes of increasingly sophisticated and representative CKD data will be collected in future to detect.

				transformation		sophisticated			
23. 2018	A Comprehensive Unsupervised Framework for Chronic Kidney Disease Prediction	Chronic kidney disease, unsupervised learning techniques, autoencoder.	To build an intelligent machine learning model that can be used reliably to establish CKD diagnosis.	*Import the data, *Impute the missing values, *Encode texts to numerical values. *Predict the CKD stages	Specific gravity, blood urea, serum creatine, pedal edema	It has been implemented for better for DB scan than for K-mean	*More variety in work, *Less healthcare costs	*More time to execute	Reduce the time and financial expenses of CKD diagnosis
24. 2018	Comprehensive Performance Assessment of Deep Learning Models in Early Prediction and Risk Identification of Chronic Kidney Disease	Artificial neural network, chronic kidney disease, classification, deep learning.	The risk of CKD is increasing rapidly, and consequently more people are suffering and dying due to a lack of proper treatment.	*Data retrieval, description *Data preprocessing, *Feature selection *Model description	Packed cell volume, diabetes mellitus, coronary artery disease, pus cell	These approaches can potentially reduce the cost and treatment of CKD and researchers can get efficient results	*Enhances better quality of life, better health, no diet and fluid intake restriction	*Unable to identify the crucial factors	Identify risk factors that are crucial for early diagnosis
25. 2018	Detection of Chronic Kidney Disease Using Machine Learning Algorithms with Least Number of Predictors	Chronic kidney disease (CKD), Random forest (RF), Gradient boosting (GB), Logistic Regression (LR), Support vector machines (SVM), Machine learning (ML), prediction.	Chronic kidney disease (CKD) is one of the most critical health problems due to its increasing prevalence.	*Data preprocessing, *Missing Values *Data Reduction *Data transformation	BP, Glucose, Pulse rate, Sodium, Potassium	The nested cross-validation approach also has been applied for the purpose of tuning the models' parameters. The experiments are conducted using Python 3.3 programming language.	*Better prediction of stages	*Not suitable larger datasets	Validate the results by using bigdata set