IAR Journal of Medical Sciences
**ISSN Online:** 2708-3594
**Frequency:** Monthly
**Language:** English
**Origin :** Kenya
**Website :** https://www.iarconsortium.org/journal-info/iarjms

OPEN ACCESS JOURNALS

Research Article

# Chronic Kidney Disease Detection Using Machine Learning Techniques

**Article History**
**Received:** 20.03.2021
**Revision:** 30.03.2021
**Accepted:** 10.04.2021
**Published:** 20.04.2021

**Author Details**
N. Vanitha[1] & S.V. Sendhuraa[2]

**Authors Affiliations**
[1]Assistant Professor, Department of Information Technology ,Dr. N. G. P. Arts and Science College, Tamil Nadu, India

[2]Student, Department of Information Technology, Dr. N. G. P. Arts and Science College, Tamil Nadu, India

**Corresponding Author**
N. Vanitha

**How to Cite the Article**
N. Vanitha & S.V. Sendhuraa (2021). Chronic Kidney Disease Detection Using Machine Learning Techniques. *IAR J Med Sci*, *2*(2), 127-133



**Abstract:** As of now, there are numerous individuals on the planet experiencing persistent kidney infections around the world. Because of the a few danger factors like food, climate and expectations for everyday comforts numerous individuals get infections out of nowhere without comprehension of their condition. Diagnosing of persistent kidney infections is for the most part obtrusive, expensive, tedious and frequently hazardous. That is the reason numerous patients arrive at late phases of it without treatment, particularly in those nations where the assets are restricted. Hence, the early recognition procedure of the illness remains significant, especially in non-industrial nations, where the sicknesses are for the most part analyzed in late stages. Finding an answer for previously mentioned issues and braving from impediments turned into a solid intention to lead this examination. In this examination study, the impacts of vector machines calculation is explored. The ongoing kidney infection dataset depends on clinical history, actual assessments, what's more, research facility tests. Exploratory outcomes appeared more than 93% of achievement rate in arranging the patients with kidney infections based on three execution measurements i.e., precision, affectability and explicitness.

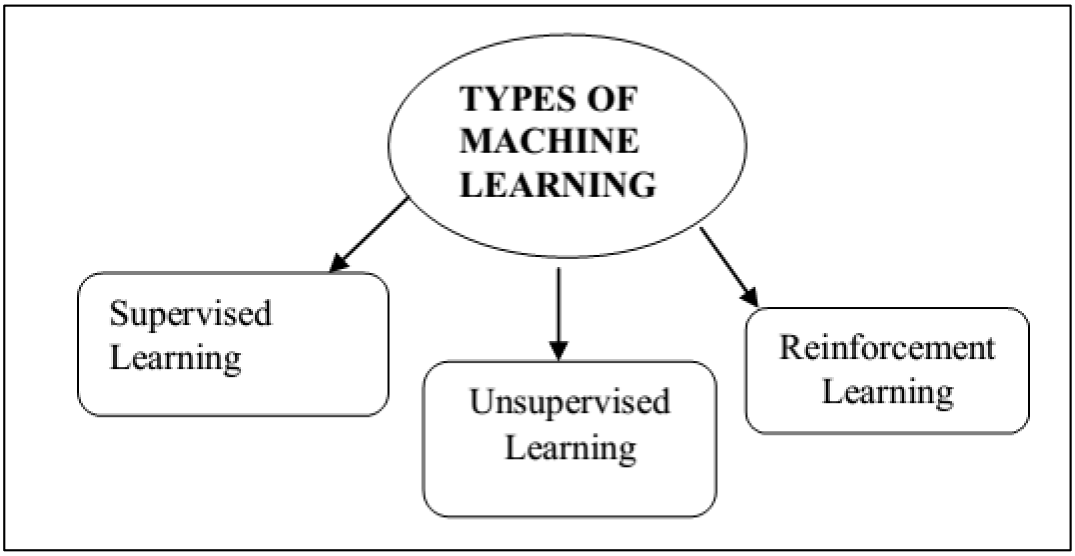
**Keywords:** Chronic Kidney Disease, machine learning, Types, Methods.

## MACHINE LEARNING

Machine learning is an application of AI that gives systems the power to automatically learn and improve from experience without being explicitly programmed. Machine learning focuses on the event of computer programs which will access data and use it to find out for themselves.

Machine learning is a branch of artificial intelligence (AI) focused on building applications that learn from data and improve their accuracy over time without being programmed to do so.

**Types of Machine Learning**



- **SupervisedLearning**
  Supervised learning is one of the most basic types of machine learning. In this type, the machine learning algorithm is trained on labeled data.

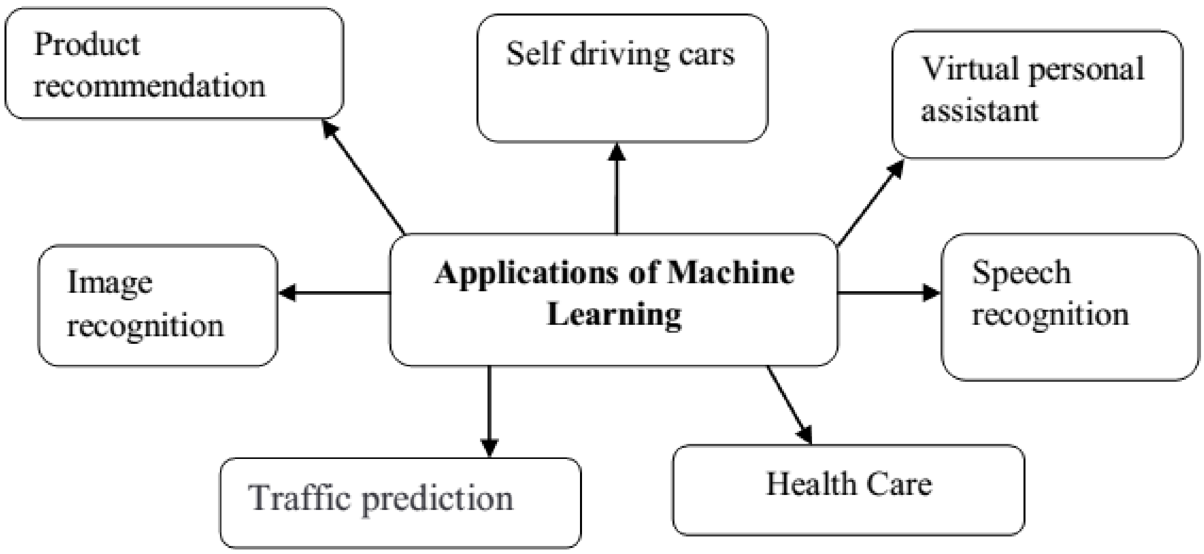
- **Unsupervised Learning**
  Unsupervised machine learning holds the advantage of being able to work with unlabeled data. This means that human labor is not required to make the dataset machine-readable, allowing much larger datasets to be worked on by the program.

- **Reinforcement Learning**
  Reinforcement learning directly takes inspiration from how human beings learn from data in their lives. It features an algorithm that improves upon itself and learns from new situations using a trial-and-error method.

**Features of Machine Learning**

i. Resurging interest in machine learning is due to the same factors that have made data mining and Bayesian analysis more popular than ever.
ii. Things like growing volumes and varieties of available data, computational processing that is cheaper and more powerful, and affordable data storage.
iii. Collect data and deliver faster, more accurate results – even on a very large scale.
iv. And by building precise models, an organization has a better chance of identifying profitable opportunities – or avoiding unknown risks.

**Application of Machine Learning**




## INTRODUCTION TO CHRONIC KIDNEY DISEASE (CKD)

Chronic kidney disease, also called chronic kidney failure, describes the gradual loss of kidney function. Your kidneys filter wastes and excess fluids from your blood, which are then excreted in your urine. When chronic kidney disease reaches an advanced stage, dangerous levels of fluid, electrolytes and wastes can build up in your body. In the early stages of chronic kidney disease, you may have few signs or symptoms. Chronic kidney disease may not become apparent until your kidney function is significantly impaired. Treatment for chronic kidney disease focuses on slowing the progression of the kidney damage, usually by controlling the underlying cause. Chronic kidney disease can progress to end-stage kidney failure, which is fatal without artificial filtering (dialysis) or a kidney transplant.

Chronic kidney disease (CKD) means your kidneys are damaged and can't filter blood the way they should. The disease is called "chronic" because the damage to your kidneys happens slowly over a long period of time.

This damage can cause wastes to build up in your body. CKD can also cause other health problems. Our kidneys have a greater capacity to do their job than is needed to keep us healthy. For example, you can donate one kidney and remain healthy. You can also have kidney damage without any symptoms because, despite the damage, your kidneys are still doing enough work to keep you feeling well. For many people, the only way to know if you have kidney disease is to get your kidneys checked with blood and urine tests.

Many people are afraid to learn that they have kidney disease because they think that all kidney disease leads to dialysis. However, most people with kidney disease will not need dialysis. If you have kidney disease, you can continue to live a productive life, work, spend time with friends and family, stay physically active, and do other things you enjoy. You may need to change what you eat and add healthy habits to your daily routine to help you protect your kidneys.

## MACHINE LEARNING ALGORITHMS

**K-nearest neighbour**

The k-nearest neighbour algorithm uses the data directly for classification without building a model first. As such, no details of model construction need to be considered, and the only adjustable parameter in the model is k, the number of nearest neighbours to include in the estimate of class membership: the value of $P(y|x)$ is calculated simply as the ratio of members of class y among the k-nearest neighbors of x. By varying k, the model can be made more or less flexible. The advantage of the k-nearest neighbor classifier is, it is robust to noisy training data and effective with large training datasets. The major drawback lies in the calculation of the case neighborhood: for this, one needs to define a metric that measures the distance between data items. In most cases it is done by trial and error.

**Random Forest**

The random forest is an ensemble approach that can also be thought of as a form of nearest neighbour predictor. Ensembles are a divide-and-conquer approach used to improve performance. The main principle behind ensemble methods is that a group of 'weak learners' can come together to form a 'strong learner'. The random forest starts with a standard machine learning technique called a 'decision tree' which, in ensemble terms, corresponds to our weak learner.

**Neural Network**

A neural network is a powerful computational data model that is able to capture and represent complex input/output relationships. The motivation for the development of neural network technology stemmed from the desire to develop an artificial system that could perform 'intelligent' tasks similar to those performed by the human brain. This model differs from the two algorithms above in the sense that it provides a functional form $f$ and parameter vector $\alpha$ to express $P(y|x)$ as $P(y|x) = f(x, \alpha)$. The parameters $\alpha$ are determined based on the data set D, usually by maximum likelihood estimation. The true power and advantage of neural networks lie in their ability to represent both linear and nonlinear relationships and in their ability to learn these relationships directly from the data being modeled. Traditional linear models are simply inadequate when it comes to modeling data that contains non-linear characteristics. The most common neural network model is the multilayer perceptron (MLP). This type of neural network is known as a supervised network because it requires a desired output in order to learn. The goal of this type of network is to create a model that correctly maps the input to the output using historical data so that the model can then be used to produce the output when the desired output is unknown.

**Support Vector Machine**

In machine learning, support-vector machines (SVMs, also support-vector networks) are supervised learning models with associated learning algorithms that analyze data for classification and multivariate analysis. Developed at AT&T Bell Laboratories by Vladimir Vapnik with colleagues, SVMs are one among the foremost robust prediction methods, being supported statistical learning frameworks or VC theory proposed by Vapnik and Chervonenkis. Given a group of coaching examples, each marked as belonging to at least one of two categories, an SVM training algorithm builds a model that assigns new examples to at least one category or the opposite , making it a non-probabilistic binary linear classifier (although methods like Platt scaling exist to use SVM during a probabilistic classification setting). An SVM maps training examples to points in space so on maximize the width of the gap between the 2 categories. New examples are then mapped into that very same space and predicted to belong to a category supported which side of the gap they fall.

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). When data are unlabelled, supervised learning isn't possible, and an unsupervised learning approach is required, which attempts to seek out natural clustering of the info to groups, then map new data to those formed groups. The support-vector clustering algorithm, created by Hava Siegelmann and Vladimir Vapnik, applies the statistics of support vectors, developed within the support vector machines algorithm, to categorize unlabeled data, and is one among the foremost widely used clustering algorithms in industrial applications. As Chronic Kidney Disease progresses slowly, early detection and effective treatment are the only cure to reduce the mortality rate. Machine learning techniques are gaining significance in medical diagnosis because of their classification ability with high accuracy rates. The accuracy of classification algorithms depends on the use of correct feature selection algorithms to reduce the dimension of datasets. In this study, Support Vector Machine classification algorithm was used to diagnose Chronic Kidney Disease. To diagnose the Chronic Kidney Disease, two essential types of feature selection methods namely, wrapper and filter approaches were chosen to reduce the dimension of Chronic Kidney Disease dataset. In wrapper approach, classifier subset evaluator with greedy stepwise search engine and wrapper subset evaluator with the Best First search engine were used. In filter approach, correlation feature selection subset evaluator with greedy stepwise search engine and filtered subset evaluator with the Best First search engine were used. The results showed that the Support Vector Machine classifier by using filtered subset evaluator with the Best First search engine feature selection method has higher accuracy rate (98.5%) in the diagnosis of Chronic Kidney Disease compared to other selected methods. Support Vector Machines are based on the Statistical Learning Theory concept of decision planes that define decision boundaries.

**Backgroundstudy**

A Deep Learning-based System for Automated Sensing of Chronic Kidney Disease published in IEEE EXPLORE in the year 2019, by the authors Navaneeth Bhaskar Sucheetha.

Diagnosis of Chronic kidney disease using Support Vector Machine and feature selection methods published in IEE EXPLORE in the year 2017 by the authors Huseyin polat,Homay Danaei Mehr,Aydin Cetin.

A Comparitive Study of classifier for chronic kidney disease prediction published in IEEE XPLORE in the year 2019 by the authors Devika, Sai vaishnavi avilala, Subramaniyaswamy

Two different Evaluators have been used for each method for filter approach Cfssubseteval with greedy stepwise search engine and filtersubsetevalwith BestFirst search algorithm Support Vector Machine, Feature selection in 2017 Diagnosis of Chronic kidney.

Analysis of Chronic kidney disease dateset by applying machine learning methods by the authors Yeldilkhan Amirgaliyev Shahriar ShamiluuluAzamat Serek in IEEE EXPLORE.

**Table 1.** Background Study

| S.NO | AUTHOR | JOURNAL | TITLE OF THE PAPER | METHOD AND YEAR | OUTCOME |
|---|---|---|---|---|---|
| 1 | Navaneeth Bhaskar Sucheetha | IEEE XPLORE | A Deep Learning-based System for Automated Sensing of Chronic Kidney Disease | Sensing module, Sample collection and testing 2019 | CKD is detected by calculating the glomerular filtration rate GFR |
| 2 | Devika, Sai vaishnavi avilala, Subramaniyaswamy | IEEE XPLORE | Comparitive Study of classifier for chronic kidney disease prediction | Navie Bayers,KNN,Random forest 2019 | A new selection Aid-Device is carried out for the prediction of CKD navie bayes random forest and KNN are implemented to locate CKD |
| 3 | Huseyin polat, | IEEE | | Support Vector Machine, | Two different |

| | | | | | |
|---|---|---|---|---|---|
| | Homay Danaei Mehr,Aydin Cetin | XPLORE | Diagnosis of Chronic kidney disease support vector machine by feature selection methods | Feature selection in 2017 | Evaluators have been used for each method for filter approach Cfssubseteval with greedy stepwise search engine and filtersubsetevalwith BestFirst search algorithm |
| 4 | Asif Salekin | IEEE EXPLORE | Detection of chronic Kidney Disease and Selecting Important predictive Atributes | Neural Network,Wrapper Embedded approach 2016 | Evaluation of three classifiers to detect CKD K-nearest neighbor random forest Neural networks. |
| 5 | Yeldilkhan Amirgaliyev Shahriar Shamiluulu Azamat Serek | IEEE EXPLORE | Analysis of Chronic kidney disease dateset by applying machine learning methods | Support Vector machine classifiers performance measure simulated program | Sensitivity value of SVM Classifier with a linear kernel is 93.100% performance measure of SVM Classifier with Linear-Kernel have been evaluated |

This table indicates that the clear analysis of literature on machine learning for Chronic Kidney Disease Detection

**Table2.** Analysis of literature on machine learning for Chronic kidney disease.

| Author | SVM | Navie Bayes | K-NN | Sensing module | Wrapper embedded | Feature selection | Random Forest |
|---|---|---|---|---|---|---|---|
| Navaneeth Bhaskar Sucheetha | - | - | - | ✓ | - | - | - |
| Devika, Sai vaishnavi avilala, Subramaniyaswamy | - | ✓ | ✓ | - | - | - | ✓ |
| Huseyin polat, Homay Danaei Mehr,Aydin Cetin | ✓ | - | - | - | - | ✓ | - |
| Asif Salekin | - | - | ✓ | - | ✓ | - | - |
| Yeldilkhan Amirgaliyev Shahriar Shamiluulu Azamat Serek | ✓ | - | - | - | - | - | - |

**Research Gaps identified**

- In the proposed paper the SVM method is used to detect the chronic kidney disease. But this is a long process and consumes a lot of time and effort.
- The techniques of deep learning have not been included in this paper. The deep learning techniques provide more accurate results. In deep learning there is no need for labeling of data.

## METHODOLOGY

The methodology used in this paper is displayed below

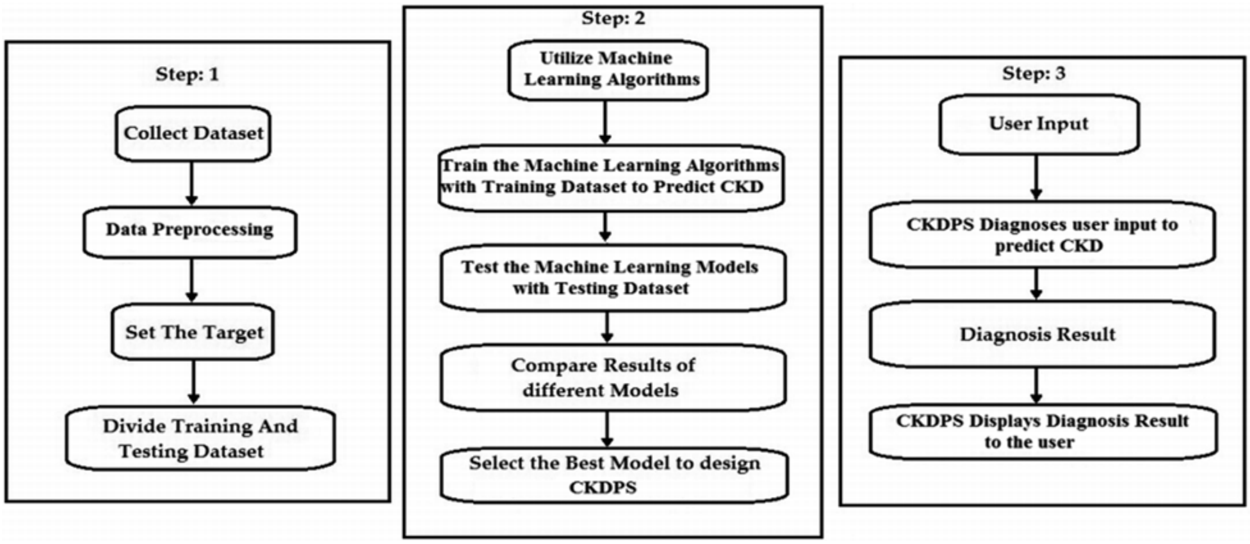



**Fig.1** Methodology used for chronic kidney disease detection

### Dataset

The chronic kidney disease dataset has been taken by UCI repository.

### Conclusion and future work

In this work, we studied many machine learning algorithm to detect chronic kidney disease .In future we will try on all features of machine learning techniques and to achieve best accuracy.The proposed SVM algorithm extracted features successfully classified the samples with an accuracy of 98.40%.In future a high accuracy can be achieved by implementing the deep learning techniques.

## REFERENCES


1. Burtis, C. A., & Bruns, D. E. (2014). *Tietz fundamentals of clinical chemistry and molecular diagnostics-e-book*. Elsevier Health Sciences.
2. Celec, P., Tóthová, Ľ., Šebeková, K., Podracká, Ľ., & Boor, P. (2016). Salivary markers of kidney function—potentials and limitations. *Clinica Chimica Acta*, *453*, 28-37.
3. Dagois, E., Khalaf, A., Sejdic, E., & Akcakaya, M. (2018). Transfer learning for a multimodal hybrid EEG-fTCD brain–computer interface. *IEEE Sensors Letters*, *3*(1), 1-4.
4. Georges, J. (1979). Determination of ammonia and urea in urine and of urea in blood by use of an ammonia-selective electrode. *Clinical chemistry*, *25*(11), 1888-1890.
5. Guo, M. F., Zeng, X. D., Chen, D. Y., & Yang, N. C. (2017). Deep-learning-based earth fault detection using continuous wavelet transform and convolutional neural network in resonant grounding distribution systems. *IEEE Sensors Journal*, *18*(3), 1291-1300.
6. Ince, T., Kiranyaz, S., Eren, L., Askar, M., & Gabbouj, M. (2016). Real-time motor fault detection by 1-D convolutional neural networks. *IEEE Transactions on Industrial Electronics*, *63*(11), 7067-7075.
7. Jayasree, T., Bobby, M., & Muttan, S. (2015). Sensor data classification for renal dysfunction patients using support vector machine. *Journal of Medical and Biological Engineering*, *35*(6), 759-764.
8. Kiranyaz, S., Ince, T., & Gabbouj, M. (2015). Real-time patient-specific ECG classification by 1-D convolutional neural networks. *IEEE Transactions on Biomedical Engineering*, *63*(3), 664-675.
9. Lekha, S., & Suchetha, M. (2017). A novel 1-D convolution neural network with SVM architecture for real-time detection applications. *IEEE Sensors Journal*, *18*(2), 724-731.
10. Navaneeth, B., & Suchetha, M. (2019). PSO optimized 1-D CNN-SVM architecture for real-time detection and classification applications. *Computers in biology and medicine*, *108*, 85-92.
11. Ng, D. K., Schwartz, G. J., Warady, B. A., Furth, S. L., & Muñoz, A. (2017). Relationships of measured iohexol GFR and estimated GFR with CKD-related biomarkers in children and adolescents. *American Journal of Kidney Diseases*, *70*(3), 397-405.
12. Saidi, T., Zaim, O., Moufid, M., El Bari, N., Ionescu, R., & Bouchikhi, B. (2018). Exhaled breath analysis using electronic nose and gas chromatography–mass spectrometry for non-

invasive diagnosis of chronic kidney disease, diabetes mellitus and healthy subjects. *Sensors and Actuators B: Chemical*, *257*, 178-188.

13. Samaranayake, L. (2007). Saliva as a diagnostic fluid. *International dental journal*, *57*(5), 295-299.
14. Wang, L., Tang, J., & Liao, Q. (2019). A study on radar target detection based on deep neural networks. *IEEE Sensors Letters*, *3*(3), 1-4.
15. Yan, K., Zhang, D., Wu, D., Wei, H., & Lu, G. (2014). Design of a breath analysis system for diabetes screening and blood glucose level prediction. *IEEE transactions on biomedical engineering*, *61*(11), 2787-2795.
16. Zhang, R., & Cao, S. (2018). Real-time human motion behavior detection via CNN using mmWave radar. *IEEE Sensors Letters*, *3*(2), 1-4.