



# CHRONIC KIDNEY DISEASE PREDICTION INTEGRATING CATEGORIZATION AND NOVEL MACHINE LEARNING TECHNIQUES

Ponnamm Lalitha<sup>1</sup>, Potta Devika<sup>2\*</sup>, Vadthya Manga<sup>3</sup>, G.Sowmya<sup>4</sup>, Deepika Borgaonkar<sup>5</sup>, Dr.Sunita.M<sup>6</sup>, Sudheer Nidamanuri<sup>7</sup>

<sup>1</sup>Assistant Professor, Department of CSE(Cys,DS) and AI&DS, VNR Vignana Jyothi Institute of Engineering and Technology, Hyderabad, India.

<sup>2\*</sup>Assistant Professor, Department of CSE(Cys,DS) and AI&DS, VNR Vignana Jyothi Institute of Engineering and Technology, Hyderabad, India.

<sup>3</sup>Assistant Professor, Department of CSE(AIML), CMR College of Engineering and Technology(CMR CET), Hyderabad, India.

<sup>4</sup>Assistant Professor, Department of CSE-AIML, MLR Institute of Technology, Hyderabad, India.

<sup>5</sup>Assistant Professor, Department of CSE(Cys,DS) and AI&DS, VNR Vignana Jyothi Institute of Engineering and Technology, Hyderabad, India.

<sup>6</sup>Associate Professor, Department of CSE (Data Science), Marri Laxman Reddy Institute of Technology and Management, Hyderabad, India.

<sup>7</sup>Assistant Professor, Department of CSE(Cys,DS) and AI&DS, VNR Vignana Jyothi Institute of Engineering and Technology, Hyderabad, India.

## Abstract

Chronic kidney disease (CKD), that causes undesired and rising volumes of hazardous liquids and wastes to accumulate in the blood and damage the body, is a typical issue with kidney function. An early screening tool to detect renal function is crucial in many circumstances because such degradation of kidney function results in kidney damage and occasionally in death. In order to diagnose and categorise diseases, Machine Learning(ML) approaches were widely applied in healthcare. In first stage, we propose Ant Colony Stacked Multi-Task Adaptive Naive Bayes (ACS-MANB) method to efficiently predict the CKDs. Initially, the CKD dataset is collected for this study. To eliminate the unwanted/duplicated information from the gathered data, normalization approach is used. In feature extraction stage, kernel-based principal component analysis (K-PCA) is used to retrieve the significant features from the normalized data. Finally, by employing the proposed method, we classify the CKDs into normal and abnormal kinds. As a second stage, we provide Genetic Fine-Tuned Regressive Support Vector Machine (GFR-SVM) technique for accurately forecasting CKDs. For this investigation, the CKD dataset is first gathered. Using a min-max normalization strategy, it is possible to pre-process the gathered data. Linear discriminant analysis is used to identify the important features from the pre-processed data. Finally, we categorise the CKDs into stage-wise by using the suggested methodology. The findings of this investigation demonstrate that the suggested approaches can categorise the data with the least amount of misclassification error. It has been discovered that the machine learning techniques utilised in this study greatly lessen the drawbacks of employing conventional approaches while also enhancing performance of the classifier. In comparison to traditional approaches, the diagnosis and classification techniques significantly improve on present system and offer more reliable predictions.

**Keywords:** Chronic kidney disease (CKD), Machine Learning(ML), Ant Colony Stacked Multi-Task Adaptive Naive Bayes (ACS-MANB), Genetic Fine-Tuned Regressive Support Vector Machine (GFR-SVM), kernel-based principal component analysis (K-PCA)



## **Introduction**

Designers and clinical professionals are attempting to create machine learning algorithms and models capable of detecting chronic kidney disease in its early stages. The issue is that data generated in the healthcare industry is large and complex, making data analysis difficult. However, using data mining technology, we can convert this data into a data format that can then be translated into machine learning algorithms.

The ineffectiveness of the kidney to do filtering action is described as kidney disease. The amount of waste products in our body precariously increases with an increase in the severity of the disease. CKD afflicts around 15% of the world population, and millions of people die each year because of this disease. According to the latest medical report on CKD, about 325 million people are affected by kidney disease globally. Regular monitoring of kidneys is essential because it is challenging to predict how kidney functioning is declining. CKD can develop heart and blood vessel problems, which can lead to severe medical problems like heart attack and strokes. The blood and proteins will be leaked to urine once the filtering system of the kidney is damaged. The kidneys work hard to remove the waste products, excess fluid, and other toxins from the body. Kidneys help control the blood pressure, regulate blood chemicals that are important for our body, and stimulate the production of red blood cells. Kidney diseases are more common in older people. CKD can progress to complete kidney failure if proper treatment measures are not taken to slow down its progression. Either dialysis or kidney transplant is needed for survival if the kidneys are completely damaged. The two significant causes of CKD are diabetes and high blood pressure. Diabetes occurs when the glucose content in the blood rises above the normal range. High blood pressure is a long-term health problem that occurs when blood pressure is elevated. High blood pressure can damage the brain, eyes, heart and kidneys. CKD develops slowly over time, and the symptoms may not appear until the kidneys are severely affected. Some of the symptoms of CKD can be caused due to some other health conditions. Therefore, it is necessary to consult health professionals to get the proper diagnosis.

## **Literature review**

By using Artificial Neural Network (ANN) and Support Vector Machine (SVM) approaches to detect CKD at a preliminary phase, the researchers of [1] hoped to aid in the management



of CKD. [2] To anticipate the existence of chronic kidney disease, they offered a deep-ANN model. A robust and efficient Adaptive Hybridized Deep Convolutional Neural Network (AHDCNN) was constructed in Study [3] for the early diagnosis of Kidney illness. A model for predicting CKD disease among individuals was created in study [4]. Both the model's performance with regard to all qualities and particular aspects was examined. The innovation in [5] comes from creating a system for diagnosing CKDs. This study aids professionals in investigating CKD prevention strategies through early identification utilising machine learning methods. Deep Neural Network for the earlier diagnosis and forecasting of CKD was revealed in research [6]. Eight distinct machine learning (ML) techniques are used in research [7] effort to propose a method that can quickly identify the infection of CKD while taking into account the patient's health status database details. The goal of the [8] was to determine whether machine learning (ML) could accurately predict individuals suffering from CKD's probability of developing end-stage kidney disease (ESKD). Utilizing K-nearest neighbour(KNN) and convolutional neural network (CNN), study [9] aims to detect and forecast individuals with more prevalent chronic illnesses. We established a productive cloud-based health care framework in [10]. In order to identify chronic renal illness, study [11] presented hybrid machine learning strategies that comprise feature selection approaches and machine learning classifying techniques rely on big data platforms (Apache Spark). In [12], an IoT multimedia data-based diagnostic framework for CKD and also its intensity is presented. A unique Hybrid Filter Wrapper Embedded (HFWE) based Feature Selection (FS) technique was presented in Study [113] for choosing the best subset of features from datasets to predict CKD datasets. The purpose of [14] would be to examine the influence of social factors on CKD in the body of available literature. To anticipate the different phases of CKD, they created and contrasted two methods, comprising J48 and random forest, in [15]. A fuzzy logic-based method for the identification of chronic renal disease was established in study [16]. A dataset that is accessible in the UCI Repository was used in paper [17] to demonstrate the forecast of renal illness in order to assist medical specialists. For the purpose of diagnosing chronic kidney disease, study [18] built a DNN model with a metaheuristic algorithm based on the best feature selection. A good feature-based framework for identifying renal disease was developed by study [19]. The Density-based Feature Selection (DFS) with Ant Colony based Optimization (D-ACO) strategy treating CKD was presented in study [20] as an intelligent prediction and classification system for the healthcare industry. Based on the symptoms or characteristics seen in a specific instance, if the state is either acute or chronic, [21] aims to predict the classification of CKD. A new deep learning-based sensing approach for the automated identification of renal illness was proposed in study [22].



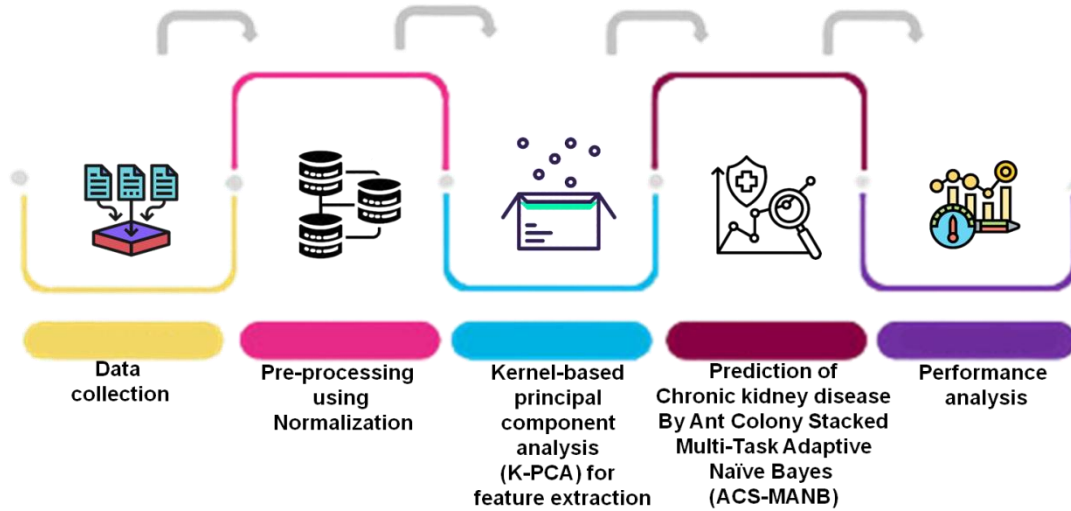
A novel deep learning framework for categorising chronic renal illness was presented in study [23] employing a stacked auto encoder system with multimedia information and a softmax classifier. CKD is diagnosed using decision tree, random forest, and support vector machine, according to [24]. A fruit fly optimization algorithm (FFOA) and a multi-kernel support vector machine (MKSVM) for disease classification were presented in study [25]. Study [26] improving the categorization of chronic renal illness using the information gain ratio and adaboost. The use of machine learning approaches to aid in the early detection of CKD in emerging regions is examined in study [27]. An artificial intelligence deep learning algorithm (DLA) was created by study [28] to identify chronic kidney disease from retinal scans. Three base learners and two ensemble strategies were reported in study [29] to increase classification accuracy for effective detection of chronic renal disease. The classification approaches, including as tree-based decision trees, random forests, and logistic regression, were the main focus of [30].

## Methodology

Stage-1: Ant Colony Stacked Multi-Task Adaptive Naive Bayes (ACS-MANB) method

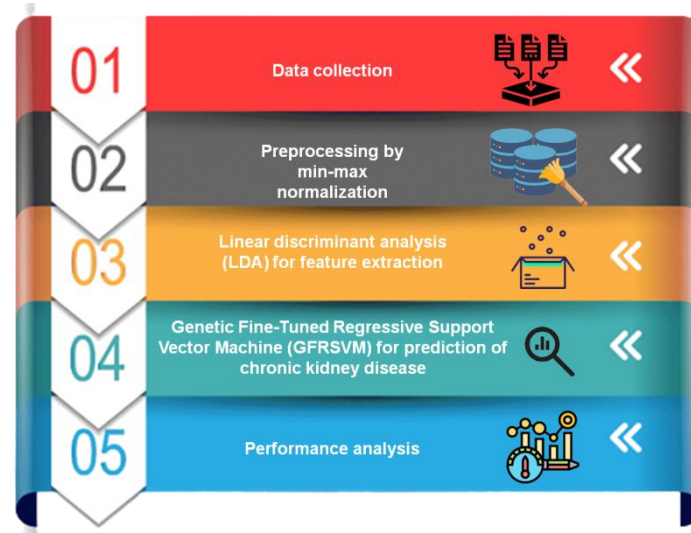
Stage-2:Genetic Fine-Tuned Regressive Support Vector Machine (GFR-SVM) method

According to the current state of research, CKD is becoming more commonplace each year. The forecasting of CKD is among the resources for subsequent treatment, and because machine learning algorithms have a high degree of classification results, they are becoming more significant in clinical diagnosis. So, this study proposes Ant Colony Stacked Multi-Task Adaptive Naive Bayes (ACS-MANB) method to predict the CKDs. Initially, the CKD dataset is collected for this study. To eliminate the unwanted/duplicated data from the raw data, normalization approach is used. In feature extraction stage, kernel-based principal component analysis is used to retrieve the significant features from the normalized data. Finally, by employing the proposed method, we classify the CKDs into normal and abnormal kinds. Analysis and contrast of the suggested technique's performance with standard approaches already in use. Our technique outclasses the other existing methods in the prediction of CKDs and also attain the greatest performance in terms of accuracy. Figure 2 depicts the flow of proposed work.



**Figure 1: Flow of proposed work**

Chronic Kidney Disease is primarily brought on by uncontrolled hypertension (CKD). Researchers from different areas of the globe use the Glomerular Filtration Rate (GFR) and kidney damage indicators to define CKD as a disorder that over time results in a decline in renal function. A patient having CKD is more likely to pass away before their time. To diagnose the various disorders that are connected to CKD during a preliminary phase and treat the condition, doctors have a challenging challenge. In this work, we provide Genetic Fine-Tuned Regressive Support Vector Machine (GFR-SVM) method for accurately forecasting CKDs. For this investigation, the CKD dataset is first gathered. Using a min-max normalization strategy, it is possible to pre-process the gathered data. Linear discriminant analysis is used to identify the important features from the pre-processed data. Finally, we categorise the CKDs into stage-wise by using the suggested methodology. Analyses and comparisons with already-used methods are done to determine how well the suggested method performs. As a result of the study's findings, we can say that the suggested method may be an effective one for nephrologists to utilise in identifying CKD. Figure 2 shows the proposed workflow.



**Figure 2: Proposed workflow**

## Results

The performance of the suggested techniques is evaluated in terms of metrics like accuracy, precision, recall, F1 score, RMSE, and MSE. The categorization models may provide four possible outputs.

**Table 1: Results of work-1**

| Methods  | Accuracy (%) | Precision (%) | Recall (%) | F1-score (%) | RMSE (%) | MSE (%) |
|----------|--------------|---------------|------------|--------------|----------|---------|
| CNN      | 97.3         | 96.9          | 99.4       | 98.2         | 25       | 19      |
| DT       | 95.8         | 98.2          | 94.3       | 92.1         | 22       | 23      |
| Proposed | 99.72        | 98.94         | 99.91      | 98.56        | 11       | 8       |

**Table 2: Results of work-2**

| Methods  | Accuracy (%) | Precision (%) | Recall (%) | F1-score (%) | RMSE (%) | MSE (%) |
|----------|--------------|---------------|------------|--------------|----------|---------|
| SVM      | 98.96        | 97.33         | 99.3       | 97.6         | 18       | 14      |
| RF       | 97.45        | 96.8          | 96.8       | 96.4         | 23       | 18      |
| Proposed | 99.91        | 98.2          | 99.5       | 98.96        | 9.6      | 6.3     |

## Summary

In this research, we developed two novel methods based on machine learning regarding the prediction of CKDs. First stage of this work, Ant Colony Stacked Multi-Task Adaptive Naive



Bayes (ACS-MANB)method was proposed while for second work, Genetic Fine-Tuned Regressive Support Vector Machine (GFR-SVM) technique was suggested regarding efficient prediction of CKDs. Our two methods attained the greatest performance than existing methods in terms of RMSE, MSE, f1-score, recall, precision, and accuracy.

## Reference

1. Almansour, N.A., Syed, H.F., Khayat, N.R., Altheeb, R.K., Juri, R.E., Alhiyafi, J., Alrashed, S. and Olatunji, S.O., 2019. Neural network and support vector machine for the prediction of chronic kidney disease: A comparative study. *Computers in biology and medicine*, 109, pp.101-111.
2. Kriplani, H., Patel, B. and Roy, S., 2019. Prediction of chronic kidney diseases using deep artificial neural network technique. In *Computer aided intervention and diagnostics in clinical and medical images* (pp. 179-187). Springer, Cham.
3. Chen, G., Ding, C., Li, Y., Hu, X., Li, X., Ren, L., Ding, X., Tian, P. and Xue, W., 2020. Prediction of chronic kidney disease using adaptive hybridized deep convolutional neural network on the internet of medical things platform. *IEEE Access*, 8, pp.100497-100508.
4. Chittora, P., Chaurasia, S., Chakrabarti, P., Kumawat, G., Chakrabarti, T., Leonowicz, Z., Jasiński, M., Jasiński, Ł., Gono, R., Jasińska, E. and Bolshev, V., 2021. Prediction of chronic kidney disease-a machine learning perspective. *IEEE Access*, 9, pp.17312-17334.
5. Senan, E.M., Al-Adhaileh, M.H., Alsaade, F.W., Aldhyani, T.H., Alqarni, A.A., Alsharif, N., Uddin, M.I., Alahmadi, A.H., Jadhav, M.E. and Alzahrani, M.Y., 2021. Diagnosis of chronic kidney disease using effective classification algorithms and recursive feature elimination techniques. *Journal of Healthcare Engineering*, 2021.
6. Singh, V., Asari, V.K. and Rajasekaran, R., 2022. A Deep Neural Network for Early Detection and Prediction of Chronic Kidney Disease. *Diagnostics*, 12(1), p.116.
7. Baidya, D., Umair, U., Islam, M.N., Shamrat, F.J.M., Pramanik, A. and Rahman, M.S., 2022, April. A Deep Prediction of Chronic Kidney Disease by Employing Machine Learning Method. In *2022 6th International Conference on Trends in Electronics and Informatics (ICOEI)* (pp. 1305-1310). IEEE.
8. Bai, Q., Su, C., Tang, W. and Li, Y., 2022. Machine learning to predict end stage kidney disease in chronic kidney disease. *Scientific reports*, 12(1), pp.1-8.



9. Alanazi, R., 2022. Identification and prediction of chronic diseases using machine learning approach. *Journal of Healthcare Engineering*, 2022.
10. Aswini, J., Yamini, B., Jatothu, R., Nayaki, K.S. and Nalini, M., 2022. An efficient cloud-based healthcare services paradigm for chronic kidney disease prediction application using boosted support vector machine. *Concurrency and Computation: Practice and Experience*, 34(10), p.e6722.
11. Abdel-Fattah, M.A., Othman, N.A. and Goher, N., 2022. Predicting Chronic Kidney Disease Using Hybrid Machine Learning Based on Apache Spark. *Computational Intelligence and Neuroscience*, 2022.
12. Hosseinzadeh, M., Koohpayehzadeh, J., Bali, A.O., Asghari, P., Souri, A., Mazaherinezhad, A., Bohlouli, M. and Rawassizadeh, R., 2021. A diagnostic prediction model for chronic kidney disease in internet of things platform. *Multimedia Tools and Applications*, 80(11), pp.16933-16950.
13. Parthiban, R., Usharani, S., Saravanan, D., Jayakumar, D., Palani, D.U., StalinDavid, D.D. and Raghuraman, D., 2021. Prognosis of chronic kidney disease (CKD) using hybrid filter wrapper embedded feature selection method. *European Journal of Molecular & Clinical Medicine*, 7(9), pp.2511-2530.
14. Burgos-Calderón, R., Depine, S.Á. and Aroca-Martínez, G., 2021. Population kidney health. A new paradigm for chronic kidney disease management. *International journal of environmental research and public health*, 18(13), p.6786.
15. Ilyas, H., Ali, S., Ponum, M., Hasan, O., Mahmood, M.T., Iftikhar, M. and Malik, M.H., 2021. Chronic kidney disease diagnosis using decision tree algorithms. *BMC nephrology*, 22(1), pp.1-11.
16. Ahmed, T.I., Bhola, J., Shabaz, M., Singla, J., Rakhra, M., More, S. and Samori, I.A., 2022. Fuzzy logic-based systems for the diagnosis of chronic kidney disease. *BioMed Research International*, 2022.
17. Saha, I., Gourisaria, M.K. and Harshvardhan, G.M., 2022. Classification System for Prediction of Chronic Kidney Disease Using Data Mining Techniques. In *Advances in Data and Information Sciences* (pp. 429-443). Springer, Singapore.
18. Lambert, J.R. and Perumal, E., 2022. Oppositional firefly optimization based optimal feature selection in chronic kidney disease classification using deep neural network. *Journal of Ambient Intelligence and Humanized Computing*, 13(4), pp.1799-1810.
19. Poonia, R.C., Gupta, M.K., Abunadi, I., Albraikan, A.A., Al-Wesabi, F.N. and Hamza, M.A., 2022, February. Intelligent Diagnostic Prediction and Classification



- 
- Models for Detection of Kidney Disease. In Healthcare (Vol. 10, No. 2, p. 371). MDPI.
20. Elhoseny, M., Shankar, K. and Uthayakumar, J., 2019. Intelligent diagnostic prediction and classification system for chronic kidney disease. Scientific reports, 9(1), pp.1-14.
  21. Saringat, Z., Mustapha, A., Saedudin, R.R. and Samsudin, N.A., 2019. Comparative analysis of classification algorithms for chronic kidney disease diagnosis. Bulletin of Electrical Engineering and Informatics, 8(4), pp.1496-1501.
  22. Bhaskar, N. and Manikandan, S., 2019. A deep-learning-based system for automated sensing of chronic kidney disease. IEEE Sensors Letters, 3(10), pp.1-4.
  23. Khamparia, A., Saini, G., Pandey, B., Tiwari, S., Gupta, D. and Khanna, A., 2020. KDSAE: Chronic kidney disease classification with multimedia data learning using deep stacked autoencoder network. Multimedia Tools and Applications, 79(47), pp.35425-35440.
  24. Revathy, S., Bharathi, B., Jeyanthi, P. and Ramesh, M., 2019. Chronic kidney disease prediction using machine learning models. International Journal of Engineering and Advanced Technology, 9(1), pp.6364-6367.
  25. Jerlin Rubini, L. and Perumal, E., 2020. Efficient classification of chronic kidney disease by using multi-kernel support vector machine and fruit fly optimization algorithm. International Journal of Imaging Systems and Technology, 30(3), pp.660-673.
  26. Lestari, A., 2020. Increasing accuracy of C4. 5 algorithm using information gain ratio and adaboost for classification of chronic kidney disease. Journal of Soft Computing Exploration, 1(1), pp.32-38.
  27. Sobrinho, A., Queiroz, A.C.D.S., Da Silva, L.D., Costa, E.D.B., Pinheiro, M.E. and Perkusich, A., 2020. Computer-aided diagnosis of chronic kidney disease in developing countries: A comparative analysis of machine learning techniques. IEEE Access, 8, pp.25407-25419.
  28. Sabanayagam, C., Xu, D., Ting, D.S., Nusinovici, S., Banu, R., Hamzah, H., Lim, C., Tham, Y.C., Cheung, C.Y., Tai, E.S. and Wang, Y.X., 2020. A deep learning algorithm to detect chronic kidney disease from retinal photographs in community-based populations. The Lancet Digital Health, 2(6), pp.e295-e302.
  29. Jongbo, O.A., Adetunmbi, A.O., Ogunrinde, R.B. and Badeji-Ajisafe, B., 2020. Development of an ensemble approach to chronic kidney disease diagnosis. Scientific African, 8, p.e00456.



- 
30. Gupta, R., Koli, N., Mahor, N. and Tejashri, N., 2020, June. Performance analysis of machine learning classifier for predicting chronic kidney disease. In 2020 International Conference for Emerging Technology (INCET) (pp. 1-4). IEEE.