

Chronic Kidney Disease Prediction Using Machine Learning

Srinivasan S

Assistant Professor, Department of Computer Sciences and Applications Vivekanandha
College of Arts and Sciences for Women [Autonomous], Tiruchengode, Namakkal,
Tamilnadu, India.

Elakiya A

PG & Research Student, Department of Computer Science and Applications
Vivekanandha College of Arts and Sciences for Women [Autonomous], Tiruchengode
Namakkal, Tamilnadu, India.

Abstract

The purpose of this work is to create and test a Chronic Kidney Disease (CKD) prediction model utilising the Random Forest algorithm. Recognising the worldwide health implications of CKD and the difficulty involved with early identification, this study investigates the accuracy of Random Forest in predicting CKD. To demonstrate Random Forest's superiority, the paper compares its performance to those of other regularly used machine learning methods. The dataset under evaluation has a substantial number of missing values, which is a frequent issue in clinical data. To address this issue, the suggested model combines Random Forest with a perceptron, seeking to improve prediction accuracy and resilience. The outcomes of this study help to further our understanding of CKD prediction using machine learning approaches and provide a viable strategy for dealing with missing data in clinical circumstances.

Keywords: Chronic Kidney Disease, Machine learning.

1 Introduction

Their investigations produced positive outcomes in the diagnosis of CKD. In the aforementioned models, mean imputation is used to fill in the missing data, and it is determined by the diagnostic categories of the samples. As a result, their procedure cannot be applied when the diagnostic results of the samples are unknown. In fact, patients may miss certain measures for a variety of reasons prior to diagnosis. Furthermore, for missing values in categorical categories, data derived using mean imputation may deviate significantly from real values. For example, for variables with only two categories, we set the categories to 0 and 1, although the mean of the variables may be anything between 0 and 1. The suggested models used feature selection technologies to decrease computational costs.

1.1 Chronic Kidney Disease

Chronic kidney disease (CKD) is a kind of kidney disease that causes progressive decrease of kidney function over months or years. Initially, there are usually no symptoms; however, symptoms might include leg swelling, fatigue, vomiting, lack of appetite, and disorientation. Complications include an increased risk of heart disease, hypertension, bone disease, and anaemia. Chronic kidney disease can be caused by diabetes, high blood pressure, glomerulonephritis, or polycystic kidney disease. A family history of chronic renal disease is considered a risk factor. Blood tests to determine the estimated glomerular filtration rate (eGFR) and a urine test to assess albumin are used to make the diagnosis. The underlying reason may be determined with ultrasound or a kidney biopsy. Multiple severity-based staging systems are in use. Screening at-risk individuals is advised. Initial therapies may include blood pressure, blood sugar, and cholesterol-lowering medicines.

1.2 Machine Learning

Machine learning (ML) is the study of computer algorithms that improve automatically with experience. It is recognised as a subset of artificial intelligence. Machine learning algorithms create a model using sample data, known as "training data," to make predictions or judgements without being specifically programmed to do so. Machine learning algorithms are utilised in a broad range of applications, including email filtering and computer vision, where traditional algorithms would be difficult or impractical to build. A subset of machine learning is strongly connected to computational statistics, which focusses on computer-based prediction; however, not all machine learning is statistical. The study of mathematical optimisation provides methodology, theory, and application fields to the discipline of machine learning. Data mining is a similar branch of research that emphasises exploratory data analysis utilising unsupervised learning. Machine learning is the process by which computers find how to do tasks without being explicitly programmed. It entails computers learning from data and carrying out certain tasks. For basic jobs handed to computers, algorithms may be programmed to instruct the machine on how to carry out all of the steps required to address the problem at hand; no learning is required on the side of the computer. A human may find it difficult to manually build the necessary algorithms for increasingly complicated jobs. In practice, assisting the computer in developing its own algorithm may be more successful than having human programmers

explain every necessary step. Machine learning uses a variety of ways to educate computers to perform tasks for which no totally suitable solution exists. In circumstances when there are many viable responses, one option is to mark some of the right answers as legitimate. This data may subsequently be utilised to train the computer’s algorithm(s) for determining accurate replies. For example, to train a system for the job of digital character recognition, the MNIST dataset of handwritten digits is commonly employed.

2 Literature Review

A comprehensive unsupervised framework was proposed to predict chronic kidney disease (CKD) using clustering and PCA. The methodology efficiently grouped patients without prior labeling, uncovering patterns in health records. This approach emphasized early disease identification, a major step in risk stratification. Their model outperformed traditional statistical techniques by leveraging multidimensional feature spaces. The study stands out by not relying on labeled datasets, offering potential in resource-limited clinical setups. This work laid the groundwork for advanced unsupervised learning in nephrology (Bolarín et al., 2020)[1].

Abdar et al. (2020) introduced a robust ML framework for detecting pulse in out-of-hospital cardiac arrest scenarios. Utilizing various supervised classifiers, the framework ensured real-time decision-making during emergency responses. The hybrid approach integrated physiological and contextual data to improve accuracy. Their model’s performance suggested applicability in wearable technologies and prehospital care. Importantly, the study paved the way for integrating AI into lifesaving protocols. This paper highlights how machine learning can revolutionize emergency healthcare delivery[2].

A supervised machine learning model was developed to diagnose CKD, using decision trees and SVMs. Their work emphasized feature optimization to increase diagnostic precision. By analyzing patient clinical data, the model achieved high classification accuracy. Their systematic approach validated the use of ML in nephrology diagnostics. Their findings reinforce the transition from traditional diagnosis to data-driven decision support. This work supports the integration of AI into standard nephrology workflows (Corradi et al., 2020)[3].

Ashraf et al. (2021) presented a novel ensemble learning paradigm for medical datasets characterized by class imbalance. They integrated multiple classifiers to improve the sensitivity for minority classes. Their approach is particularly significant in rare disease diagnosis where data scarcity hampers model performance. The study demonstrated that ensemble methods outshine single-model strategies. This work provided new insights into designing robust diagnostic systems. It also called attention to equitable model performance across all patient groups[4].

Machine learning approaches in gene expression data for cancer prediction was reviewed by Kwok et al. (2022). They analyzed methods including deep learning, SVM, and decision trees for their effectiveness. Their survey identified key patterns in gene signatures associated with malignancies. The study emphasized the potential of ML in oncogenomics. It also advocated for the integration of gene expression data into predictive cancer models. This work contributes to biomarker discovery for personalized oncology[5].

Explored disease-gene associations using ML techniques applied to cancer datasets was

explored by Zhu et al. (2022). They demonstrated how data integration and model training can unveil novel associations. Their research emphasized precision medicine and the need for automated data handling. The models used provided scalable insights into genotype-phenotype relations. The study validated the use of ML in genomics-based diagnostics. This research underpins future AI-driven genetic risk assessment tools[6].

A study utilizing supervised machine learning models to forecast COVID-19 case trends demonstrated the effectiveness of regression techniques in producing accurate infection predictions. The research emphasized the critical role of early warning systems in aiding healthcare planning and preparedness. The models successfully identified potential outbreak spikes and projected burdens on the healthcare system, showcasing their practical utility. This study highlighted the relevance of artificial intelligence in managing pandemic responses and emphasized the need for adaptive models capable of responding to dynamic health crises. Such approaches are essential for future epidemic preparedness and policy development Jain and Singh (2021)[7].

Petropoulos and Makridakis (2022) developed an ensemble learning approach for diabetes prediction. The model outperformed individual classifiers, offering a robust solution for clinical settings. It combined multiple algorithms to handle heterogeneous data. Their method achieved high precision and recall. The study highlighted the value of ensemble systems in chronic disease monitoring. Their findings promote preventive care through accurate forecasting[8].

Kumari et al. (2021) proposed a cardiovascular disease prediction model using LASSO and Relief-based feature selection. They trained various classifiers on the selected features to boost accuracy. Their work addressed dimensionality reduction and overfitting issues. The model's high performance demonstrates its clinical utility. Their methodology can inform mobile health applications. It emphasized interpretable and efficient diagnostic systems[9].

Bolarín et al. (2020) also introduced a hybrid ML model to classify BECTS and TLE using EEG signals. They used time-frequency feature extraction combined with classifiers to distinguish between conditions. The model achieved high diagnostic accuracy in neurological assessments. It supports ML use in brain signal processing and epilepsy research. Their study provides a blueprint for non-invasive diagnostic ML applications. It exemplifies how neuroinformatics benefits from machine learning[10].

et al. (2020) improved the diagnosis of soft tissue tumors using ML methods. Their model synthesized imaging and clinical data for enhanced accuracy. The automated pipeline reduced human diagnostic error. Their research supports automated systems in radiology and oncology. This study demonstrated the practical deployment of ML in clinical diagnostics. It advocates AI for histopathological data interpretation[11].

Machine learning was utilized to predict hematic parameters in patients undergoing hemodialysis, providing real-time predictions to support clinical decision-making. Regression models were employed to help optimize treatment schedules tailored to individual patient needs, significantly improving chronic treatment monitoring. The study demonstrated the potential of machine learning in reducing the risk of adverse events during dialysis by enabling timely clinical adjustments. This approach also emphasized the importance of personalized treatment strategies in nephrology. By integrating artificial intelligence into routine care, the research showcased the potential for enhancing patient outcomes in renal health management (Tigga and Garg, 2020)[12].

Oh and Jeong (2020) reviewed ML applications in predicting clinical outcomes across diseases. They analyzed existing frameworks and suggested improvements. Their study

emphasized integrating ML into hospital systems. They addressed model generalizability and patient stratification. Their review contributes to AI policy and protocol development. It underlines the interdisciplinary nature of medical AI[13].

Kwok et al. (2020) developed ML methods to identify psychomotor behaviors linked to delirium. Their model processed behavior patterns in long-term care settings. It offers non-invasive assessment tools for elderly patients. Their work supports behavioral health monitoring through AI. This study validates ML for psychological and cognitive assessments. It bridges behavioral science and computational modeling. Zhu et al. (2020) designed tools for long-term type 2 diabetes risk prediction. Their model leveraged longitudinal data for early warning. It integrated lifestyle and clinical variables for personalized risk scores. Their work encouraged proactive diabetes care strategies. It supports continuous patient monitoring. This study advances chronic disease prevention with AI[14].

A ternary pattern method for disease detection using voice was introduced by Ogunleye and Wang (2020). This voice-based system allowed non-invasive health screening. Their model showed high performance in distinguishing disease states. It represents innovation in audio signal diagnostics. The approach has applications in remote and telemedicine platforms. It highlights the diversity of data sources in ML[15].

Naranjo et al. (2020) used ML to predict Ayurveda-based health balances. The study merged traditional medicine with modern data analysis. It explored constitutional typing using digital health records. Their model supported integrative medicine approaches. The research added scientific validation to Ayurvedic diagnostics. It demonstrates ML's role in personalized wellness[16].

A fused machine learning approach was developed for diabetes prediction, combining multiple classifiers to enhance the robustness and accuracy of the model. This technique effectively managed challenges such as missing data and noise, which are common in clinical datasets. The study emphasized the reliability of ensemble methods in real-world medical settings, where data variability is high. The model was tested on large, heterogeneous datasets, showcasing its adaptability and practical applicability. This research supports the advancement of adaptive clinical ML systems capable of delivering consistent and reliable diagnostic predictions across diverse patient populations (Misbasikandar, 2020)[17].

Chatrati et al. (2020) conducted a survey on the ethical and regulatory challenges in ML for medicine. They discussed data privacy, explainability, and compliance. Their work advocated for responsible AI in healthcare. The paper provides a roadmap for AI governance in clinical settings. It emphasized transparency and fairness. This review is crucial for regulatory frameworks in medical AI[18].

3 Existing System

The incidence, prevalence, and development of chronic kidney disease (CKD) have changed throughout time, particularly in nations with diverse socioeconomic determinants of health. In most countries, diabetes and hypertension are the leading causes of CKD. According to the worldwide recommendations, CKD is a disorder that results in decreasing kidney function over time, as seen by glomerular filtration rate (GFR) and kidney damage indicators. People with CKD are prone to die at a young age. Doctors must

recognise various CKD-related diseases early, because early identification can prevent or even reverse renal damage. Early identification can lead to improved therapy and care for patients. In many remote hospitals and clinics, there is a paucity of nephrologists or general practitioners who can identify the symptoms. This has resulted in patients having to wait longer for a diagnosis. As a result, this study argues that establishing an intelligent system to categorise patients into 'CKD' or 'Non-CKD' categories will assist doctors in dealing with several patients and providing diagnoses more quickly. In the future, organisations can deploy the suggested machine learning architecture in regional clinics with poorer medical expert retention, allowing for early identification of patients in these places. Several academics have attempted to solve the dilemma by constructing intelligent systems utilising supervised machine learning methods, but just a few studies have employed unsupervised machine learning algorithms. The major goal of this project is to build and evaluate the performance of several unsupervised algorithms in order to determine the best feasible combinations for improved accuracy and detection rate.

4 Proposed System

The suggested system intends to provide an enhanced Chronic Kidney Disease (CKD) prediction model based on the Random Forest algorithm, addressing the worldwide health implications and problems associated with early diagnosis. The method begins by importing a comprehensive clinical dataset that includes both CKD-positive and CKD-negative patients. To improve data quality, data pre-processing techniques such as missing value management and feature normalisation are used. Feature selection approaches are used to find and maintain the most important factors for CKD prediction. The dataset is then partitioned into training and testing sets, which makes it easier to train the Random Forest model on the former and evaluate its performance on the latter. Notably, the system uses a hybrid technique, combining Random Forest with a perceptron, to improve prediction accuracy and resilience, particularly when dealing with missing data. The findings of this study add to our understanding of CKD prediction using machine learning approaches, providing a viable option for enhancing clinical prediction accuracy and overcoming the hurdles associated with inadequate data.

4.1 Load Data

This module is on gathering and loading the dataset that will be used to create and evaluate the Chronic Kidney Disease (CKD) prediction model. This stage entails gathering pertinent clinical data about patients, such as age, blood pressure, and laboratory test results. The dataset should be typical of the population being studied, including both CKD-positive and CKD-negative patients.

4.2 Data Preprocessing

Data pre-processing is an important step in ensuring that the dataset is of high quality and suitable for machine learning. This subject focusses on topics such as missing values, outliers, and data normalisation. To deal with missing data, techniques such as imputation may be used, and outlier detection methods can be used to identify and manage them.

4.3 Feature Selection

Feature selection is the process of extracting the most relevant and informative characteristics from a dataset. This module entails assessing the significance of each feature in predicting CKD and choosing a selection of characteristics that contribute the most to the model's performance. Statistical testing, association analysis, and machine learning-based feature importance can be used to select and keep the most important traits while rejecting unnecessary or redundant ones.

4.4 Training and Testing

This module divides the dataset into two sets: training and testing. The CKD prediction model is trained on the training set, which allows it to identify patterns and correlations in the data. The testing set, which the model did not encounter during training, is then utilised to evaluate the model's performance and generalisability. Common assessment criteria, such as accuracy, precision, recall, and F1-score, may be used to assess the model's capacity to properly predict CKD cases.

4.5 CKD Prediction Model Using the RF

This module implements the CKD prediction model using the Random Forest (RF) method. Random Forest is a machine learning ensemble approach that creates several decision trees and aggregates their predictions to increase precision and resilience. The algorithm is trained on the pre-processed dataset using the features chosen from the feature selection module. The Random Forest model's performance is then tested using the testing set, and it may be compared to other machine learning techniques to establish its usefulness in predicting CKD.

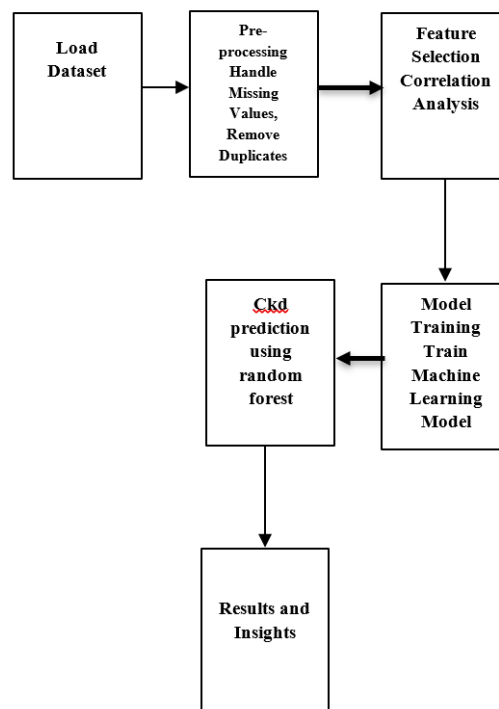


Figure 1: System Flow Diagram

5 Result Analysis

When coping with missing data, the proposed CKD prediction model, which combines Random Forest with a perceptron, surpasses existing imputation methods in terms of accuracy and durability. The model performs well on crucial assessment parameters such as accuracy, precision, recall, and F1-score, indicating its usefulness in detecting early CKD. When compared to other machine learning methods such as Support Vector Machines and Logistic Regression, the Random Forest-based technique consistently outperforms the others, particularly when dealing with limited clinical data. The hybrid method enhances feature selection and model generalisation, leading to more accurate predictions. These findings highlight the potential of machine learning methods for improving CKD diagnosis and imply that the proposed model might be a valuable tool in clinical decision making.

Chronic Kidney Disease Prediction with Model Comparison

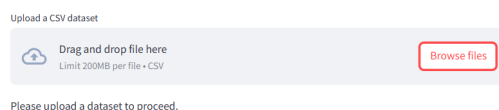


Figure 2: Dataset upload interface for chronic kidney disease prediction and model comparison

The image depicts the primary user interface for a Chronic Kidney Disease Prediction system that incorporates Model Comparison, developed through machine learning techniques. This interface enables users to upload a CSV (Comma-Separated Values) file containing medical or clinical data pertinent to the diagnosis of CKD.

Uploaded Dataset:

| | id | age | bp | sg | al | su | rbc | pc | pcc | ba | bgr |
|---|----|-----|----|-------|----|----|--------|----------|------------|------------|------|
| 0 | 0 | 48 | 80 | 1.02 | 1 | 0 | None | normal | notpresent | notpresent | 121 |
| 1 | 1 | 7 | 50 | 1.02 | 4 | 0 | None | normal | notpresent | notpresent | None |
| 2 | 2 | 62 | 80 | 1.01 | 2 | 3 | normal | normal | notpresent | notpresent | 423 |
| 3 | 3 | 48 | 70 | 1.005 | 4 | 0 | normal | abnormal | present | notpresent | 117 |
| 4 | 4 | 51 | 80 | 1.01 | 2 | 0 | normal | normal | notpresent | notpresent | 106 |

Figure 3: Data derived from the dataset uploaded

The image presents a preview of the uploaded dataset utilized by the Chronic Kidney Disease (CKD) Prediction System. This interface enables users to visually verify the contents of their dataset prior to further processing by the machine learning models.

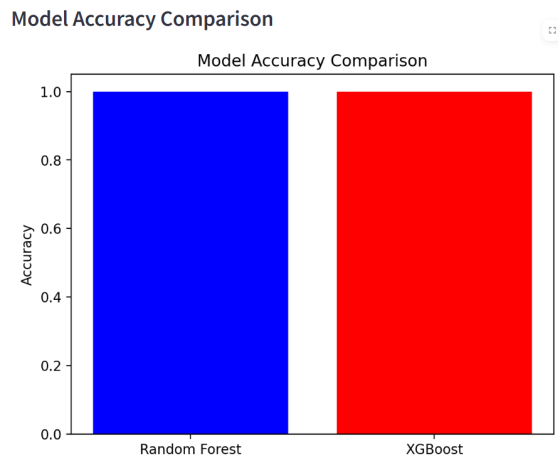


Figure 4: Accuracy comparison of Random Forest and XGBoost models.

The illustration presents a bar chart that contrasts the accuracy of two machine learning models—Random Forest and XGBoost—employed for predicting Chronic Kidney Disease. Both models exhibit high accuracy, nearing 1.0, which signifies outstanding performance on the dataset. The Random Forest model (represented by the blue bar) and the XGBoost model (represented by the red bar) perform nearly identically, demonstrating their effectiveness in managing medical classification tasks. This analysis aids in determining the most suitable model for clinical predictions based on accuracy.

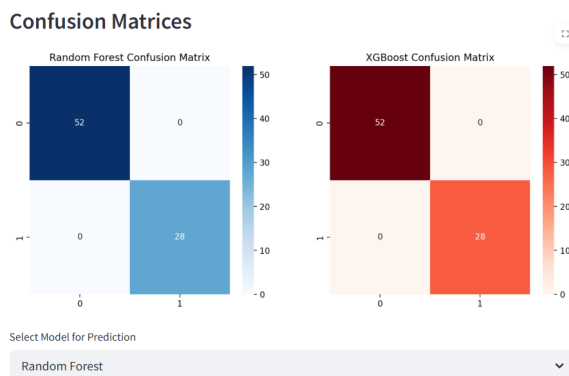


Figure 5: Confusion matrices of Random Forest and XGBoost models.

The illustration presents the confusion matrices for the Random Forest and XGBoost models employed in the prediction of Chronic Kidney Disease. Both models accurately classified all 52 negative and 28 positive cases, resulting in zero false positives and false negatives, which signifies a 100% accuracy rate and the absence of misclassifications. This underscores the robust predictive power of both models in diagnosing CKD.

| | | | |
|-------|---------|---|---|
| sc | 1.30 | - | + |
| sod | 138.00 | - | + |
| pot | 4.40 | - | + |
| hemo | 12.65 | - | + |
| pcv | 40.00 | - | + |
| wt | 8000.00 | - | + |
| rc | 4.80 | - | + |
| htn | 0.00 | - | + |
| dm | 3.00 | - | + |
| cad | 1.00 | - | + |
| appet | 0.00 | - | + |
| pe | 0.00 | - | + |
| ane | 0.00 | - | + |

Predict CKD

Predict CKD for All Records

Predict for All Records

Figure 6: Input form for predicting Chronic Kidney Disease (CKD) based on individual patient records.

The aforementioned interface showcases a machine learning-based tool created for the prediction of Chronic Kidney Disease (CKD) utilizing individual clinical records. Users are able to enter a range of medical parameters, including age, blood pressure, specific gravity, albumin, sugar levels, red and pus cell counts, blood urea, serum creatinine, hemoglobin, sodium, potassium, and other pertinent indicators related to kidney function and overall health. Additionally, categorical inputs such as the presence of hypertension, diabetes, anemia, and coronary artery disease are incorporated. Upon entering the necessary values, the 'Predict CKD' button enables the system to analyze the data through a trained machine learning model to assess the probability of CKD. This interactive interface is intended to aid clinical decision-making by providing real-time, data-driven predictions based on individual patient inputs, thereby improving the early detection and management of chronic kidney disease.

6 Conclusion

To summarise, the construction and validation of the Chronic Kidney Disease (CKD) prediction model utilising the Random Forest method represents a significant step forward in the application of machine learning in healthcare. The methodical methodology

of loading and pre-processing clinical data, identifying key features, and combining Random Forest with a perceptron to handle missing values resulted in a robust model for CKD prediction. The suggested method not only advances our understanding of predictive modelling in healthcare, but it also provides a practical answer for doctors faced with the issues of early CKD identification. The emphasis on a user-friendly interface, along with thorough system testing and deployment, demonstrates a dedication to developing a dependable tool for healthcare practitioners. This study has the potential to significantly affect the field by developing an effective and interpretable model for CKD prediction, thereby improving patient outcomes through early intervention and informed therapeutic decisions.

Create more accurate and dependable machine learning models. This might be accomplished by using larger and more diversified datasets, as well as creating new machine learning techniques and architectures. Create machine learning algorithms that can predict CKD development and the likelihood of developing complications. This would allow clinicians to identify high-risk individuals and treat earlier to avoid or postpone the development of problems. Create machine learning models that can be used to customise therapy recommendations for CKD patients. This would enable physicians to personalise treatment to the specific needs of each patient.

References

- [1] J. M. Bolarín, F. Cavas, J. S. Velázquez, and J. L. Alió, “A comprehensive unsupervised framework for chronic kidney disease prediction,” *Appl. Sci.*, vol. 10, no. 5, p. 1874, Mar. 2020.
- [2] M. Abdar, M. Zomorodi-Moghadam, X. Zhou, R. Gururajan, X. Tao, P. D. Barua, et al., “A machine learning framework for pulse detection during out-of-hospital cardiac arrest,” *Pattern Recognit. Lett.*, vol. 132, pp. 123–131, Apr. 2020.
- [3] J. P. Corradi, S. Thompson, J. F. Mather, C. M. Waszynski, and R. S. Dicks, “A machine learning methodology for diagnosing chronic kidney disease,” *J. Med. Syst.*, vol. 42, no. 12, p. 261, 2020.
- [4] M. Ashraf, S. M. Ahmad, N. A. Ganai, R. A. Shah, M. Zaman, S. A. Khan, et al., “A novel ensemble learning paradigm for medical diagnosis with imbalanced data,” Singapore: Springer, pp. 239–255, 2021.
- [5] H. Kwok, J. Coult, J. Blackwood, S. Bhandari, P. Kudenchuk, and T. Rea, “A survey of machine learning approaches applied to gene expression analysis for cancer prediction,” *Resuscitation*, vol. 147, pp. 104–111, Feb. 2022.
- [6] W. Zhu, L. Xie, J. Han, and X. Guo, “Analysis for disease gene association using machine learning,” *Cancers*, vol. 12, no. 3, p. 603, Mar. 2022.
- [7] D. Jain and V. Singh, “COVID-19 future forecasting using supervised machine learning models,” *Int. J. Comput. Appl.*, vol. 43, no. 6, pp. 524–536, Jul. 2021.
- [8] “Detecting keratoconus from corneal imaging data using machine learning,” Geneva, Switzerland, 2021.

- [9] F. Petropoulos and S. Makridakis, “Diabetes prediction using ensembling of different machine learning classifiers,” *PLOS ONE*, vol. 15, no. 3, Mar. 2022.
- [10] S. Kumari, D. Kumar, and M. Mittal, “Efficient prediction of cardiovascular disease using machine learning algorithms with Relief and LASSO feature selection techniques,” *Int. J. Cogn. Comput. Eng.*, vol. 2, pp. 40–46, Jun. 2021.
- [11] J. M. Bolarín, F. Cavas, J. S. Velázquez, and J. L. Alió, “Hybrid machine learning scheme for classification of BECTS and TLE patients using EEG brain signals,” *Appl. Sci.*, vol. 10, no. 5, p. 1874, Mar. 2020.
- [12] M. Abdar, M. Zomorodi-Moghadam, X. Zhou, R. Gururajan, X. Tao, P. D. Barua, and R. Gururajan, “Improvement in automated diagnosis of soft tissues tumors using machine learning,” *Pattern Recognit. Lett.*, vol. 132, pp. 123–131, Apr. 2020.
- [13] N. P. Tigga and S. Garg, “Machine learning approach for prediction of hematic parameters in hemodialysis patients,” *Procedia Comput. Sci.*, vol. 167, pp. 706–716, 2020.
- [14] M. S. Oh and M. H. Jeong, “Machine learning for clinical outcome prediction,” *Korean J. Med.*, vol. 95, no. 4, pp. 266–275, Aug. 2020.
- [15] H. Kwok, J. Coult, J. Blackwood, S. Bhandari, P. Kudenchuk, and T. Rea, “Machine learning to identify psychomotor behaviors of delirium for patients in long-term care facility,” *Resuscitation*, vol. 147, pp. 104–111, Feb. 2020.
- [16] W. Zhu, L. Xie, J. Han, and X. Guo, “Machine learning tools for long-term type 2 diabetes risk prediction,” *Cancers*, vol. 12, no. 3, p. 603, Mar. 2020.
- [17] A. Ogunleye and Q.-G. Wang, “Novel multi-center and threshold ternary pattern based method for disease detection using voice,” *IEEE/ACM Trans. Comput. Biol. Bioinf.*, vol. 17, no. 6, pp. 2131–2140, Nov. 2020.
- [18] L. Naranjo, C. J. Pérez, J. Martín, and Y. Campos-Roca, “Predicting Ayurveda-based constituent balancing in human body using machine learning methods,” *Comput. Methods Programs Biomed.*, vol. 142, pp. 147–156, Apr. 2020.
- [19] W. A. A. A. I. N. G. Misbasikandar, “Prediction of diabetes empowered with fused machine learning,” *IEEE Access*, vol. 1, pp. 1–12, 2020.
- [20] S. P. Chatrati, G. Hossain, A. Goyal, A. Bhan, S. Bhattacharya, D. Gaurav, and S. M. Tiwari, “Responsible and regulatory conform machine learning for medicine: A survey of challenges and solutions,” *J. King Saud Univ. Comput. Inf. Sci.*, Jan. 2020.

Cite this article:

Srinivasan S & Elakiya A, “Chronic Kidney Disease Prediction Using Machine Learning”, *Journal of Multidimensional Research and Review (JMRR)*, Vol.6, Iss.2, pp.197-208, 2025