A Comprehensive Review of Chronic Kidney Disease Using Machine Learning

Kaushal Sharma[1,*], Saurabh Shah[2], Snehlata Barde[3]

[1]Department of Computer Engineering, Parul University, Vadodara, India
[2]Talent & Career Development Cell, Parul University, Vadodara, India
[3]Department of Computer Engineering, Parul University, Vadodara, India
kaushal.sharma33479@paruluniveristy.ac.in

Abstract. Chronic Kidney Disease (CKD) is a significant public health issue leading to considerable morbidity and mortality and necessitating early diagnosis for its prevention from advancing to the end stage of renal failure. Recently, machine learning (ML) has also been applied as an effective technique to predict CKD by analyzing complicated correlations in medical data. This paper reviews the recent developments on ML based techniques for CKD prediction including supervised, unsupervised and hybrid methods. The performance of various models like logistic regression, support vector machine (SVM), decision tree (DT), random forest (RF), gradient boosting (GB) and deep learning (DL) models is discussed in terms of their accuracy, feature selection methods, data pre-processing methods, interpretability. Clinically important features such as blood pressure, serum creatinine, hemoglobin, albumin and demographic characteristics are particularly focused on, as they are well-established indicators of disease progression and have significant impact on model performance. It also discusses challenges like data imbalance, missing values, and the demand for explainable models to increase clinical trust. Experiments show that ensemble and hybrid learning approaches obtain the best prediction accuracy and robustness. In conclusion, we discuss future avenues and emphasize the need for incorporating data-driven, interpretable ML models into healthcare systems for early and accurate CKD prediction and clinical recommendations.

Keywords: chronic kidney disease, machine learning, classification algorithm, decision tree, explainable ai

1 Introduction

Chronic Kidney Disease (CKD) is a chronic disorder that causes progressive kidney function decline and leads to permanent end-stage kidney disease along with enormous social and economic burdens. From the epidemiologic perspective, its prevalence is rising worldwide and exceeds 13% in some populations due to an increasing incidence of diabetes, hypertension, and obesity, as well as aging populations. Laboratory assessments such as glomerular filtration rate (GFR), albuminuria and multiple other clinical surrogates are traditionally employed for diagnosing and staging CKD, but are complicated by inter individual variability, underestimation at early asymptomatic stages and potential errors in measurement. Consequently, substantial numbers of

J. Saxena et al. (eds.), *Proceedings of the International Conference on Sustainable Micro-Nano Materials & Innovative Technology (ICSUMMIT 2026)*, Atlantis Highlights in Engineering 44,
https://doi.org/10.2991/978-94-6239-727-9_16

patients are not diagnosed until the disease is already well advanced, limiting the potential for treatment and raising the odds of complications [1][2].

Recent progress in AI and ML has dramatically influenced medical prediction, enabling the construction of high-accuracy data-driven disease diagnosis and prognosis models. ML methods (e.g., "support vector machines (SVM), random forests, decision trees and recent ensemble, deep learning-based approaches) are anticipated to reveal hidden knowledge from complex and heterogeneous clinical data providing potential solutions for early detection and personalized risk- profiling of CKD. Several XAI tools have been developed and they have increased the interpretability and have explained to clinicians why some biomarkers and demographic features influence model predictions [3][4][5][6][7].

Despite there is an increasing interest, several issues still prevent the generalization of this machine learning based CKD prediction applications. These are among others, issues related to missing or incomplete data, no external face validation and incomplete consideration of real-world clinical constraints, algorithmic bias and interpretability of composite models. Closing these gaps will require complete literature surveys and feature discussions as well comparative evaluations for systematically assessing the performance, generalizability and clinical relevance of the current state of the art. In this review, we make an effort to (1) provide a comprehensive survey of the recent ML based models in the field of CKD prediction, (2) highlight the strengths and limitations of the employed methodology, (3) identify important factors influential to the prediction capability, and (4) draw an integrative and critical view to provide suggestions for procuring future research works as well as for clinical application[8][9].

2 Literature Review

The application of ML and DL methods for predictions and diagnosis of chronic kidney disease (CKD) has been extensive, providing good results in recent years. Yang et al. [1] synthesized a multi-algorithm based on a series of ML and DL techniques for the UCI CKD dataset, traditional ML models including Random Forest (RF), Support Vector Machine (SVM), AdaBoost and eXtreme Gradient Boosting (XGBoost) were found to be much more effective with higher accuracy and lower training time as compared to the DL models for small clinical datasets. Similar findings are reported in scathing review studies on prediction of chronic disease in which efficient data pre-processing, handling of missing values and feature selection were the most predictive performance contributing factors [6]. In addition, coevolutionary feature engineering with ensemble learning can greatly improve the early detection of chronic diseases like CKD by increasing the robustness and generalization of models [7].

Different studies have concentrated on optimized, hybrid and deep learning-based models to further enhance the prediction accuracy for CKD detection. Neural networks based on optimization, such as physics informed neural networks, particle swarm optimization with neural networks, Boolean logic with back propagation and cuckoo search, have shown better performance than traditional classifiers [8]. Deep learning–based risk models for unfavourable CKD outcomes also exist, which enable the discernment of interpretable risk predictors and corresponding insights into disease progression and severity [9]. Furthermore, data-driven early diagnosis models with

explainable AI methodologies demonstrated enhanced prediction performance with facilitating transparent decision-making in the medical field [10]. This implies that hybrid and DL based models may also be promising with the adequate data, optimization techniques and explainability in place.

In addition to enhancing predictive performance, a growing body of the literature highlights interpretability/explainability and data privacy as key factors for fostering clinical trust and real-world adoption. Jawad et al. [4] utilized Explainable AI (XAI) on ensemble models to predict CKD and extracted important clinical biomarkers such as albumin, serum creatinine, blood pressure and age, thereby making the model interpretable. Zhao et al. [3] developed an interpretable DBRB diagnostic model based on expert knowledge and hierarchical rule reasoning that achieves high diagnostic accuracy and is interpretable. With the privacy implication addressed, Puri et al. [2] presented a distributed privacy-first machine learning framework over block chain and IPFS for collaborative learning among healthcare institutes without sharing sensitive patient data. Overall, these studies are indicative of a transition towards precise, explicable, and privacy-centric AI techniques for trustworthy CKD prediction and diagnosis.

Overall, the above-discussed literature indicates that ML-based methods, especially ensemble and optimized models are suitable for early prediction of chronic kidney disease. Current advances for interpretable and explainable models have enhanced clinical transparency, but are typically assessed on small datasets and without being validated in real-world scenarios. In addition, privacy-preserving schemes are still at the conceptual level, and they introduce complex process and computational overheads without standard performance evaluation. Thus, an integrated, scalable and explainable CKD prediction framework is required to offer a good trade-off among prediction accuracy, computational complexity and data privacy, to meet the demands of practical clinical utilization.

Table 1 Summary of the Method Used for Prediction

Authors (year)	Dataset	Methods Used	Performance (Proposed model)
Khan et al., 2020	UCI-CKD (Apollo Hospital, India)	NBTree, J48, SVM, LR, MLP, Naïve Bayes, CHIRP	CHIRP:99.75%, SVM: 98.25%
Islam et al., 2023	UCI-CKD (Apollo Hospital, India)	12 ML classifiers, incl. XGBoost	XGBoost: 98.3% (Precision/Recall/F1 = 0.98)
Debal & Sitote, 2022	St. Paul's Hospital, Addis Ababa,Ethiopia	RF, SVM, DT + Recursive feature elimination	RF + RFE: 99.75%
Jawad et al., 2025	UCI-CKD (Apollo Hospital, India)	Ensemble ML + Explainable AI	RF Best Among All
Antony et al., 2021	CKD clinical dataset	K-Means, DB-Scan, Isolation Forest	K-Means + feature selection: 99%

Vellela et al., 2023	UCI-CKD (Apollo Hospital, India)	Random Forest, Logistic Regression, KNN	RF best among tested
Khade et al., 2023	Real hospital dataset (800 patients)	Hybrid SAG- BPSO + Cuckoo Search + BPNN	98.07%
Moreno-Sánchez, 2023	UCI-CKD (Apollo Hospital, India)	XGBoost + Explainable AI	99.2% ± 0.8 (CV); 97.5% (test)
Wibawa et al. (2017)	UCI-CKD (Apollo Hospital, India)	KNN + AdaBoost	98.1%

3 Methodology

Supervised learning, unsupervised learning, ensemble learning as well as hybrid methods that combine two or more of the above exist in the methodologies applied to study CKD prognosis. All of these methodologies share the same ultimate aim, which is the refining of diagnostic accuracy, interpretability and clinical impact. There is a strong focus on CKD research and this is mainly done with supervised learning techniques. Conventional classifiers Logistic Regression, Decision Trees, Support Vector Machines and Naïve Bayes were extensively experimented (Khan et al., 2020). In these, Support Vector Machines (SVM) and Random Forest (RF) have proven to be the best for early-stage classification (Debal & Sitote, 2022). Logistic Regression provided interpretable results, but failed to achieve high overall accuracy. Ensemble methods achieved significant performance gains.

Islam et al. (2023), twelve classifiers were employed and it has been shown that (XGBoost) has achieved a 98.3% accuracy, which is better than that of single learners in preserving prediction power and explanation capability. Unsupervised learning has also been studied.

Antony et al. (2021) applied clustering algorithms like KMeans, DBScan and auto encoders and obtained maximum accuracy of 99% with feature selection. This will highlight the potential of label-free methods for CKD screening. Neural networks were improved upon with optimization techniques in hybrid models.

Khade et al. (2023) the accuracy of the proposed SCGB-BPNN-CS model was 98.07% on the real hospital data. Such frameworks involve exploitation-exploration trades, semi regularization overfitting and further generalizability. Explainable AI (XAI) approaches finally define a new trend.

Moreno-Sánchez (2023) demonstrated that with three features alone XGBoost can reach an extremely high accuracy (99.2%), highlighting cost-effectiveness and interpretability. In summary, approaches to predicting CKD vary from simple classifiers to complex hybrid ensemble models, and are showing growing interest in interpretability and clinical confidence. The move away from accuracy-focused architectures to more interpretable and resource-efficient models is indicative of how mature the field has become.

3.1 Dataset

The accuracy of the ML models in predicting CKD is heavily dependent on the quality and quantity of data used. Most of the research has been conducted using the UCI CKD dataset consisting of 400 instances, with 24 clinical signs, demographic and laboratory details such as serum creatin, hemoglobin, albumin and specific gravity (Khan et al., 2020; Islam et al., 2023). Although very common and accessible to check the robustness, occasionally work also cover real hospital dataset. Khade et al (2023). Indian clinical data collected from Indian as well as American patients) to the literature. Healthcare institutions should adopt it to make it clinically applicable. Even use hospital data on a stage-by-stage basis and can be stage by stage. Similarly, Jawad et al. (2025) conducted experiments on various local datasets in addition to UCI to test whether their XAI ensemble models can be extended, improving generalizability. Furthermore, despite Antony et al. (2021) utilized the Chandigarh University CKD dataset and Moreno-Sánchez (2023) demonstrated good interpretability using XGBoost on the UCI dataset with merely 3 features.

Together these works demonstrate that the quality of models depends on the dataset Variety. The preponderance of UCI based evaluations further reinforces the overwhelming large- and meta-registers of trials.

Table 2 Classification of Features in the UCI CKD Dataset.

Feature Type	Attribute Name	Abbr.	Unit / Values
Numerical (11)	Age	age	Years
	Blood Pressure	bp	mm/Hg
	Blood Glucose Random	bgr	mgs/dl
	Blood Urea	bu	mgs/dl
	Serum Creatinine	sc	mgs/dl
	Sodium	sod	mEq/L
	Potassium	pot	mEq/L
	Hemoglobin	hemo	gms
	Packed Cell Volume	pcv	%
	White Blood Cell Count	wc	cells/cumm
	Red Blood Cell Count	rc	millions/cmm
Nominal (13)	Specific Gravity	sg	1.005 - 1.025
	Albumin	al	0, 1, 2, 3, 4, 5
	Sugar	su	0, 1, 2, 3, 4, 5
	Red Blood Cells	rbc	Normal, Abnormal
	Pus Cell	pc	Normal, Abnormal
	Pus Cell Clumps	pcc	Present, Not Present
	Bacteria	ba	Present, Not Present

	Hypertension	htn	Yes, No
	Diabetes Mellitus	dm	Yes, No
	Coronary Artery Disease	cad	Yes, No
	Appetite	appet	Good, Poor
	Pedal Edema	pe	Yes, No
	Anemia	ane	Yes, No
Target	Class	class	CKD, Not CKD

3.2 Data Pre-Processing

Data cleaning has been crucial for learning models to predict chronic kidney disease (CKD). Due to the presence of noise, missing values and class imbalance in clinical data, preprocessing can have a great impact on the performance of the constructed models. The issue of missing values was addressed in the majority of solutions, especially for medical data. For instance, Khan et al. (2020) filled in missing features with the mean or mode of features to maintain the consistency of the dataset. In a similar fashion, Islam et al. (2023) proposed imputation methods to prevent classifier biased due to missing values. In addition, feature selection and dimensionality reduction have also been widely applied to enhance the efficiency and interpretability. The first extant (25 features) more selectively reduced the clinical features than the second extant (= top 30%) which was reduced (122%) of the size of feature space that enhanced the CA of the T2/3 images and reduced the computational time load.

Moreno-Sánchez (2023) raised the bar even higher by demonstrating that 3 variables only hemoglobin, date arteria blood pressure and age with XGBoost sufficed to obtain a remarkable 99.2 % accuracy. Various normalization and standardization techniques were applied to the features to reduce the effect of the followed scales on the heterogeneous features. Debal & Sitote (2022) highlighted the significance of univariate scaling on the continuous predictors to regularize the classification boundaries in RF and SVM so also on the latter. Finally, the problem of high CKD staging class imbalance was aided with-in sampling techniques. Khade et al. (2023) put forth an optimization-based preprocessing technique to correct bias in minority classes for training hetero-generous models. In general, preprocessing enables the models to predict CKD robustly while being interpretable and clinically generalizable. One separate but consistent finding from these studies is that imputation, feature reduction and normalization are critical in designing a robust ML framework for CKD.

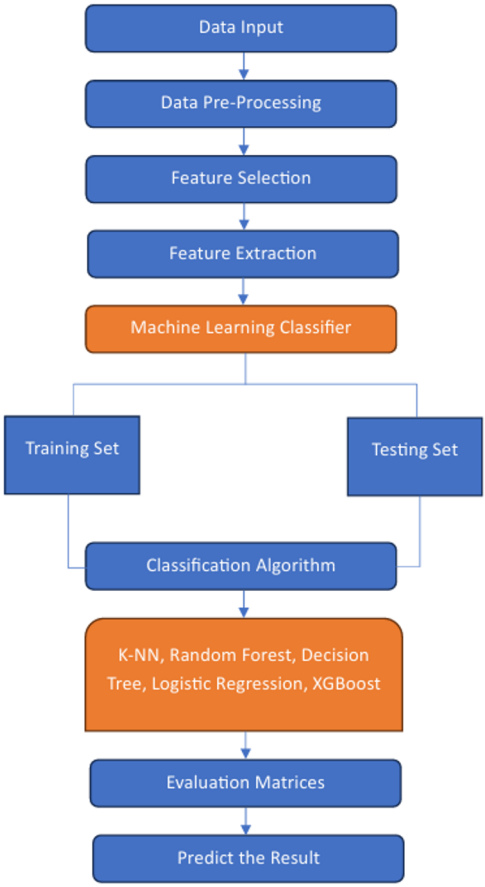

Fig. 1. Common Methodological Steps in ML-Based CKD Prediction Systems

3.3 Feature Selection

Retrospectively, some features of the studies appear to be or could be direct predictors of CKD. Hemoglobin is the strongest single predictor variable of CKD status and is also associated with progression of disease and complications, such as anemia. A routine measurement is urine specific gravity as an indication of renal concentrating ability and among the earliest markers of renal impairment. Hypertension is also a well-established leading clinical predictor and like iq is automatically selected in an identical fashion from each feature importance/model explainability tool further supporting its emerging from the noise [1][3][7][10][11].

Additional factors, such as albuminuria, serum creatinine and age are also highly predictive. It may be based on feature selection algorithms (e.g., PCA, RFE and SHAP) and cohorts. The inclusion of laboratory variables (PCV, Na, K and urea) and comorbidity (diabetes, coronary artery disease) enhance prediction in few models, yet could introduce complexity and compromise external validity if not appropriately

validated. The interpretability of a model on its own merit is also preferred by a concise, high importance feature set, as demonstrated in studies with explainable AI or model selection strategies [2][4][7][10][11].
The performance of the model in a clinical setting may be affected by the selection of the feature selection methods. Filter based (correlation, Gain Ratio), wrapper based (RFE), embedded based (SHAP, Random Forest) reduce input variables by a factor of two or three without any loss in prediction power. Clinical consultation and validation remain essential especially when the features prioritized by the models appear to be pathophysiological relevant [3].

4 Discussion

While accuracy is an important reference standard, the clinical relevance in real life of prediction models for CKD is influenced by other aspects. Explainability is important, as some models (e.g., Logistic Regression, DT, NB) are inherently explainable, while in ensemble and kernel-based models (e.g., RF, XGBoost, SVM, and KNN) it is necessary to apply post-hoc explainability methodologies such as SHAP or LIME. A well calibrated probability is just as important in making clinical decisions as it is in other domains. Logistic Regression is known to provide well-calibrated probabilities, but Random Forest and XGBoost have the potential to produce overconfident predictions, which needs to be addressed by further calibration. Furthermore, missing data handling is still a challenge, with SVM and KNN being affected by missing entries, while tree based classifier is more robust.
In terms of deployment, shallow models' linear regression and naïve bayes are better suited for real time clinical adoption due to their low computational cost and KNN is not convenient since it is computationally intensive and requires storing the entire dataset during the inference stage. Moreover, there is a prominent reliance on the UCI CKD dataset for both training and testing purposes across studies, raising concerns about generalizability since external validation as well as evaluation under feature constraint scenarios are missing to a great extent. Future work may consider the emphasis on multi-dataset validation, better calibration and Electronic Health Record integration to further increase the applicability in the real world.

5 Conclusions

In summary, the use of machine learning for CKD prediction holds great promise for improving early screening, risk stratification and healthcare resource management. Over the past two decades, hybrid and ensemble methods have shown superior performance compared to single model methods by combining classical methods with metaheuristics, feature selection methods, and explainable AI techniques. Important clinical factors, like haemoglobin, specific gravity and hypertension. It recurred as the most influential factors which result in the development of highly accurate, parsimonious and interpretable diagnostic models [3][6][7][8][11]. Despite this progress, there are still challenges to be tackled, such as robust external validation,

appropriate dealing with missing and unbalanced data as well as the integration of heterogeneous data. Model interpretability must also be feasible to achieve in order to obtain clinician trust and real-world implementation.

Concerning the future, focus should be put on the conduction of multicentre and longitudinal studies for model validation in heterogeneous populations. In addition, there remains considerable potential for the inclusion of more complex data types such as genomic data and health monitoring data in real time to enhance predictive accuracy and robustness. More advanced methods, such as deep learning and ensembles, are to be investigated for the analysis of complex high dimensional data. Solving key issues such as data privacy, transparency and user centric design will be critical to converting machine learning innovations into scalable clinical solutions. In the end, these advances may pave the way to transform CKD management by enabling earlier detection, tailored therapy and improved patient outcomes.

Disclosure of Interest: The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Evans, M., Lewis, R. D., Morgan, A. R., Whyte, M. B., Hanif, W., Bain, S. C., Davies, S., Dashora, U., Yousef, Z., Patel, D. C., & Strain, W. D. (2022). A Narrative Review of Chronic Kidney Disease in Clinical Practice: Current Challenges and Future Perspectives. Advances in therapy, 39(1), 33–43.

2. Zsom, L., Zsom, M., Salim, S. A., & Fülöp, T. (2022). Estimated Glomerular Filtration Rate in Chronic Kidney Disease: A Critical Review of Estimate-Based Predictions of Individual Outcomes in Kidney Disease. Toxins, 14(2), 127.

3. K. M. Tawsik Jawad, A. Verma, F. Amsaad and L. Ashraf, "A Study on the Application of Explainable AI on Ensemble Models for Predictive Analysis of Chronic Kidney Disease," in IEEE Access, vol. 13, pp. 23312-23330, 2025, doi: 10.1109/ACCESS.2025.3535692.

4. L. Antony et al., "A Comprehensive Unsupervised Framework for Chronic Kidney Disease Prediction," in IEEE Access, vol. 9, pp. 126481-126501, 2021, doi: 10.1109/ACCESS.2021.3109168.

5. S. S. Vellela, L. R. Vuyyuru, K. B. S. K, N. MalleswaraRaoPurimetla, L. Dalavai and M. V. Rao, "A Novel Approach to Optimize Prediction Method for Chronic Kidney Disease with the Help of Machine Learning Algorithm," 2023 6th International Conference on Contemporary Computing and Informatics (IC3I), Gautam Buddha Nagar, India, 2023, pp. 1677-1681, doi: 10.1109/IC3I59117.2023.10397974.

6. A. Khade, A. V. Vidhate and D. Vidhate, "Design of an Optimized Self-Acclimation Graded Boolean PSO with Back Propagation Model and Cuckoo Search Heuristics for Automatic Prediction of Chronic Kidney Disease," in Journal of Mobile Multimedia, vol. 19, no. 6, pp. 1395-1414, November 2023, doi: 10.13052/jmm1550-4646.1962.

7. P. A. Moreno-Sánchez, "Data-Driven Early Diagnosis of Chronic Kidney Disease: Development and Evaluation of an Explainable AI Model," in IEEE Access, vol. 11, pp. 38359-38369, 2023, doi: 10.1109/ACCESS.2023.3264270.

8. B. Khan, R. Naseem, F. Muhammad, G. Abbas and S. Kim, "An Empirical Evaluation of Machine Learning Techniques for Chronic Kidney Disease Prophecy," in IEEE Access, vol. 8, pp. 55012-55022, 2020, doi: 10.1109/ACCESS.2020.2981689.
9. Islam, M. A., Majumder, M. Z. H., & Hussein, M. A. (2023). Chronic kidney disease prediction based on machine learning algorithms. Journal of pathology informatics, 14, 100189. https://doi.org/10.1016/j.jpi.2023.100189
10. Dritsas, E., & Trigka, M. (2022). Machine Learning Techniques for Chronic Kidney Disease Risk Prediction. Big Data and Cognitive Computing, 6(3), 98. https://doi.org/10.3390/bdcc6030098
11. Debal, Dibaba & Sitote, Tilahun. (2022). Chronic kidney disease prediction using machine learning techniques. Journal of Big Data. 9. 10.1186/s40537-022-00657-5.
12. Satria Wibawa, Made & May Sanjaya, I Made & Putra, I Made. (2017). Boosted classifier and features selection for enhancing chronic kidney disease diagnose. 1-6. 10.1109/CITSM.2017.8089245.