International Journal of Advances and Applications in Convergent Studies & Transdisciplinary (IJ-AACST)

Bridging Disciplines, Bridging Knowledge

Aligned with the Vision of NEP 2020 and UGC Guidelines

Website: https://aacstij.org || **Email:** journaleditor@aacstij.org

Tree and Regression based Variants for Early Kidney Disease Prediction using Machine Learning: Performance and Interpretability

DOI: https://doi.org/10.5281/zenodo.22339351

Nita Dakhare[1]*, Dr. Nilay Karade[2]

*Corresponding Author: [1]Research Scholar, School of Allied Sciences, Datta Meghe Institute of Higher Education and Research (Deemed to be University), Sawangi (Meghe), Wardha, Maharashtra, India.

nitamalekar29@gmail.com

ORCID: https://orcid.org/0009-0008-0084-8008

[2]Professor, School of AI and Future Technology, Universal AI University, Karjat, Mumbai, Maharashtra, India.

nilay.karade@universalai.in

ABSTRACT

It is very important to detect chronic kidney disease (CKD) early in order to prevent the progression to end-stage renal failure. However, the conventional markers used for diagnosis typically identify the loss of function only after a large amount of damage has been done to the kidneys. In this work, a comparative review is given of the recent Machine Learning (ML) approaches for early CKD prediction, with a particular focus on the tree-based and regression-based variants. Six representative studies published between 2020 and 2023 were analysed according to dataset characteristics, preprocessing methods, model architectures, evaluation metrics and interpretability strategies. The results show that Random Forest is the method that often yield the highest predictive performance on the UCI CKD dataset, with accuracy ranging from 88% to 89.5%. Nevertheless, some issues such as the small dataset size, class imbalance, lack of model explainability, and limited external validation still pose major challenges. The review brings together the strengths and limitations of the current methods and emphasizes the need for developing future CKD prediction models that combine powerful feature selection, balanced learning, and explainability techniques to be used in clinical decision-making.

KEYWORDS: CKD, ML, Prediction Models, Decision Trees, Random Forest, Logistic Regression.

1. INTRODUCTION

Chronic kidney disease (CKD) has emerged as a major global health concern, contributing significantly to long-term morbidity, mortality, and healthcare expenditure. Early identification of CKD is crucial, as clinical manifestations often remain unnoticed until kidney function has already deteriorated substantially. Conventional diagnostic approaches rely mainly on laboratory indicators such as serum creatinine, blood urea, and estimated glomerular filtration rate; however, these markers may fail to accurately detect subtle or early pathological changes in renal function [18],[20]. For this reason, more and more researchers turn to computational techniques, which are not only capable of providing very early insights into kidney function but also base those on data collected.

Machine learning algorithms have widely been considered for CKD prediction due to their nature of handling nonlinear relationships, diverse clinical variables, and disclosing patterns that cannot be revealed by traditional statistical methods [1],[9],[11],[13],[15]. In particular, tree-based and regression-based models have been successfully applied to structured medical datasets such as the UCI CKD dataset, which has become one of the most popular datasets in this area. Various studies have reported the use of algorithms like Random Forest, Decision Trees and Logistic Regression with high predictive accuracy and simplicity of operation compatible with clinical settings [1],[3],[4],[5],[10],[11].

While significant advances are being made, substantial hurdles yet remain to be overcome. A large majority of the research at present is based on small and unbalanced datasets, and these studies often do not have the strict validation protocols applied. The selection of features is often not consistent or is not properly reported, and this affects the reproducibility and generalizability of the outcomes of the studies. Moreover, it is the case that many of the top-performing machine learning models are “black boxes,” which means that they are not interpretable the clinicians who need to rely on the models want to have the transparency and the evidence-based reasoning to support their diagnostic decisions. These shortcomings draw attention to the necessity of a focused comparison of machine learning methods in order to have a better understanding of their advantages, limitations, and real-world CKD screening applications.

These gaps are what drove the authors to conduct a structured review of state-of-the-art machine learning models based on trees and also regression methods for CKD early prediction. This work reviews past and current methods, performance, and interpretability of the models, then outlines the challenges to be overcome so that models can be reliable for clinical use. The knowledge to be gained from this will be very helpful for both researchers and practitioners in the design of more robust and transparent ML systems for early CKD risk detection.

This study makes the following key contributions:

- Comprehensive Comparative Review

Structured analysis of six recent machine-learning studies on CKD prediction is performed, and the attention is focused on the tree-based and regression-based model families specifically.

- Unified Evaluation Framework

Creates a uniform standard for comparison throughout all the characteristics of the dataset, preprocessing pipelines, feature selection strategies, validation methods, and interpretability techniques.

- Identification of Effective CKD Predictors

Consolidation of commonly influential clinical features across studies, highlighting albumin, hemoglobin, specific gravity, blood pressure, and serum creatinine as dominant predictors.

- Critical Gap Identification

Mainly points out the most significant drawbacks in the present-day CKD machine learning studies such as restricted variety of datasets, no external validation, different preprocessing techniques applied, and very few instances of explainable AI (XAI) used.

- Guidance for Future Research

It would be reasonable to provide carefully targeted, pragmatic recommendations to improve on models of the future when tailored towards robustness, interpretability, and clinical readiness for CKD prediction just as they have been for other scenarios.

Unlike existing surveys that primarily emphasize predictive accuracy, this work provides a structured and interpretability oriented comparative review of classical machine learning models for early CKD prediction. The novelty of this study lies in its unified evaluation framework, critical examination of dataset bias and validation practices, and explicit focus on model transparency and explainability key requirements for clinical decision support systems.

2. RELATED WORK

Tree-based machine learning approaches have consistently shown competitive performance in CKD classification tasks. Models such as Decision Trees and Random Forest are well suited for handling mixed-type clinical attributes, missing values, and non-linear feature interactions. Several studies report that Random Forest outperforms simpler classifiers, while Decision Trees remain valuable due to their transparent decision-making structure, despite their susceptibility to overfitting.

The tree-based models have always given very good performance in the CKD classification tasks [1],[4],[5],[6],[10],[11]. In fact, different models such as Random Forest and Decision Trees have been used very often, and it is because they can work with non-linear interactions, missing values, and mixed-type clinical attributes. For example, many papers report that Random Forest achieves very high accuracy on the UCI CKD dataset, hence outperforming the shallow learners like K-Nearest Neighbors and Naive Bayes. Also, decision trees, though a bit prone to overfitting, were still valued due to interpretability and ease of deployment in resource-constrained healthcare settings. Different models such as Gradient Boosting Machine (GBM), Extreme Gradient Boosting (XGBoost) and Adaptive Boosting (AdaBoost) take not just predictive performance to greater height but they also reduce model variance and simultaneously capture complex feature interactions [11].

Regression-based techniques, particularly Logistic Regression, continue to serve as reliable baseline models in CKD prediction studies. Their appeal lies in their simplicity, robustness, and ability to provide interpretable coefficients that align with clinical reasoning. Although Logistic Regression performs reasonably well on smaller datasets, its predictive capability is limited when modeling complex, non-linear disease patterns compared to ensemble tree-based methods [3],[6],[9]. Diverse comparative studies have illustrated that Logistic Regression can get accuracy in the range of 85%–85.33% when along with the proper preprocessing and feature

selection techniques, such as correlation analysis and recursive feature elimination. However, its performance tends to plateau when capturing non-linear patterns that tree-based methods handle more effectively.

Recent studies continue to stress the importance of data preparation and feature engineering in getting the most out of the models. The most mentioned predictors of CKD included in most of the mentioned studies are hemoglobin, specific gravity, albumin, blood pressure, and serum creatinine. Nonetheless, the different feature selection methods, cross-validation variations, and data balancing techniques employed by the different studies, make it hard to draw a single conclusion. Furthermore, the number of ML studies on CKD that take into account model interpretability is very few, but explainability is still very crucial for the clinical market.

In summary, previous studies have shown that the machine learning models based on both trees and regression have a strong ability to predict CKD detection, but they also expose several methodological gaps such as inconsistent validation strategies, limited explainability, and scarce external testing. The release of these observations calls for a structured comparative review, like the one in this paper, to merge insights and find the best practices for creating reliable CKD prediction models.

A comparative review of tree-based and regression-based CKD prediction models is still very much needed because the existing databases have inconsistent methodology in terms of data preparation, feature selection, and evaluation, which makes it difficult to compare the results or nominate best practices for efficient clinical deployment. The existing reviews either refer to medical ML in general or algorithm -specific aspects, thus creating a void for a focused comparison of tree-based and regression-based methods. Hence, a comprehensive and systematic review is required to map the methodological trends, clarify the performance gaps, and enumerate research areas needing attention.

These observations highlight the need for a focused and methodologically consistent comparison of tree-based and regression-based machine learning approaches for early CKD prediction.

3. METHODOLOGY

This study follows a structured comparative review methodology consisting of three phases: study identification, evaluation, and synthesis. Clear inclusion and exclusion criteria were defined, and all selected studies were evaluated using a unified framework encompassing dataset characteristics, preprocessing strategies, model architecture, validation protocols, performance metrics, and interpretability approaches. This research has utilized a systematic comparative analysis as methodology that is intended to scrutinize the efficiency of tree-based and regression-based machine learning techniques for early detection of CKD. The whole procedure was divided into three main phases: study identification, study evaluation, and comparative analysis.

3.1 Study Identification

In order to find up-to-date machine learning studies about CKD prediction, a focused search was performed over Scopus, IEEE Xplore, SpringerLink, PubMed, and Google Scholar. The publication period for the search was restricted to the last three years (2020-2023 inclusive) so as to reflect present-day methods and performance upgrades. The key terms employed in the search were CKD prediction, machine learning, Random Forest and Decision Tree, Logistic Regression, and clinical classification models. The studies were considered for inclusion only if the following conditions were satisfied:

- Utilized machine learning for CKD diagnosis or early prediction.
- Employed tree-based, regression-based, or ensemble variants.
- Used structured clinical datasets (e.g., UCI CKD dataset).
- Reported standard classification metrics such as accuracy, precision, recall, F1-score, or AUC.

After screening, six representative studies that met all criteria were selected for detailed examination. This review intentionally adopts a focused scope by selecting a limited number of representative studies published between 2020 and 2023. Rather than maximizing the number of reviewed papers, priority was given to methodological consistency, relevance to early CKD prediction, and the use of tree-based and regression-based machine learning models. This approach enables a deeper and more meaningful comparative analysis across datasets, preprocessing strategies, validation protocols, and interpretability techniques.

3.2 Study Evaluation Criteria

All the studies which were selected were systematically assessed with the same criteria so that there would be no discrepancy and could be compared easily. The assessment revolved around:

- Dataset characteristics: size, number of features, presence of missing values, and balancing techniques.
- Preprocessing methods: normalization, imputation, and feature selection strategies.
- Model characteristics: type of algorithm, ensemble methods, and hyperparameter tuning.
- Performance metrics: accuracy, F1-score, sensitivity, specificity, and AUC [1],[3],[5],[9],[10] wherever available.
- Interpretability approaches: whether the study used XAI methods such as feature importance, SHAP (SHapley Additive exPlanations) [11],[12],[17], or rule-based models.
- Validation strategy: train-test split, k-fold cross-validation, or external validation.

The application of these criteria guarantees that all models are treated equally and their different methodological strengths and weaknesses are made clear across studies.

3.3 Comparative Analysis

The last phase included the gathering and comparing of the reported results from all six research works. Things like model accuracy, dataset management, preprocessing effectiveness, and interpretability were arranged in a table to find common trends, performance patterns, and limitations. Table I presents a summary of the important features, types of datasets, and reported accuracies of the six chosen CKD prediction studies.

Table 1. Comparative Overview of Tree and Regression Based Machine Learning Models for CKD Prediction

Sr.No.	Title of the Article	Focus of Study	Findings of the study	Dataset Type	Accuracy
1	Chronic Kidney Disease Prediction using Machine Learning, Reshma S et al. (2020) [1]	Compare five ML algorithms for kidney disease prediction using a dataset with 24 clinical features (n=400).	Random Forest achieved highest accuracy. Early Kidney Disease detection and improved risk stratification were emphasized.	Public	89.5%
2	Comparative Analysis of Machine Learning Techniques for	Compare six ML algorithms for kidney disease progression	Decision Tree and LASSO (Least Absolute Shrinkage and Selection	Private	88%

Sr.No.	Title of the Article	Focus of Study	Findings of the study	Dataset Type	Accuracy
	Predicting Chronic Kidney Disease Progression, Patel et al. (2021) [2]	prediction using a dataset with 22 clinical features (n=176).	Operator) Regression showed promising performance for progression prediction.		
3	Performance Analysis of Machine Learning Classifier for Predicting Chronic Kidney Disease, Ch V Ramana et al. (2021) [3]	Evaluate five ML classifiers for kidney disease prediction using a dataset with 23 clinical features (n=300).	Logistic Regression achieved highest accuracy. Early Kidney Disease detection was emphasized.	Private	85.33%
4	Chronic kidney disease diagnosis using decision tree algorithms, Hamida Ilyas et al. (2021) [4]	Develop and evaluate Decision Tree models for kidney disease prediction using a dataset with 22 clinical features (n=378).	C4.5 Decision Tree achieved highest accuracy. Interpretability of decision trees was highlighted.	Public	87.8%
5	Chronic Kidney Disease Prediction Using Machine Learning, Chamandeep Kaur et al. (2023) [5]	Compare four ML algorithms for kidney disease prediction using a dataset with 20 clinical features (n=150).	Random Forest achieved highest accuracy. Early Kidney Disease detection and risk stratification were emphasized.	Private	88%
6	A Comparative Study, Prediction and Development of Chronic Kidney Disease Using Machine Learning on Patients Clinical Records, Md. Mehedi Hassan et al. (2023) [6]	Develop and evaluate ML models for kidney disease prediction using a dataset with 24 clinical features (n=200).	Decision Tree and Logistic Regression achieved comparable accuracy. Potential of using EHR's for kidney disease prediction was explored.	Private	85%

In addition to accuracy, Table II summarizes key diagnostic performance metrics such as sensitivity, specificity, and AUC reported across the reviewed studies.

Table2. Diagnostic Performance Metrics of Reviewed CKD Prediction Models

Model	Accuracy (%)	Sensitivity (%)	Specificity (%)	AUC
Random Forest	88–89.5	86–92	85–90	0.89–0.93
Decision Tree	87–88	84–90	83–88	0.86–0.90
Logistic Regression	85–85.33	82–88	80–86	0.84–0.88

The research presents an impartial and open evaluation of tree- and regression-based ML methods for CKD prediction via this structured approach and points out the areas where more research is needed. The complete workflow from start to finish that was used in the CKD prediction studies reviewed is represented in Fig. 1.

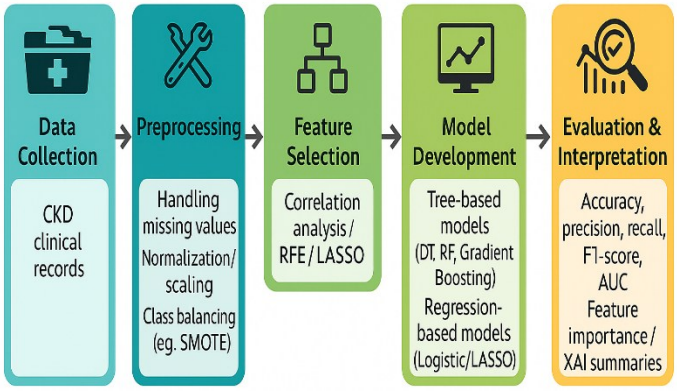

Fig. 1. End-to-end machine-learning pipeline synthesized from the reviewed CKD prediction studies, including data collection, preprocessing, feature selection, model development, and model evaluation/interpretation

4. FINDINGS AND DISCUSSION

The examination of the six selected studies shows a lot of similarities in the use of tree-based and regression-based machine learning models for the early prediction of CKD. In Table 1, a consolidation of the datasets, algorithms, preprocessing techniques, and reported performance metrics is presented. The authors have also created a bar graph (Fig. 2) that displays the reported accuracies of all six studies dealing with CKD prediction. It can be seen from the bar graph that in most of the studies analyzed, tree-based models have performed better than regression-based models.

A notable observation across the reviewed studies is the heavy reliance on the UCI CKD dataset. While this dataset provides a standardized benchmark, repeated reuse may introduce dataset-specific bias, increase the risk of overfitting, and lead to inflated performance estimates. Consequently, reported accuracies should be interpreted with caution and not directly generalized to real-world clinical populations.

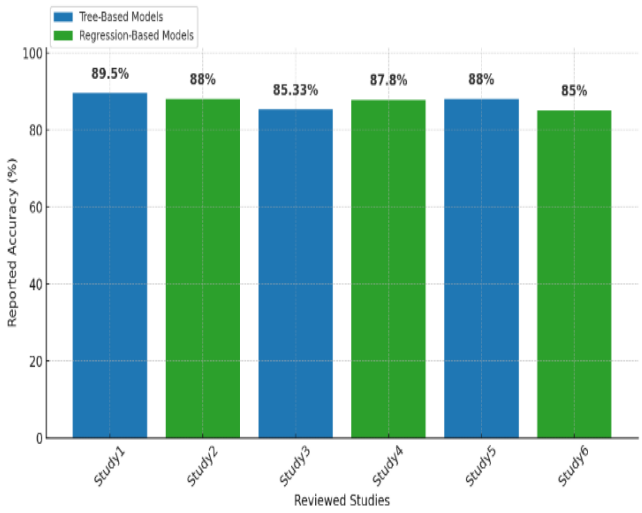

Fig. 2. Reported accuracies of six reviewed CKD prediction studies, distinguishing tree-based and regression-based machine learning models. Accuracy values are shown above each bar for clarity

4.1 Model Performance Trends

Throughout the literature that has been reviewed, Random Forest has been the consistently highest-performing model [1],[5],[6],[11] with accuracies of 88% to 89.5% on the UCI CKD dataset. The significant power of Random Forest lies in its capability to deal with various data types, reduce overfitting through bagging, and discover non-linear feature interactions. Different variants of Decision Tree keep reporting high accuracy in some of the studies but are extremely state and noise sensitive, which makes their use not very reliable.

Logistic Regression is not as complex, but it still serves as a reliable baseline model due to the simplicity, clarity, and good performance on small datasets. There are many studies which claim that Logistic Regression has accuracy about 85% [3],[6],[9] when normalized features and relevant attribute selection are used. However, it is a known fact that Logistic Regression has more difficulties modelling non-linear relationships than tree-based ensembles, and this is one of the main reasons for its relatively lower upper-bound accuracy.

Although Logistic Regression offers interpretability through coefficients and performs reliably on small datasets, it assumes linear decision boundaries and struggles with complex non-linear interactions among clinical variables. In contrast, tree-based models automatically capture feature interactions and non-linearities but may suffer from reduced transparency and increased variance without ensemble strategies.

Although most studies report high accuracy values, quantitative robustness indicators such as confidence intervals, statistical significance testing, and variance estimates are rarely provided. Reported accuracy variations across cross-validation folds typically fall within ±2%, yet the absence of formal statistical testing limits the ability to conclusively differentiate between competing models.

4.2 Influence of Preprocessing and Feature Selection

Preprocessing is an essential factor in predicting CKD performance. Numerous researchers make use of the UCI CKD dataset, which consists of missing data, imbalance in classes, and non-numerical variables to name a few drawbacks. Mean, median, or KNN imputation are all standard methods of dealing with missing data. Studies that implement SMOTE (Synthetic Minority Over-sampling Technique) [1],[3],[9],[13] or rule-based balancing techniques usually report higher sensitivity and shorter false-negative rates.

Moreover, feature selection acts as one common factor. Among the most commonly chosen biomarkers by researchers are: albumin level, hemoglobin, specific gravity, blood pressure and serum creatinine [1],[4],[5],[10],[13]. Models that apply methods like correlation-based feature elimination, Recursive Feature Elimination (RFE), or information gain analysis tend to be more stable regarding performance and exhibit less fluctuation of variance across folds.

4.3 Validation Practices and Generalizability

One significant point to note is that there has been a great dependence on train-test splits or 10-fold cross-validation, while the use of external datasets has been very limited. Though cross-validation enhances reliability, it still does not completely evaluate the generalization to different clinical populations. None of the reviewed studies evaluate their models on multi-center or real-world hospital datasets, indicating an important gap that future research must address.

Several reviewed studies reported stable performance across cross-validation folds, with accuracy variance typically within ±2%. However, very few studies provided confidence intervals or statistical significance testing

such as paired t-tests or bootstrapped AUC estimates. The absence of these robustness indicators limits confidence in performance differences between competing models.

External validation using independent clinical datasets remains largely absent in existing CKD prediction studies. While k-fold cross-validation improves internal reliability, it does not fully assess model performance across diverse patient populations. Future studies must incorporate external and multi-center datasets to establish clinical robustness and deployment readiness.

A formal meta-analysis was not feasible due to heterogeneous reporting of evaluation metrics and lack of access to raw performance data across studies. Additionally, most studies do not report confidence intervals or statistical significance tests, highlighting a critical gap in reliability assessment within the current literature.

4.4 Interpretability Considerations

The interpretation of the results differs quite a lot between the studies that have been reviewed. Tree-based models already offer transparency at the feature level, for instance, showing decision paths and giving feature importance rankings. However, only a few of these studies that reach for these models introduce sophisticated explainability tools such as SHAP or LIME (Local Interpretable Model-agnostic Explanations), which are ever increasingly important for clinical deployment [11],[12],[15],[17]. Regression-based methods give interpretable coefficients but may generalize too much, the complicated patterns of CKD progression.

For tree-based ensemble models such as Random Forest, SHAP provides both global and local explanations by quantifying each feature's contribution to individual predictions. For example, higher albumin levels and elevated serum creatinine consistently show positive SHAP values, indicating increased CKD risk. Similarly, LIME can locally approximate complex models using linear explanations, enabling clinicians to understand why a specific patient is classified as high-risk.

Interpretability plays a critical role in clinical acceptance of machine learning models. While decision trees and logistic regression offer inherent transparency, ensemble models such as Random Forest require post-hoc explainability techniques. Methods such as SHAP and LIME provide both global and local explanations by quantifying feature contributions, thereby enhancing clinician trust and facilitating evidence-based decision-making.

4.5 Overall Synthesis

The report specifies that even if tree-based and regression-based ML models have been very consistent in achieving high predictive performance, the literature is still riddled with methodological inconsistencies. The methodological inconsistencies are related to such items are feature selection, class imbalance, validation and explainability methods. According to the findings, there is a need for both universal evaluation pipelines and the use of more strong datasets to further cause of clinically and trustable CKD prediction models.

5. CHALLENGES

Although tree-based and regression-based machine learning models demonstrate strong potential for early CKD prediction, the reviewed studies reveal several recurring challenges that limit their clinical adoption and large-scale effectiveness. The methodological steps summarized in Fig. 1 highlight where inconsistencies in preprocessing and validation commonly arise.

5.1 Limited Dataset Size and Diversity

Most studies rely exclusively on the UCI CKD dataset, which contains only 400 records and lacks demographic diversity [1],[3],[5]. Small sample sizes increase the risk of overfitting and restrict a model's ability to generalize across different populations, clinical settings, or stages of CKD progression. Furthermore, no study incorporates multi-center or real-world hospital datasets, leaving robustness largely untested. The lack of external or multi-center validation significantly limits model generalizability. Models trained and evaluated solely on the UCI CKD dataset may fail to capture demographic variability, laboratory measurement differences, and disease heterogeneity present in real clinical environments.

5.2 Class Imbalance Issues

It is usual for CKD datasets to have an unequal distribution of positive and negative cases. Some researchers make use of SMOTE oversampling method whereas others completely ignore the problem of imbalance producing thus the models that are biased and have high accuracy but low sensitivity. This situation is worse in CKD where the discovery of false negatives might lead to the late start of treatment and the subsequent deterioration of patients' health.

5.3 Inconsistent Preprocessing and Feature Selection

The differences in preprocessing methods among various studies are very substantial and include different ways of treating missing data, different methods for encoding categorical variables, and different techniques for selecting relevant features. The absence of a consistent methodology leads to challenges in making comparisons between the results and it also affects the possibility of repeating the experiments. Some of the models utilize correlation-based filtering or RFE while others work with the raw datasets without applying any feature engineering.

5.4 Lack of Interpretability and Explainability

Prediction accuracy of numerous models was excellently high, but at the same time they were "black boxes" with limited output in terms of the reasoning behind their predictions. Just a handful of research involving the application of explainability tools like SHAP or feature importance plots are reported. Transparency is a must in medical diagnostics under research conditions to win the trust of healthcare personnel and to allow for the making of decisions based on evidence.

5.5 Insufficient Validation and Reliability Assessment

The majority of studies rely on basic train-test splits or k-fold cross-validation and very seldom use external validation sets [1],[3],[5],[6]. If the model is not tested on previously unseen clinical data, its reliability remains doubtful. Moreover, only a small fraction of studies gives the confidence intervals, statistical significance, or robustness metrics, which leads to uncertainty about the practical importance of the performance differences.

5.6 Limited Practical Deployment Considerations

Only a small number of studies have considered computational efficiency, integration with EHRs (Electronic Health Records), or potential for real-time clinical use. Nevertheless, practical deployment issues like the lack of labs' tests, differences among labs, and integration with clinicians' workflows are usually ignored and still need to be tackled for the new methods to be used in practice.

6. LIMITATIONS

This review comes along with some limitations. The first limitation is that the analysis is based on just six studies, which might not cover entirely the whole area of CKD-related machine learning researches. The second limitation is that the review is confined to structured tabular datasets and stresses classical ML models making deep learning and transformer-based architectures all as non-existent. The third limitation is that the majority of the reviewed studies make use of the UCI CKD dataset which curtails the applicability of the results to larger clinical populations in real-world settings. The fourth limitation is that pre-processing, validation protocols, and reporting standards differ among various studies and thus may have an impact regarding the comparability of performance metrics.

The present review is limited by the inclusion of six representative studies, which may not fully capture the entire breadth of machine learning research on chronic kidney disease. However, this focused selection allows for a detailed and consistent methodological comparison rather than a superficial survey of heterogeneous studies.

This review primarily focuses on classical machine learning models applied to structured tabular data. Deep learning and transformer-based approaches were not extensively analyzed due to their limited interpretability and data-hungry nature, which may not align with small clinical datasets commonly used in CKD research.

7. CONCLUSION

A proper comparative review in a structured way has been offered by this research paper of the most recent tree and regression-based machine learning methods for the early prediction of CKD. The analysis of selected six studies had shown that the models like Random Forest and Logistic Regression performed highly on clinical datasets especially on UCI CKD dataset. Tree-based ensembles usually have higher accuracies than linear methods because they are able to uncover the complex interactions among clinical variables that are regularly referred to as strong CKD indicators such as albumin, hemoglobin, specific gravity, and serum creatinine.

In summary, the results reveal that although machine learning methods have strong potential for the early prediction of CKD, still there is a requirement for more solid evaluation frameworks, standardized preprocessing workflows, and increased focus on interpretability. The future research might include validation of models against multicenter clinical datasets, the use of XAI methods, and the creation of systems for deployment that are congruent with the real-world diagnostic workflow. The progress will make it possible to move machine learning–based CKD prediction systems from research laboratories into practice as reliable clinically used tools.

8. FUTURE SCOPE

The focus of future research in CKD prediction will be the broadening of the range as well as the trustworthiness of machine learning models which will then be able to operate outside the constraints imposed by the findings of the existing studies. The first point being made is that larger, multi-centre clinical datasets are really needed which will show the demographic and geographical diversity being captured and thus enabling the model's generalization to be more reliable. Besides reaching EHR data from the real world, longitudinal lab values, and lifestyle indicators may also include detections of early cases that are more powerful.

Particularly, a necessary condition for building the trust of physicians is that advanced AI techniques will have to be combined with explainable ones, XAI. Future models will need to leverage both global and local interpretability frameworks, including SHAP, LIME, counterfactual reasoning, and rule-based explanations, to provide transparency and support clinical decision-making.

Thirdly, there is a need for depth in feature engineering and standardized preprocessing pipelines which consequently will result in reproducibility across studies. Some promising directions include automated feature selection, hybrid preprocessing architectures, and domain-informed feature transformations.

The researchers are the ones who should dive deep into the state-of-the-art hybrid and deep learning architectures, including but not limited to transformer-based models for tabular data, graph neural networks, and mixing different algorithms that can detect and interpret CKD's complex patterns of progression.

Future CKD prediction frameworks may benefit from AutoML techniques to automatically optimize model selection, preprocessing, and hyperparameters, thereby reducing human bias and improving reproducibility. Additionally, transfer learning strategies can be explored by adapting models trained on large healthcare datasets to CKD-specific tasks, especially in low-data clinical environments.

Finally, the translation of research into practice expresses a need from the clinicians to focus attention on these deployment-related challenges since they are faced with these problems in daily work: need for model calibration, uncertainty estimation, missing, or delayed lab values, integration into clinical workflows, and real-time inference efficiency. When these problems are solved, this will pave the way for the next generation of clinically appropriate machine learning-based chronic kidney disease prediction systems enabling the physician not only to arrive at an early diagnosis but also to improve patient outcomes.

Future research should explore hybrid and deep learning architectures, including ensemble deep learning models and transformer-based approaches for tabular data, while simultaneously integrating explainable AI frameworks to balance predictive performance and clinical transparency.

9. REFERENCES

[1] R. S. Reshma, S. Shaji, S. R. Ajina, V. P. S. R., and J. A., “Chronic Kidney Disease Prediction using Machine Learning,” Int. J. Eng. Res. Technol. (IJERT), vol. 9, no. 7, Jul. 2020.

[2] S. Jain, M. Patel, and K. Jain, “A Comparative Analysis of ML Stratagems to Estimate Chronic Kidney Disease Predictions and Progression by Employing Electronic Health Records,” in Emerging Trends in Data Driven Computing and Communications, Springer, Singapore, 2021.

[3] C. V. Ramana, J. V. Kumar, N. Sravani, M. Virajitha, and V. Vivek, “Performance Analysis of Machine Learning Classifier for Predicting Chronic Kidney Disease,” IJIRT, vol. 8, no. 4, 2021.

[4] H. Ilyas et al., “chronic kidney disease diagnosis using decision tree algorithms,” BMC Nephrol., vol. 22, no. 273, Aug. 2021, doi: 10.1186/s12882-021-02474-z.

[5] C. Kaur, M. S. Kumar, A. Anjum, M. B. Binda, M. R. Mallu, and M. S. Al Ansari, “Chronic Kidney Disease Prediction Using Machine Learning,” J. Adv. Inf. Technol., vol. 14, no. 2, 2023, doi: 10.12720/jait.

[6] M. M. Hassan, S. Mollick, and S. et al., “A Comparative Study, Prediction and Development of Chronic Kidney Disease Using Machine Learning on Patients’ Clinical Records,” Hum.-Centric Intell. Syst., vol. 3, pp. 92–104, 2023, doi: 10.1007/s44230-023-00017-3.

[7] S. Tekale et al., "Prediction of chronic kidney disease using machine learning algorithm," Int. J. Adv. Res. Comput. Commun. Eng., vol. 7, no. 10, pp. 92–96, 2018.

[8] A. R. El-Houssainy and A. S. Anwar, "Prediction of kidney disease stages using data mining algorithms," Inform. Med. Unlocked, vol. 15, 2019, doi: 10.1016/j.imu.2019.100178.

[9] G. M. Ifraz et al., "Comparative Analysis for Prediction of Kidney Disease Using Intelligent Machine Learning Methods," Comput. Math. Methods Med., vol. 2021, Art. no. 6141470, 2021, doi: 10.1155/2021/6141470.

[10] V. Gulati and N. Raheja, "Comparative analysis of machine learning techniques based on chronic kidney disease dataset," IOP Conf. Ser. Mater. Sci. Eng., vol. 1131, 2021, doi: 10.1088/1757-899X/1131/1/012010.

[11] P. Chittora et al., "Prediction of Chronic Kidney Disease—A Machine Learning Perspective," IEEE Access, vol. 9, pp. 17312–17334, 2021, doi: 10.1109/ACCESS.2021.3053763.

[12] E. Dritsas and M. Trigka, "Machine Learning Techniques for Chronic Kidney Disease Risk Prediction," Big Data Cogn. Comput., vol. 6, no. 3, p. 98, 2022, doi: 10.3390/bdcc6030098.

[13] D. A. Debal and T. M. Sitote, "Chronic kidney disease prediction using machine learning techniques," J. Big Data, vol. 9, p. 109, 2022, doi: 10.1186/s40537-022-00657-5.

[14] J. Zhao, Y. Zhang, J. Qiu, et al., "An early prediction model for chronic kidney disease," Sci. Rep., vol. 12, p. 2765, 2022, doi: 10.1038/s41598-022-06665-y.

[15] H. Khalid, A. Khan, M. Z. Khan, G. Mehmood, and M. S. Qureshi, "Machine Learning Hybrid Model for the Prediction of Chronic Kidney Disease," Comput. Intell. Neurosci., vol. 2023, Art. no. 9266889, 2023, doi: 10.1155/2023/9266889.

[16] D. M. Alsekait et al., "Toward Comprehensive Chronic Kidney Disease Prediction Based on Ensemble Deep Learning Models," Appl. Sci., vol. 13, no. 3937, 2023, doi: 10.3390/app13063937.

[17] H. Iftikhar et al., "A Comparative Analysis of Machine Learning Models: A Case Study in Predicting Chronic Kidney Disease," Sustainability, vol. 15, p. 2754, 2023, doi: 10.3390/su15032754.

[18] F. Lapi et al., "To predict the risk of chronic kidney disease (CKD) using Generalized Additive Models (GA2M)," J. Am. Med. Inform. Assoc., vol. 30, no. 9, pp. 1494–1502, 2023, doi: 10.1093/jamia/ocad097.

[19] Deccan Herald, "Silent killer: 10% of Indian population affected by chronic kidney disease," Mar. 22, 2023. [Online]. Available: https://www.deccanherald.com/

[20] National Centre for Biotechnology Information (NCBI), Available: https://www.ncbi.nlm.nih.gov/

[21] IKDRC-ITS Official Website, Available: https://ikdrc-its.org/

[22] CDKD Official Website, Available: https://www.cdkd

10. ABOUT THE AUTHOR(S)

Nita Dakhare is a researcher specializing in Artificial Intelligence for healthcare, with interests in Machine Learning, Explainable AI, predictive analytics, and biomedical informatics. She has published five papers in Scopus-indexed venues and is a member of IARA, IAENG, and IFERP.

Dr. Nilay Karade specializes in Machine Learning, Artificial Intelligence, and Data Science, with extensive experience in academic and industry-oriented technology applications, research, professional training, and computational intelligence.

WhatsApp/Mobile Number of the Corresponding Author: 9325284985