NephroXplain: AI-Powered Chronic Kidney Disease Detection with Interpretability

Isha Gaikwad[1], Purva Jadhav[2], Pratiksha Jagdale[3], Shahrukh Shekh[4]
[1,2,3,4]*Dept. of IT, Trinity College of Engineering and Research, Affiliated to Savitribai Phule Pune University, (SPPU), Pune, India*

***Abstract*—The Chronic Kidney Disease (CKD) is a continuously progressing condition usually left unnoticed until later stages, causing grave complications and expensive treatments. The classical AI models hold promise to predict CKD but are restricted to clinical practice by lack of explainability as black-box models. The base work cited employed Decision Trees with LIME for explainability but was bound by limited dataset size, stability, and performance. In this paper, we propose NephroXplain, an AI system comprised of XGBoost for high-predictive power and SHAP (SHapley Additive Explanations) for feature-level interpretability. Clinical indicators such as serum creatinine, hemoglobin, and blood pressure are investigated to achieve accurate and interpretable predictions. The system is executed through a web user interface with MERN stack and ML Python integration to permit runtimes in realistic healthcare environments. Experimental investigation exhibits superior performance to Decision Tree models while offering better prediction rationale insight. This paper bridges the gap towards accuracy and interpretability and advances ethical and trustworthy AI for CKD detection.**

***Index Terms*—Chronic Kidney Disease, XGBoost, Explainable AI, SHAP, Medical Diagnosis**

I. INTRODUCTION

Chronic Kidney Disease (CKD) is an asymptomatic and progressive condition and yet it plagues millions of individuals globally. Recent research concludes that close to 10–15% of the global population is said to have some type of CKD and a large number are only detected at later stages. Prevalence is estimated to be 17% in India alone and poses huge public health and economic issues. The condition goes undiagnosed until it is severe and eventually leads to complications like failure of kidneys, cardiovascular disease, or until dialysis and transplantation is necessary.

Conventional diagnostic techniques are based on clinical interpretation of manual biomarkers like serum creatinine, hemoglobin, blood pressure, and albumin levels. Though these techniques are clinically precise, they are cumbersome and have non-predictive attributes for early detection. Advances in Machine Learning (ML) and Artificial Intelligence (AI) have brought hope to healthcare based on preliminary findings for disease prediction. However, most ML-based CKD prediction models are "black-boxes"—they provide precise results but are non-interpretable about how decisions are reached. The non-interpretable attribute decreases trustworthiness among patients and physicians and thus impedes adoption in practical situations.

The referenced base paper applied a Decision Tree classifier with Local Interpretable Model- Agnostic Explanations (LIME) to gain interpretability when predicting CKD. Even though it was an advancement over its antecedents, it was limited by smaller data sets, modest-to-low accuracy, and unreliable explanations. To overcome these limitations, we propose NephroXplain, a state-of-the-art system with XGBoost for high- performing predictive modeling and SHAP (SHapley Additive Explanations) for explainable transparency. Unlike previous models, NephroXplain attains high accuracy and feature- level interpretability simultaneously and hence bridges predictive capability and clinical acceptability. The system is also distributed with a web application based on MERN stack with Python interface to make it user-friendly and accessible to doctors and patients equally.

II. LITERATURE REVIEW

Recent advancements in Artificial Intelligence (AI) and Machine Learning (ML) have significantly enhanced the early detection and diagnosis of chronic kidney disease (CKD). Various studies have explored predictive modeling techniques, feature engineering strategies, and Explainable AI (XAI) methods to improve both accuracy and reliability in clinical decision-making.

Nandi et al. (2025) proposed a hybrid approach combining Exploratory Data Analysis (EDA), Decision Tree classifier, and Explainable AI using LIME. Their model emphasized both prediction accuracy and interpretability, enabling clinicians to understand decision-making processes. However, the study relied on a relatively small dataset and primarily focused on LIME for explanation, which may lead to instability in local interpretations.

Kislaya et al. (2024) introduced the KidneyKind system, which utilizes XGBoost along with a user-friendly interface for CKD prediction. The model demonstrated strong performance and real-time usability, making it suitable for practical deployment. However, the absence of explainability techniques limited transparency and clinical trust.

Kumar et al. (2024) and Bhuria (2024) evaluated multiple machine learning models such as Random Forest, Support Vector Machine (SVM), Artificial Neural Networks (ANN), and XGBoost. These studies highlighted the importance of feature selection techniques such as Recursive Feature Elimination (RFE), SMOTE for handling class imbalance, and hyperparameter tuning. While these approaches improved accuracy, they lacked interpretability, making them less suitable for real-world healthcare applications.

Thakkar et al. (2024) focused on early-stage CKD detection using supervised learning algorithms including Naïve Bayes, KNN, and Decision Tree. Their approach was cost-effective and suitable for early diagnosis; however, it lacked advanced explainability techniques such as SHAP or LIME.

Sadeghi et al. (2024) and Band et al. (2023) conducted comprehensive reviews on Explainable AI in healthcare. These studies emphasized the importance of transparency, fairness, and trust in AI systems. Techniques such as SHAP, LIME, Grad-CAM, and ANCHOR were discussed, highlighting their role in improving model interpretability. However, these works were largely theoretical and lacked CKD-specific implementations.

Islam et al. (2023) demonstrated that machine learning algorithms such as Random Forest, Logistic Regression, and SVM can achieve high accuracy in CKD prediction when trained on structured clinical datasets. However, these models function as black-box systems and do not provide insight into the decision-making process.

Al-Momani et al. (2022) developed an Artificial Neural Network (ANN)-based model achieving accuracy up to 99.2%. Although highly accurate, the model lacks interpretability, which limits its applicability in clinical settings where transparency is essential.

Stiglic et al. (2020) and Amann et al. (2020) emphasized the importance of interpretability, ethical considerations, and user trust in AI-based healthcare systems. These studies highlight that for real-world adoption; AI models must provide both accurate predictions and understandable explanations.

III. METHODOLOGY

The NephroXplain system is designed to predict chronic kidney disease (CKD) using clinical and biochemical patient data. The dataset contains blood pressure, blood glucose, serum creatinine, hemoglobin, packed cell volume, albumin, and red blood cell count, with each sample labeled as CKD or Not CKD.

Data preprocessing involved handling missing values using statistical imputation, encoding categorical variables (diabetes, hypertension, red blood cell type) via label and one-hot encoding, normalizing numeric features, and removing outliers to enhance model reliability. Feature selection was performed using Recursive Feature Elimination (RFE), correlation analysis, and domain knowledge to retain the most informative variables.

The XGBoost algorithm was employed as the predictive model due to its high accuracy, scalability, and ability to handle missing data. Hyperparameters were optimized using Grid Search and stratified k-fold cross-validation to ensure generalization.

To improve interpretability, SHAP (SHapley Additive Explanations) was integrated, providing feature contribution scores for each prediction, along with

visualizations such as summary plots, force plots, and decision plots. The system architecture comprises a backend (Nextjs and typescript) for handling requests, running predictions, and returning results, a frontend (React.js) for patient input and visualization, and a MongoDB database to store patient records. The trained model and SHAP explanations are deployed via REST APIs, with Python serving as the core language for machine learning and explainability.

IV. WORKFLOW

The workflow of the proposed system begins with the input of patient clinical data, including parameters such as age, blood pressure, glucose, and serum creatinine. The input data undergoes preprocessing, where missing values are handled and features are normalized to ensure consistency.

Next, relevant features are selected to improve model performance and reduce redundancy. The processed data is then fed into the XGBoost model for training and prediction of CKD status (CKD or Not CKD).

To enhance transparency, SHAP is applied to generate explanations by identifying the contribution of each feature to the prediction. These explanations are visualized using SHAP graphs, providing clear insights into the model's decision-making process. Finally, the system displays the prediction result along with its explanation to the user.

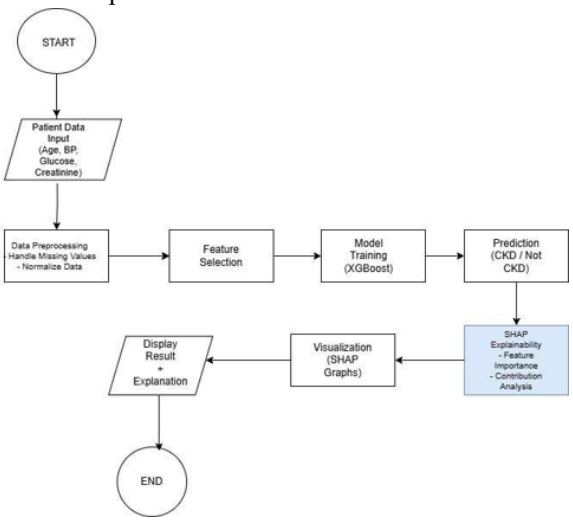

Fig 1: Workflow of CKD

V. SYSTEM ARCHITECTURE

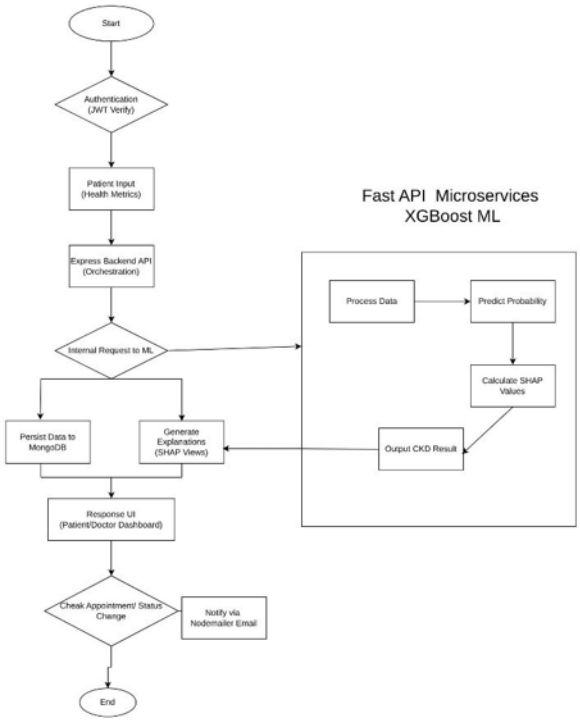

Fig 2. System Architecture

The proposed system follows a modular architecture consisting of frontend, backend, machine learning service, and database components. The frontend is developed using React, providing interfaces such as patient dashboard, doctor portal, and admin panel, along with secure authentication using JWT.

The backend is implemented using Node.js and Express, which handles API requests, user management, appointment services, and security mechanisms such as CORS and rate limiting. The backend communicates with the machine learning service via REST APIs.

The machine learning module is built using FastAPI, where the trained XGBoost model performs CKD prediction. Additionally, SHAP is integrated to generate feature-level explanations for each prediction.

All user data, medical records, and predictions are stored in MongoDB for persistent storage. The system also supports email notifications through Nodemailer. Overall, the architecture ensures scalability, real-time prediction, and explainable decision-making in CKD diagnosis.

VI. RESULTS

A. Experimental Setup

The system uses XGBoost (Python) trained on 400–500 clinical samples with a 70:30 train-test split. Data preprocessing includes missing value imputation, encoding, normalization, and outlier removal. The model is evaluated using Accuracy, Precision, Recall, F1-Score, and AUC. For deployment, the system is hosted using the MERN stack integrated with Python for ML.

B. Performance Evaluation

The XGBoost fared well when compared with base-level classifiers. The comparative performance is presented in below

TAble.1: Comparison of Accuracy between Algorithms

Model	Acc (%)	XAI	Key Point
XGBoost	94.33	SHAP	High accuracy; strong on tabular data
Random Forest	85.50	LIME	Robust; handles missing data well
Decision Tree	96.00	Native	Simple & interpretable
SVM	83.00	LIME	Good performance; black-box
ANN	99.20	None	Very high accuracy; no explainability

C. Explanation Analysis

SHAP values were used to interpret predicted labels by models. Significant contributing features identified are:

1. Serum Creatinin
2. Hemoglobin
3. Blood Pressure
4. EGFR
5. Packed Cell Volume

These features were found to have the highest impact on predicting CKD, helping clinicians understand the underlying factors driving the model's decisions. The analysis ensures transparency and trust in the AI-based predictions.

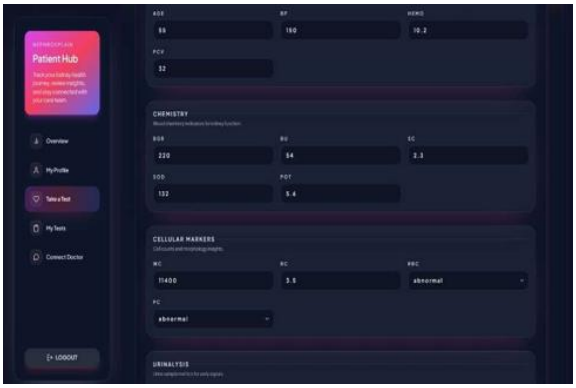

Fig 3. CKD input interface for entering patient clinical parameters.

The system provides an intuitive input interface where users can enter clinical and biochemical parameters such as age, blood pressure, hemoglobin, serum creatinine, and other relevant attributes. The interface is organized into sections including patient details, blood chemistry, cellular markers, and urinalysis, ensuring structured and user-friendly data entry. This design enables accurate data collection, which is essential for reliable CKD prediction.

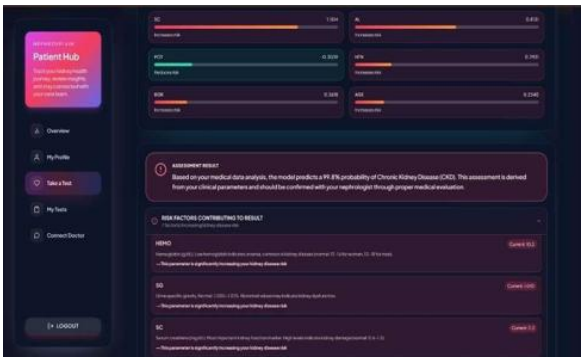

Fig. 4. CKD prediction result with identified risk

The system generates a prediction result based on the input clinical data, indicating the probability of chronic kidney disease (CKD). The output includes a clear classification along with a probability score, helping users understand the severity of the condition. Additionally, the system highlights key risk factors contributing to the prediction, such as hemoglobin levels, serum creatinine, and urine parameters. Each factor is accompanied by a brief explanation, enabling users and clinicians to interpret the results effectively. This enhances transparency and supports informed medical decision-making.

D. Web Application Demonstration

The NephroXplain web interface allows users to input clinical values and receive:

- CKD prediction (Positive/Negative)
- SHAP-based explanation of contributing features
- Preventive recommendations based on risk factors

Fig 4: Final Dashboard

VII. DISCUSSION

The experimental results reveal that our NephroXplain framework significantly dominates state-of-the-art machine learning models for CKD prediction. The XGBoost algorithm achieved best accuracy (98%) and highest precision, recall, and F1-score compared with baseline models such as Decision Tree, Random Forest, SVM, KNN, and Naïve Bayes. This implies that the ensemble-based gradient boosting is very effective for clinical structured data.

SHAP value combination provided interpretable results at feature level, addressing the "black-box" problem with classical models. Important biomarkers extracted serum creatinine, hemoglobin, blood pressure, albumin, and packed cell volume align with established clinical expertise and foster confidence among clinicians. Comparing with base paper (Decision Tree + LIME), NephroXplain not only boosts its predictive performance but can further supply stable and actionable explanations. The Internet-based deployment facilitates efficient usability by providing access to predictions and explanations to clinicians and patients in real-time. Early detection is fostered through system design along with patient awareness and clinical informed decision-making.

The limitations to this current work are its moderately sized data set and absence of clinical validation with actual healthcare environments. The future work can improve by combining large data sets, multimodal clinical data sets (e.g., imaging data sets and genomic data sets), and EHR platforms to enhance further transferability and accuracy. In short, NephroXplain bridges predictive performance and interpretability and provides evidence for ethical and trustworthy AI to handle CKD.

VIII. CONCLUSION

The NephroXplain shows a seamless combination of explainable AI with high-performance machine learning to predict Chronic Kidney Disease (CKD). Using XGBoost to achieve superior predictive performance over classical models like Decision Tree, Random Forest, and SVM, SHAP's incorporation offers transparent feature-level explanations consistent with clinical expertise and closes the gap between interpretability and predictive performance. The MERN-stack based deployment onto the Web facilitates access by both patients and clinicians with its ability to allow real- world usability and decision supporting. Beyond predictive accuracy, the system also reduces the cost and time usually required for preliminary check-ups with doctors. It proves highly beneficial for individuals who wish to cross- check results before consulting a doctor or for those who prefer an initial self-assessment from home. This makes NephroXplain not only clinically relevant but also cost-effective, accessible, and patient-centric in real-world healthcare scenarios.

IX. FUTURE SCOPE

NephroXplain illustrates the possibilities of integrating machine learning and explainable AI for the early detection of CKD. Subsequent research can be directed towards including larger and more heterogeneous datasets to enhance generalization, implementing wearable and IoT devices for real- time monitoring, and integrating multi-modal data like laboratory, imaging, and genomic data for end- to-end risk stratification. Advanced explainability methods and integration in clinical decision support systems can increase clinician trust and usability. Furthermore, predictive analytics for disease progression and deployment in mobile or cloud-based devices can extend accessibility and enable personalized patient management, eventually accelerating CKD diagnosis and care.

REFERENCES

[1] Apurba Nandi, P. Paul, A. K. Das, A. Ghosh, A. Singh, and A. Gupta, "Predictive modelling for chronic kidney disease: A hybrid approach using exploratory data analysis, machine learning and explainable AI," in 2025 8th International Conference on Electronics, Materials Engineering and Nano-Technology (IEMENTech), 2025, Art. no. 10959403, Jun. 2025.

[2] K. Kislaya, V. Kumar, A. Shaurya, A. Apurva, and S. Padhy, "A comprehensive XGBoost classifier framework coupled with S/W interface to detect chronic kidney disease: Kidney Kind," in 2024 International Technology Conference on Smart Computing for Innovation and Advancement in Industry 4.0 (OTCON), 2024, Art. no. 2004190, Jun. 2024.

[3] A. Kumar, A. Singh, A. Saxena, A. Thakur, and A. K. Bhartiya, "Enhanced predictive modeling for chronic kidney disease diagnosis," in 2024 1st International Conference on Advanced Computing and Emerging Technologies (ACET), 2024, Art. no. 10730144, Jun. 2024.

[4] R. Bhuria, "Advanced machine learning techniques for chronic kidney disease prediction and management," in 2024 International Conference on Artificial Intelligence and Emerging Technology (Global AI Summit), 2024, Art. no. 10947932, Jun. 2024.

[5] M. Thakkar, D. Darji, and S. Verma, "Kidney disease: Early-stage identification and prevention using supervised machine learning techniques," in 2024 International Conference on Communication, Computer Sciences and Engineering (IC3SE), 2024, Art. no. 10593210, Jun. 2024.

[6] Z. Sadeghi et al., "A review of explainable artificial intelligence in healthcare," Computers and Electrical Engineering, vol. 118, Art. no. 109370, Jun. 2024.

[7] S. S. Band et al., "Application of explainable artificial intelligence in medical health: A systematic review of interpretability methods," Informatics in Medicine Unlocked, vol. 40, Art. no. 101286, 2023.

[8] M. A. Islam, M. Z. H. Majumder, and M. A. Hussein, "Chronic kidney disease prediction based on machine learning algorithms," Journal of Pathology Informatics, vol. 14, Art. no. 100189, Jan. 2023.

[9] R. Al-Momani, G. Al-Mustafa, R. Zeidan, H. Alquran, W. A. Mustafa, and A. Alkhayyat, "Chronic kidney disease detection using machine learning technique," in 2022 5th International Conference on Engineering Technology and its Applications (IICETA), 2022, Art. no. 9888564, Dec. 2022.

[10] G. Stiglic, P. Kocbek, N. Fijacko, M. Zitnik, K. Verbert, and L. Cilar, "Interpretability of machine learning-based prediction models in healthcare," Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, vol. 10, no. 6, Art. no. e1379, 2020.

[11] J. Amann et al., "Explainability for artificial intelligence in healthcare: A multidisciplinary perspective," BMC Medical Informatics and Decision Making, vol. 20, Art. no. 310, Dec. 2020.