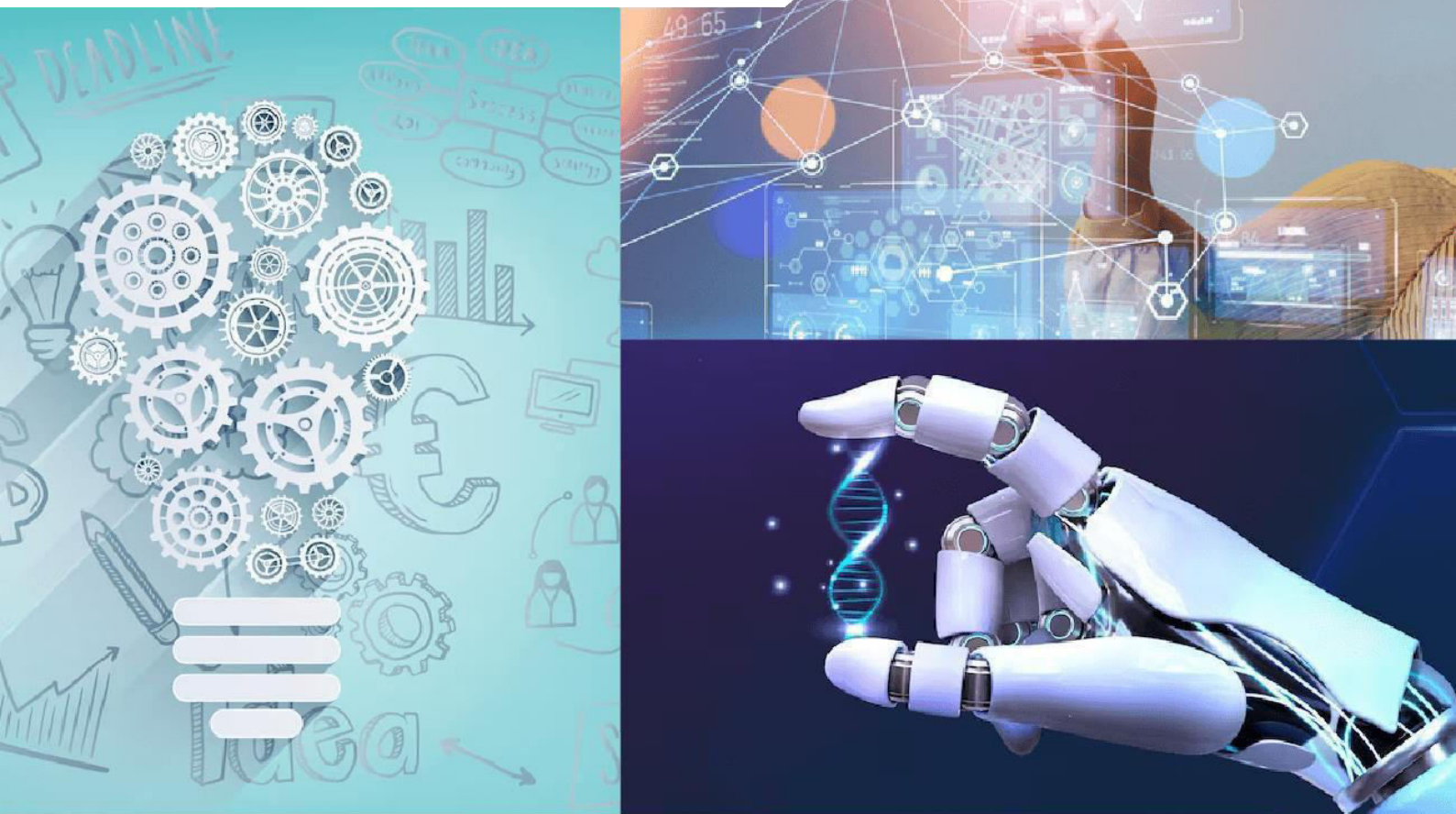




International Journal of Advanced Research in Education and Technology (IJARETY)

Volume 13, Issue 3, May - June 2026

Impact Factor: 9.971



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



NephroGuard AI: An Intelligent System for Early Chronic Kidney Disease Detection

Rakshith Uday Bandikatte

Dept. of Master of Computer Applications, CMR Institute of Technology, Bengaluru, India

ABSTRACT: Chronic Kidney Disease (CKD) is a progressive medical condition that often develops without noticeable symptoms, leading to late diagnosis and severe health complications. Traditional diagnostic methods rely heavily on isolated biomarker thresholds and manual interpretation of laboratory reports, which limits early detection and comprehensive analysis.

This study presents NephroGuard AI, an intelligent, AI-driven system designed to predict and explain CKD risk using multi-parameter clinical data. The system employs machine learning classifiers, specifically Random Forest and XGBoost, trained on clinical datasets containing 24 biomarkers. To address the lack of transparency in conventional AI systems, SHAP (Shapley Additive Explanations) is integrated to provide interpretable insights into model predictions.

Additionally, the system incorporates clinically validated eGFR calculations based on the CKD-EPI 2021 standard to ensure medical reliability. A Streamlit-based web interface enables real-time interaction, visualization, and risk assessment, while a conversational AI assistant enhances user understanding through personalized explanations.

Experimental outcomes demonstrate that the system effectively identifies early CKD risk patterns, offering a practical, low-cost, and explainable solution suitable for real-world healthcare applications.

I. INTRODUCTION

Chronic Kidney Disease (CKD) represents a major global health concern, characterized by its silent progression and delayed symptom manifestation. Early detection is critical, as timely intervention can significantly reduce the risk of kidney failure and associated complications. However, current diagnostic practices rely on fragmented lab reports and isolated biomarkers, making it difficult for patients and practitioners to interpret overall kidney health effectively.

Existing tools such as medical calculators and hospital portals primarily focus on data storage or single-metric evaluation, lacking integrated analysis and predictive capabilities. Furthermore, patients often struggle to understand complex medical data without expert guidance, which delays proactive healthcare decisions.

To address these challenges, NephroGuard AI is proposed as an intelligent system that combines machine learning, clinical validation, and explainable AI to provide comprehensive CKD risk assessment. The system analyzes 24 clinical parameters simultaneously, identifying hidden patterns and correlations that are not apparent through traditional methods.

The platform integrates predictive modeling with SHAP-based explainability, enabling users to understand not only the risk score but also the contributing factors. Additionally, the inclusion of eGFR calculations ensures alignment with established nephrology standards. A user-friendly Streamlit interface and AI-powered assistant further enhance accessibility and usability.

This research aims to bridge the gap between raw clinical data and actionable insights by providing an accurate, interpretable, and scalable solution for early CKD detection.

II. LITERATURE REVIEW

2.1 Traditional CKD Detection Approaches

Conventional diagnostic methods rely on laboratory reports and threshold-based evaluation of biomarkers such as creatinine and blood urea. While clinically reliable, these approaches often fail to capture complex relationships between multiple parameters, limiting early detection capabilities.

2.2 Medical Calculators and Hospital Systems

Digital tools such as eGFR calculators and Electronic Health Record (EHR) systems provide structured storage and basic analysis of patient data. However, they lack predictive intelligence, visualization, and explainability, making them insufficient for proactive healthcare management.

2.3 AI-Based Detection Systems

Recent advancements in machine learning have enabled predictive models for CKD detection. Algorithms such as Random Forest and XGBoost demonstrate high accuracy in classification tasks. However, these models often function as “black boxes,” limiting trust and interpretability in clinical environments.

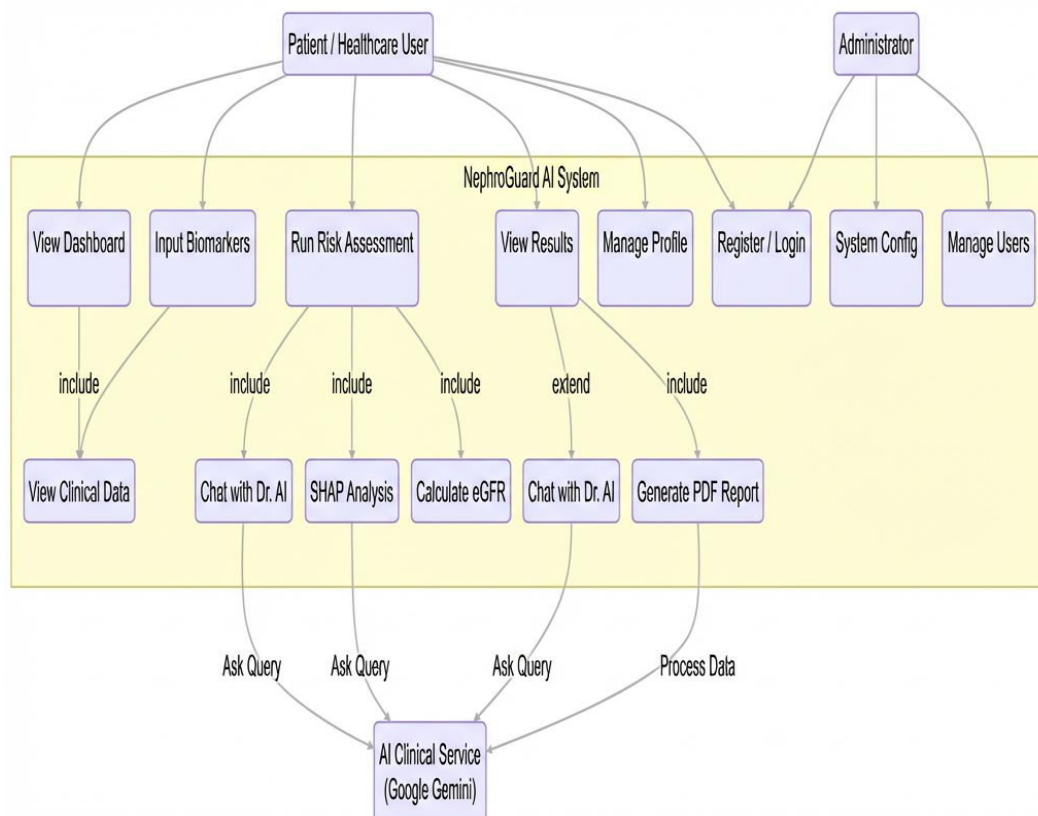
2.4 Limitations of Existing Systems

Despite technological progress, current systems face several challenges:

- Lack of integrated multi-parameter analysis
- Absence of explainability in AI predictions
- Limited user-friendly visualization
- No real-time interaction or clinical guidance

III. METHODOLOGY

3.1 System Architecture



The proposed system consists of the following components:

- Data Input Module (24 clinical parameters)
- Machine Learning Engine (Random Forest, XGBoost)
- Explainability Layer (SHAP)
- Clinical Validation Module (eGFR Calculation)
- User Interface (Streamlit Dashboard)
- AI Assistant (Google Gemini Integration)

The architecture follows a modular design, ensuring scalability and maintainability while supporting real-time prediction and analysis.

3.2 Data Processing and Feature Engineering

The dataset undergoes preprocessing steps including:

- Handling missing values using KNN imputation
- Feature scaling and normalization
- Encoding categorical variables

A total of 24 biomarkers are used, including creatinine, hemoglobin, blood pressure, and glucose levels, ensuring comprehensive clinical representation.

3.3 Machine Learning Models

Two classification models are implemented:

- Random Forest: Provides robust performance through ensemble learning
- XGBoost: Enhances prediction accuracy through gradient boosting

These models are trained on clinical datasets and optimized for performance and reliability.

3.4 Explainable AI (SHAP)

To overcome the black-box nature of ML models, SHAP is used to:

- Identify feature importance
- Explain individual predictions
- Provide transparency for clinical decision-making

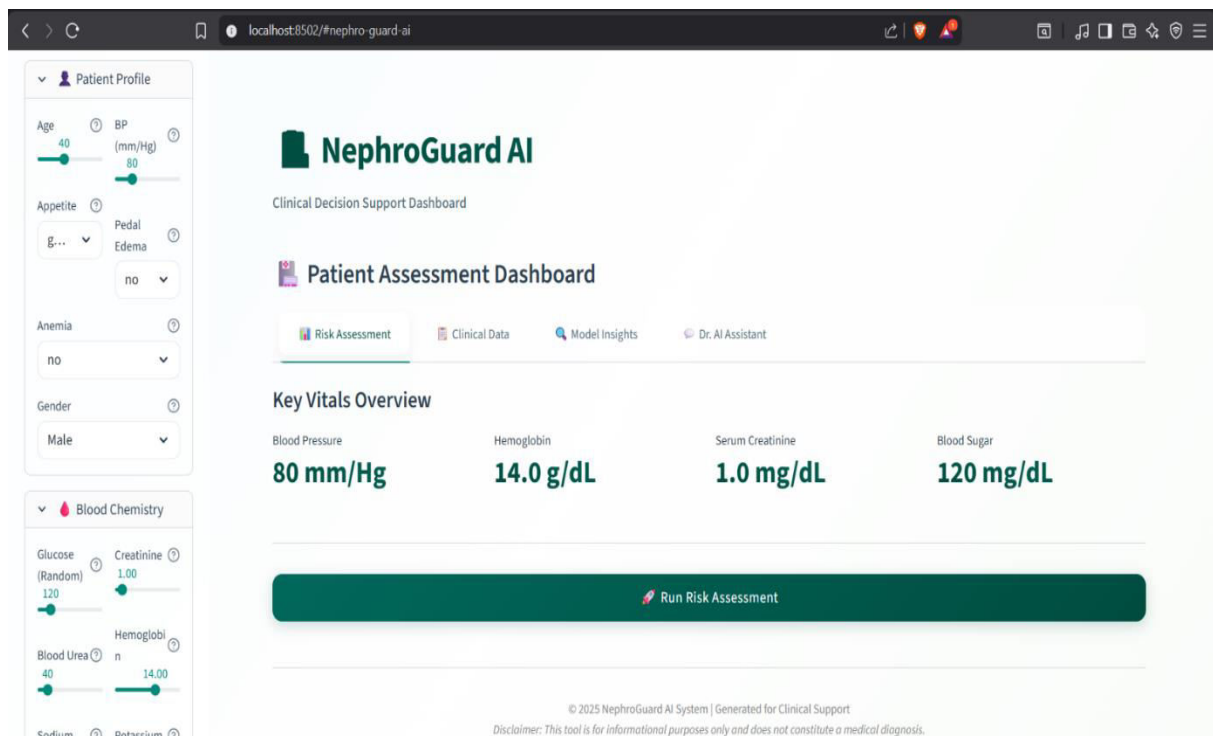
This enhances trust and interpretability in healthcare applications.

3.5 Clinical Validation using eGFR

The system incorporates eGFR calculation based on the CKD-EPI 2021 formula, ensuring:

- Clinical accuracy
- Cross-verification of AI predictions
- Detection of discrepancies between AI and medical standards

3.6 User Interface and Interaction



A Streamlit-based web application provides:

- Interactive dashboards
- Real-time risk visualization
- Radar charts and heatmaps
- AI-powered chatbot for clinical guidance

The interface ensures accessibility for both patients and healthcare professionals.

IV. CONCLUSION

This paper presents NephroGuard AI, an intelligent system for early prediction of Chronic Kidney Disease (CKD) using machine learning techniques. By combining Random Forest and XGBoost models with clinically relevant biomarkers, the system achieves reliable and accurate risk prediction. The integration of SHAP-based explainability further enhances transparency by clearly identifying the contributing factors behind each prediction.

The inclusion of clinically validated eGFR calculations ensures alignment with established medical standards, improving the credibility of the system. Additionally, the web-based dashboard enables interactive visualization and real-time analysis, making the system accessible and user-friendly for both patients and healthcare professionals. Unlike traditional approaches that rely on isolated metrics or manual interpretation, the proposed solution provides a comprehensive, data-driven, and cost-effective method for early CKD detection.

Future improvements may include integration with IoT-based health monitoring devices, mobile application deployment, and the incorporation of advanced deep learning models to further enhance prediction accuracy and scalability.

REFERENCES

- [1] T. Levey, L. A. Stevens, C. H. Schmid et al., "A New Equation to Estimate Glomerular Filtration Rate," *Annals of Internal Medicine*, 2009.
- [2] National Kidney Foundation, "KDIGO 2021 Clinical Practice Guideline for Chronic Kidney Disease," 2021.
- [3] J. R. Kshirsagar et al., "Chronic Kidney Disease: Detection and Evaluation," *American Family Physician*, 2017.
- [4] L. Breiman, "Random Forests," *Machine Learning*, 2001.
- [5] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," *KDD Conference*, 2016.
- [6] S. Lundberg and S. Lee, "A Unified Approach to Interpreting Model Predictions (SHAP)," *NeurIPS*, 2017.



International Journal of Advanced Research in Education and Technology

ISSN: 2394-2975

Impact Factor: 9.971