Explainable Deep Learning For Chronic Kidney Disease Prediction

Masith Jayarathna
Software Engineering Teaching Unit
University of Kelaniya
Dalugama, Sri Lanka
pramudithajayarathna@gmail.com

Isuru Hewapathirana
Software Engineering Teaching Unit
University of Kelaniya
Dalugama, Sri Lanka
ihewapathirana@kln.ac.lk

Keywords—CKD, Deep Learning, Explainable AI, LLM, MLP

I. Introduction

Chronic Kidney Disease (CKD) is a progressive and incurable condition affecting over 10% of the global population (Kovesdy, 2022). The disease imposes a significant financial burden on healthcare systems due to the high costs of medications, dialysis, and specialized treatments. Early detection and continuous monitoring are essential in mitigating CKD's effects, but traditional diagnostic methods, relying on clinical tests and expert analysis, can be time-consuming and subjective. This research aims to develop an explainable deep-learning model to predict CKD progression by analyzing historical patient data. Unlike conventional black-box models, this approach will incorporate explainability techniques such as Local Interpretable Model-agnostic Explanations (LIME) and (SHapley Additive exPlanation) SHAP to provide transparent insights into risk factors, enabling medical professionals to make data-driven decisions.

II. Literature Review

Various studies have explored deep learning for CKD prediction, with Arafat et al. (2021) achieving 99% accuracy using a Neural Network, while Akter et al. (2021) found ANN, MLP, and SimpleRNN to be the most effective, reaching 97% accuracy. Elkholy et al. (2021) used a Deep Belief Network (DBN) for early-stage CKD detection, and Jhumka et al. (2022) demonstrated a Deep Neural Network (DNN) outperforming Random Forest with 98.8% accuracy. Zhang et al. (2022) identified key CKD predictors using a DNN, achieving 82.66% accuracy, while Saif et al. (2024) developed an ensemble model combining CNN and LSTM for CKD prediction. Explainable AI (XAI) has been integrated into CKD studies, with Arif et al. (2024) using LIME for interpretability, achieving 100% accuracy. They used the LIME technique to offer lucid explanations for the outputs produced by their proposed machine learning model.

Recent studies have demonstrated the effectiveness of deep learning in CKD prediction, achieving high accuracy rates (Arafat et al., 2021; Akter et al., 2021; Jhumka et al., 2022). However, most models function as ‘black boxes’, limiting their usability in clinical decision-making. This study addresses this limitation by integrating Explainable AI (XAI) into deep learning for CKD prediction. By leveraging techniques such as LIME and SHAP ensuring interpretability while maintaining high accuracy.

III. Method

The study utilizes a dataset $D = \{(x_i, y_i)\}_{i=1}^{n}$ where each x_i represents a feature vector containing patient demographic information, laboratory test results, and clinical history, while $y_i \in \{0,1\}$ is a binary label indicating the presence (1) or absence (0) of CKD. The objective is to develop a predictive model $f: X \rightarrow Y$ that achieves high accuracy while providing interpretable outputs to help clinicians understand the contribution of each feature to the prediction.

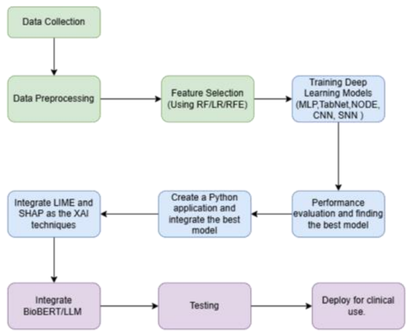

Figure 1: Proposed Solution

Figure 1 represents the proposed solution of this study. The solution involves developing an explainable deep learning framework that balances predictive accuracy with interpretability.

A. Preprocessing

The CKD dataset was obtained from the UCI Machine Learning Repository which consisted of 400 instances and 24 features. Standard preprocessing steps were applied, including the removal of invalid characters and handling of missing values using mode imputation for categorical features and cross-validation to determine the best imputation strategy (mean, median, or KNN) for numerical features. Categorical variables were encoded using LabelEncoder. Logarithmic transformation was applied to numerical features to reduce the impact of outliers without losing data.

Feature selection was performed using Random Forest, Lasso Regression, and Recursive Feature Elimination (RFE), with the final set of ten features selected based on their importance across at least two methods. The selected features were Albumin, Serum Creatinine, Hemoglobin, Packed Cell

https://doi.org/10.31705/ADScAI.2025.14

Volume, Red Blood Cell Count, Diabetes Mellitus, Specific Gravity, Blood Glucose Random, Hypertension, and Appetite

B. Model Training

Following feature selection, the dataset was split into training (70%) and testing (30%) sets and scaled using MinMaxScaler. Five deep learning models including Multilayer Perceptron (MLP), TabNet, Neural Oblivious Decision Ensembles (NODE), Convolutional Neural Network (CNN), and Self-Normalizing Network (SNN) were developed and optimized for classification, ensuring robustness and generalization. These models were trained using validation techniques and saved for further evaluation and comparison.

IV. RESULT AND DISCUSSION

Model	**Evaluation Metrics**				
	Accuracy	***Precision***	***Recall***	***F1 Score***	***AUC-ROC***
MLP	0.9833	0.9787	0.9866	0.9823	1.0000
TabNet	0.8000	0.9636	0.7067	0.8154	0.8815
NODE	0.7833	0.7426	1.0000	0.8523	0.9991
CNN	0.6250	0.6250	1.0000	0.7692	0.9964
SNN	0.9667	1.0000	0.9467	0.9726	0.9997

Table 1: Models' Evaluation Scores

The evaluation of the proposed models reveals that the MLP model outperforms all others across various metrics demonstrating strong performance in correctly classifying CKD and Not CKD cases. It achieved a performance similar to the MLP model of Akter et al. (2021), which reported 96% accuracy.

The MLP model of the study consisted of three dense layers (128, 64, 32 neurons) with ReLU activation and 30% Dropout to prevent overfitting. It was trained with the Adam optimizer and categorical crossentropy loss, with EarlyStopping to halt training if validation loss stagnated for 5 epochs.

V. APPLICATION DEVELOPMENT

Figure 2: Initial UI

Figure 3: SHAP

Figure 4: LIME

A user-friendly Python Streamlit application was developed allowing users to input relevant features and receive predictions (Figure 2). To enhance interpretability, the application integrates both LIME and SHAP techniques. Figures 3 and 4 represent the explanations of the model prediction using SHAP and LIME. Gemini is integrated by feeding the user inputs, outputs of MLP model and SHAP to provide medical insights and guidance (Figure 5).

Figure 5: Medical Insights

VI. CONCLUSION

This research evaluated five deep-learning models for CKD prediction, with MLP achieving the highest accuracy. The trained model is integrated into a Streamlit application for user-based predictions. SHAP and LIME were used for explainability, while Gemini provided personalized risk factor insights and medical guidance. Hosted on the Streamlit community cloud, the application showcases MLP's effectiveness in CKD prediction, demonstrating the benefits of explainable AI and personalized insights for improved diagnostics and patient care.

REFERENCES

[1] Akter, S., Habib, A., Islam, M. A., Hossen, M. S., Fahim, W. A., Sarkar, P. R., & Ahmed, M. (2021). Comprehensive Performance Assessment of Deep Learning Models in Early Prediction and Risk Identification of Chronic Kidney Disease. *IEEE Access*, *9*, 165184–165206.

[2] Arafat, F., Khan, T., Bapon, A. Das, Khan, M. I., & Noori, S. R. H. (2021). A Deep Learning Approach to Predict Chronic Kidney Disease in Human. *2021 IEEE 12th Annual Information Technology, Electronics and Mobile Communication Conference, IEMCON 2021*, 1010–1015.

[3] Arif, M. S., Rehman, A. U., & Asif, D. (2024). Explainable Machine Learning Model for Chronic Kidney Disease Prediction. *Algorithms*, *17*(10).

[4] Elkholy, S. M. M., Rezk, A., & Saleh, A. A. E. F. (2021). Early Prediction of Chronic Kidney Disease Using Deep Belief Network. *IEEE Access*, *9*, 135542–135549.

[5] Jhumka, K., Auzine, M. M., Casseem, M. S., Khan, M. H.-M., & Mungloo-Dilmohamud, Z. (2022). Chronic Kidney Disease Prediction using Deep Neural Network. *2022 3rd International Conference on Next Generation Computing Applications (NextComp)*, 1–5.

[6] Kovesdy, C. P. (2022). Epidemiology of chronic kidney disease: an update 2022. *Kidney International Supplements*, *12*(1), 7–11.

[7] Saif, D., Sarhan, A. M., & Elshennawy, N. M. (2024). Early prediction of chronic kidney disease based on ensemble of deep learning models and optimizers. *Journal of Electrical Systems and Information Technology*, *11*(1), 17.

[8] Zhang, Q., Liang, P., Yang, J., Wang, W., Yuan, G., Han, M., & Li, Z. (2022). *Deep Learning Identifies Intelligible Predictors of Poor Prognosis in Chronic Kidney Disease*.

[9] Rubini, L., Soundarapandian, P., & Eswaran, P. (2015). *Chronic Kidney Disease* [Dataset]. UCI Machine Learning Repository.