



An Exploratory Data Analysis-Guided Multilayer Perceptron Framework for Chronic Kidney Disease Classification Using Clinical Tabular Data

Vineeta Gulati¹, Arun Kumar Rana², Sumit Kumar Rana³, Amit Kumar Kesarwani², Olubunmi Ajala⁴,
Ayodeji Olalekan Salau^{5,6}, Abdulfattah Adedolapo Biliameen⁴, Sepiribo Lucky Braide⁷, Kaung Myat Paing⁸,
Tin Tin Ting^{8,9*}

¹ Department of Computer Science and Engineering (CSE), Maharishi Markandeshwar Engineering College (MMEC), Maharishi Markandeshwar (Deemed to be University), Mullana 133207, India

² Galgotias College of Engineering and Technology, Greater Noida 201310, India

³ University Institute of Engineering, Chandigarh University, Mohali 140413, India

⁴ National Centre for Artificial Intelligence and Robotics (NCAIR), National Information Technology Development Agency (NITDA), Abuja 564, Nigeria

⁵ Department of Electrical/Electronics and Computer Engineering, Afe Babalola University, Ado-Ekiti 360101, Nigeria

⁶ Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai 600124, India

⁷ Department of Electrical and Electronics Engineering, Rivers State University, Port Harcourt 5080, Nigeria

⁸ Faculty of Data Science and Information Technology, INTI International University, Persiaran Perdana BBN Putra Nilai, Nilai 71800, Malaysia

⁹ School of Information Technology, UNITAR International University, Petaling Jaya 47301, Malaysia

Corresponding Author Email: Ting.tintin@researcher.unitar.my

Copyright: ©2026 The authors. This article is published by IETA and is licensed under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

<https://doi.org/10.18280/isi.310813>

ABSTRACT

Received: 4 August 2025

Revised: 23 April 2026

Accepted: 24 August 2026

Available online: 31 August 2026

Keywords:

chronic kidney disease, multilayer perceptron, exploratory data analysis, clinical tabular data, deep learning, medical classification, health risk

Chronic kidney disease (CKD) is a progressive disorder that requires early identification to reduce the risk of severe complications. This study proposes an exploratory data analysis (EDA) guided multilayer perceptron (MLP) framework for CKD classification using clinical tabular data. The publicly available University of California, Irvine (UCI) CKD dataset containing 400 patient records and 24 clinical attributes was used for model development and evaluation. EDA was first conducted to investigate feature distributions, correlations, and clinically relevant variables, identifying important indicators including serum creatinine, hemoglobin, blood urea, and blood pressure-related features. Data preprocessing involved missing-value imputation, categorical feature encoding, feature normalization, and class balancing using the Synthetic Minority Over-sampling Technique (SMOTE). The developed MLP model consisted of two hidden layers with Rectified Linear Unit (ReLU) activation and was trained using the Adam optimizer. Its performance was compared with Gaussian Naïve Bayes (GNB), Support Vector Classifier (SVC), K-Nearest Neighbors (KNN), and Logistic Regression (LR) models. Experimental results showed that the proposed MLP achieved an accuracy of 98.75%, sensitivity of 98.07%, specificity of 100%, and an area under the curve (AUC) of 99.03%. These results indicate that an EDA-guided MLP framework can effectively classify CKD cases from structured clinical data. The proposed approach provides a lightweight deep learning solution for CKD prediction research and offers a basis for further validation using larger and multi-center clinical datasets.

1. INTRODUCTION

Chronic kidney disease (CKD) is a progressive condition characterized by the gradual decline of kidney function over time [1]. It affects more than 850 million people worldwide and is projected to become the fifth leading cause of death by 2040 [2, 3]. Beyond increasing the risk of premature mortality, CKD substantially impairs patients' quality of life and is frequently associated with serious cardiovascular and metabolic complications [4].

Early detection of CKD is important since early treatment can prevent, or at least slow, disease progression, decrease the

associated complications, and improve the outcomes for patients with CKD. However, in the majority of cases, particularly in low- and middle-income countries, tools for early diagnosis and possibilities of treatment are restricted [5, 6]. This highlights the importance of sensitive, specific, user-friendly, and efficient tools for early identification of CKD [7]. Figure 1 shows the implications of healthy and diseased kidneys [8].

Machine learning (ML) is a powerful technique in healthcare that can analyse complex health data and identify patterns not easily uncovered by traditional statistical methods [9, 10]. These techniques have been successfully applied to a

wide range of medical applications, including disease prediction, clinical diagnosis, prognosis, and treatment optimization [10]. In the context of CKD, ML models have shown considerable potential for early prediction by extracting informative features from patient records that capture subtle indicators of disease onset and progression.

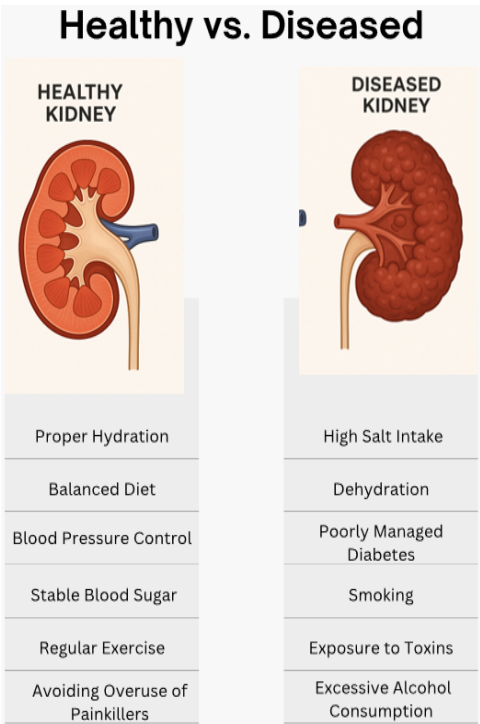


Figure 1. Healthy vs. diseased kidney

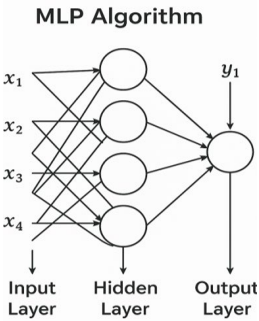


Figure 2. Multilayer perceptron architecture

MLP techniques are appropriate for non-linear relations over high-dimensional data [11, 12]. As shown in Figure 2, the MLPs consist of interconnected nodes that can learn fine-grained patterns from data. These can be applied to classification problems [13]. As shown in Table 1, EDA with proper preprocessing can improve the performance of the model.

The advantage of choosing the MLP is that it can model nonlinear relationships, but the architecture is relatively simple. This makes it applicable to clinical tabular data like data from the CKD UCI dataset. Moreover, MLPs can learn complex feature relationships without the need to train on very large datasets. The attributes make the use of an MLP within an EDA-guided approach for the classification of CKD in clinical tabular data appealing. In this context, the present study addresses an EDA-inspired MLP classification framework for the classification of CKD, focusing on the data preprocessing, feature selection, and comparison with the standard machine-learning classifiers.

Table 1. Summary of review of related work

References	Methods Used	Key Findings	Role of EDA/Feature Selection
Debal and Sitote [14]	RF, SVM, DT, XGBoost	RF with RFECV was able to predict 99.8% for binary classification.	Employed RFECV to enhance model performance by selecting the most predictive features.
Khan et al. [15]	SVM, GNB, DT, RF, Logistic Regression, KNN, Gradient Boosting, Adaptive Boosting, XGBoost	RF achieved 100% accuracy in CKD detection without data leakage.	Developed an innovative hybrid feature selection approach and a strong data formatting pipeline.
Haque et al. [16]	RF, MLP, Logistic Regression, Fine-tuned CatBoost	CatBoost achieved 98.75% accuracy with an AUC of 0.9993; MLP also demonstrated high performance.	Utilized Simulated Annealing and Cuckoo Search for feature selection and outlier adjustment, enhancing model accuracy.
Ghosh et al. [17]	XGBoost, RF, Logistic Regression, AdaBoost, Hybrid Model	Hybrid Model achieved 94.99% accuracy, outperforming individual algorithms.	Highlighted the integration of multiple models and feature selection to improve prediction accuracy.
Arifuzzaman et al. [18]	EfficientNetV2, InceptionNetV2, MobileNetV2, ViT, Ensemble Modeling	The ensemble model achieved 96% accuracy in early CKD detection.	Employed transfer learning and ensemble techniques to enhance predictive capabilities.
Song et al. [19]	Logistic Regression, RF, XGBoost, SVM, GNB	The RF classifier with SMOTE attained the most successful predictive power, especially for early (/advanced)-stage CKD.	Highlighted the role of feature selection and preprocessing in model quality.
Sawhney et al. [20]	ANN, RF, DT, SVM, Logistic Regression	ANN achieved 100% accuracy; RF and DT achieved 96% accuracy.	Demonstrated the effectiveness of ANN in CKD prediction, highlighting the role of model selection.
Firdaus et al. [21]	DTs, SVM, Neural Networks	ML models outperformed traditional methods in CKD classification and prediction.	Noted that feature selection approaches improve model efficiency and interpretability.

Note: RF = Random Forest; SVM = Support Vector Machine; DT = Decision Tree; XGBoost = Extreme Gradient Boosting; RFECV = Recursive Feature Elimination with Cross-Validation; MLP = multilayer perceptron; ViT = Vision Transformer; ANN = Artificial Neural Network; CKD = chronic kidney disease; ML = machine learning; EDA = exploratory data analysis; SMOTE = Synthetic Minority Over-sampling Technique; KNN = K-Nearest Neighbors; GNB = Gaussian Naïve Bayes; AUC = area under the receiver operating characteristic curve.

1.1 Research objective

The specific objectives of this research study are:

- (1) Analyse the UCI CKD dataset by performing EDA to perceive data distribution patterns or any anomalies.
- (2) Preprocess the dataset by addressing missing data through mean imputation, converting categorical data to numeric format using label encoding, normalizing features, and balancing classes if necessary.
- (3) Develop an MLP model with appropriate hidden layers and activation functions to predict CKD presence.
- (4) Calculate accuracy, precision, recall, and F1-score. Visualize the results using receiver operating characteristic (ROC) curves and training loss curves for model effectiveness analysis.
- (5) Interpret and visualize results to enhance understanding of model predictions and feature relationships, aiding in the development of early CKD recognition and intervention methods.

The rest of the paper is organized as follows: Section 2 presents the materials and methods applied in the proposed approach. Results and analysis are given in Section 3, and conclusions and future work are presented in Section 4.

2. MATERIALS AND METHODS

This study employed a systematic methodology encompassing data acquisition, preprocessing, model development, training, and performance evaluation to develop an accurate CKD prediction model. The Standard CKD dataset was obtained from the UCI ML Repository and comprised records from 400 patients, each characterized by 24 clinical features and one target class attribute. Data preprocessing was conducted using the Jupyter Notebook integrated development environment (IDE) within the Anaconda Navigator platform. During this stage, the dataset was cleaned by identifying and removing irrelevant or redundant information, handling inconsistencies, and preparing the data for subsequent ML model training and evaluation.

The sample was divided into 70% training, 15% validation, and 15% testing. The training set was the only one to which the SMOTE algorithm was applied to avoid information leakage into data in the validation or test sets. Adam optimizer was used with default parameters (learning rate = 0.001, b1 = 0.9, b2 = 0.999), and the MLP model was trained for 500 epochs with a batch size of 32. The other hyperparameters, such as the size of the hidden layer and the activation functions, have been clearly given in the revised manuscript.

2.1 Exploratory data analysis

To analyze the dataset and identify important features for CKD prediction, the researchers performed EDA. This study provided valuable insights into data distribution, relationships between features, and potential predictors for developing a valid ML model [22, 23].

2.1.1 Class distribution analysis

The first step was observing the class distribution of the target variable 'classification'. The count plot revealed a significant imbalance between CKD and non-CKD cases, with CKD cases being more common, as shown in Figure 3. This finding justified the application of class balancing strategies in

the preprocessing phase to prevent model bias towards the majority class [24].

2.1.2 Correlation heatmap

A correlation heatmap revealed relationships between numerical data in the dataset, specifically 'hemo', 'sod', and 'classification'. The heatmap showed a strong positive correlation between 'hemo' and non-CKD status, while 'sod' exhibited a moderate positive correlation with the target variable, as shown in Figure 4. These findings indicate higher levels of hemoglobin and sodium in non-CKD groups compared to CKD groups, highlighting the significance of these parameters [25-28].

2.1.3 Pairplot analysis

A pairplot illustrated the relationship between 'hemo' and 'bp' (blood pressure) across CKD classifications. Scatter plots revealed distinct clustering patterns, with CKD patients generally showing lower hemoglobin levels and higher blood pressure than non-CKD individuals, as shown in Figure 5. This visual disparity underscores the importance of these features in distinguishing between the two classes.

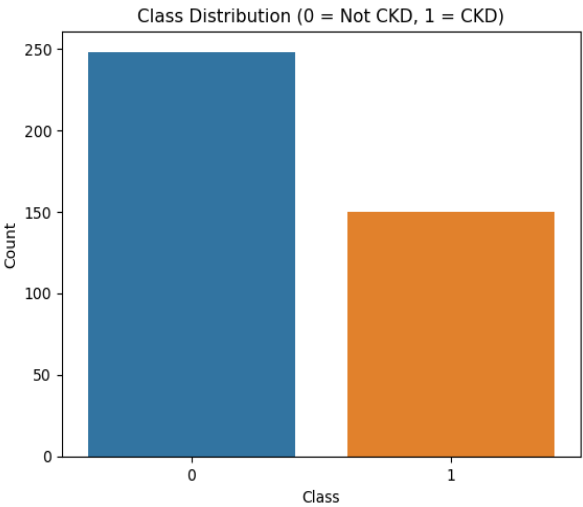


Figure 3. Class distribution analysis

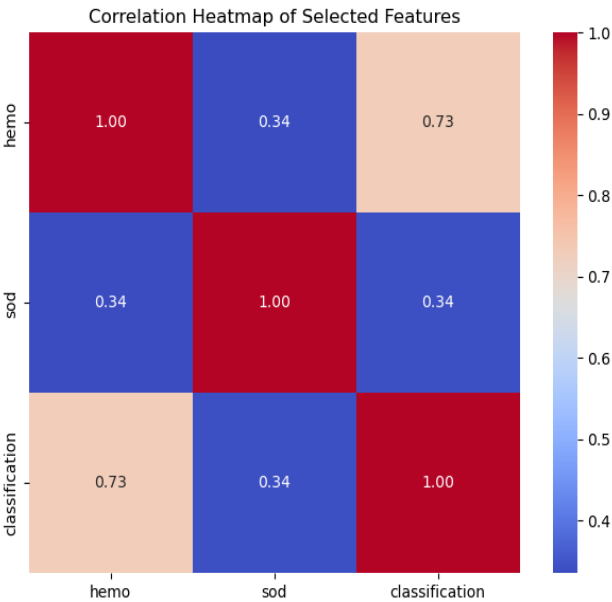


Figure 4. Correlation heatmap

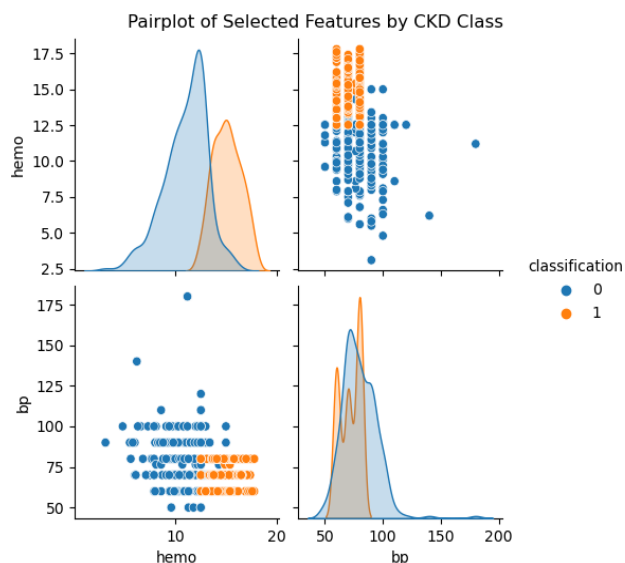


Figure 5. Pairplot analysis

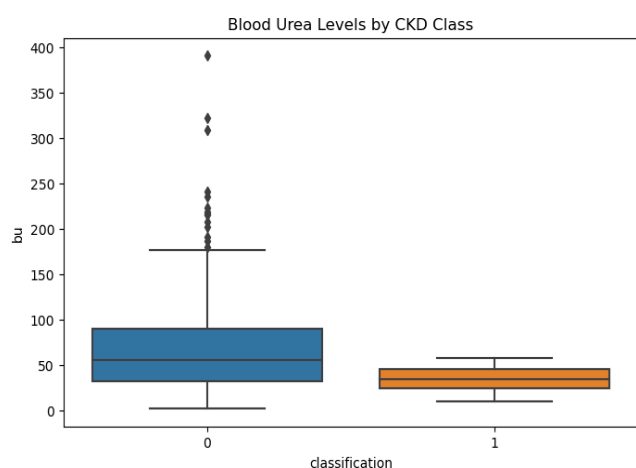


Figure 6. Boxplot examination of blood urea

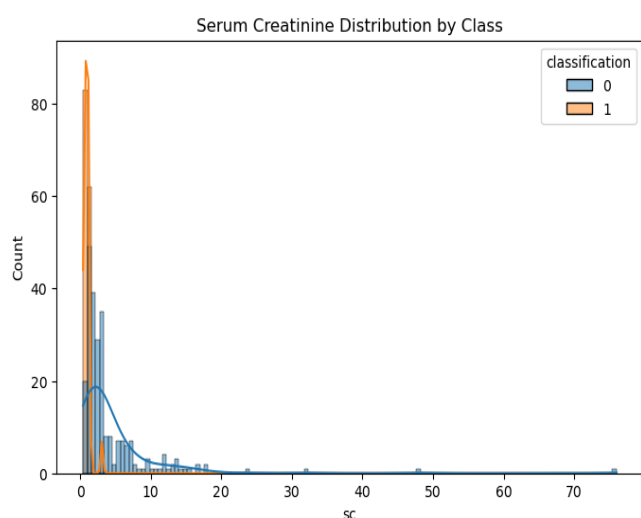


Figure 7. Serum creatinine distribution

2.1.4 Boxplot examination

A boxplot was used to compare the distribution of “blood urea” levels between CKD and non-CKD patients. Patients with CKD had higher blood urea concentrations with a broader interquartile range, suggestive of more variability, as shown in

Figure 6. This finding is consistent with the medical view that elevated blood urea is usually indicative of kidney malfunction and other pathological states.

2.1.5 Serum creatinine distribution

Furthermore, the distribution of serum creatinine levels also demonstrated the difference between CKD and non-CKD patients. The serum creatinine level in the CKD patients was significantly elevated compared to non-CKD patients and was frequently elevated beyond the normal range, as shown in Figure 7. This discrepancy in ratios underscores the relevance of serum creatinine as a biomarker for CKD diagnosis [26].

In conclusion, the EDA identified key features (hemoglobin, sodium, systolic blood pressure, blood urea, and serum creatinine) that demonstrate significant differences between CKD and non-CKD populations. These discoveries influenced feature selection and emphasized the required pre-processing steps, such as dealing with missing features, encoding classification features, scaling numerical features, and balancing classes.

3. RESULTS AND DISCUSSION

The study used Node-Net for predicting chronic CKD. Before training the model, we performed EDA to examine patterns and distributions in the data. During data pre-processing, we applied mean imputation to handle missing values, transformed categorical features into numerical data using label encoding, and mitigated class imbalance using the SMOTE algorithm. Feature normalization was also applied to create normalized data, so as to enhance the effect of the MLP model.

Integrating these EDA results with the model development pipeline enhanced the MLP classifier performance, further emphasizing the importance of EDA in the wider context of machine-learning healthcare diagnostic applications.

In this section, we compare the MLP model against the four classical ML models: GNB, SVC, KNN, and LR with Cross-Validation in terms of predicting CKD. The performance measures used are Accuracy, root mean square error (RMSE), Sensitivity, Precision, Specificity, and AUC.

The MLP achieved the highest accuracy of 98.75%, outperforming the other classifiers as shown in Figure 8. It also showed the lowest RMSE (11.18%), which signifies its best predictive ability, and the MLP had the highest sensitivity (98.07%), which means that it could identify the CKD-positive cases well. Although Linear SVC showed slightly better precision (98.03%) compared to MLP's 96.55%, the overall trade-off between sensitivity and precision favored MLP. MLP, KNN, and Logistic Regression all achieved 100% specificity, perfectly classifying CKD-negative cases. MLP attained the best AUC of 99.03%, as shown in Table 2, indicating its outstanding performance in discriminating between CKD-positive and CKD-negative cases across various threshold levels.

The confusion matrix confirms these results, showing 51 true negatives, 28 true positives, 1 false positive, and 0 false negatives. This demonstrates the model's strong predictive ability with minimal misclassification, as shown in Figure 9. These results support the MLP model's good performance and its promising potential as a useful tool for early CKD diagnosis.

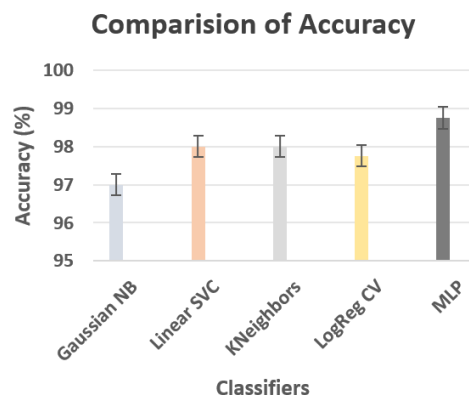


Figure 8. Accuracy comparison of multilayer perceptron with other classifiers

Each of the experiments has been trained with the identical standardized training pipeline and the hyperparameters that are described in the Methods section and allow complete reproducibility. The performance measures included in this paper are based on the analysis of the held-out test set, which allows an objective estimation of the model.

In recent studies, the prediction accuracy of CKD has been reported to be high with complex models like CatBoost and ensemble deep-learning architectures, but these models are typically complex, require extensive tuning, and are not easily interpretable. In contrast, this study shows that a relatively lightweight and explainable MLP network can achieve competitive accuracy. The EDA-guided feature selection and preprocessing pipeline also supports a simple, reproducible,

and potentially clinically useful approach. This makes the suggested model an effective alternative to provide high predictive performance without being overly complex due to resource-intensive approaches.

The comparison shown in Figure 8 demonstrates that the MLP model is both robust and reliable for CKD prediction. Its competitive performance across multiple evaluation metrics, as shown in Table 2, indicates the model's potential for early CKD detection, which is crucial for timely intervention and improved patient outcomes.

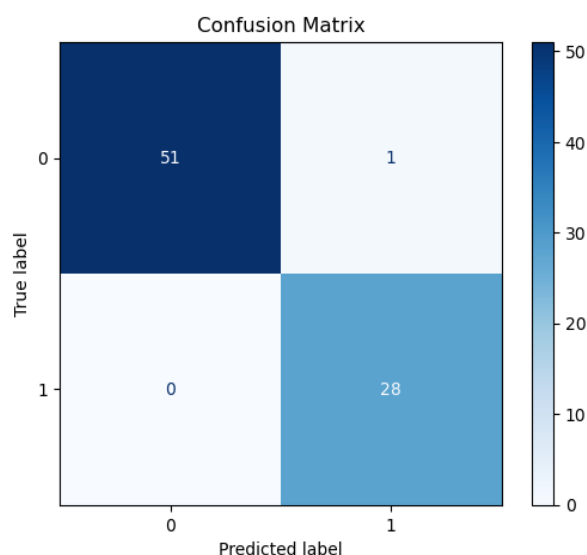


Figure 9. Confusion matrix

Table 2. Performance of classification algorithm on chronic kidney disease (CKD) dataset

Classifier	Accuracy (%)	RMSE (%)	Sensitivity (%)	Precision (%)	Specificity (%)	AUC (%)
GNB	97.00	17.32	95.91	96.15	98.09	96.97
Linear SVC	98.00	14.14	97.95	98.03	98.03	97.95
KNN	98.00	14.14	95.91	96.22	100	97.95
LogReg CV	97.75	15.14	95.91	96.22	100	97.95
MLP	98.75	11.18	98.07	96.55	100	99.03

Note: GNB = Gaussian Naïve Bayes; SVC = Support Vector Classifier; KNN = K-Nearest Neighbors; LogRegCV = Logistic Regression with Cross-Validation; MLP = multilayer perceptron; AUC = area under the receiver operating characteristic curve; RMSE = root mean square error.

4. CONCLUSION AND FUTURE SCOPE

In this study, a strong model based on an MLP neural network for early CKD prediction was developed. Comprehensive EDA was used to guide data preprocessing and feature selection. The proposed model has shown promising overall classification accuracy and is competitive with the tested conventional ML classifiers. It has high sensitivity, specificity, and discriminative ability, which means that it is able to separate CKD-positive and CKD-negative cases reliably. Furthermore, the model attained a sensitivity of 98.07%, a specificity of 100%, and an AUC of 99.03%, demonstrating its excellent capability to accurately distinguish between CKD-positive and CKD-negative cases. In general, the results indicate that EDA-informed preprocessing and feature selection combined with an MLP can be a useful method for the classification of CKD and aid in the exploration of clinically oriented prediction systems.

Building on this study, several areas for future research emerge:

- Future studies should use Explainable Artificial

Intelligence (XAI) techniques, including SHapley Additive exPlanations (SHAP) or integrated gradients, to interpret the most predictive biomarkers detected by the MLP model, thus giving the clinician more insight into why certain features are relevant predictors of CKD.

- Design a prototype clinical decision-support interface to include the identified feature patterns associated with this study, provide real-time warnings of CKD risk, and enable clinicians to visualize trends of biomarkers and model explanations.

- Add longitudinal and multi-center CKD datasets to increase variability of the study population and help develop temporal models, instead of using only static baseline data. Monitor time-varying biomarkers to develop progression-sensitive and early-detection models.

- Investigating hybrid modelling strategies, such as ensemble methods or combining MLP with other ML models, to improve predictive power and resilience to data fluctuations.

Addressing these challenges could lead to more interpretable and clinically relevant CKD prediction models,

ultimately benefiting patients and healthcare utilization.

ACKNOWLEDGMENT

This study was conducted under the NCAIR (NITDA) research syndicate and was funded by INTI International University and UNITAR International University.

REFERENCES

- [1] Francis, A., Harhay, M.N., Ong, A.C., et al. (2024). Chronic kidney disease and the global public health agenda: An international consensus. *Nature Reviews Nephrology*, 20(7): 473-485. <https://doi.org/10.1038/s41581-024-00820-6>
- [2] Becker, A. (2019). Artificial intelligence in medicine: What is it doing for us today? *Health Policy and Technology*, 8(2): 198-205. <https://doi.org/10.1016/j.hlpt.2019.03.004>
- [3] Gulati, V., Raheja, N., Kumar, R. (2024). A hybrid feature extraction-based prediction system (HFEPS) for early diagnosis of chronic kidney disease. In 2024 International Conference on Advances in Computing Research on Science Engineering and Technology (ACROSET), Indore, India, pp. 1-6. <https://doi.org/10.1109/ACROSET62108.2024.10743328>
- [4] Greener, J.G., Kandathil, S.M., Moffat, L., Jones, D.T. (2022). A guide to machine learning for biologists. *Nature Reviews Molecular Cell Biology*, 23(1): 40-55. <https://doi.org/10.1038/s41580-021-00407-0>
- [5] Sanmarchi, F., Fanconi, C., Golinelli, D., Gori, D., Hernandez-Boussard, T., Capodici, A. (2023). Predict, diagnose, and treat chronic kidney disease with machine learning: A systematic literature review. *Journal of Nephrology*, 36(4): 1101-1117. <https://doi.org/10.1007/s40620-023-01573-4>
- [6] Jankowski, J., Floege, J., Fliser, D., Böhm, M., Marx, N. (2021). Cardiovascular disease in chronic kidney disease: Pathophysiological insights and therapeutic options. *Circulation*, 143(11): 1157-1172. <https://doi.org/10.1161/CIRCULATIONAHA.120.050686>
- [7] Stevens, P.E., Ahmed, S.B., Carrero, J.J., et al. (2024). KDIGO 2024 clinical practice guideline for the evaluation and management of chronic kidney disease. *Kidney International*, 105(4): S117-S314. <https://doi.org/10.1016/j.kint.2023.10.018>
- [8] Denic, A., Rule, A.D., Glasscock, R.J. (2022). Healthy and unhealthy aging on kidney structure and function: Human studies. *Current Opinion in Nephrology and Hypertension*, 31(3): 228-234. <https://doi.org/10.1097/MNH.0000000000000780>
- [9] Khan, B., Naseem, R., Muhammad, F., Abbas, G., Kim, S. (2020). An empirical evaluation of machine learning techniques for chronic kidney disease prophecy. *IEEE Access*, 8: 55012-55022. <https://doi.org/10.1109/ACCESS.2020.2981689>
- [10] Gulati, V., Rana, A.K. (2024). Research applications in healthcare with machine learning approaches. In *Internet of Things and Big Data Analytics-Based Manufacturing*, pp. 195-206. <https://doi.org/10.1201/9781032673479-15>
- [11] Safar, A.A., Salih, D.M., Murshid, A.M. (2023). Pattern recognition using the multi-layer perceptron (MLP) for medical disease: A survey. *International Journal of Nonlinear Analysis and Applications*, 14(1): 1989-1998. <https://doi.org/10.22075/ijnaa.2022.7114>
- [12] Hamedan, F., Orooji, A., Sanadgol, H., Sheikhtaheri, A. (2020). Clinical decision support system to predict chronic kidney disease: A fuzzy expert system approach. *International Journal of Medical Informatics*, 138: 104134. <https://doi.org/10.1016/j.ijmedinf.2020.104134>
- [13] Abdel-Jaber, H., Devassy, D., Al Salam, A., Hidaytallah, L., El-Amir, M. (2022). A review of deep learning algorithms and their applications in healthcare. *Algorithms*, 15(2): 71. <https://doi.org/10.3390/a15020071>
- [14] Debal, D.A., Sitote, T.M. (2022). Chronic kidney disease prediction using machine learning techniques. *Journal of Big Data*, 9(1): 109. <https://doi.org/10.1186/s40537-022-00657-5>
- [15] Tusar, M.T.H.K., Islam, M.T., Raju, F.I. (2022). Detecting chronic kidney disease (CKD) at the initial stage: A novel hybrid feature-selection method and robust data preparation pipeline for different ML techniques. In 2022 5th International Conference on Computing and Informatics (ICCI), New Cairo, Cairo, Egypt, pp. 400-407. <https://doi.org/10.1109/ICCI54321.2022.9756094>
- [16] Haque, M.E., Islam, S.J., Maliha, J., Sumon, M.S.H., Sharmin, R., Rokoni, S. (2025). Improving chronic kidney disease detection efficiency: Fine tuned catboost and nature-inspired algorithms with explainable AI. In 2025 IEEE 14th International Conference on Communication Systems and Network Technologies (CSNT), Bhopal, India, pp. 811-818. <https://doi.org/10.1109/CSNT64827.2025.10968421>
- [17] Ghosh, B.P., Imam, T., Anjum, N., et al. (2024). Advancing chronic kidney disease prediction: Comparative analysis of machine learning algorithms and a hybrid model. *Journal of Computer Science and Technology Studies*, 6(3): 15-21. <https://doi.org/10.32996/jcsts.2024.6.3.2>
- [18] Arifuzzaman, M., Ahmed, I., Chowdhury, M.J.U., et al. (2024). A novel ensemble-based deep learning model with explainable AI for accurate kidney disease diagnosis. *arXiv preprint arXiv:2412.09472*. <https://doi.org/10.48550/arXiv.2412.09472>
- [19] Song, W.Z., Zhou, X.S., Duan, Q., et al. (2022). Using random forest algorithm for glomerular and tubular injury classification in chronic kidney disease. *Frontiers in Medicine*, 9: 911737. <https://doi.org/10.3389/fmed.2022.911737>
- [20] Sawhney, R., Malik, A., Sharma, S., Narayan, V. (2023). A comparative assessment of artificial intelligence models used for early prediction and evaluation of chronic kidney disease. *Decision Analytics Journal*, 6: 100169. <https://doi.org/10.1016/j.dajour.2023.100169>
- [21] Firdaus, S., Alam, S., Navya, S., Raviteja, J. (2025). Machine learning algorithm for chronic kidney disease prediction. *International Advanced Research Journal in Science, Engineering and Technology*, 12(5): 29-32. <https://doi.org/10.17148/IARJSET.2025.12505>
- [22] Ren, J. (2023). A systematic analysis of data diversity in machine learning for EDA. In *Proceedings of the 6th International Conference on Information Technologies and Electrical Engineering*, Association for Computing

- Machinery, New York, NY, USA, pp. 283-287. <https://doi.org/10.1145/3640115.3640160>
- [23] Vijayaprabakaran, K., Reddy, P., Reddy, T.P.K., Munnaf, K., Prasad, G. (2021). Chronic kidney disease diagnosis using machine learning. *International Research Journal of Engineering and Technology*, 8(6): 4029-4033.
- [24] Salau, A.O., Markus, E.D., Assegie, T.A., Omeje, C.O., Eneh, J.N. (2023). Influence of class imbalance and resampling on classification accuracy of chronic kidney disease detection. *Mathematical Modelling of Engineering Problems*, 10(1): 48-54. <https://doi.org/10.18280/mmep.100106>
- [25] Badrouchi, S., Ahmed, A., Bacha, M.M., Abderrahim, E., Abdallah, T.B. (2021). A machine learning framework for predicting long-term graft survival after kidney transplantation. *Expert Systems with Applications*, 182: 115235. <https://doi.org/10.1016/j.eswa.2021.115235>
- [26] Silveira, A.C.D., Sobrinho, Á., Silva, L.D.D., Costa, E.D.B., Pinheiro, M.E., Perkusich, A. (2022). Exploring early prediction of chronic kidney disease using machine learning algorithms for small and imbalanced datasets. *Applied Sciences*, 12(7): 3673. <https://doi.org/10.3390/app12073673>
- [27] Mohammad, A.S., Mohammad, S.I., Vasudevan, A., et al. (2026). Exploring the intersection of robotics, artificial intelligence, and healthcare: A bibliometric analysis. In *Studies in Systems, Decision and Control*, pp. 107-119. https://doi.org/10.1007/978-3-032-02056-7_9
- [28] Zhao, Y., Chaw, J.K., Liu, L., et al. (2024). Systematic literature review on reinforcement learning in non-communicable disease interventions. *Artificial Intelligence in Medicine*, 154: 102901. <https://doi.org/10.1016/j.artmed.2024.102901>