

Privacy-Preserving Prediction of Chronic Kidney Disease Using Ensemble Machine Learning with Laplacian Differential Privacy and Explainable AI.

Author

Mohammad Mamun

Mohammed Ibrahim Hussain

Mohammed Sowket Ali

Md.Shafiul Alam Chowdhury

Safiul Haque Chowdhury

Muhammad Minoar Hossain

Abstract

Chronic Kidney Disease (CKD) is a progressive and potentially life-threatening condition that demands early and accurate prediction for effective clinical intervention. This study proposes a robust and privacy-preserving framework for CKD prediction by integrating multiple Machine Learning (ML) algorithms with advanced privacy techniques. Our approach involved the use of five high-performing ML models: Extreme Gradient Boosting (XGB), Random Forest (RF), Bagging, Gradient Boosted Decision Tree (GBDT), and Stacking Ensemble (SE), to ensure predictive robustness. Among these, Random Forest achieved the highest accuracy of **99.75%** without any privacy constraints. To ensure data privacy in sensitive healthcare applications, we incorporated Differential Privacy (DP) using the Laplacian mechanism (LM), focusing on various privacy budgets ranging from $\epsilon = 0.25$ to $\epsilon = 2.5$, evaluated at intervals of 0.25. The optimal privacy-accuracy balance was achieved at $\epsilon = 1.0$, where the Laplacian-DP-enhanced Random Forest model attained an impressive accuracy of **85.75%**, ensuring both strong privacy protection and reliable prediction. Furthermore, we evaluated all models using 10-fold cross-validation to validate consistency and robustness. To enhance the interpretability of our results and support clinical decision-making, we integrated Explainable Artificial Intelligence (XAI) techniques, including Shapley Additive explanations (SHAP) and Local Interpretable Model-Agnostic Explanations (LIME). These tools provided critical insights into feature importance and model behavior, aiding healthcare professionals in understanding the predictive patterns. The approach ensures high prediction accuracy with strong patient privacy, making it suitable for real-world CKD diagnosis and critical condition prediction.

Publication:

Communications in Computer and Information Science

ISSN: 18650929

Volume: 2681 CCIS

Pages: 346 - 361

Conference Paper2025

EID: 2-s2.0-105027572853