Review of: "Implementing Machine Learning to predict the 10-year risk of Cardiovascular Disease"

Shokhan M. Al-Barzinji[1]

1 University of Anbar, Iraq

Potential competing interests: No potential competing interests to declare.

The study makes a compelling case for using ML to predict CVD risk, demonstrating improved performance over traditional models. However, it falls short in addressing methodological and practical aspects that would make the findings more actionable and relevant to clinical practice. By addressing these limitations, the research could better position itself as a pivotal contribution to the field of personalized healthcare. The abstract outlines the study's objectives, methodology, and significance in addressing cardiovascular disease (CVD) risk prediction using machine learning (ML). It highlights the use of various ML models, optimization techniques, and a comparative analysis with traditional models. The introduction contextualizes CVD as a global health concern and discusses the limitations of traditional risk models while presenting ML as a promising alternative. The study utilizes the UCI Heart dataset and applies multiple ML algorithms, including Logistic Regression (LR), Random Forest (RF), Support Vector Machines (SVM), Artificial Neural Networks (ANN), K-Nearest Neighbors (KNN), and Decision Trees (DT). It employs techniques such as cross-validation, normalization, and hyperparameter tuning for performance optimization. The results indicate that Adaboost (AdB) and Random Forest (RF) achieved the best performance (AUC-ROC: 0.8734), outperforming both traditional models and simpler ML approaches like Decision Trees and KNN.

General Improvements:

1. Limited exploration of why specific models performed better or worse.
2. Over-reliance on AUC-ROC as the primary evaluation metric without discussing the clinical implications of these results.
3. Figures and tables, such as Fig 6 comparing AUC-ROC values, are referenced but not included in the provided text, reducing clarity.