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

K. Y. Obiwusi

Potential competing interests: No potential competing interests to declare.

The following section provides detailed comments on the manuscript titled "Machine Learning Models for Cardiovascular Disease Risk Prediction." The review aims to provide constructive feedback and critical evaluation of the study's methodology, results, and conclusions. Each comment is intended to assist the authors in enhancing the quality and impact of their research.

1. The abstract is clear but lacks specific quantitative results, novelty, and mentions of limitations. Including performance metrics and highlighting the research's unique contributions would enhance it. Additionally, briefly addressing potential future directions could improve the abstract's completeness.
2. The introduction provides a concise overview of cardiovascular disease (CVD) as a significant global health issue and the limitations of traditional risk prediction models. It introduces the potential of machine learning (ML) in improving CVD risk prediction. The mention of conducting a comprehensive literature review adds credibility to the research. However, it could benefit from a clearer statement of the study's specific objectives and contributions.
3. The section effectively reviews traditional and ML-based CVD risk prediction methods, highlighting their strengths and limitations. It offers valuable insights into the potential of ML models to overcome traditional model shortcomings. However, more specific examples of ML model performance and challenges, as well as data sources, could enhance the comprehensiveness of this review.
4. This methodology section is comprehensive, detailing the data source, feature selection, data transformation, machine learning algorithms used, and performance metrics. The use of hyperparameter optimization is a notable strength. However, providing more context on the rationale behind feature selection and explaining why certain hyperparameters were chosen could enhance clarity. Additionally, it would benefit from a brief mention of ethical considerations in data use.
5. The results section provides a clear summary of the performance metrics of various ML algorithms in predicting CVD risk. It effectively communicates which models performed well, such as AdB, RF, LR, SVM, and XGB, with high accuracy and discrimination ability. However, it would be beneficial to include statistical tests, such as p-values or confidence intervals, to assess the significance of performance differences between models.
6. The discussion provides a robust analysis of the study's findings, emphasizing the advantages of ML models over traditional approaches for CVD risk prediction. However, it could benefit from discussing the potential ethical considerations, biases, and challenges associated with implementing ML models in clinical practice to provide a more comprehensive perspective.

7. The conclusions emphasize the superiority of ML models in CVD risk prediction, underlining their potential to transform the field. However, the discussion of interpretability could be expanded, addressing the challenge of understanding ML model decisions in clinical practice. Additionally, mentioning potential ethical considerations, such as data privacy and bias, would provide a more comprehensive outlook.