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

D. Lakshmi Padmaja[1]

1 Anurag College of Engineering

Potential competing interests: No potential competing interests to declare.

Key Aspects:

1.The article examines a range of machine learning algorithms, including AdaBoost, Decision Trees, K-Nearest Neighbours, Logistic Regression, Random Forests, Support Vector Machines, and XGBoost, in the context of the study.

2. Evaluation of Performance measures: The article assesses the performance of machine learning models by employing commonly used measures such as Accuracy, Precision, Recall/Sensitivity, F1 score, and AUC-ROC.

3. Machine Learning for Cardiovascular Disease (CVD) Risk Prediction: This article highlights the use of machine learning to predict the risk of CVD and highlights its potential for early detection and intervention.

4.The article discusses the application of both Hyperparameter Optimization (HPO) and Default Hyperparameter (DHP) in order to improve performance metrics.

5. Open-Source Code: The code used in the study is publicly available on GitHub, which contributes to the reproducibility of the research.

Changes We Recommend:

1. Proposal for the Inclusion of a Highlights part: The incorporation of a specialized part providing a concise overview of the principal contributions and conclusions would prove advantageous for both readers and reviewers.

2. Elucidation on Algorithms: Although the article references several algorithms, it would be advantageous to provide a more comprehensive elucidation about the rationale for the selection of these particular algorithms and their comparative analysis.

3. Comprehensive Analysis: The article has the potential to offer a more thorough examination of the performance metrics by incorporating a comparative table or chart that effectively illustrates the performance of each algorithm across different measures.

4. Clinical Significance: In order to enhance the article's comprehensiveness, it would be beneficial to have a discussion on the clinical significance and prospective healthcare implementations of the machine learning models, considering their emphasis on cardiovascular disease risk prediction.

5. Ethical Considerations: The inclusion of a section dedicated to examining the ethical ramifications, particularly in relation to healthcare data, would enhance the scholarly nature of the work.

6. Literature analysis: The inclusion of a more extensive literature analysis in the publication would enhance its academic value by establishing connections between the current study and prior research conducted in the fields of machine learning and healthcare.

7. Methodological Rigour: Additional details pertaining to the dataset employed, such as its dimensions, characteristics, and any preparation procedures undertaken, might be supplied.

8.External Validation: If feasible, evaluating the generalizability of the models by applying them to an external dataset could yield valuable insights.

9. Conclusion and Future Work: A more comprehensive conclusion that summarises the most important results and outlines potential directions for further study might offer a more satisfying finish to the paper.