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

Ariful Islam Arif[1]

1 Daffodil International University

Potential competing interests: No potential competing interests to declare.

The paper titled "**Implementing Machine Learning to predict the 10-year risk of Cardiovascular Disease**" presents a comprehensive study on the influence of data preprocessing techniques on the accuracy of cardiovascular disease prediction models. The paper is well-written and structured, making it easy for readers to follow the research methodology and results. The language is clear and concise, and the use of figures and tables aids in understanding the presented information.

The study's results demonstrate a notable improvement in the accuracy of cardiovascular disease prediction models after applying various data preprocessing techniques. The authors can describe how they handled missing values and outliers. It would be beneficial if the paper could provide more details on the specific methods used for these tasks, such as imputation methods for missing data and outlier detection algorithms.

- Provide more detailed explanations of data preprocessing techniques, including data cleaning, feature selection, and scaling.
- Offer a more comprehensive discussion of the handling of imbalanced data if relevant.
- Explain ROC curves and AUC values in more depth, relating them to the paper's results and their significance in cardiovascular disease prediction. To enhance the clarity and usefulness of these results, the authors should provide a detailed explanation of ROC curves, including the concepts of true positive rate (sensitivity) and false positive rate (1-specificity). Additionally, the AUC (Area Under the Curve) values should be discussed in the context of model performance, including what constitutes a good AUC score.