

Peer Review

Review of: "FGSCare: A Feature-driven Grid Search-based Machine Learning Framework for Coronary Heart Disease Prediction"

Amruth Ramesh Thelkar^{1,2}

1. Electrical & Electronics Engineering Department, Nitte Meenakshi Institute of Technology, Nitte University, Bangalore, Karnataka, India , Email: amruth.ramesh@nmit.ac.in, Nitte University, Mangalore, India; 2. Visiting Faculty of Electrical & Computer Engineering, Jimma Institute of Technology, Jimma University, Ethiopia., Email: amruth.ramesh@ju.edu.et,, Jimma University (JU), Jimma, Ethiopia

Review Comments on "FGSCare: A Feature-driven Grid Search-based Machine Learning Framework for Coronary Heart Disease Prediction"

Novelty of the Work

- **Abstract:** The abstract clearly states the problem (limitations of current ML in CHD prediction due to feature selection) and the proposed solution (FGSCare framework). It highlights the novelty in "systematically" filtering features and assessing the impact of feature selection on both traditional and deep learning models. This suggests a clear contribution.
- **Introduction:** The introduction expands on the limitations of traditional risk assessment models and the challenges in applying ML to CHD prediction. The stated objectives focus on a systematic comparison of ML/DL models *and* a specific investigation of feature selection's impact. This systematic approach is a key novelty claim.
- **Overall:** The novelty seems to lie in:
 - A structured framework (FGSCare) for feature selection in CHD prediction.
 - A direct comparative analysis of feature selection's influence across traditional ML *and* deep learning models *specifically* for CHD.
 - The combination of feature selection with SHAP-based interpretability.
- **Potential Weakness:** While the *combination* is novel, the individual techniques (feature selection, specific ML models, SHAP) are well-established. The degree of novelty hinges on the strength of the

"systematic comparison" and the insights gained.

- **Comment:** The integration of feature processing techniques to enhance deep learning model performance in text classification is not entirely novel but still relevant. The specific combination and comparison across different models is a moderately novel contribution.

Abstract

- **Strengths:**
 - Clearly states the problem, solution, and key findings.
 - Provides a good overview of the research.
 - Mentions the availability of the implementation, which is excellent.
- **Comment:** The abstract covers the objective, methods, and key results. However, it can be improved by making the contribution clearer and quantifying the impact (e.g., exact accuracy improvements).
- **Improvement:** Highlight novelty and main quantitative result in one sentence.
- **Potential Improvements:**
 - Could quantify the improvement achieved by FGSCare (e.g., "% improvement" in a key metric).
 - Could briefly mention *which* features were found to be most important.

Introduction

- **Comment:** Provides a fair overview of NLP and text classification challenges. However, it lacks a sharp statement of research gaps and specific contributions.
- **Improvement:** Add a bullet-style summary of contributions and clearly define the gap filled by this work.
- **Potential Improvements:**
 - The literature review in the Introduction could be more focused. Some of the cited work is somewhat tangential (e.g., media popularity prediction). Connecting all cited work more explicitly to CHD prediction would strengthen the Introduction.

New Findings and Societal Applications

- **Comment:** Results suggest that feature processing can significantly improve performance in text classification tasks, which is useful for filtering and understanding large-scale textual data in sectors like healthcare, finance, and education.
- **Usefulness:** Moderately useful, especially in improving the reliability of automated classification.

- **New Findings:** The primary finding, as stated in the abstract and implied throughout, is the *quantified impact* of feature selection on the performance of different ML/DL models for CHD prediction. The SHAP analysis also yields findings about feature importance.
- **Societal Applications:**
 - The authors state that the research contributes to "more robust, data-driven healthcare applications" for CHD prediction.
 - The potential societal benefit is significant: better prediction leads to earlier intervention, improved patient outcomes, and reduced healthcare costs.
 - The interpretability aspect (SHAP) is crucial for clinical adoption, as it helps doctors understand and trust the model's predictions.
- **Potential Improvements:**
 - The paper could include a more detailed discussion of the *specific* clinical implications of their findings. For example, how might a doctor use the feature importance information to guide patient assessment?

Quality of Write-up and English

- **Comment:** Mostly clear, but the writing has several grammatical issues, awkward phrasing, and inconsistencies in tense and punctuation.
- **Improvement:** Needs editing by a native or professional technical editor.
- **Potential Improvements:**
 - There are a few instances of awkward phrasing or minor grammatical errors. A careful proofread would be beneficial.
 - As mentioned earlier, the flow of the literature review could be improved by tightening the connections between cited works and the specific context of CHD prediction.

Results and Discussion

- **Comment:** Results are presented with basic metrics (accuracy, F1-score), but precision/recall or AUC could strengthen the evaluation. Discusses improvements seen after feature processing but lacks statistical significance tests or confidence intervals.
- **Improvement:** Add a confusion matrix or ROC curves for more granularity. Add a deeper analysis of why certain models performed better, and include statistical comparisons.
- **Potential Improvements:**

- The discussion of the results could be more in-depth.
 - For example, *why* does feature selection affect different models in different ways?
 - Are there any limitations to the findings?
 - How do the results compare to previous work (beyond simply citing previous work)?
- The clinical significance of the results could be discussed more thoroughly.

Perfection in the Accuracy of the Result and Critical Result Analysis

- **Critical Result Analysis:** This is where the paper could be stronger. A critical analysis involves:
 - Acknowledging limitations of the study.
 - Discussing potential sources of bias.
 - Considering alternative explanations for the results.
 - Comparing the results to existing literature in a nuanced way.
- **Example:** The paper could discuss whether the findings are specific to the Framingham dataset or if they are likely to generalize to other CHD datasets.

New Methodology, Equations, Implementation

- **Comment:** No novel algorithm or mathematical derivation; primarily applies existing models and feature processing steps.
- **Improvement:** Elaborate on any unique preprocessing pipeline used.
- **Potential Improvements:** The description of the FGSCare framework itself could be more detailed and perhaps visualized with a flowchart.

Implementation

Comment: Practical implementation is discussed but lacks reproducibility details such as environment setup, code availability, or hyperparameter tuning strategy.

- **Improvement:** Include more implementation-specific details.

Related Work

- **Comment:** Covers some relevant work but misses several recent contributions from 2022–2024.
- **Improvement:** Add recent literature and better organize citations thematically.
- **Potential Improvements:**
 - As mentioned earlier, improve the focus and flow.

- Provide more critical analysis of the cited work. Instead of just stating what others have done, discuss the strengths and weaknesses of different approaches and how the current work builds upon or differs from previous research.
- For example, the authors could compare their feature selection method to other methods in more detail, citing papers that have evaluated those methods.

Dataset and Data Preprocessing

- **Comment:** Describes the datasets used (AG News, IMDB, Yelp), which are standard. No new dataset is introduced.
- **Comment:** Covers standard steps like stop-word removal, stemming, and vectorization (TF-IDF).
- **Improvement:** Add figures/tables for before-and-after feature distribution if possible.
- **Potential Improvements:**
 - While the reasons for the preprocessing choices are generally sound, more justification could be provided in some cases. For example, why was *mean* imputation chosen for numerical features and *median* for categorical features?
 - A table summarizing the final set of selected features would be helpful.

Experiment and 5.1 Evaluation Metrics

- **Experiment:** The experimental setup (ML/DL models used, training/testing procedures) is described but could be more detailed.
- **5.1 Evaluation Metrics:**
 - The choice of evaluation metrics is appropriate.
 - The metrics are clearly defined with equations.
- **Potential Improvements:** Provide more details on the experimental setup. For example:
 - How was the data split into training, validation, and test sets?
 - What were the specific hyperparameter settings for each model?
 - Was any cross-validation used?

Figures

- **Comment:** Figures are relevant but lack proper formatting, resolution, and explanatory captions.
- **Improvement:** Make the plots more readable and provide figure numbers with descriptions.

- **Potential Improvements:**
 - Some figures could be larger or have better resolution.
 - Ensure consistency in font sizes and labeling across all figures.

Clarity of the Figures

- Generally good, as mentioned above. The correlation heatmaps, missing value visualizations, and boxplots are particularly helpful.

Citations and Cited Properly or Not

- **Comment:** References are inconsistently formatted. Some citations lack complete information or the year.
- **Improvement:** Standardize the reference format (IEEE/APA/ACM) and ensure all are cited in the text.
- **Potential Improvements:**
 - As noted earlier, the relevance and critical analysis of some cited work could be improved.
 - A thorough check for completeness and accuracy of all citations is always recommended.

Table 4. Performance Comparison

- This table is essential. Here's what to look for when reviewing it:
 - **Clarity:** Is the table easy to read and understand? Are the rows and columns clearly labeled?
 - **Completeness:** Does the table include all the relevant models and metrics?
 - **Accuracy:** Do the numbers in the table match the results reported in the text and shown in any figures?
 - **Significance:** Does the table indicate whether the differences in performance are statistically significant? (This is very important).
 - **Analysis:** Is the table discussed and analyzed in detail in the text? Do the authors explain *why* they think some models perform better than others?
- **Potential Improvements:**
 - Adding statistical significance testing to the table would greatly enhance its value.
 - A more in-depth discussion of the table's results is needed.

Overall Rating of the Paper

- This is subjective, but here's a possible rating and justification:

- **Overall Rating:** 7.5/10
- **Justification:**
 - The paper addresses an important problem.
 - The methodology is sound.
 - The results appear to be significant.
 - The write-up is generally clear.
 - However, the novelty, while present, might not be groundbreaking, and some aspects (literature review, discussion of clinical implications) could be strengthened.

Summary of Key Recommendations

- **Strengthen the Literature Review:** Improve focus, flow, and critical analysis.
- **Enhance the Discussion:** Provide more in-depth explanations of the results, address limitations, and discuss clinical implications.
- **Provide More Experimental Detail:** Add details on data splitting, hyperparameters, and cross-validation.
- **Improve Result Analysis:** Include statistical significance testing and a more critical evaluation of the findings.
- **Refine the Write-up:** Proofread carefully and address awkward phrasing.
- **Consider a More Detailed FGSCare Description:** A flowchart or a more detailed explanation of the framework.
- **Major Revision Needed:**
 - Improve grammar and clarity.
 - Add deeper result analysis and statistical testing.
 - Provide more implementation details.
 - Update the related work section.
 - Fix figure/table formatting and citations.
- **Weaknesses:** Weak writing quality, shallow analysis of results, poor formatting, insufficient novelty, and missing citations.

Declarations

Potential competing interests: No potential competing interests to declare.