**An Efficient Algorithm for Diseases Classification based on Hemogram Blood Test Samples**

Thank you for taking the time to review our paper titled An Efficient Algorithm for Diseases Classification based on Hemogram Blood Test Samples. We appreciate your thoughtful feedback and suggestions, which have ndoubtedly enhanced the quality of our work. In response to your comments, we have made several revisions to address the concerns raised and improve the overall clarity and coherence of the paper. Below, we outline our responses to each of your points:

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Review 1** | **Review 2** | **Response** |
| Title | 4: (good)  The title is complete | **3**: (fair)  The title is somewhat generic and does not specify which diseases are being classified. Adding more specificity about the types of diseases being addressed would make the title more informative.  The title mentions an "efficient algorithm," but it doesn't provide any details about the algorithm itself. Incorporating a bit more information about the nature of the algorithm could generate more interest from readers. | **New title: Improvements in the Imbalanced Hemogram Data Classification** |
| Abstract | 4: (good)  Kindly, add the contribution statement after the AIM | **3**: (fair)  The abstract doesn't provide a clear and specific problem statement. It mentions that the growth of hospital information systems has led to vast medical data. Still, it doesn't explicitly state what problem or challenge this abundance of data poses.  The contribution is stated as proposing an efficient algorithm for disease classification, but it lacks details on how this algorithm is innovative or different from existing approaches. It doesn't highlight why this contribution is significant for medical data analysis.  The abstract doesn't discuss the implications of the study's findings for the medical field. It should address questions like: How could this algorithm be integrated into clinical practice? What are the potential benefits and challenges of doing so? | Abstract problem statement: We agree that the abstract should provide a clear and specific problem statement. In the revised version, we have added a concise and explicit statement highlighting the challenge posed by the abundance of medical data in hospital information systems. We emphasize the need for efficient algorithms to handle this data and improve disease classification accuracy. In additon, we acknowledge the need to provide more details on the innovative aspects of our proposed algorithm. In the revised abstract, we have included a brief description of the unique features and techniques employed in our algorithm that differentiate it from existing approaches. We emphasize its effectiveness in handling imbalanced hemogram data and achieving improved classification performance. Moreover, we appreciate your comment regarding the need to discuss the implications of our findings for the medical field. In the revised abstract, we have included a paragraph highlighting the potential integration of our algorithm into HIS. |
| Introduction | 4: (good)  At the end of the introduction, I suggest you to add the contribution statements for 3-4 statements | **3**: (fair)  While the introduction discusses the importance of medical data analysis and mentions the significance of hematological data, it doesn't explicitly state the specific problem or gap in the field that the proposed research aims to address.  The references are mostly mentioned briefly without providing a summary of their contributions or relevance to the current study. Including a brief summary of the key findings or methodologies of these references would help readers understand their context.  The introduction mentions a lack of research on hematological data and briefly touches on the use of machine learning for disease classification, but it doesn't clearly define the gap in existing literature that the current study is meant to fill. Additionally, the justification for addressing this gap could be more explicit in terms of the potential impact on medical practice. | In the revised version, we have added 3-4 concise and specific contribution statements. In addition, we have provided a more explicit and focused statement of the research gap related to the analysis of hematological data and the classification of diseases using machine learning techniques. We emphasize the limited research in this specific area and the need for improved methods to handle the challenges posed by imbalanced hemogram data. |
| Materials and Method | 3: (fair)  1. This is too general "The database contained a total of 1,766 samples from patients", author should explain it more detail  2. Author did not use the springer template for caption for FIGURE and TABLE. | **4**: (good)  The section mentions that data was collected from a laboratory information system, but it lacks details about the types of diseases covered, the attributes recorded, and any specific details about the patients. Providing more context about the data collection process would be informative.  You mention comparing your approach with other algorithms, but you don't provide any insight into how these comparisons were carried out or what the results were. Sharing a summary of the comparison results would be informative. | In the revised version, we have provided a more comprehensive description of the database in Table 1 and Fig. 2. We have rectified this issue by ensuring the captions now follow the appropriate format as per the template guidelines. We appreciate your suggestion to provide insight into the comparison process and the results obtained. In the revised version, we have included a summary of the comparison results, highlighting the performance metrics achieved by our proposed algorithm compared to the other algorithms considered. |
| Results | 4: (good)  1. Author should put the name and unit for x and y axis in the Graphichs (2, 3, 4, 5, 6), for example Weight (kg), Time (seconds) so on | **3**: (fair)  The section uses tables and figures to present results, which is great. However, the tables are referenced before they are introduced, and the figures aren't included in the provided text. To improve readability, make sure tables and figures are presented in the appropriate order and included in the text.  The section jumps straight into discussing the two experiments without providing an introductory paragraph that explains the purpose or context of the experiments. Providing a brief overview of why these experiments were conducted would improve clarity.  The text references figures (Figure 4 and Figure 5) without explaining their contents. Providing a brief description or interpretation of what these figures illustrate would help readers understand their significance. | In the revised version, we have added the appropriate labels to each of the graphics. In addition, we have reorganized the tables and figures to ensure they are introduced and referenced in the correct sequence. We add Table 4 to explain results. |
| Discussion | 3: (fair)  Discussion should be added in this section. Discussion should consist of: a) interpretation of the finding, b) comparison to other studies, c) weaknesses or limitation, d) implication to society. | **3**: (fair)  While the section presents the classification results, it lacks an in-depth discussion of the implications of these results. What do the higher values of CA and AUC mean for disease prediction and diagnosis? How do these improvements align with the initial goals of the research?  While you mention the limitations of the study in terms of dataset source and disease variety, you could further elaborate on the potential impact of these limitations. How might the findings change if a more diverse dataset were used? How does the single-hospital dataset affect the applicability of the proposed approach?  Compare the results of this study with other similar studies | In the revised version, we have expanded the discussion to include the implications of the higher values of classification accuracy (CA) and area under the curve (AUC) for disease prediction and diagnosis in Table 4.. In the revised version, we have provided a more detailed discussion of these limitations. We appreciate your suggestion to compare the results of our study with other similar studies. However, this real dataset is collected to local hospital, so that we only use various algorithms to compare. |
| Conclusion | 4: (good)  complete | **3**: (fair)  The conclusion does not provide a concise recap of the main findings and contributions of the study. This is important to remind the reader of the study's significance.  While you mention that the results demonstrate the potential of machine learning, the impact of the findings on the medical field and patient care could be more explicitly stated. How exactly does the high accuracy and efficiency of disease diagnosis and treatment improve patient outcomes and reduce healthcare costs? |  |
| References | 3: (fair)  1. AUthor should use springer citattion style, change your citation style: IEEE to springer computer science 2. Use at least 30 references | **3**: (fair)  The conclusion does not provide a concise recap of the main findings and contributions of the study. This is important to remind the reader of the study's significance.  While you mention that the results demonstrate the potential of machine learning, the impact of the findings on the medical field and patient care could be more explicitly stated. How exactly does the high accuracy and efficiency of disease diagnosis and treatment improve patient outcomes and reduce healthcare costs? | We have used 30 references and changed citation style |