Review of: "Comparative Analysis of Machine and Deep Learning Techniques for Text Classification with Emphasis on Data Preprocessing"

Amirmohammad Paksaz[1]

1 University of Iowa

I've reviewed the paper and found several areas where improvements are needed. While the analysis is thorough, there are critical issues that must be addressed to strengthen the work and make it more impactful. Below are specific comments and suggestions for improvement:

1. The methodology and results presented in the paper are well-structured and straightforward, making it easy to comprehend. However, enhancing reproducibility metrics with a more detailed description of data sets and model configurations would help other scientists use this research as a foundation for their own.

2. While the comparative analysis is comprehensive, a deeper exploration of the reasons behind the performance of particular models and preprocessing techniques could benefit the discussion section. Providing insights into why some approaches work better than others would further bolster the overall contribution of the paper.

3. The paper effectively compares established machine learning and deep learning methods, but it would benefit from introducing novel insights or approaches. Consider exploring less common preprocessing techniques or proposing innovative combinations of methods to add uniqueness to this study.

4. If the author wants to make the research more relevant to real-world situations, then it would be beneficial for him to discuss how his findings can be applied and what the trade-offs would be between the performance of the model and the amount of computational power needed. It would be useful for people who want to use these methods in practice.

5. The author might want to include a section about the computational complexity of different models he has examined. In real applications, it is important to find a trade-off between performance and resource usage. This is particularly true when one is faced with a decision on whether to use conventional machine learning algorithms or use deep learning ones.

Thank you.