

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

Muna Hadi Saleh¹

¹ University of Baghdad

Potential competing interests: No potential competing interests to declare.

The author was successful in choosing the algorithms (BiLSTM, CRF, and RWA) for Deep Learning and (NB, GB, and SVM) for Machine Learning, and the results are good, but please consider these questions to get more accurate research

- Will you please apply the table for comparison with the rest of the previous works?
- Can you provide more details on why the execution time for each algorithm is important for the study?
- Could you specify what data preprocessing steps were used, and how they compare to those used in previous works?
- Can you elaborate on the specific advantages and disadvantages of each algorithm used in the study?
- Could you provide examples or references to previous works that highlight the importance of execution time in text classification studies?
- Can you suggest specific metrics or criteria that could be included in the comparison table to enhance its usefulness?
- Could you recommend any specific studies or papers that the authors should reference when discussing execution time and its impact on text classification?
- Can you provide specific examples of how the data preprocessing steps you mentioned (eliminating duplicates, filling in blanks) impacted the results?
- Could you suggest additional metrics or criteria (e.g., precision, recall) that would be useful to include in the comparison table to provide a more comprehensive evaluation of the algorithms?
- Can you provide specific examples of how execution time differences impacted the performance of the algorithms?
- Could you suggest additional data preprocessing techniques that could improve the accuracy and efficiency of the models?
- Can you provide specific examples of previous works that used similar data preprocessing techniques and how they impacted the results?
- Could you suggest how the inclusion of additional metrics such as F1 score or AUC could provide a more comprehensive evaluation of the algorithms?
- Can you provide specific examples or case studies that demonstrate the impact of data preprocessing on model performance?
- Could you include a brief explanation of how each algorithm's performance was measured and compared in the study?
- Can you provide a specific example of a study where execution time significantly impacted the choice of algorithm?

- Could you explain how the accuracy of 98.5% for BiLSTM compares to the accuracy rates of the other algorithms tested?