

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

Robert Pavur¹

¹ University of North Texas

Potential competing interests: No potential competing interests to declare.

This study illustrates the advantage of using BiLSTM in classifying observations. The literature review and explanation of the deep learning and machine learning methods were informative. The authors note that BiLSTM models can discard irrelevant information and retain the order of the remaining information. The illustration of its performance with the well-known Titanic Disaster data was helpful in understanding its performance.

Some recommendations for improving the quality of the paper are as follows.

1. Some paragraphs are way too long, particularly in Section 2. Breaking them up into coherent smaller paragraphs will make the paper more readable.
2. Spell out all acronyms the first time that they are presented. For example, CRF appears in the text and is not spelled out. Check each acronym.
3. Mention that deep learning is actually a subset of machine learning, but in your study, you are separating it from other traditionally used machine learning methods.
4. Figures 9 and 10 show that BiLSTM's accuracy is higher for the test data than for the training data. It should be the opposite. This would be unusual and appears to be an error in the labelling of the graphs.