

Peer Review

# Review of: "Advancing Sentiment Analysis: A Novel LSTM Framework with Multi-head Attention"

K.India Sudheer Reddy<sup>1</sup>

1. Independent researcher

This paper, "Advancing Sentiment Analysis: A Novel LSTM Framework with Multi-head Attention," proposes a novel sentiment classification model based on an LSTM architecture enhanced with a multi-head attention mechanism and TF-IDF feature extraction. The authors claim significant improvements in performance on benchmark datasets over standard LSTM and transformer-based models like BERT and RoBERTa. The model outperforms common baselines (BERT, RoBERTa) in classification accuracy, with a test accuracy of **80.28%**, indicating a well-optimized architecture.

1. Consider quantifying improvements more precisely in the abstract (e.g., "achieves 12% gain in accuracy over baseline LSTM").
2. Consider elaborating more on *why* the combination of TF-IDF and multi-head attention resolves feature conflicts better than other embedding methods. The authors should strengthen the theoretical justification and possibly add empirical insights on **why** this hybrid combination improves performance.
3. Introduction - Briefly mention related works using multi-head attention with RNNs to position the novelty more clearly.
4. Justify with pseudocode to prove how TF-IDF vectors interact with the LSTM inputs and attention layers.
5. Performance is measured on both training and test sets. However, the **training accuracy (99.64%)** is much higher than the **test accuracy (80.28%)**, indicating **possible overfitting**. Justify with **cross-validation results** or confidence intervals to verify generalization. **Training Accuracy vs. Test Accuracy Gap**

**Training Accuracy: 99.64%, Test Accuracy: 80.28%, Gap: 19.36 percentage points**

*This large discrepancy strongly suggests the model has **memorized** the training data rather than learning generalizable patterns. While some performance drop is expected, a gap of nearly 20% is typically indicative of **overfitting**. Adding dropout layers, using early stoppers, and so on is recommended.*

1. Add scientific details for the fusion mechanism and attention integration.
2. While applying regularization techniques like dropout or early stopping, mention them if already used.
3. The Discussion and Conclusion sections are somewhat repetitive and could be merged.
4. Some places have inconsistent phrasing (e.g., "necessity and necessity," "attention-LSTM integration... to clean hidden states"). Rephrase these.
5. Adding statistical significance (p-values) in Table 2 would strengthen the claims.
6. Figures 3-5 are referred to in the text, but actual images were not embedded in the PDF version reviewed. Please ensure proper formatting in the final version.
7. Some citations are informal (e.g., "<sup>^</sup>Zhang Y, et al.")—ensure consistency with referencing style guidelines.
8. Improve transitions between subsections to enhance logical flow.
9. Compare against additional models or perform statistical tests to confirm significance.

## **Declarations**

**Potential competing interests:** No potential competing interests to declare.