

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

Sarah Shafqat¹

¹ International Islamic University, Islamabad

Potential competing interests: No potential competing interests to declare.

The concept of paper is good and well documented.

I'm recommending some recent references to cite for big data text in tabular form categorization using bi-LSTM, deep learning heuristics, RNNs, and CNNs. The below-mentioned papers have analyzed Big Complex EHR data for multi-class multi-label diagnoses using bi-LSTM and deep learning heuristics, reaching 100% accuracy using a custom auto ML/DL technique. It would greatly strengthen your findings.

[1] Shafqat, S., Fayyaz, M., Khattak, H.A. *et al.*, Leveraging Deep Learning for Designing Healthcare Analytics Heuristics for Diagnostics. *Neural Process Lett* **55**, 53–79 (2023). <https://doi.org/10.1007/s11063-021-10425-w>

[2] Shafqat, S., Anwar, Z., Javaid, Q., & Ahmad, H. F. (2024). NER Sequence Embedding of Unified Medical Corpora to Incorporate Semantic Intelligence in Big Data Healthcare Diagnostics.

[3] Sarah Shafqat, Zahid Anwar, Raihan Ur Rasool, Qaisar Javaid, et al. (2024). Rules Extraction, Diagnoses, and Prognosis of Diabetes and its Comorbidities using Deep Learning Analytics with Semantics on Big Data. Qeios. doi:10.32388/67KZ7S.3.