# Review of: "Empowering Dysarthric Speech: Leveraging Advanced LLMs for Accurate Speech Correction and Multimodal Emotion Analysis"

Dr. Narina Thakur[1]

1 Independent researcher

Potential competing interests: No potential competing interests to declare.

The paper addresses a significant communication challenge faced by individuals with dysarthria by proposing an advanced solution using state-of-the-art language models (OpenAI's Whisper, LLaMa 3.1, and Mistral). Leveraging AI accelerators and advanced language models for both speech correction and emotion analysis is impressive, highlighting the paper's technical sophistication. The framework has strong potential to improve communication for individuals with speech disorders, thereby enhancing their quality of life and social interaction. The proposed approach not only corrects speech but also identifies emotional context, a critical component for meaningful communication.

The paper may be accepted, subject to major revisions.

1. The literature review lacks depth, with only six references, and lacks recent studies that could better contextualize and justify the proposed methodology.
2. Table 1 - Comparison Issues: The models are compared based on individual sentence prediction accuracy rather than on a uniform standard, which reduces the clarity and reliability of the comparison.
3. The methodology and results analysis section is well-presented, offering clear and detailed explanations of the approach and its outcomes. The algorithm used is robust and effectively supports the research objectives.
4. Paper Structure and Formatting: The paper structure needs refinement; some equations are missing numbers, affecting readability and presentation.
5. Dataset Clarity: While the TORGO dataset and Google speech data are mentioned, further detail on the dataset construction and labeling process would enhance transparency.
6. The study makes a meaningful contribution to dysarthric speech recognition by incorporating both advanced language models and emotion recognition, which can notably improve user experience. However, mentioning specific improvements in accuracy or user testing results in the conclusion would reinforce the work's practical utility and impact.
7. Clarity and Completeness is required in the conclusion: While the conclusion captures the study's intent and achievements, adding specific insights or quantitative results from the research findings would provide readers with a clearer understanding of the effectiveness and limitations of the current model.
8. Holistic Communication Enhancement: The inclusion of emotion recognition in the model demonstrates a thoughtful approach to enhancing communication for dysarthric individuals. Further detail on how emotion analysis aids

communication in practical settings would add depth to the conclusion, underscoring the model's holistic benefits for users.