

# Review of: "Current Trends in the Use of Machine Learning for Error Correction in Ukrainian Texts"

Arash Heidari<sup>1</sup>

<sup>1</sup> Institute of Electrical and Electronics Engineers (IEEE)

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

Provide a more detailed explanation of the methodologies used for the analysis of error correction tools, including the criteria for tool selection and the specific metrics used for comparison.

Elaborate on the specific characteristics and size of the existing data corpora analyzed. Include statistical information about the corpus, such as the number of sentences, types of errors annotated, and the annotation process's inter-annotator agreement.

Include a comprehensive description of the machine learning model architectures evaluated, specifying hyperparameters, training procedures, and the rationale behind choosing certain architectures over others.

Perform a thorough error analysis to categorize the types of errors most frequently encountered in Ukrainian-language texts and discuss the performance of different models on each error category.

Compare the proposed models against strong baseline models in addition to state-of-the-art models, providing a clear performance benchmark.

Conduct ablation studies to understand the contribution of different components of the proposed models. This would help in identifying which parts of the model are most crucial for performance improvements.

Test the models on texts from various domains (e.g., literary, technical, colloquial) to assess their robustness and generalizability across different types of Ukrainian texts.

Discuss how the models handle context when correcting errors. Provide examples of context-sensitive corrections and analyze situations where the models might fail due to lack of context.

Include a section comparing human error correction performance with that of the models. This could highlight the current gap between human expertise and automated systems and suggest areas for improvement.

Expand the discussion on future research directions, including scalability issues for large-scale deployment and potential integration with other linguistic resources and tools. Discuss possible advancements in model interpretability and user feedback mechanisms to improve correction accuracy.