

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

Ahmed Hamdi¹

¹ Université de La Rochelle

Potential competing interests: No potential competing interests to declare.

The article provides an in-depth examination of the challenges in identifying and correcting errors in Ukrainian-language texts, offering a comprehensive analysis of current research, tools, and data corpora. It emphasizes the necessity of developing a large annotated corpus and creating machine learning algorithms specifically designed for morphologically complex languages like Ukrainian.

The paper thoroughly reviews the state-of-the-art efforts in natural language processing (NLP) for Ukrainian, covering a range of approaches from traditional machine learning models to the latest advancements using transformers. Various existing systems have been evaluated using the UA-GEC dataset, with results indicating that precision generally surpasses recall. This discrepancy raises intriguing questions about the consistently lower recall rates and the choice of using the F0.5 score instead of the F1-score.

Additionally, including examples with translations would enhance clarity for non-Ukrainian speakers. While translating non-correct phrases might be challenging, providing descriptions of the errors could significantly aid readers' understanding.

The paper lacks a detailed discussion with valuable insights to help guide the development of more effective correction algorithms dealing with specific linguistic challenges unique to Ukrainian.

The article is highly valuable as a survey of NLP efforts in the Ukrainian language, providing a comprehensive summary of existing research. I believe it should be oriented in this direction.