

ІНТЕРНЕТ-СЕРВІСИ ДЛЯ ПЕРЕВІРКИ ОРФОГРАФІЧНИХ НОРМ УКРАЇНСЬКОЇ МОВИ

¹ Вінницький національний технічний університет

Анотація

У роботі проаналізовано інтернет-сервіси для перевірки орфографічних норм української мови. Було протестовано роботу сервісів на звичайних алгоритмах та на основі ШІ. Виявлено найбільш ефективні універсальні текстові моделі нейромереж.

Ключові слова: орфографічні норми, інтернет-сервіси, штучний інтелект.

Abstract

The paper analyzes online services for checking the spelling of the Ukrainian language. The services were tested using conventional algorithms and AI-based ones. The most effective universal text models of neural networks were identified.

Keywords: orthographic norms, Internet services, artificial intelligence.

Вступ

Сучасний користувач інтернет-мережі прагне дедалі більше скоротити витрати часу на поточні завдання й, де можливо, делегувати їхнє виконання інструментам штучного інтелекту. Можливість автоматичної підготовки презентацій та текстів за допомогою штучного інтелекту із мінімальним втручанням автора стає все популярнішою з кожним днем. Проте унаслідок цього часто виникає проблема коректності таких текстів, адже перекладення частини власної роботи на цю технологічну новинку може як зекономити час, так і погіршити становище. Зокрема, це стосується й генерування текстів українською мовою, де часто трапляються різні мовні помилки.

Метою роботи є дослідження інтернет-сервісів для перевірки орфографічних норм, аналіз ефективності їхньої роботи і визначення найбільш оптимального сервісу для редагування українськомовних текстів.

Результати дослідження

Серед популярних інтернет-сервісів для перевірки орфографії [3] виявлено лише три, які підтримують українську мову, третина ресурсів розрахована винятково для мобільних пристроїв, що може бути проблемою, коли користувач надає перевагу працювати з ПК. Потрібні інші сервіси, які б викликати довіру в користувача, який не має часу або здібностей, щоб перевірити власні помилки самостійно. Одним із ключових питань є якість орфографічної перевірки.

Для тестування сервісів було створено текст із помилками – частина тексту згенерована за допомогою ШІ, щоб максимально відтворити його стиль типових помилок, інша частина містить навмисні помилки, зроблені людиною.

Перший сервісом «Мова – ДНК нації» (див. перелік використаних інтернет-сервісів у списку використаної літератури [1; 4-9]), попри перше місце в рейтингах багатьох сайтів, продемонстрував низьку здатність до виявлення помилок (25). Ідентичні результати показав сервіс UkrainianCorrection.com, підтвердивши загальну тенденцію – ці сервіси користуються схожими алгоритмами, що робить їх недостатньо надійними.

Перспективним рішенням у сфері перевірки тексту на помилки виявився сервіс LanguageTool, що застосовує елементи штучного інтелекту для аналізу граматики та орфографії. Цей сервіс продемонстрував дещо кращу ефективність – 30 виявлених помилок. Проте варто зазначити, що навіть цей сервіс не зміг розпізнати зайві великі літери в окремих словах та виправити звичайні одруківки (літеру «с» замінити на «з»).

Аналогічні результати продемонстрував інший сервіс — istio.com. На відміну від LanguageTool, він не використовує штучний інтелект для аналізу тексту. Хоч istio.com і виявився ефективним у виявленні значної кількості помилок, проте так само, як і LanguageTool, не «побачив» усі помилки.

Сервіс aspose.ai. також застосовує штучний інтелект, початкові результати його роботи були обнадійливими. На початку він впевнено розпізнавав і виправляв численні помилки, що виникали під час тестування. Проте сервіс просто видалив значну частину тексту, трактуючи його як «помилку». Подібних випадків не траплялося під час тестування попередніх сервісів. Очевидно, що навіть ІІІ-базовані рішення мають обмеження в контексті нових змін української мови.

Позитивною реалізацією потенціалу штучного інтелекту став сервіс WordCount. Він успішно проаналізував увесь текст, розпізнавши і виправивши всі помилки без винятків, забезпечивши бездоганний результат.

Для більшої доказовості експерименту було вирішено випробувати також універсальні можливості стандартної моделі ChatGPT 4.0. Це дало змогу порівняти роботу вузькоспеціалізованого сервісу, який через свою специфіку може бути обмеженим, з більш гнучкою текстовою моделлю. Кінцеві результати роботи обох систем виявилися ідентичними – кожна виявила 43 помилки. Різниця була настільки незначною, що єдиною зміною стало рішення ChatGPT замінити слово «їм» на «для них» в уривку «ім головне свято», що, за його словами, мало надати тексту більш природного стилю. Результати тестування розміщено на рисунку 1:

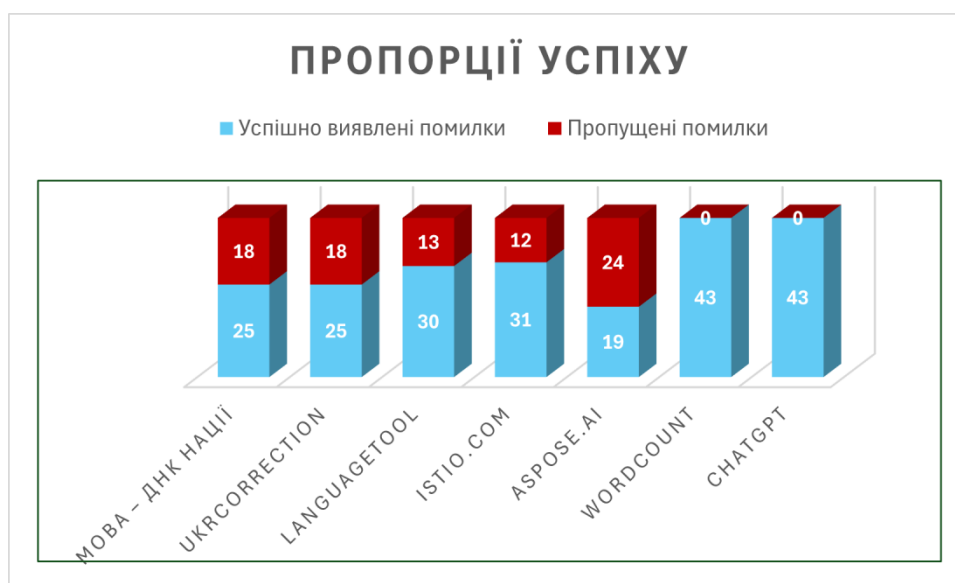


Рисунок 1. Результати тестування інтернет-сервісів для перевірки орфографічних помилок

Висновки

Отже, стандартні алгоритми перевірки орфографічних помилок в українськомовних текстах є застарілими і не зовсім ефективними. Окремі сервіси, які використовують ІІІ або текстові моделі, виявилися більш результативними. Варто звертати увагу, що довіряти ІІІ таку критичну роботу, як перевірку помилок, потрібно з обачністю, адже, як зрозуміло в ході дослідження, далеко не всі сервіси здатні якісно працювати саме з українськомовними текстами.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Мова – ДНК нації. URL: <https://ukr-mova.in.ua/perevirka-tekstu>
2. Радомська Л., Вітковський М. Інформаційні ресурси для перевірки й коригування українськомовних текстів. *Філологічний часопис*. 2024. 2(24). С. 80–96. [https://doi.org/10.31499/2415-8828.2\(24\).2024.317042](https://doi.org/10.31499/2415-8828.2(24).2024.317042).
3. ТОП-10 сервісів для перевірки орфографії та пунктуації URL: <https://web-promo.ua/ua/blog/top-10-servisiv-dlya-perevirki-orfografiyi-ta-punktuaciyi/>
4. Aspose.ai. URL: <https://products.aspose.ai/words/uk/grammar/> Gamma app. URL : <https://gamma.app/>
5. Chat-GPT 4o. URL: <https://chatgpt.com>.
6. Istio.com. URL: <https://istio.com/uk/text/orfo>.
7. LanguageTool: Перевірка граматики. URL: <https://languagetool.org/uk>
8. UkrainianCorrection. URL: <https://www.ukrainiancorrection.com/>
9. WordCount: Grammar checker. URL: <https://wordcount.com/uk/grammar-checker>

Радомська Людмила Анатоліївна – кандидат філологічних наук, доцент кафедри мовознавства, Вінницький національний технічний університет, м. Вінниця, e-mail: ludarad9@gmail.com.

Вітковський Микола Романович – студент групи 5ПІ-236, Вінницький національний технічний університет, м. Вінниця. e-mail: vitkovskijm@gmail.com

Radomska Lyudmyla A. – Cand. Sc. (Eng), Assistant Professor, Department of Linguistics, Vinnytsia National Technical University, Vinnytsia, e-mail: ludarad9@gmail.com.

Vitkovskiy Mykola R. – student of 5PE-23b group, Vinnytsia National Technical University, Vinnytsya. e-mail: vitkovskijm@gmail.com