УДК 004.415.2:004.912:004.8

ПРОЄКТУВАННЯ ТА РОЗРОБКА ЗАСТОСУНКУ ДЛЯ ВЕДЕННЯ НОТАТОК З АВТОМАТИЧНИМ УЗАГАЛЬНЕННЯМ

Шевченко Д. А. [1]; Новікова Т.О. [2]
[1,2] Національний університет «Одеська політехніка»
[1,2] Україна, Одеса;
[1] 8084054@stud.op.edu.ua; [2] novikova.t.o@op.edu.ua

Анотація. *Під час дослідження спроєктовано, розроблено та здійснено експериментальне оцінювання вебзастосунку Notach для ведення текстових нотаток з автоматичним узагальненням. Визначено функціональні вимоги, обґрунтовано хмарну архітектуру та засоби реалізації інтелектуальних функцій. Реалізовано контроль версій і захист даних. Проведено експериментальне оцінювання якості та швидкодії автоматичного узагальнення нотаток.*
Ключові слова: *вебзастосунок; ведення нотаток; автоматичне узагальнення; штучний інтелект; хмарна архітектура; Firebase; Firestore.*

Вступ. Зростання кількості навчальних конспектів, робочих записів і персональних нотаток ускладнює пошук та повторне використання інформації. Традиційні редактори забезпечують зберігання й упорядкування текстів, але не завжди дозволяють швидко визначити головні положення довгих документів. Методи абстрактивного узагальнення на основі трансформерів дають змогу формувати зв’язний стислий виклад [1]. Водночас генеративні моделі можуть додавати непідтверджені твердження або пропускати важливі факти [2]. Тому актуальним є проєктування та розробка застосунку, в якому автоматичне узагальнення поєднується з повноцінним керуванням нотатками, перевіркою результатів і збереженням першоджерела.

Метою роботи є проєктування та розробка вебзастосунку Notach для створення, редагування, зберігання та пошуку нотаток із можливістю автоматичного формування короткого змісту, основних думок і тематичних тегів. Для досягнення мети сформовано вимоги, порівняно варіанти розміщення компонента штучного інтелекту, визначено структуру даних, спроєктовано програмні модулі, розроблено алгоритм оброблення тексту та проведено експериментальне оцінювання якості й швидкодії інтелектуального компонента.

Основна частина. Під час аналізу вимог виділено основні функції застосунку (автентифікація, створення, редагування, видалення, пошук і групування нотаток) та інтелектуальні функції (узагальнення, виділення ключових думок і генерація тегів). На архітектурному етапі порівняно чотири варіанти розміщення мовної моделі: на персональному комп’ютері, на телефоні, у власній хмарній інфраструктурі та у віддаленому AI-сервісі через захищений серверний адаптер. Для розробки Notach обрано останній варіант – хмарне виконання інтелектуальних функцій з відокремленням їх від редагування та зберігання нотаток. Це дозволяє оновлювати AI-компонент без змін клієнтського інтерфейсу.

Вебзастосунок Notach розроблено з використанням React і TypeScript для клієнтської частини, Node.js/Express для серверної частини та сервісів Firebase. Firebase Authentication відповідає за автентифікацію, Cloud Firestore – за зберігання нотаток, метаданих і результатів узагальнення, Firebase Hosting – за розміщення клієнтської частини. Реалізовано функції створення, редагування, зберігання, пошуку та групування нотаток, а також взаємодію з інтелектуальним компонентом. Окремий серверний маршрут приймає запит на узагальнення,

перевіряє токен користувача та його право доступу до нотатки, виконує валідацію тексту й передає запит адаптеру мовної моделі. Секретний ключ зовнішнього AI-сервісу не передається браузеру. Доступ до документів Firestore додатково обмежується правилами безпеки [3].

Інтелектуальну підсистему спроєктовано та реалізовано на основі окремого програмного інтерфейсу, незалежного від конкретного постачальника моделі. На вхід подаються текст нотатки, мова результату, бажана довжина та ознака формування тегів. На виході система формує структуровану відповідь із коротким змістом, переліком основних думок і запропонованими тегами. Сервер перевіряє формат отриманих даних перед відображенням. Для довгих записів передбачено поділ тексту на фрагменти, узагальнення окремих частин та об'єднання проміжних результатів із контролем повноти обробки.

Під час проєктування моделі даних передбачено сутності користувача, нотатки, узагальнення та оцінки результату. Оригінальний текст і сформований короткий зміст зберігаються окремо. Поле revision у нотатці та sourceRevision у результаті дозволяють встановити, для якої редакції створено узагальнення. Якщо текст змінено під час генерації, результат не підміняє актуальної редакції. Користувач може переглянути, відредагувати та оцінити узагальнення, не змінюючи вихідної нотатки. Відмова AI-сервісу не перешкоджає базовим операціям зберігання.

Для перевірки розробленого вебзастосунку сформовано сценарії функціонального тестування реєстрації, збереження, пошуку, узагальнення та редагування результатів. Експериментальне оцінювання інтелектуального компонента проведено на наборі з 60 текстових нотаток: 20 навчальних конспектів, 20 робочих протоколів і 20 особистих записів. Тексти розподілено на три групи за обсягом: 300–800, 801–2000 та 2001–5000 слів. Оцінювання здійснено за показниками покриття ключових думок, фактичної узгодженості, швидкодії, корисності результатів та релевантності тематичних тегів. Методика поєднує кількісні вимірювання з експертним оцінюванням, що відповідає висновкам дослідження SummEval щодо обмежень автоматичних метрик якості узагальнення [4].

За результатами експериментального оцінювання середнє покриття ключових думок становило 84,7 %, фактична узгодженість – 94,0 %, а середній час формування узагальнення – 6,1 с. Для коротких нотаток середній час оброблення становив 3,4 с, для середніх – 5,6 с, для довгих – 9,2 с. Значення 95-го процентиля часу відповіді (p95) становили відповідно 5,1 с, 8,4 с та 13,8 с. Середня експертна оцінка корисності узагальнень досягла 4,17 бала з 5, а точність визначення тематичних тегів – 78,3 %. Отримані показники свідчать про досягнення встановлених критеріїв якості та швидкодії для текстів обсягом до 2000 слів. Для довгих нотаток виявлено збільшення часу оброблення та зниження повноти збереження ключових думок, що обґрунтовує необхідність подальшої оптимізації алгоритму узагальнення.

Висновки. У роботі спроєктовано та розроблено вебзастосунок Notach для створення, редагування, зберігання та пошуку текстових нотаток із можливістю автоматичного формування короткого змісту, основних думок і тематичних тегів. Сформовано функціональні вимоги, обґрунтовано хмарну архітектуру та реалізовано основні програмні модулі застосунку. Розроблено інтелектуальну підсистему із захищеним серверним адаптером, реалізовано роздільне зберігання нотаток і результатів узагальнення та контроль редакцій тексту. Проведено експериментальне оцінювання розробленого рішення на наборі з 60 нотаток різного обсягу й тематики. Середнє покриття ключових думок становило 84,7 %, фактична узгодженість – 94,0 %, середній час формування узагальнення – 6,1 с, а середня експертна оцінка корисності – 4,17 бала з 5. За результатами експериментального оцінювання виявлено тенденцію до збільшення часу формування узагальнення та зниження повноти збереження

ключових думок зі зростанням обсягу вхідного тексту. Отримані результати підтверджують працездатність розробленого вебзастосунку та визначають напрями його подальшого вдосконалення.

Використання інструментів ШІ. Під час підготовки публікації використано мовну модель OpenAI ChatGPT (модель GPT-6) для покращення академічних формулювань у вступній, основній частинах і висновках та перекладу англомовних анотацій (розділ «Abstract»).

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

[1] Lewis, M., Liu, Y., Goyal, N., Ghazvininejad, M., Mohamed, A., Levy, O., Stoyanov, V., & Zettlemoyer, L. (2020). BART: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension. In D. Jurafsky, J. Chai, N. Schluter, & J. Tetreault (Eds.), *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* (pp. 7871–7880). Association for Computational Linguistics.
https://doi.org/10.18653/v1/2020.acl-main.703

[2] Maynez, J., Narayan, S., Bohnet, B., & McDonald, R. (2020). On faithfulness and factuality in abstractive summarization. In D. Jurafsky, J. Chai, N. Schluter, & J. Tetreault (Eds.), *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* (pp. 1906–1919). Association for Computational Linguistics.
https://doi.org/10.18653/v1/2020.acl-main.173

[3] Firebase. (n.d.). *Get started with Cloud Firestore Security Rules*.
https://firebase.google.com/docs/firestore/security/get-started

[4] Fabbri, A. R., Kryściński, W., McCann, B., Xiong, C., Socher, R., & Radev, D. (2021). SummEval: Re-evaluating summarization evaluation. *Transactions of the Association for Computational Linguistics, 9*, 391–409.
https://doi.org/10.1162/tacl_a_00373

Shevchenko D. A.; Novikova T. O.

Design and development of an application for note-taking with automatic summarization

Abstract. *This paper presents the design, development, and experimental evaluation of Notach, a web application for managing text notes with automatic summarization. The functional requirements are defined, and the cloud-based architecture and technologies used to implement intelligent features are justified. Version control and data protection mechanisms are implemented. An experimental evaluation of the quality and performance of automatic note summarization was conducted.*

Keywords: *web application; note-taking; automatic summarization; artificial intelligence; cloud architecture; Firebase; Firestore.*