

WEB-РЕСУРС ДЛЯ ВЕДЕННЯ БЛОГІВ З МОДУЛЕМ ПЕРСОНАЛІЗОВАНИХ РЕКОМЕНДАЦІЙ

Вінницький національний технічний університет

Анотація

Розглянуто проблему відсутності прозорих механізмів персоналізації контенту у сучасних блогівих платформах. Запропоновано та реалізовано WEB-ресурс для ведення блогів із клієнт-серверною архітектурою та модулем персоналізованих рекомендацій на основі методу TF-IDF і косинусної міри подібності. Описано архітектурні рішення системи, функціональні модулі та три режими формування стрічки публікацій.

Ключові слова: WEB-ресурс, блог, персоналізовані рекомендації, TF-IDF, косинусна міра подібності, FastAPI, React, PostgreSQL.

Abstract

The problem of the absence of transparent content personalization mechanisms in modern blogging platforms is considered. A web resource for blogging with a client-server architecture and a personalized recommendation module based on the TF-IDF method and cosine similarity has been proposed and implemented. The architectural solutions of the system, functional modules, and three feed formation modes are described.

Keywords: web resource, blog, personalized recommendations, TF-IDF, cosine similarity, FastAPI, React, PostgreSQL

Вступ

За даними аналітичної платформи Statista, станом на 2024 рік у світі налічується понад 600 мільйонів активних блогів, а щодня публікується понад 6 мільйонів нових дописів [1]. Попри масштаб галузі, більшість популярних платформ використовують різні алгоритмічні підходи до формування стрічки контенту. Частина таких механізмів є закритою або лише частково доступною для налаштування користувачем, що обмежує можливості прямого керування відображенням матеріалів.

Метою роботи є розширення функціональних можливостей по керуванні стрічкою у веб-ресурсах для ведення блогів шляхом розробки модуля персоналізованих рекомендацій на основі аналізу текстових даних і взаємодій користувачів.

Результати дослідження

На першому етапі проведено аналіз існуючих платформ для публікації контенту, зокрема Instagram, Reddit та персональних блогів. Встановлено, що сучасні платформи використовують різні підходи до формування стрічки: від алгоритмічних рекомендацій до сортування за новизною, популярністю або підписками користувача. Наприклад, Reddit демонструє можливість гнучкішого керування відображенням матеріалів, тоді як в Instagram основний акцент зроблено на алгоритмічному доборі контенту. Це стало підставою для розробки власного WEB-ресурсу для ведення блогів, у якому користувач може самостійно обирати режим формування стрічки публікацій.

Систему побудовано за клієнт-серверною архітектурою. Серверна частина реалізована на основі фреймворку FastAPI (Python) з використанням реляційної СУБД PostgreSQL та ORM-бібліотеки SQLAlchemy. Авторизація користувачів забезпечується механізмом JWT-токенів. Клієнтська частина розроблена на базі бібліотеки React із застосуванням React Router та axios, що реалізує підхід Single Page Application. Обмін між частинами системи здійснюється через REST API у форматі JSON.

Ключовим елементом системи є модуль персоналізованих рекомендацій. На основі зафіксованих взаємодій користувача із публікаціями формується профіль інтересів. Текстовий вміст матеріалів векторизується методом TF-IDF (Term Frequency–Inverse Document Frequency), після чого тематична схожість між публікаціями визначається за допомогою косинусної міри подібності. Для нових

користувачів, у яких ще не накопичено достатньо взаємодій, передбачена fallback-логіка – система відображає хронологічну стрічку нових публікацій до моменту персоналізації.

Веб-ресурс підтримує три режими формування стрічки публікацій: «Нові» – хронологічний порядок відображення матеріалів; «Рекомендації» – персоналізована підбірка на основі TF-IDF та косинусної подібності; «Змішаний» – поєднання рекомендованих і нових публікацій. Перемикач між режимами доступний безпосередньо в інтерфейсі, що надає користувачу явний контроль над принципом формування контентної стрічки.

Функціональні можливості системи охоплюють реєстрацію та авторизацію користувачів, перегляд і створення публікацій, коментування матеріалів, а також фіксацію взаємодій для подальшої персоналізації стрічки. Перевірка серверної частини здійснювалась через автоматично згенеровану Swagger-документацію FastAPI; клієнтська частина верифікувалась у браузері за основними сценаріями використання.

Висновки

Розроблено WEB-ресурс для ведення блогів із підтримкою персоналізованої стрічки публікацій. Реалізовано клієнт-серверну архітектуру на основі FastAPI, React та PostgreSQL, механізм авторизації на JWT-токенах, а також модуль рекомендацій на основі TF-IDF та косинусної міри подібності. Впроваджено три режими формування стрічки – «Нові», «Рекомендації» та «Змішаний» – які забезпечують користувачу більш керований досвід взаємодії з контентом. На відміну від платформ, де логіка стрічки часто визначається переважно алгоритмічно, розроблений ресурс надає користувачу явний вибір режиму відображення матеріалів. Мета досягається за рахунок того, що на відміну від конкурентних рішень, має додаткові можливості по керуванні стрічкою.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Авраменко В. С. Проектування інформаційних систем : навч. посіб. / В. С. Авраменко, А. С. Авраменко. – Черкаси : Черкаський національний університет ім. Б. Хмельницького, 2017. – 434 с.
2. Гужва В. М. Інформаційні системи і технології на підприємствах : навч. посіб. / В. М. Гужва. – К. : КНЕУ, 2001. – 400 с.
3. Сухий О. Л. Алгоритми пошуку в інформаційних системах : методичні рекомендації / О. Л. Сухий. – К. : Інститут інформаційних технологій і засобів навчання НАПН України, 2015. – 36 с.
4. Яскевич В. О. JavaScript-бібліотека React : навч. посіб. / В. О. Яскевич. – Київ : Київський університет імені Бориса Грінченка, 2023. – 122 с.
5. Ricci F. Recommender Systems Handbook / F. Ricci, L. Rokach, B. Shapira. – 3rd ed. – New York : Springer, 2022. – 1050 p.
6. Manning C. D. Introduction to Information Retrieval / C. D. Manning, P. Raghavan, H. Schütze. – Cambridge : Cambridge University Press, 2008. – 482 p.
7. FastAPI Documentation [Електронний ресурс]. – Режим доступу: <https://fastapi.tiangolo.com/>
8. scikit-learn. Feature extraction: TF-IDF [Електронний ресурс]. – Режим доступу: https://scikit-learn.org/stable/modules/feature_extraction.html

Куца Дарина Вадимівна – студентка групи ЗКН-226, факультет інтелектуальних інформаційних технологій та автоматизації, Вінницький національний технічний університет, Вінниця, e-mail: dasha.kutsa05@gmail.com;

Науковий керівник: Сілагін Олексій Віталійович – доцент кафедри комп'ютерних наук, кандидат технічних наук, Вінницький національний технічний університет, доцент, м. Вінниця, email: avsilagin@vntu.edu.ua

Kutsa Daryna V. – Faculty of Intellectual Information Technologies and Automation, Vinnytsia National Technical University, Vinnytsia, e-mail: dasha.kutsa05@gmail.com;

Supervisor: Silahin Olesiy V. – Associate Professor of the Department of Computer Science, Candidate of Technical Sciences, Associate Professor, Vinnytsia National Technical University, Vinnytsia, email: avsilagin@vntu.edu.ua