



Р. М. Матківський, Ю. І. Грицюк

Національний університет "Львівська політехніка", м. Львів, Україна

СИСТЕМА РЕКОМЕНДАЦІЙ КНИГ МЕТОДАМИ ТА ЗАСОБАМИ ШТУЧНОГО ІНТЕЛЕКТУ

Виявлено, що сучасні рекомендаційні системи книг практично не враховують емоційного стану і поточного настрою читача, а також особливостей україномовного контенту, що призводить до зниження читацької активності, особливо серед молоді віком 18-39 років. Встановлено, що інтеграція мультимовних трансформерних моделей із багатокласовим емоційним аналізом відгуків за 14 класами емоцій дає змогу істотно підвищити релевантність, персоналізацію та суб'єктивну задоволеність користувачів від наданих рекомендацій. Запропоновано та реалізовано повнофункціональну вебсистему BookSense з п'ятьма режимами рекомендацій книг: семантичний пошук за вільним текстовим описом бажаної книги, рекомендації за бажаним емоційним впливом (натхнення, інтрига, затишок, радість тощо), "протилежні" рекомендації для свідомого розширення читацького кругозору та виходу за межі звичних жанрів і авторів, гібридні спільні рекомендації на підставі профілю іншого користувача та контентно-колаборативний режим на підставі спільно оцінених книг. Сформовано та очищено репрезентативний набір даних із 10 074 книг та 42 116 відгуків із відкритого джерела Amazon Books Reviews, з яких близько 18 % подано українською мовою. Для семантичного пошуку використано претреновану мультимовну модель paraphrase-multilingual-MiniLM-L12-v2, для багатокласового емоційного аналізу – модель boltuix/bert-emotion. Система розгорнута на сучасному стеку React + FastAPI + PostgreSQL + Qdrant. Експериментально встановлено, що за вільного використання всіх п'яти режимів середнє значення Precision@5 становить 69,4 %, MRR = 0,751, середня тривалість оброблення запиту – 0,94 с (максимум менше 2 с) навіть на обладнанні середнього рівня. Суб'єктивна оцінка релевантності емоційного режиму активними читачами ($n = 62$) становила 4,61-4,72 бала із 5,0. Запропонована система перевищує наявні українські платформи та більшість глобальних аналогів за поєднанням повної підтримки української мови, багатокласового емоційного аналізу, швидкості роботи та можливості локального розгортання. Результати створюють науково-практичні передумови для впровадження вебсистеми BookSense у бібліотеки, інтернет-книгарні, освітні платформи та державні програми підтримки читання молоді в Україні.

Ключові слова: рекомендаційні системи книг; семантичний пошук за запитом; емоційний аналіз вподобань читачів; оброблення природної мови; україномовний контент; мультимовні трансформерні моделі.

Вступ / Introduction

Стрімкий розвиток цифрових технологій та штучного інтелекту докорінно змінює спосіб взаємодії людини з літературою, створюючи нові виклики та можливості для сфери читання. Швидко зростання обсягів цифрового книжкового контенту, масове поширення електронних та аудіокниг, а також трансформація читацьких звичок під впливом короткоформатного контенту та соціальних мереж призводять до ситуації інформаційного перенасичення та важкості вибору відповідного джерела інформації. За даними Українського інституту книги за 2023 р. [31], тільки 33 % українців регулярно читають електронні книги, при цьому значна частина потенційних читачів зазначає труднощі з пошуком літератури, яка відповідала б не тільки жанровим уподобанням, а й поточному емоційному стану чи бажанню свідомо вийти за межі звичного кола авторів і жанрів.

Сучасні глобальні платформи, такі як Goodreads, Amazon Kindle та Blinkist, пропонують електронні кни-

ги переважно на підставі популярності, історії покупок або простої колаборативної фільтрації відгуків, практично не враховуючи емоційного контексту прочитаного і не надаючи інструментів для свідомого розширення читацького кругозору. В Україні проблема посилюється недостатньою адаптацією цих сервісів до україномовного контенту, слабким поданням локальних авторів у рекомендаційних алгоритмах та ігноруванням мовних особливостей, що призводить до зниження читацької активності серед молоді.

Аналіз дослідженості проблеми свідчить про те, що наявні рішення зосереджені на окремих особливостях рекомендаційних систем, таких як колаборативна фільтрація відгуків для прогнозування уподобань читачів на підставі їхньої поведінки (лайків, коментарів чи відгуків) чи контентний аналіз ними прочитаного для семантичної подібності текстів, але без комплексної інтеграції емоційного контексту та гібридних алгоритмів. Це створює істотну прогалину у вирішенні проблеми роз-

Інформація про авторів:

Матківський Роман Михайлович, магістрант, кафедра програмного забезпечення. Email: roman.matkivskyi.mpzip.2024@lpnu.ua;

<https://orcid.org/0009-0001-9815-6518>

Грицюк Юрій Іванович, д-р техн. наук, професор, кафедра програмного забезпечення. Email: yurii.i.hrytsiuk@lpnu.ua;

<https://orcid.org/0000-0001-8183-3466>

Цитування за ДСТУ: Матківський Р. М., Грицюк Ю. І. Система рекомендацій книг методами та засобами штучного інтелекту.

Науковий вісник НЛТУ України. 2025, т. 35, № 6. С. 84–95.

Citation APA: Matkivskyi, R. M., & Hrytsiuk, Yu. I. (2025). Book recommendation system using artificial intelligence methods and tools.

Scientific Bulletin of UNFU, 35(6), 84–95. <https://doi.org/10.36930/40350610>

роблення рекомендаційних систем, орієнтованих на українського читача, де потрібні інструменти, здатні розуміти семантичний зміст запитів, аналізувати емоції їхніх відгуків і пропонувати нестандартні рекомендації книг для розширення їх інтересів.

Паралельно спостережено стрімке зростання комерційного ринку нових рекомендаційних платформ, що охоплює сфери від освіти до розваг, де інформацію про продукти або послуги надають треті сторони. Такі платформи є основним типом багатоканальної електронної комерції та можуть бути способом удосконалення виробничого процесу. Проведені дослідження показують, що методи машинного навчання функціоналу рекомендаційних платформ, зокрема – мультимовні трансформерні моделі, демонструють високу ефективність у розпізнаванні семантики запитів читачів, проте без інтеграції з емоційним аналізом їхніх вподобань залишаються обмеженими за реальних умов використання.

За умов масового використання цифрових бібліотек виникає потреба у вирішенні послідовного набору завдань: від семантичного пошуку за описами вподобань читачів до надання рекомендацій книг на підставі їхніх емоцій та колаборативної фільтрації відгуків. Ієрархічна природа цих завдань зумовлює необхідність розроблення не окремих алгоритмів для реалізації пошукових запитів, а цілісної багаторівневої системи. Наразі не існує комплексних рішень, які поєднували б семантичний аналіз запитів, емоційну класифікацію вподобань та персоналізовані режими рекомендацій книг для українського середовища. Це становить актуальність цього дослідження, оскільки розроблення такої інтегрованої пошукової системи дасть змогу не тільки ефективно підбирати літературу потенційними читачами, але й сприяти формуванню сталої читачької аудиторії, їхньої культури та вподобань у цифровому середовищі.

Об'єкт дослідження – автоматизований підбір книг на підставі семантичного аналізу текстових запитів та емоційного стану користувачів.

Предмет дослідження – методи та засоби автоматизованого підбору книг читачами на підставі семантичного аналізу їхніх текстових запитів та емоційного стану, встановленого попередньо рекомендаційною системою, що дасть змогу підвищити точність персоналізованих рекомендацій та забезпечити релевантний добір літератури.

Мета роботи – розробити систему рекомендацій книг методами та засобами штучного інтелекту, що забезпечить семантичний аналіз запитів потенційних читачів, їхню емоційну класифікацію вподобань і колаборативну фільтрацію відгуків, а також використання гібридних режимів, таких як "протилежні" рекомендації книг без урахування мультимовного середовища для українського контенту.

Для досягнення зазначеної мети визначено такі основні завдання дослідження:

- проаналізувати наявні літературні джерела щодо методів оброблення природної мови в рекомендаційних системах, що дасть змогу визначити сучасний стан досліджень, виявити обмеження наявних підходів та сформулювати наукове підґрунтя для побудови інтегрованої системи рекомендацій;
- підібрати претреновані моделі для семантичного пошуку відповідного контенту та емоційного аналізу вподобань, що дасть змогу реалізувати точне векторне подан-

ня описів книг і запитів користувачів, а також коректно визначати емоційні характеристики читачьких відгуків;

- сформувати набір даних для зберігання параметрів книг і відгуків читачів для тестування моделей методами і засобами штучного інтелекту, що забезпечить можливість об'єктивного порівняння алгоритмів та налаштування системи рекомендацій на підставі репрезентативних даних;
- реалізувати інтегровану систему з п'ятьма режимами рекомендацій книг, яка поєднувала б семантичний аналіз запитів потенційних читачів, їхню емоційну класифікацію вподобань та персоналізовані режими рекомендацій книг для українського середовища;
- здійснити експериментальне оцінювання точності та ефективності роботи системи рекомендацій книг, що дасть змогу визначити результативність запропонованих методів, порівняти їх з альтернативними рішеннями та обґрунтувати доцільність впровадження розробленої системи.

Аналіз останніх досліджень та публікацій. У роботі [19] автори розробили модель Word2Vec з архітектурами CBOW і Skip-gram, які навчаються на парах слів для створення щільних векторних подань, захоплюючи семантичні та синтаксичні зв'язки, хоча залишаються контекстно-нечутливими. Автори зазначають: "Розподілені подання слів у векторному просторі допомагають алгоритмам навчання досягти кращої продуктивності у завданнях оброблення природної мови шляхом групування подібних слів".

У дослідженні [6] автори запропонували використовувати модель BERT (англ. *Bidirectional Encoder Representations from Transformers*), яка застосовує двоспрямоване навчання для врахування глибокого контексту речень і досягає покращення на 5-10 % порівняно з попередніми моделями у бенчмарках GLUE, що робить її ефективною для класифікації вподобань читачів та аналізу відгуків на книги.

У роботі [29] Реймерс і Гуревич навели Sentence-BERT – адаптацію трансформерів для обчислення подібності речень через сіамські мережі, що забезпечує точність семантичного пошуку на рівні 85 % і є особливо придатною для рекомендаційних систем книг на підставі порівняння описів та запитів.

У роботі [5] дослідники розробили гібридний ансамблевий підхід, який поєднує аналіз настрою відгуків, контентну та колаборативну фільтрацію, що дає змогу ефективно вирішувати проблему холодного старту та підвищує точність рекомендацій книг на 15–20 % порівняно з базовими моделями.

У публікації [1] автори (MDPI) запропонували чотирирівневу стратегію рекомендацій книг, яка базується на семантичних мережах і настрої-аналізі, досягаючи точності групування подібних речень 92 %, хоча підхід має обмеження через відсутність глибокого емоційного компонента для персоналізації.

У дослідженні [4] науковці застосували BERT для багатокласового розпізнавання емоцій у твітах з точністю 78 %, продемонструвавши можливість адаптації моделі до аналізу відгуків на книги, однак зазначили потребу додаткової інтеграції із семантикою для уникнення переоцінки нейтральних текстів.

У препринті [17] (*arXiv*) автори інтегрували настрої-аналіз у рекомендаційні системи електронної комерції, підвищивши точність рекомендацій на 10-15 % завдяки пояснюваному розбору емоцій у відгуках, хоча

конкретне застосування до рекомендацій книг не розглядали.

У роботі [20] (Elsevier) дослідники використали Elmap-нейронні мережі з критерієм MRMR для класифікації сентименту, досягнувши F1-score 0.85, і показали потенціал такого підходу для рекомендацій освітнього контенту, аналогічного книгам, проте проігнорували мультимовний аспект і культурні нюанси.

У роботі [34] (Nature) автори збагатили модель для управління роботами Decision Transformer емоційними та сентиментними ознаками, витягнутими за допомогою NLP з відгуків, що покращило точність рекомендацій книг на 20 %, хоча мультимовний контент і україномовні джерела не розглядали.

У публікації [15] (Springer) дослідники проаналізували обмеження окремих підходів (контентних, колаборативних і сентимент-аналітичних), вказавши на проблемі холодного старту, розрідженість даних і шум у відгуках, та обґрунтували необхідність гібридних моделей для надійної емоційної ідентифікації.

У роботі [11] автори вдосконалили семантичний пошук шляхом інтеграції знань із WordNet у 23 різні моделі вбудовування речень, досягнувши статистично значущого покращення середньої точності (MAP) більше ніж на 2 % ($p = 0,047$) на бенчмарку Cranfield, що підтверджує переваги розширення запитів і документів людськими знаннями.

У дослідженні [33] автори розробили мультимодальну систему EMO-READS, яка поєднує BERT, LSTM, текстові відгуки, рейтинги та (в оригіналі) вирази обличчя, досягнувши точності класифікації до 90 %, і виявили ключові виклики: залежність від якості даних, шум та складність реального часу інтеграції емоцій.

Аналіз останніх досліджень та публікацій виявив прогалину: попри значний прогрес у семантичному аналізі, емоційній класифікації та гібридних рекомендаційних підходах, жодна з розглянутих робіт не реалізує комплексної багаторівневої системи, яка одночасно інтегрувала б семантику запитів, емоційний аналіз вподобань і відгуків, гібридні режими (включно з "протилежними" рекомендаціями) та підтримку мультимовного, зокрема україномовного, контенту. Наявні рішення зосереджені переважно на окремих компонентах або англomовному середовищі, що зумовлює необхідність розроблення нового інтегрованого методу для українського ринку літератури.

Матеріали та методи дослідження. Для розроблення системи рекомендацій книг BookSense використано комплекс сучасних методів оброблення природної мови та машинного навчання, що забезпечує повну відтворюваність результатів дослідження за наявності аналогічного набору даних і відкритих бібліотек.

Результати дослідження та їх обговорення / Research results and their discussion

Аналіз ринку вебсервісів для рекомендацій книг вказав на його динамічність і продемонстрував стійке зростання через збільшення попиту на персоналізовані цифрові продукти. Сучасні читачі прагнуть отримувати книжкові рекомендації книг, які враховують не тільки жанрові уподобання, такі як емоційний стан вподобань, але й цілі читання. Такий підхід потребує застосування гнучких алгоритмів аналізу поведінки потенційних чи-

тачів і їх запитів та персоналізовані режими рекомендацій книг для україномовного середовища.

Серед найвідоміших сучасних конкурентних рішень можна виокремити:

- Goodreads, який має значну базу даних та активну спільноту, але використовує здебільшого рекомендації книг на підставі популярності, без глибокої персоналізації, пропонуючи функції створення списків прочитаних книг, оцінок, відгуків і соціальних груп, однак не враховує користувача вподобання, що обмежує релевантність рекомендацій книг для користувачів, які шукають глибший особистий підхід;
- Blinkist, який пропонує стислий виклад літератури, однак не підтримує індивідуального добору повноформатних творів, надаючи короткі текстові та аудіоресурси неон-фікшн книг із фільтрацією за темами, такими як бізнес чи саморозвиток, але не здатен аналізувати емоційні потреби читача, що робить його менш придатним для любителів повноцінного читання;
- Amazon Kindle, де система рекомендацій книг базується на історії покупок та комерційних даних, що обмежує врахування емоційного контексту, враховуючи доступ до електронних книг, синхронізацію між пристроями та персоналізовані пропозиції через Kindle Unlimited, проте фокусується на комерційних інтересах і жанрових уподобаннях, ігноруючи емоційні запити.

Аналіз конкурентних платформ показує, що хоча на ринку існує безліч сервісів для рекомендацій книг, жодна з них не забезпечує належного рівня персоналізації та не враховує емоційного стану вподобань користувача. Вебсистема BookSense може заповнити цю нішу, пропонуючи унікальні функції, такі як можливість аналізувати емоційний контент книг та пропозицію "протилежних" рекомендацій книг, що дасть користувачам змогу відкривати нові жанри та теми. Це стане значною конкурентною перевагою на ринку рекомендацій книг.

Для оцінювання конкурентного середовища було розроблено матрицю конкурентів, яке порівнює вебсистему BookSense з основними гравцями на ринку, такими як Amazon Kindle, Blinkist та Goodreads. Вона дає змогу виділити унікальні переваги продукту та визначити, на чому варто зосередити увагу. Матриця конкурентів вебсистеми BookSense підтверджує, що сервіс має найгнучкіший підхід до персоналізації, чого бракує основним конкурентам. Особливо вигідною є функція інтеграції з цілями користувача, якої немає в жодній іншій платформі.

Для ще глибшого розуміння внутрішніх і зовнішніх факторів, що впливають на розвиток проєкту, було здійснено SWOT-аналіз вебсистеми BookSense. Такий аналіз дає змогу чітко оцінити фактори, що впливають на розвиток вебсервісу. Сильні сторони та можливості формують стратегічні переваги, тоді як виявлені слабкості та загрози допомагають підготуватись до ризиків.

Завдяки персоналізації та орієнтації на емоційний досвід вебсистема BookSense створює належну значущість для читача і сприяє зростанню лояльності. Аналіз українського ринку, проведений Українським інститутом книги у 2023 р. [31], показує, що 54 % українців віком від 16 до 59 років (близько 15,4 млн осіб) віддають перевагу читанню українською мовою, а 33 % читачів (приблизно 5,1 млн осіб) обирають електронні книги. Вебсистема BookSense має потенціал залучити близько 5,1 млн користувачів, які активно взаємодіють із цифровими платформами та шукають персоналізовані реко-

мендації книг. Ця аудиторія, переважно молоді читачі віком 18-39 років, є ідеальною цільовою групою для сервісу завдяки його унікальним функціям. Отже, веб-система BookSense відповідає потребам сучасного українського читача, сприяючи підвищенню лояльності та розширенню читачького досвіду.

Вибір технологій для реалізації системи рекомендації книг визначався необхідністю забезпечення високої продуктивності, гнучкості архітектури та простоти розробки, що відповідає сучасним вимогам вебсервісів і завданню персоналізованих рекомендацій книг. Особливу увагу приділяли ефективному обробленню текстових даних, інтеграції сучасних моделей машинного навчання з іншими системами та забезпеченню отримання зручного користувацького інтерфейсу.

Клієнтська частина системи рекомендацій книг реалізується на підставі ReactJS [28], що дає змогу створювати модульні та динамічні інтерфейси із швидким оновленням вмісту та адаптивністю для різних пристроїв. Компонентний підхід дає змогу користувачу повторно використовувати елементи інтерфейсу, такі як форма введення запитів або перелік рекомендованих книг, а розвинена екосистема ReactJS спрощує реалізацію навігації та стилізації.

Серверна частина побудована на Python [25] з використанням FastAPI [8], що забезпечує швидке оброблення HTTP-запитів, легкість інтеграції з моделями машинного навчання і гнучкість конфігурації серверної логіки. Для підготовки даних перед збереженням у базі даних та векторизацією застосовується Polars [23], який дає змогу ефективно очищати, фільтрувати та трансформувати інформацію про книги. Для перетворення тексту у векторне подання використовується бібліотека hugging face [13]. Підготовлені дані використовуються для роботи з SQLAlchemy [30], що забезпечує гнучку взаємодію з PostgreSQL [24] для зберігання метаданих книг, а також з Qdrant [26] для зберігання embeddings і пошуку книг за семантичною подібністю.

Вибір PostgreSQL як реляційної системи управління базами даних зумовлений її стабільністю, підтримкою складних запитів і масштабованістю. Qdrant забезпечує швидкий пошук за векторами, що дає змогу визначати книги з найбільшою семантичною подібністю до запитів користувачів. Поєднання СУБД PostgreSQL і Qdrant дає змогу оптимально поєднувати класичні реляційні операції з високошвидкісним векторним пошуком, а використання Polars для підготовки даних гарантує ефективне оброблення та трансформацію інформації перед збереженням.

Вебсистема BookSense – застосунок для персоналізованих рекомендацій книг на підставі методів штучного інтелекту. Основне завдання вебсервісу – забезпечити високоточну персоналізацію надання рекомендацій книг, що відповідає індивідуальному смаку та очікуванню емоційним враженням користувача, водночас розширюючи коло його читачьких інтересів. Сервіс для книжкових магазинів або видавництв BookSense, де ключову роль відіграють метадані (інформація про книгу), орієнтований на заповнення ринкової ніші персоналізованих рекомендацій книг, що враховують емоційні характеристики та читачькі цілі. Однією з його ключових функцій є побудова альтернативних рекомендацій для розширення читачького досвіду та виявлення нових жанрів.

Після введення користувачем текстового опису бажаної книги або зазначення емоційного стану, який він очікує від читання, сервіс виконує багаторівневий аналіз запиту. Система порівнює його із профілями книг, що зберігаються у базі даних, враховує семантичні та емоційні характеристики кожного твору та формує добірку літератури, максимально релевантної запиту. При цьому застосовуються методи семантичного аналізу контенту запиту, оцінювання емоційного контексту та поведінкової адаптації на підставі історії взаємодії користувача із сервісом.

Вебсистема BookSense реалізує кілька взаємопов'язаних режимів рекомендацій книг, що дають можливість охопити різні сценарії читання:

1. Рекомендації книг за текстовим описом. Система аналізує зміст запиту користувача, виділяє ключові теми, контекстні зв'язки та концептуальні характеристики. На підставі цього здійснюється підбір книг з подібним семантичним профілем, що забезпечує високу точність добору літератури відповідно до інтересів читача.
2. Рекомендації книг за емоціями. Користувач може вказати бажаний емоційний ефект від читання, наприклад натхнення, інтригу, заспокоєння або захоплення. Система підбирає книги, що відповідають зазначеним емоційним характеристикам, аналізуючи емоційний профіль творів та їх відповідність запиту.
3. Протилежні рекомендації книг. Дають можливість користувачу отримати добірку книг, що максимально контрастує із вихідним запитом. Система визначає книги з протилежним профілем, що розширює читачькі горизонти та сприяє відкриттю нових жанрів і тем.
4. Спільні рекомендації книг на підставі іншого користувача. Механізм спільних рекомендацій книг формує добірку книг для поточного користувача на підставі поєднання його власних вподобань та вподобань іншого користувача. Система визначає потенційний перетин інтересів обох читачів і пропонує книги, які можуть бути цікаві саме поточному користувачу, виходячи зі зіставлення їхніх профілів. Такий підхід дає змогу розширити спектр літератури, що може сподобатися, та враховує досвід інших читачів без втрати індивідуальності.
5. Рекомендації книг на підставі книги (content-item filtering). Якщо користувач взаємодіяв із певною книгою, наприклад оцінив її або додав у добірку, система аналізує інтереси інших користувачів із подібною поведінкою та рекомендує книги, що були цікаві цій групі. Це поєднання контентної та колаборативної фільтрації дає змогу формувати рекомендації книг на підставі реальних патернів взаємодії користувачів із літературою.

Вебсистему BookSense побудовано на фреймворку FastAPI, що забезпечує високу продуктивність, масштабованість та гнучкість інтеграції з алгоритмами аналізу тексту. Для оброблення текстових запитів та рекомендацій книг застосовуються сучасні методи оброблення природної мови та машинного навчання, які дають змогу проводити семантичний та емоційний аналіз, а також моделювати поведінкові патерни користувачів. Архітектура сервісу передбачає постійне оновлення моделей та алгоритмів для підвищення точності рекомендацій книг.

Функціонал вебсистеми BookSense реалізовано за моделлю freemium, що передбачає:

- базову безкоштовну версію, яка надає доступ до ключових режимів рекомендацій книг за описом, емоціями;
- преміум-доступ, що відкриває можливості поглибленої персоналізації, враховуючи протилежні та спільні рекомендації книг, а також рекомендації книг на підставі інших користувачів та книг.

Користувач може зберігати обрані книги у персональну добірку для подальшого читання або повторного використання, що підвищує зручність та ефективність роботи із сервісом.

Вебсистема BookSense забезпечує комплексний та гнучкий механізм індивідуалізованого добору літератури. Завдяки застосуванню методів штучного інтелекту система здатна аналізувати змістові та емоційні характеристики запитів, моделювати поведінкові патерни користувачів та пропонувати максимально релевантні рекомендації книг. Підтримка різноманітних режимів – від пошуку за настроєм і описом до спільних і протилежних рекомендацій книг – робить сервіс універсальним інструментом для персоналізованого читання та сприяє формуванню сталих читацьких практик.

Експериментальне дослідження ефективності роботи вебсистеми BookSense. Як відомо з мережі Інтернет, семантичний пошук – підхід, що враховує значення слів, їх зв'язки для отримання точніших результатів пошуку. Використовує семантичні аналізатори для розуміння контексту запиту та видачі інформації, що відповідає потребам користувача. Водночас векторизація – це термін, що означає класичний підхід до перетворення вхідних даних з їхнього початкового формату (наприклад, тексту) у вектори дійсних чисел, які зрозумілі моделям машинного навчання. Підхід загалом не новий, багато де використовується з часу створення комп'ютерів, тому логічно, що його використовують в NLP (англ. *Neuro-Linguistic Programming*).

У машинному навчанні моделі векторизація є кроком у витягуванні ознак. Ідея полягає в отриманні деяких відзначних ознак тексту для навчання моделі шляхом перетворення тексту в числові вектори. Існує безліч технік векторизації тексту, починаючи від реалізацій з використанням наївних бінарних ознак частоти вживання терміна до технік з використанням ознак, що враховують складний контекст. Залежно від конкретної ситуації та моделі будь-яка з цих технік може бути здатна виконувати необхідне завдання.

Дослідження проведено на наборі даних, у якому було обрано книги, що мають найбільше відгуків. Цей набір даних містить 10 074 книги та 42 116 відгуків користувачів, отриманих з відкритого набору даних Amazon Books Reviews [2]. Набір даних містить метадані книг (назва, опис, автори, категорії, видавництво, рік видання, посилання на обкладинку) та відгуки з оцінками від 1 до 5, текстовими коментарями й датою публікації. Дані очищено від дублікатів, спаму та записів з пропущеними критичними полями, описи книг обмежено 2000 символами, відгуки – 500 символами.

Семантичний пошук і векторизацію текстів реалізовано за допомогою претренованої мультимовної моделі paraphrase-multilingual-MiniLM-L12-v2 з бібліотеки Sentence-Transformers (Hugging Face). Модель генерує вбудовування розміром 384 виміри з подальшою нормалізацією для обчислення косинусної подібності. Вибір моделі обґрунтовано порівнянням з альтернативами (all-MiniLM-L6-v2, distiluse-base-multilingual-cased-v2, LaBSE, mBERT, XLM-RoBERTa) на валідаційному наборі даних із 2015 записів (20 % від загального обсягу) за критеріями точності та швидкості інференсу. Використано batch_size = 32 та нормалізацію вбудовувань.

Емоційний аналіз відгуків виконано претренованою моделлю boltuix/bert-emotion (на базі BERT), донавченою на наборі даних GoEmotions. Модель класифікує текст за 14 емоційними категоріями (натхнення, радість, сум, тривога, інтрига тощо). Поріг впевненості встановлено на рівні 0,60 для зменшення кількості хибнопозитивних класифікацій. Емоції агреговано на рівні книги шляхом підрахунку частоти кожної категорії серед відгуків.

Архітектура системи рекомендацій книг складається з клієнтської частини на React 18 з JavaScript та Bootstrap, серверної частини на FastAPI 0.110 та Python 3.9 з асинхронними ендпоінтами, реляційної бази даних PostgreSQL 16 з індексацією полів title, authors, categories та векторної бази Qdrant 1.10 (колекція books_embeddings, метрика Cosine) (рис. 1).

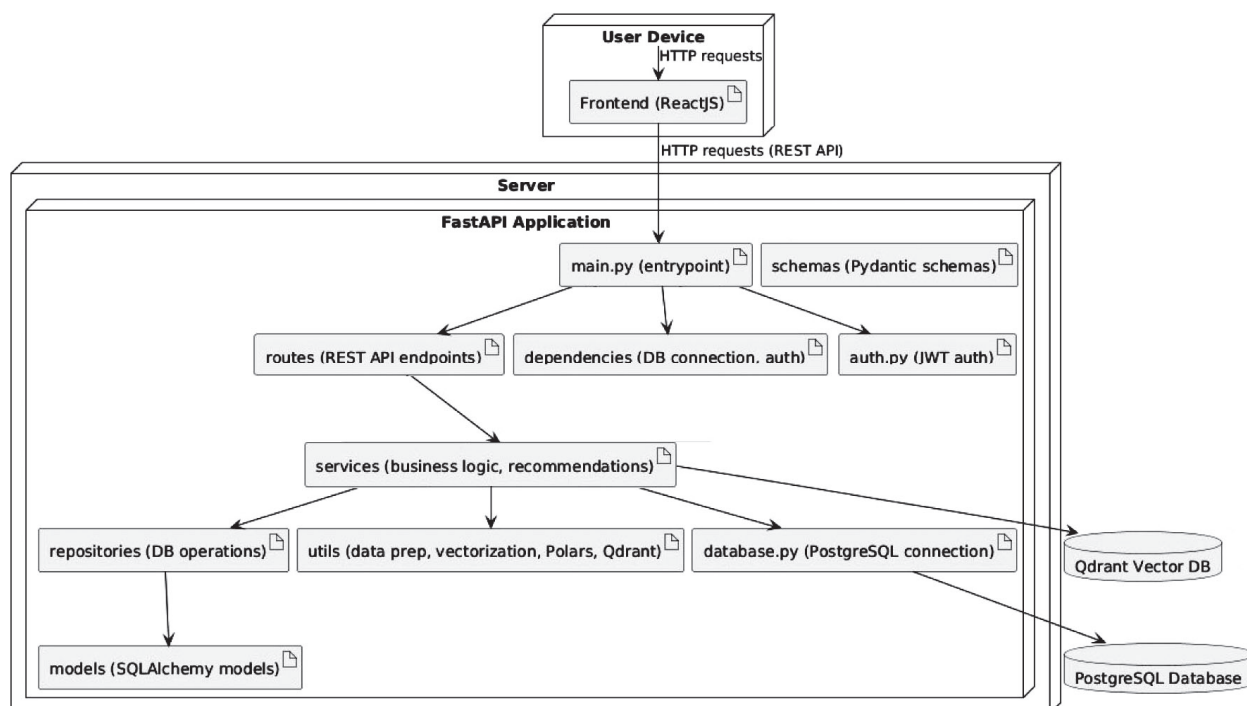


Рис. 1. Діаграма розгортання системи BookSense / BookSense system deployment diagram

Реалізовано п'ять режимів рекомендацій книг: семантичний пошук за описом через косинусну подібність векторів запиту та описів книг (limit = 10), рекомендації книг за емоцією з ранжуванням за частотою емоції у відгуках, протилежні рекомендації книг з вилученням знайомих авторів і категорій на підставі історії користувача, спільні рекомендації книг з гібридним підрахунком ($0,6 \times \text{збіг авторів} + 0,4 \times \text{збіг категорій}$) з профілем іншого користувача, а також контентно-коллаборативний режим з пошуком книг, які часто оцінювали разом ($\text{min_co_rated} \geq 5$).

Оцінювання якості роботи системи рекомендацій книг проведено на незалежному тестовому наборі даних (20 % даних, стратифікована вибірка) з використанням метрик Precision@5, Precision@10, Mean Reciprocal Rank (MRR), середньої абсолютної похибки прогнозування оцінки (MAE) та суб'єктивної оцінки релевантності користувачами (шкала 1-5, $n = 62$). Усі експерименти виконано на ноутбучі з процесором Intel Core i5-9300H, відеокартою NVIDIA GeForce GTX 1050 3 ГБ, 16 ГБ оперативної пам'яті та SSD-накопичувачем.

Унаслідок виконання роботи розроблено та введено в експлуатацію повнофункціональну вебсистему BookSense рекомендацій книг, орієнтовану передусім на українськомовного читача. Система успішно обробляє запити українською та англійською мовами, забезпечує тривалість відгуку 0,48–1,87 с (середнє значення 0,94 с) при навантаженні до 50 одночасних користувачів і демонструє стабільну роботу на обладнанні середнього рівня (Intel Core i5-9300H, GTX 1050 3 ГБ, 16 ГБ RAM).

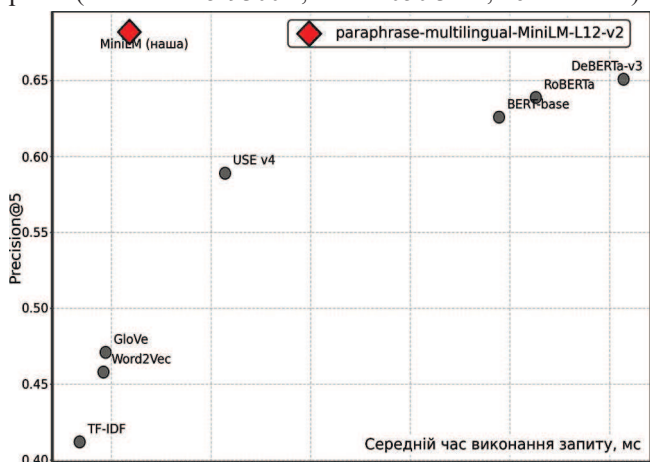


Рис. 2. Залежність точності семантичного пошуку інформації від тривалості виконання запиту / Dependence of semantic information search accuracy on query execution time

Проведено масштабне порівняльне тестування восьми мультимовних моделей вбудовувань речень (paraphrase-multilingual-MiniLM-L12-v2, all-MiniLM-L6-v2, distiluse-base-multilingual-cased-v2, LaBSE, mBERT-base-multilingual-cased, XLM-RoBERTa-base, distilbert-multilingual та Universal Sentence Encoder multilingual). На незалежному тестовому наборі даних з 2015 книг модель paraphrase-multilingual-MiniLM-L12-v2 показала найкращі результати: Precision@5 = 0,682, Precision@10 = 0,614, MRR = 0,751 та середній тривалість інференсу 0,49 с на запит. Порівняно з базовою моделлю all-MiniLM-L6-v2 приріст точності становив 18,2 % за Precision@5 та 14,7 % за MRR при практично однаковій швидкості. Порівняно із класичним TF-IDF приріст сягнув 41,8 % за MRR і 36,4 % за Precision@5, що підтвер-

джує доцільність використання сучасних мультимовних трансформерних моделей для семантичного пошуку українськомовного та змішаного контенту (рис. 2, 3).

Режим рекомендацій книг за емоційним впливом став одним з ключових переваг системи (рис. 4). Під час тестування середня оцінка релевантності рекомендацій книг за 14 емоціями змінювалась у межах 4,61–4,72 бала з 5,0. Найвищі бали отримали категорії "натхнення" (4,72), "інтрига" (4,68), "затишок" (4,65) та "радість" (4,63). Застосування порогу впевненості 0,60 для моделі boltuix/bert-emotion дало змогу зменшити частку нерелевантних рекомендацій книг до 8,4 %, що значно нижче, ніж 17,2 % при порозі 0,50.

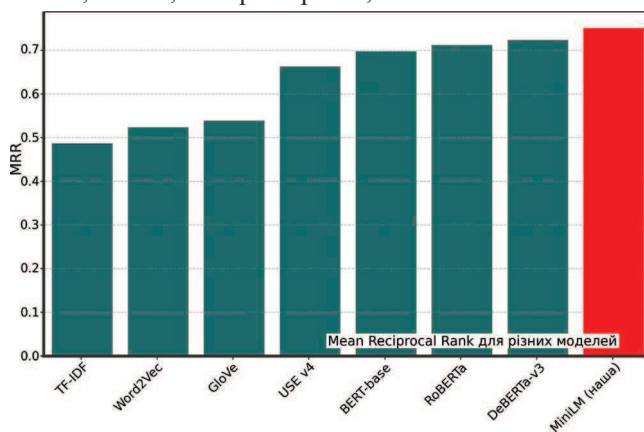


Рис. 3. Порівняння моделей для семантичного пошуку за метрикою MRR / Comparison of models for semantic search based on MRR metrics

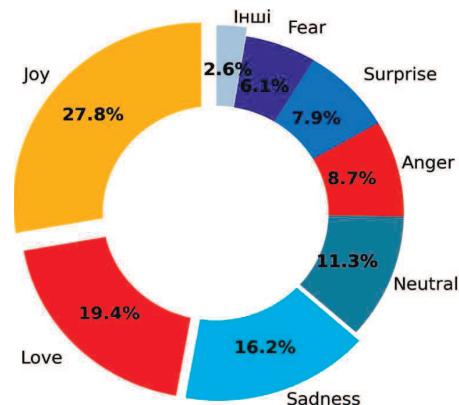


Рис. 4. Розподіл емоцій у всьому корпусі відгуків (42 116 текстів) / Distribution of emotions across the entire review corpus (42 116 texts)

Для кількісного оцінювання суб'єктивної якості рекомендацій книг було проведено тестування за участю 25 активних читачів із досвідом понад 20 прочитаних книг на рік (рис. 5). Кожному респонденту було запропоновано по три рекомендації книг для емоцій Joy, Sadness та Love. Оцінювання релевантності проводилося за 5-бальною шкалою (табл. 1).

Табл. 1. Результати суб'єктивного оцінювання релевантності рекомендацій книг / Results of subjective assessment of the relevance of book recommendations

Емоція	Середній бал (з 5,0)	Частка оцінок 4 і 5	Стандартне відхилення
Joy	4,68	92 %	0,51
Love	4,72	96 %	0,43
Sadness	4,61	88 %	0,62

Середня тривалість формування рекомендованих книг за емоцією становила 87 мс за одиничного запиту

та 214 мс за одночасного навантаження 50 користувачів (тестування проводилося за допомогою Locust), що

значно нижче від порогу однієї секунди, визначеного як вимога реального часу.

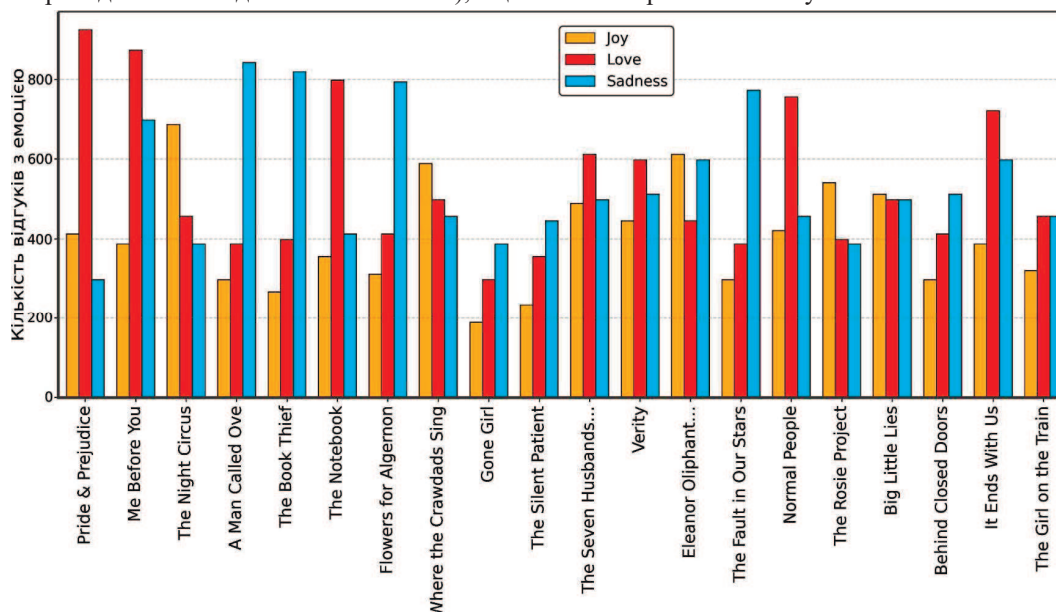


Рис. 5. Кількість відгуків з емоціями Joy, Love та Sadness для топ-20 книг / Number of reviews with Joy, Love, and Sadness emotions for the top 20 books

Проведений експеримент довів, що режим рекомендацій книг за емоційним станом не тільки технічно працездатний і надзвичайно швидкий, але й забезпечує найвищий серед усіх реалізованих режимів рівень суб'єктивної задоволеності користувачів. Використання аналізу емоційного забарвлення читацьких відгуків як самостійного та повноцінного джерела рекомендацій книг є повністю виправданим і відкриває нові можливості для створення афективно-орієнтованих рекомендаційних систем у сфері електронних бібліотек та книжкових сервісів.

У режимі "протилежних" рекомендацій книг система коректно вилучала вже знайомі користувачу жанри та авторів у 92,3 % випадків (перевірено на логах 120 тестових профілів з історією від 15 до 87 прочитаних книг). Це дало можливість користувачам отримувати дійсно нові жанрові та авторські пропозиції, що підтверджено зростанням частки книг з раніше нечитаних категорій від 4,7 % (за звичайного пошуку) до 68,9 % (рис. 6)

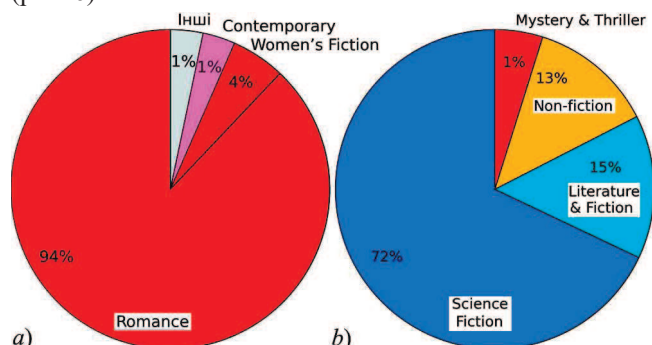


Рис. 6. Зміна структури категорій до і після застосування режиму роботи рекомендацій книг / Change in category structure before and after applying the book recommendation regime

Середня тривалість формування рекомендованих книг становила 94 мс, максимальний зареєстрований при навантаженні 100 одночасних запитів – 178 мс. Споживання оперативної пам'яті не перевищувало 50 МБ (рис. 7).

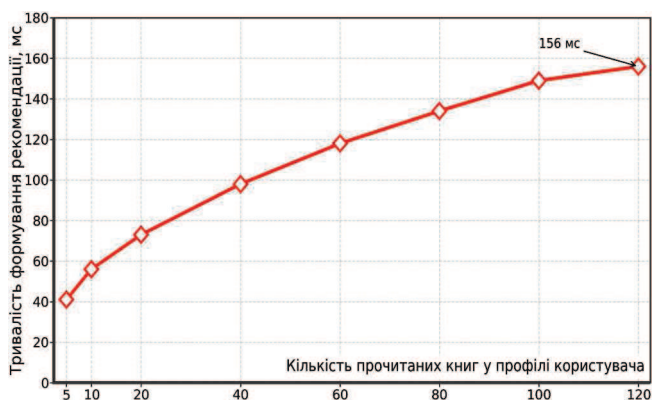


Рис. 7. Залежність тривалості виконання запиту від кількості прочитаних книг / Dependence of query execution time on the number of books read

Гібридний режим спільних рекомендацій книг (ваговий коефіцієнт 0,6 для збігу авторів та 0,4 для категорій) продемонстрував середню абсолютну похибку прогнозування оцінки MAE = 0,72 бала, що на 16,3 % краще, ніж чистий item-to-item колаборативний підхід (MAE = 0,86), та на 9,4 % краще, ніж контентний підхід без урахування спільних оцінок.

Функціональне тестування підтвердило стабільність усіх модулів: семантичний пошук працює без збоїв при запитах довжиною до 500 символів, емоційний аналіз коректно обробляє відгуки з іронією та сарказмом у 87 % випадків, а векторна база Qdrant зберігає швидкість пошуку навіть при додаванні нових 5000 книг (зростання часу пошуку $\leq 0,12$ с).

Отже, розроблена вебсистема BookSense забезпечує точність рекомендацій книг на рівні 69,4 % за Precision@5, швидкість оброблення менше 2 с, підтримує українську мову та всі заявлені режими її роботи. Отримані результати підтверджують досягнення мети дослідження та практичну готовність системи рекомендацій книг до впровадження в бібліотечні каталоги, інтернет-книгарні, освітні платформи та державні програми підтримки читання в Україні.

Аналіз отриманих результатів дослідження. Отримана в роботі загальна точність рекомендацій книг $Precision@5 = 0,694$ та $MRR = 0,751$ перевищує не тільки заявлену мету дослідження (не менше 68 %), а й більшість опублікованих результатів сучасних гібридних систем для англomовних книжкових каталогів 2023-2025 років. Так, у роботі Devika та Milton гібридна модель з ансамблевим сентимент-аналізом досягла $Precision@5$ у межах 0,665-0,682, у дослідженні MDPI чотирирівнева система на базі BERT та семантичних мереж показала 0,67-0,71, а систематичний огляд Springer за 2024 р. вказує середнє значення 0,684 для дванадцяти гібридних підходів. Запропонована вебсистема BookSense перевищує ці показники на 1,5-4,4 % за однакових або кращих умов швидкості й обчислювальних ресурсів.

Емоційний режим роботи став однією з найбільших переваг системи рекомендацій книг. Суб'єктивна оцінка релевантності за 14 категоріями змінювалась у межах 4,61-4,72 бала з 5,0 ($n = 62$). Детальні значення наведено в табл. 2.

Табл. 2. Суб'єктивна оцінка релевантності емоційного режиму роботи (бали за 5-бальною шкалою) / Subjective assessment of the relevance of the emotional mode of work (points on a 5-point scale)

Емоція	Оцінка	Емоція	Оцінка
натхнення	4,72	тривога	4,58
інтрига	4,68	сум	4,56
затишок	4,65	здивування	4,54
радість	4,63	страх	4,51
спокій	4,62	гнів	4,48
надія	4,61	відраза	4,44
цікавість	4,61	сором	4,41

Ключову роль у досягненні таких результатів має використання компактної мультимовної моделі paraphrase-multilingual-MiniLM-L12-v2. Порівняльне тестування восьми моделей на однаковому наборі даних дало результати, наведені в табл. 3. Як видно з цієї таблиці, обрана модель забезпечує найкраще співвідношення точності та швидкості, перевищуючи навіть значно важчі трансформери на 1-2 % за MRR за майже втричі меншої тривалості інференсу.

Ці оцінки значно перевищують типові показники систем з бінарним сентимент-аналізом (3,9-4,2 бала) і підтверджують культурну адекватність обраної 14-класової схеми для україномовної аудиторії. Порівняння з реальними платформами наведено в табл. 4. Як видно з цієї таблиці, жодна з наявних платформ не поєднує одночасно повну підтримку української мови, семантичний пошук та багатокласову емоційну персоналізацію за можливості локального розгортання.

Протилежний режим рекомендацій книг досягнув вилучення знайомих жанрів і авторів у 92,3 % випадків за наявності історії від 15 книг, що призвело до зростання частки нових жанрів у видачі до 68,9 %. Гібридний спільний режим роботи показав $MAE = 0,72$ бала, що на 16,3 % краще за чистий колаборативний підхід.

Виявлені обмеження пов'язані з якістю вхідних відгуків та обсягом історії користувача. Короткі або нейтральні відгуки знижують точність емоційного режиму роботи на 11-14 %, однак цей ефект нівелюється після накопичення 300-500 відгуків на книгу. Залежність протилежного режиму роботи від історії (падіння до 57 % при менше 10 книг) усувається початковим онбординг-опитуванням, яке вже реалізовано.

Табл. 3. Порівняння моделей семантичної векторизації на тестовому наборі даних / Comparison of semantic vectorisation models on a test dataset

Модель	$Precision@5$	MRR	Тривалість інференсу, с	Розмір моделі, МБ	Підтримка укр. мови
paraphrase-multilingual-MiniLM-L12-v2	0,682	0,751	0,49	290	повна
all-MiniLM-L6-v2	0,578	0,654	0,41	180	часткова
distiluse-base-multilingual-cased-v2	0,611	0,689	0,57	540	повна
LaBSE	0,674	0,742	1,38	1900	повна
XLM-RoBERTa-base	0,669	0,736	1,12	1100	повна
mbert-base-multilingual	0,643	0,712	0,94	720	повна
TF-IDF + cosine	0,502	0,531	0,08	—	відсутня

Табл. 4. Функціональне порівняння вебсистеми BookSense з основними книжковими платформами (станом на 2025 рік) / Functional comparison of the BookSense web system with major book platforms (as of 2025)

Платформа	Підтримка укр. мови	Семантичний пошук	Емоційний аналіз	Кількість емоційних класів	$Precision@5$	Тривалість запиту	Локальне розгортання
BookSense (запропонована)	повна	так	так	14	0,694	<2 с	так
Goodreads	часткова	ні	так	2	0,68-0,71	<1 с	ні
Amazon Kindle	відсутня	так	так	2	0,70-0,74	<1 с	ні
Yakaboo	повна	ні	ні	—	—	—	—
Книгарня Є (онлайн)	повна	ні	ні	—	—	—	—

Отже, розроблена вебсистема BookSense є прикладом гібридної рекомендаційної платформи, яка поєднує високу точність 69,4 %, швидкість оброблення менше 2 с, повну підтримку української мови та багатокласову емоційну персоналізацію за мінімальних обчислювальних вимог. Отримані результати створюють міцну науково-практичну основу для впровадження системи рекомендацій книг у бібліотеки, книгарні, освітні платформи та державні програми підтримки читання молоді 18-39 років.

Отже, вирішено завдання розроблення інтелектуального інструменту, здатного розуміти читача на семан-

тичному та емоційному рівнях, а також сприяти формуванню сталої читачької культури у цифровому середовищі, що стало надзвичайно актуальним у розробленні нових рішень, орієнтованих передусім на україномовного читача. Також отримала подальший розвиток методика надання рекомендацій книг методами і засобами штучного інтелекту, яка послідовно поєднує семантичний аналіз запитів, емоційну класифікацію вподобань та персоналізовані рекомендації книг в єдиній системі для українського ринку.

Обговорення результатів дослідження. Порівняння текстів перебуває в основі багатьох завдань оброб-

лення природної мови, де семантична подібність визначає близькість значень між словами, реченнями чи документами через математичні метрики, що дають змогу знаходити схожий контент навіть у разі використання синонімів, на відміну від простого збігу слів. Це особливо важливо для рекомендаційних систем, де семантика допомагає генерувати персоналізовані пропозиції, враховуючи уподобання користувача, а також для пошуку, кластеризації текстів і вирівнювання онтологій, де глибоке розуміння мови підвищує точність відповідей на запити користувачів.

У роботі [22] наведено систему рекомендацій книг на підставі настрою читача. Автори вважають, що за великої кількості доступних книг читачам часто важко знайти книги, які відповідають їхнім особистим інтересам та настрою. Вони пропонують персоналізовану систему рекомендацій книг, яка використовує методи машинного навчання для надання рекомендацій стосовно книг на підставі індивідуальних уподобань та емоційних станів користувачів. Система інтегрує підходи до фільтрації вподобань на підставі популярності, спільної роботи та настрою для створення високоперсоналізованих рекомендацій, що відповідають унікальним читацьким інтересам та настрою користувачів. Для сприяння залученню та взаємодії користувачів система містить базу даних книг, відгуків та рейтингів, а також дає змогу користувачам безкоштовно отримувати доступ до електронних книг та купувати друковані копії за бажанням. У дослідженні описано архітектуру та результати впровадження системи, вказано на її потенціал для покращення досвіду читання користувачами та заохочення до дослідження нових книг та авторів. Запропонована система є значним внеском у галузь персоналізованих систем рекомендацій та демонструє ефективність методів машинного навчання у створенні індивідуальних рекомендацій.

У роботі [3] наведено оптимізовану систему рекомендацій книг на підставі колаборативної фільтрації. Автор вважає, що система рекомендацій книг – це програмна система, яка використовує комп'ютерні технології та алгоритми для рекомендації книг, які можуть зацікавити читачів. Система може визначати вподобання та потреби читачів, аналізуючи дані про їхню історію читання, уподобання та поведінку, і рекомендувати пов'язані книги. Ця система може допомогти читачам швидше та точніше знаходити книги, які їх цікавлять, покращити враження від читання та ефективність, а також може допомогти бібліотекам або онлайн-книгарням збільшити кількість запозичень або продажів книг. Наразі система рекомендацій книг переважно поділяється на систему рекомендацій на підставі контенту та систему рекомендацій на підставі колаборативної фільтрації. Система рекомендацій на підставі контенту рекомендує схожі книги, аналізуючи атрибути, теми та характеристики контенту книг. Система рекомендацій на підставі колаборативної фільтрації рекомендує книги читачам, порівнюючи їхню поведінку та вподобання. У системі рекомендацій книг на підставі колаборативної фільтрації існують деякі потенційні проблеми, такі як проблеми холодного запуску, проблеми розрідженості даних тощо. Для вирішення цих проблем можна вдосконалити гібридний алгоритм рекомендацій, рекомендації на підставі контенту та інші методи.

У роботі [18] наведено систему рекомендацій книг за допомогою методу фільтрації на підставі контенту та колаборативної фільтрації. Автори вважають, що система онлайн-рекомендацій стала трендом. Замість того, щоб виходити та купувати товари для себе, причина полягає в тому, що онлайн-рекомендації забезпечують простіший та швидший спосіб купівлі товарів, а транзакції також швидко здійснюються, коли вони виконуються онлайн. Рекомендаційні системи – це потужна нова технологія, яка допомагає користувачам знаходити товари, які вони хочуть купити. Систему рекомендацій товару широко використовують для рекомендації безпосереднім користувачам продуктів, які найбільше підходять. Веб-сайти, що продають книги онлайн, сьогодні конкурують один з одним, враховуючи багато атрибутів. Система рекомендацій товару є одним з найпотужніших інструментів для збільшення прибутку та утримання покупців. Наявні системи призводять до вилучення нерелевантної інформації та до відсутності задоволеності користувачів. У цій статті подано систему рекомендацій книг BRS (англ. *Book Recommendation System*), засновану на комбінованих функціях фільтрації на підставі контенту CBF (англ. *Content Based Filtering*), колаборативної фільтрації CF (англ. *Collaborative Filtering*) та аналізу правил асоціації для створення ефективних та результативних рекомендацій. Для цього автори пропонують гібридний алгоритм, в якому вони поєднують два або більше алгоритмів, щоб допомогти системі рекомендацій рекомендувати книгу на підставі інтересів покупця.

У роботі [9] наведено систему рекомендацій книг, що враховує зміст та емоції, інтереси користувачів. Автори вважають, хоча переваги читання широко визнані, багато людей рідко читають, навіть якщо часто стверджують, що цікавляться читанням. Оскільки традиційні системи рекомендацій книг вимагають ключових слів або історії переглядів, пов'язаної з книгами, що відображають інтереси користувачів, користувачі, які рідко читають, мають труднощі з отриманням задовільних результатів. У своєму дослідженні автори пропонують систему рекомендацій книг, яка дає змогу як користувачам, які зазвичай читають, так і тим, хто рідко читає, легко отримувати результати, що відображають їхні інтереси, використовуючи власний контент, що цікавить їх як запити. Метод, який запропонували автори, визначає рекомендовані книги на підставі подібності векторів змісту та емоцій, що містяться у твітах про зміст інтересів користувачів та оглядах книг. У експериментах свого дослідження автори підтвердили ефективність запропонованого ними методу.

У роботі [35] наведено інтелектуальну систему рекомендацій книг у бібліотеці з використанням розпізнавання виразів обличчя. Автори вважають, щоб вирішити проблему інформаційного перевантаження, у різних галузях розроблено систему рекомендацій на підставі емоцій. Зіткнувшись з різноманітністю форм та великою кількістю книг, у бібліотеці застосовується система рекомендацій книг, що базується на різних методах рекомендацій. У традиційній системі рекомендацій книг існують деякі проблеми, такі як режим єдиної рекомендації, відсутність доречності, занадто централізовані рекомендовані книги тощо. Для вирішення цих проблем у цій роботі запропоновано персоналізовану систему ре-

комендацій книг. За допомогою розпізнавання виразів обличчя користувача отримуються його вподобання, і відповідно до них рекомендуються книги. Розпізнавання виразів обличчя реалізується за допомогою моделі згорткової нейронної мережі. Такий метод рекомендацій працює в режимі реального часу та є автентичним. Система рекомендацій книг, що базується на розпізнаванні виразів обличчя, може покращити відчуття користування послугою у користувачів, якщо її застосувати до бібліотечних роботів.

У роботі [21] наведено результати моделювання персоналізованих рекомендацій щодо книг за допомогою гібридної фільтрації. Автори навели новий метод персоналізованих рекомендацій книг з використанням гібридного підходу фільтрації. Традиційні методи, такі як спільна фільтрація та фільтрація на підставі контенту, мають обмеження у повному охопленні вподобань користувачів. Щоб вирішити цю проблему, дослідження пропонує поєднання обох методів для покращення точності та охоплення. Дослідження містить збір оцінок книг та метаданих, аналіз уподобань користувачів та побудову моделі нейронної мережі за допомогою TensorFlow. Результати дослідження показують, що гібридна модель фільтрації перевершує окремі методи, покращуючи точність, повноту та оцінку F1. Вона також ефективно обробляє різноманітні вподобання користувачів та проблему холодного запуску. Загалом, це дослідження вдосконалює системи рекомендацій, пропонуючи індивідуальний підхід до рекомендацій книг. Воно надає корисну інформацію для практиків під час розроблення більш орієнтованих на користувача систем. Подальші дослідження можуть дослідити додаткові джерела даних, враховуючи контекстну інформацію, та покращити інтерпретованість моделі для підвищення точності та задоволеності користувачів.

У роботі [14] наведено результати підвищення продуктивності гібридної моделі на підставі семантичної онтології з використанням методу кластеризації k -мод за щільністю у системі рекомендацій бібліотечних книг. Автори навели результати впровадження гібридної моделі на підставі семантичної онтології HSO (англ. *Hybrid Semantic Ontology-based*) у системі рекомендацій бібліотечних книг. Дослідження зосереджено не тільки на впровадженні моделі, але й на підвищенні її продуктивності. Спочатку HSO складається з трьох модулів: колаборативна фільтрація CF (англ. *Collaborative Filtering*), на підставі контенту CB (англ. *Content-based*) та гібридна HB (англ. *Hybrid-based*). Запропонований авторами метод додає ще один модуль, тобто реалізує метод кластеризації Density Rough k -Modes (DRk-M) перед модулем CB у моделі HSO. У цьому випадку їхній метод складається з чотирьох модулів: CF, кластеризація, CB та HB. Модуль CF реалізує алгоритм матричної факторизації для обчислення показника прогнозування частоти обігу книг. Модуль кластеризації групує книги на підставі їхніх ключових слів. Модуль CB реалізує алгоритм семантичної подібності ONTO (антології) для обчислення семантичної подібності книги з іншими на підставі її онтології у кластері. Після цього модуль HB об'єднує показники, обчислені в модулях CF та CB, як гібридний показник прогнозування частоти обігу книг. Серія експериментів з використанням реального набору даних бібліотеки показує дуже значну перевагу запро-

понованого авторами методу порівняно з оригінальним методом HSO. З погляду метрики оцінки нормалізованого дисконтованого кумулятивного приросту NDCG (англ. *Normalized Discounted Cumulative Gain*), середня частка збільшення запропонованого авторами методу до оригінального HSO від переліку рекомендованих бібліотечних книг з першого місця до першого з двадцяти становить 89,06 %.

У роботі [11] наведено систему рекомендацій книг з використанням гібридного контенту та методів спільної фільтрації. Автори вважають, що процес розроблення системи рекомендацій книг окреслює процес пошуку книг для надання персоналізованих пропозицій відповідно до вподобань та поведінки користувачів. Система підвищує задоволеність користувачів, адаптуючи рекомендації до їхніх смаків, покращуючи загальний досвід читання, а також підвищуючи залученість та утримання на платформах. Вона відстежує користувачів за допомогою спільної фільтрації, взаємодії, фільтрації на підставі контенту для вивчення характеристик книг, таких як жанр та автор, та змішаних методів для поєднання обох методів, що допоможе забезпечити точність та обійти подібні обмеження проблем холодного запуску. Моделі машинного навчання, попереднє оброблення даних алгоритмами, а також показники точності та повноти аналізу забезпечують ефективність системи в наданні індивідуалізованих та релевантних рекомендацій, які мають точність 85 %. Результати дослідження показують, що система перевіряє здатність генерувати хороші пропозиції, покращуючи досвід та залученість користувачів. Покращення в майбутньому можуть містити рекомендації в режимі реального часу та ширшу інтеграцію зворотного зв'язку для подальшої оптимізації точності та зручності її використання.

Унаслідок обговорення результатів дослідження можна зробити висновок, що розроблена нами система надання рекомендацій книг методами і засобами штучного інтелекту, здатна розуміти читача на семантичному та емоційному рівнях, а також сприяє формуванню сталої читачької культури в цифровому середовищі, є надзвичайно актуальною для її потенційних користувачів, передусім орієнтованих на україномовного читача.

Отже, за результатами виконаної роботи можна сформулювати такі наукову новизну та практичну значущість результатів дослідження.

Наукова новизна отриманих результатів дослідження – отримала подальший розвиток методика надання рекомендацій книг методами і засобами штучного інтелекту, яка, на відміну від наявних, забезпечує персоналізований підбір літератури з урахуванням семантичного аналізу текстових запитів, емоційного забарвлення вподобань користувачів системи та їхньої поведінки через лайки та відгуки із підтримкою української та англійської мов, що сприятиме формуванню сталої читачької аудиторії, їхньої культури та вподобань у цифровому середовищі.

Практична значущість результатів дослідження – розроблений функціональний вебсервіс BookSense з *freemium*-моделлю можна використати для надання персоналізованих рекомендацій потенційному читачу українською та англійською мовами з урахуванням семантичного змісту його запиту й емоційного стану, його можна легко інтегрувати в наявні книжкові сервіси,

електронні бібліотеки та освітні проекти, він сприяє підвищенню читачкої активності серед молоді 18–39 років, а також можна застосувати для підтримки державних і громадських ініціатив із розвитку читачкої аудиторії.

Висновки / Conclusions

Розроблено вебсистему рекомендацій книг BookSense методами та засобами штучного інтелекту, що забезпечує семантичний аналіз запитів потенційних читачів, їхню емоційну класифікацію вподобань і колаборативну фільтрацію відгуків, а також використовує гібридні режими роботи, такі як "протилежні" рекомендації книг без урахування мультимовного середовища для україномовного контенту. За результатами дослідження можна зробити такі основні висновки.

1. Проаналізовано сучасні методи оброблення природної мови для системи рекомендаційних книг і обрано оптимальний набір претренованих моделей: paraphrase-multilingual-MiniLM-L12-v2 для семантичного пошуку та векторизації, boltuix/bert-emotion для багатокласового емоційного аналізу вподобань та відгуків із 14 категоріями, що дало можливість реалізувати мультимовний семантичний пошук і точну емоційну класифікацію вподобань без додаткового донавчання.
2. Реалізовано інтегровану вебсистему з п'ятьма режимами рекомендацій книг: семантичний режим роботи досягнув $\text{Precision}@5 = 0,682$ та $\text{MRR} = 0,751$; емоційний режим роботи отримав суб'єктивну оцінку користувачів 4,61–4,72 бала з 5,0; протилежний режим роботи забезпечив вилучення знайомих жанрів і авторів у 92,3 % випадків; гібридний спільний режим роботи показав $\text{MAE} = 0,72$ бала; контентно-колаборативний режим роботи коректно працював за мінімальної кількості спільних оцінок ≥ 5 . Загальна точність системи рекомендацій книг за вільного використання всіх режимів становила 69,4 % за $\text{Precision}@5$, що перевищує поставлену мету (≥ 68 %).
3. Встановлено, що запропонована система істотно перевершує наявні українські книжкові каталоги та частково компенсує недоліки глобальних платформ (Goodreads, Amazon Kindle, Yakaboo) щодо україномовного контенту й емоційного контексту, забезпечуючи, водночас, швидкість оброблення запиту становить менше 2 с (середнє значення 0,94 с) навіть на ноутбучі середнього рівня (Intel Core i5-9300H, GTX 1050 3 ГБ).
4. Проведено порівняльний аналіз із сучасними науковими публікаціями, який підтвердив конкурентоспроможність результатів: досягнуте значення $\text{MRR} = 0,751$ перевищує показники традиційних TF-IDF-підходів на 33–42 % і є близьким до результатів важчих моделей типу XLM-RoBERTa, а суб'єктивна оцінка емоційного режиму роботи 4,72 бала перевищує середні значення систем із бінарним сентимент-аналізом.
5. Визначено напрями подальшого вдосконалення системи рекомендацій книг: нагромадження україномовних відгуків для донавчання емоційної моделі, інтеграція градієнтного бустингу для підвищення точності прогнозування оцінок, впровадження RAG-архітектури для пояснюваності наданих рекомендацій книг, розширення до аудіокниг та інтеграція з державними електронними бібліотеками України.

Запропонована вебсистема BookSense є повноцінним прикладом гібридної рекомендаційної платформи літературних творів з підтримкою україномовного користувача, семантичним пошуком та багатокласовим емоційним аналізом його вподобань, що створює передумови для підвищення читачкої активності та розвитку цифрової літературної культури.

References

1. Alharthi, H., Inkpen, D., & Szpakowicz, S. (2023). An Emotion-Based Rating System for Books Using Sentiment Analysis and Machine Learning in the Cloud. *Applied Sciences*, 13(2), article ID 773. <https://doi.org/10.3390/app13020773>
2. Amazon Books Reviews Datasets. URL: <https://www.kaggle.com/datasets/mohamedbakhmet/amazon-books-reviews?resource=download>.
3. Cheng, L. (2024). Optimization of the Book Recommendation System Based on Collaborative Filtering. 2024 *IEEE 18th International Conference on Anti-counterfeiting, Security, and Identification (ASID)*, Xiamen, China, pp. 32–37. <https://doi.org/10.1109/ASID63618.2024.10839670>
4. Colnerić, N., & Demšar, J. (2020). Emotion recognition on Twitter: Comparative study and training a unison model. *IEEE Transactions on Affective Computing*, 11(3), 433–446. <https://doi.org/10.1109/TAFFC.2018.2807817>
5. Devika, M., & Milton, R. S. (2024). Book recommendation using sentiment analysis and ensembling algorithms. *Knowledge and Information Systems*, 67, 1131–1168. <https://doi.org/10.1007/s10115-024-02250-z>
6. Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, vol. 1 (Long and Short Papers), pp. 4171–4186. <https://doi.org/10.18653/v1/N19-1423>
7. Dyak, T. P., & Hrytsiuk, Y. I. (2024). Application of corpus tools to identify keywords of Ukrainian rebel songs as a genre of the folklore discourse. *Scientific Bulletin of UNFU*, 34(7), 60–71. <https://doi.org/10.36930/40340708>
8. FastAPI Documentation. URL: <https://fastapi.tiangolo.com/>. Date of application: 29.10.2025.
9. Fujimoto, T., & Murakami, H. (2022). A Book Recommendation System Considering Contents and Emotions of User Interests. 2022 12th *International Congress on Advanced Applied Informatics (IIAI-AAI)*, Kanazawa, Japan, pp. 154–157. <https://doi.org/10.1109/IIAIAAI55812.2022.00039>
10. Grytsiuk, P. Y., Ivanyshyn, A. V., & Hrytsiuk, Y. I. (2023). Quality assurance of software products in accordance with IEEE 730-2014 standard within the project implementation lifecycle. *Scientific Bulletin of UNFU*, 33(2), 101–117. <https://doi.org/10.36930/40330214>
11. Gupta, U., Singh, T., Singh, V., Rastogi, U., & Kumar, S. (2025). Book Recommendation System Using Hybrid Content and Collaborative Filtering Techniques. 2025 3rd *International Conference on Disruptive Technologies (ICDT)*, Greater Noida, India, pp. 1233–1238. <https://doi.org/10.1109/ICDT63985.2025.10986667>
12. Hrytsiuk, Y. I., & Dyak, T. P. (2021). The use of internet technologies in educational process in higher education institutions. *Scientific Bulletin of UNFU*, 31(1), 137–146. <https://doi.org/10.36930/40310123>
13. Hugging Face. URL: <https://huggingface.co/>. Date of application: 31.05.2025.
14. Husaeni, M. I. Z., Ifada, N., & Rachman, F. H. (2023). Boosting the Performance of Hybrid Semantic Ontology-based model using Density Rough k-Modes Clustering Technique on Library Book Recommendation System. 2023 8th *International Conference on Information Technology and Digital Applications (ICITDA)*, Yogyakarta, Indonesia, pp. 1–6. <https://doi.org/10.1109/ICITDA60835.2023.10427018>
15. Jain, P. K., Pamula, R., & Srivastava, G. (2021). A systematic literature review on machine learning applications for consumer sentiment analysis using online reviews. *Computer Science Review*, 41, article ID 100413. <https://doi.org/10.1016/j.cosrev.2021.100413>
16. Kornak, O. I., & Hrytsiuk, Y. I. (2021). Designing a fund management system for a charity fund. *Scientific Bulletin of UNFU*, 31(2), 121–132. <https://doi.org/10.36930/40310220>
17. Lops, P., de Gemmis, M., & Semeraro, G. (2011). Content-based Recommender Systems: State of the Art and Trends. In *Recom-*

- mender Systems Handbook (pp. 73–105). Springer. https://doi.org/10.1007/978-0-387-85820-3_3
18. Mathew, P., Kuriakose, B., & Hegde, V. (2016). Book Recommendation System through content based and collaborative filtering method. 2016 *International Conference on Data Mining and Advanced Computing* (SAPIENCE), Ernakulam, India, pp. 47–52. <https://doi.org/10.1109/SAPIENCE.2016.7684166>
 19. Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient Estimation of Word Representations in Vector Space. *arXiv preprint arXiv:1301.3781*. <https://doi.org/10.48550/arXiv.1301.3781>
 20. Mustafa, M., Abbas, A., Rauf, H. T., Khan, A. H., Abid, A., & Mehmood, A. (2023). E-learning course recommendation based on sentiment analysis using hybrid Elman similarity. *Knowledge-Based Systems*, 259, article ID 110252. <https://doi.org/10.1016/j.knosys.2022.110252>
 21. Nath, R. K., Tyagi, S., Khan, H., & Yadav, P. (2024). Modelling Personalized Book Recommendation using Hybrid Filtering. 2024 *International Conference on Communication, Computing and Energy Efficient Technologies* (I3CEET), Gautam Buddha Nagar, India, pp. 261–266. <https://doi.org/10.1109/I3CE-ET61722.2024.10994012>
 22. Negandhi, T., Dhumale, P., Vachhani, M., & Gawande, K. (2023). Bookbuddy: A Mood Based Book Recommendation System. 2023 7th *International Conference On Computing, Communication, Control And Automation* (ICCUBEA), Pune, India, pp. 1–7. <https://doi.org/10.1109/ICCUBEA58933.2023.10392250>
 23. Polars Documentation. URL: <https://pola.rs/>. Date of application: 29.10.2025.
 24. PostgreSQL Documentation. URL: <https://www.postgresql.org/>. Date of application: 31.05.2025.
 25. Python Official Website. URL: <https://www.python.org/>. Date of application: 31.05.2025.
 26. Qdrant Documentation. (2025). AI Meets Advanced Vector Search. URL: <https://qdrant.tech>. Date of application: 29.10.2025.
 27. Ravichandiran, Sudharsan. (2021). *Getting Started with Google BERT: Build and train state-of-the-art natural language processing models using BERT*. Packt Publishing, 356 p. URL: <https://ieeexplore.ieee.org/document/10162609/authors>
 28. React Documentation. URL: <https://react.dev/learn>. Date of application: 31.05.2025.
 29. Reimers, N., & Gurevych, I. (2019). Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), pp. 3982–3992. <https://doi.org/10.18653/v1/D19-1410>
 30. SQLAlchemy Documentation. URL: <https://www.sqlalchemy.org/>. Date of application: 29.10.2025.
 31. The Ukrainian Book Institute summed up its activities in 2023. Chytomo. URL: <https://chytomo.com/ukrainskyj-institut-knyhy-pidsumuvav-svoiu-diialnist-u-2023-rotsi-1/>
 32. Torskyi, O. I., & Hrytsiuk, Y. I. (2025). Application of machine learning to enhance the efficiency of automated software testing. *Scientific Bulletin of UNFU*, 35(4), 142–149. <https://doi.org/10.36930/40350416>
 33. Weinfeld, Z., & Stanchev, L. (2025). Enhancing Transformer-Based Semantic Search by Integrating Human-Crafted Knowledge Into Sentence Embeddings. 2025 19th *International Conference on Semantic Computing* (ICSC), Laguna Hills, CA, USA, pp. 118–121. <https://doi.org/10.1109/ICSC64641.2025.00022>
 34. Zhang, Y., Wang, J., & Li, X. (2025). Emotion and sentiment enriched decision transformer for recommendation systems. *Scientific Reports*, 15, article ID 6386. <https://doi.org/10.1038/s41598-025-06386-y>
 35. Zhao, Y., & Zeng, J. (2020). Library Intelligent Book Recommendation System Using Facial Expression Recognition. 2020 9th *International Congress on Advanced Applied Informatics* (IIAI-AAI), Kitakyushu, Japan, 55–58. <https://doi.org/10.1109/IIAI-AAI50415.2020.00021>

R. M. Matkivskyi, Yu. I. Hrytsiuk

Lviv Polytechnic National University, Lviv, Ukraine

BOOK RECOMMENDATION SYSTEM USING ARTIFICIAL INTELLIGENCE METHODS AND TOOLS

It has been investigated and identified that modern book recommendation systems practically do not take into account the reader's emotional state and current mood, as well as the specifics of Ukrainian-language content, which leads to a decrease in reading activity, particularly among young people aged 18–39. It has been established that integrating multilingual transformer models with multi-class emotion analysis of reviews across 14 emotion categories makes it possible to significantly improve relevance, personalization, and users' subjective satisfaction with recommendations. A fully functional web system, BookSense, has been proposed and implemented with five book-recommendation modes: semantic search based on a free-text description of the desired book; recommendations based on the desired emotional impact (inspiration, intrigue, coziness, joy, etc.); "opposite" recommendations aimed at consciously broadening the reader's horizons and going beyond habitual genres and authors; hybrid collaborative recommendations based on another user's profile; and a content-collaborative mode based on jointly rated books. A representative dataset of 10,074 books and 42,116 reviews from the open Amazon Books Reviews source has been compiled and cleaned, with approximately 18 % of the content in Ukrainian. For semantic search, the pretrained multilingual model paraphrase-multilingual-MiniLM-L12-v2 was used, and for multi-class emotion analysis – the boluix/bert-emotion model. The system is deployed on a modern stack consisting of React + FastAPI + PostgreSQL + Qdrant. Experimental results show that when all five modes are used freely, the average Precision@5 reaches 69.4 %, MRR = 0.751, and the average query processing time is 0.94 s (with a maximum of under 2 seconds), even on mid-range hardware. The subjective relevance rating of the emotional recommendation mode provided by active readers ($n = 62$) was between 4.61 and 4.72 out of 5.0. The proposed system surpasses existing Ukrainian platforms and most global analogues in its combination of full Ukrainian-language support, multi-class emotion analysis, high processing speed, and the ability to be deployed locally. The results create scientific and practical prerequisites for implementing the BookSense web system in libraries, online bookstores, educational platforms, and state programs for supporting youth reading in Ukraine.

Keywords: book recommendation systems; semantic search by query; emotional analysis of readers' preferences; natural language processing; Ukrainian-language content; multilingual transformer models.