

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Центр _____ Післядипломної освіти
(повна назва)

Кафедра _____ Штучного інтелекту
(повна назва)

КВАЛІФІКАЦІЙНА РОБОТА Пояснювальна записка

рівень вищої освіти _____ перший (бакалаврський)

_____ AI-орієнтована система «CineMind» для персоналізованих
_____ рекомендацій фільмів
(тема)

Виконав:
здобувач _____ другого _____ року навчання,
групи _____ ІТШП-23-1

_____ Сергій Салтан
(власне ім'я, прізвище)

Спеціальність 122 Комп'ютерні науки
_____ (код і повна назва спеціальності)

Тип програми _____ освітньо-професійна
Освітня програма _____ Штучний інтелект
_____ (повна назва освітньої програми)

Керівник _____ ас. Сергій Гриньов
(посада, власне ім'я, прізвище)

Допускається до захисту

Завідувач кафедри ШІ _____
(підпис)

_____ Олег ЗОЛОТУХІН
(власне ім'я, прізвище)

2025 р.

Харківський національний університет радіоелектроніки

Центр _____ Післядипломної освіти _____

Кафедра _____ Штучного інтелекту _____

Рівень вищої освіти _____ перший (бакалаврський) _____

Спеціальність _____ 122 Комп'ютерні науки _____
(код і повна назва)

Тип програми _____ освітньо-професійна _____

Освітня програма _____ Штучний інтелект _____
(повна назва)

ЗАТВЕРДЖУЮ:

Зав. кафедри _____

(підпис)

« _____ » _____ 20 ____ р.

ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ

здобувачеві _____ Салтану Сергію Олександровичу _____
(прізвище, ім'я, по батькові)

1. Тема роботи _____ AI-орієнтована система «CineMind» для персоналізованих рекомендацій фільмів _____

затверджена наказом університету від 19 травня 2025 р. № 387Ст

2. Термін подання студентом роботи до екзаменаційної комісії 24 червня 2025 р.

3. Вихідні дані до роботи Література з поясненнями використання рекомендаційних систем та машинного навчання, Науково-технічні публікації, матеріали присвячені навчанню моделей для узагальнення і скорочення текстових даних, документація з мови програмування Python та використання необхідних бібліотек _____

4. Перелік питань, що потрібно опрацювати в роботі _____

1) Аналіз предметної галузі та постановка задачі _____

2) Проектування системи рекомендацій _____

3) Реалізація вебдодатку _____

4) Аналіз ефективності системи _____

РЕФЕРАТ

Пояснювальна записка: 61 с., 6 рис., 6 табл., 1 додаток, 17 джерел.

АДАПТАЦІЯ ДО УПОПОДОБАНЬ, АНАЛІЗ ТОНАЛЬНОСТІ, ІНТЕЛЕКТУАЛЬНА РЕКОМЕНДАЦІЙНА СИСТЕМА, ОБРОБКА ПРИРОДНОЇ МОВИ, САМАРІЗАЦІЯ ВІДГУКІВ, BART, BERT.

Об'єкт дослідження – інтелектуальні системи рекомендацій кіноконенту, що використовують методи обробки природної мови (NLP) та гібридні підходи до фільтрації даних.

Предмет дослідження – методи та алгоритми генерації.

Мета роботи – розробка вебдодатку, який забезпечує аналіз відгуків користувачів та професійних критиків на фільми, формує персоналізовані рекомендації з урахуванням віку, статі, індивідуальних уподобань користувача, а також автоматично узагальнює ключові теми та тональність рецензій, оптимізуючи процес пошуку релевантного кіноконенту.

Методи дослідження – методи аналізу вимог, архітектурного проектування, інтеграції API, обробки природної мови (NLP), побудови гібридних моделей рекомендацій та тестування функціональних компонентів. У роботі застосовано методи аналізу вимог, архітектурного проектування, інтеграції API, обробки природної мови (NLP), побудови гібридних моделей рекомендацій та тестування функціональних компонентів.

У роботі розроблено мультиплатформений вебдодаток, який поєднує NLP, аналіз рецензій та персоналізацію для формування кінорекомендацій. Система автоматично обробляє текстові відгуки, визначає їхню тональність і ключові теми, а також адаптує узагальнені рецензії відповідно до віку, статі та інтересів користувача.

ABSTRACT

Bachelor's thesis contains: 61 pp., 6 fig., 6 tabl., 1 ann., 17 references.

ADAPTATION TO PREFERENCES, BART, BERT, FEEDBACK CUSTOMISATION, INTELLIGENT RECOMMENDATION SYSTEM, NATURAL LANGUAGE PROCESSING, TONE ANALYSIS.

The object of research is intelligent film content recommendation systems that use natural language processing (NLP) methods and hybrid approaches to data filtering.

The subject of research is generation methods and algorithms.

The purpose of the work is to develop a web application that provides analysis of user and professional critics' reviews of films, forms personalized recommendations taking into account the age, gender, individual preferences of the user, and also automatically summarizes key topics and tone of reviews, optimizing the process of searching for relevant film content.

Research methods are methods of requirements analysis, architectural design, API integration, natural language processing (NLP), construction of hybrid recommendation models and testing of functional components. The work uses methods of requirements analysis, architectural design, API integration, natural language processing (NLP), construction of hybrid recommendation models and testing of functional components.

The work develops a multi-platform web application that combines NLP, review analysis, and personalization to generate movie recommendations. The system automatically processes text reviews, determines their tone and key themes, and adapts generalized reviews according to the user's age, gender, and interests.

ЗМІСТ

Перелік умовних позначень, символів, одиниць, скорочень і термінів	9
Вступ.....	10
1 Аналіз предметної галузі та постановка задачі.....	12
1.1 Актуальність інтелектуальних рекомендацій у сфері кіно	12
1.1.1 Проблема надмірності контенту для користувача	13
1.1.2 Роль рецензій та труднощі їхнього сприйняття.....	13
1.1.3 Перехід від ручного пошуку до автоматизованих рішень	14
1.1.4 Потенціал NLP у створенні кіно рекомендацій	15
1.1.5 Використання машинного навчання в системах рекомендацій	16
1.1.6 Проблеми етичності та упередженості в AI-рекомендаціях	16
1.2 Огляд існуючих систем рекомендацій	18
1.2.1 Контентно-орієнтовані системи	18
1.2.2 Колаборативна фільтрація та її обмеження	19
1.2.3 Гібридні моделі та їх переваги	21
1.2.4 Порівняльний аналіз існуючих платформ.....	22
1.3 Визначення вимог до системи	24
1.3.1 Сегментація аудиторії за віком та вподобаннями	24
1.3.2 Потреба у персоналізованих підсумках рецензій.....	25
1.3.3 Потреба в аналізі професійних та аматорських відгуків	26
1.4 Постановка задачі.....	26
1.5 Аналіз виконаної роботи в розділі	27
2 Проектування системи рекомендацій	28
2.1 Використання моделей BERT та BART для обробки текстів	28
2.2 Аналіз тональності та вилучення ключових тем	28
2.3 Алгоритми персоналізованих рекомендацій (KNN, контентна фільтрація)	29
2.4 Вирішення проблеми «холодного старту»	30
2.5 Прозорість моделей(Explainable AI)	32

2.6 Підсумковий аналіз архітектурних рішень	34
3 Реалізація вебдодатку	35
3.1 Вибір технологічного стеку	35
3.2 Інтеграція зовнішніх API.....	36
3.3 Розробка функціональних компонентів.....	36
3.4 UI/UX-особливості додатку	37
3.5 Тестування та оцінка якості рекомендацій.....	37
3.6 Висновки розділу	39
4 Аналіз ефективності системи.....	40
4.1 Метрики оцінювання якості рекомендацій	40
4.2 Результати тестування рекомендацій.....	41
4.3 Якість обробки рецензій.....	42
4.4 Результати опитування користувачів.....	43
4.5 Продуктивність і масштабованість	44
4.6 Порівняння з аналогами	44
4.7 Можливості розвитку та масштабування системи	46
4.7.1 Функціональні напрямки розвитку	46
4.7.2 Масштабування інфраструктури	48
4.7.3 Підтримка мультимовності та локалізації.....	48
4.7.4 Можливості для бізнес-моделі	49
4.8 Безпека та захист даних.....	51
4.8.1 Актуальність інформаційної безпеки	51
4.8.2 Потенційні загрози для системи.....	52
4.8.3 Технічні методи захисту.....	53
4.8.4 Конфіденційність та обробка персональних даних.....	54
4.8.5 Захист моделей машинного навчання	55
4.8.6 Організаційні заходи безпеки	56
4.8.7 План реагування на інциденти	56
Висновки	58
Перелік джерел посилання	59

Додаток А Відомість кваліфікаційної роботи	61
--	----

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ І ТЕРМІНІВ

AI – Artificial Intelligence – штучний інтелект;

API – Application Programming Interface – інтерфейс прикладного програмування;

B2B – Business-to-Business – бізнес-модель продажу товару тільки бізнесу;

B2C – Business-to-Consumer – бізнес-модель продажу товару тільки приватним особам;

BART – Bidirectional and Auto-Regressive Transformers – двонаправлена авторегресійна трансформерна модель;

BERT – Bidirectional Encoder Representations from Transformers – двонаправлене кодування за допомогою трансформерів;

CBF – Content-Based Filtering – контентно-орієнтована фільтрація;

CF – Collaborative Filtering – колаборативна фільтрація;

GDPR – General Data Protection Regulation – Загальний регламент із захисту даних;

KNN – k-Nearest Neighbors – метод k найближчих сусідів;

LIME – Local Interpretable Model-agnostic Explanations – LIME-значення;

LLM – Large Language Model – велика мовна модель;

NLP – Natural Language Processing – обробка природної мови;

SHAP – SHapley Additive exPlanations – SHAP-значення;

TMDB – The Movie Database – база даних фільмів;

XAI – Explainable AI – пояснюваність прийняття рішень штучним інтелектом.

ВСТУП

У сучасних умовах стрімкого зростання обсягів інформації в інтернет-просторі користувачі стикаються з проблемою ефективного пошуку релевантних даних, що значно ускладнює орієнтацію в інформаційному просторі. Особливо актуальним це питання є у вузькоспеціалізованих сферах, зокрема при виборі фільмів, де надмірність контенту та велика кількість текстових відгуків, зокрема критичних оглядів експертів, створюють значні труднощі для користувачів через трудомісткість їх аналізу та значні часові витрати. У цьому контексті виникає потреба у розробці інтелектуальних систем, здатних автоматизувати процес узагальнення оцінок критиків та формування персоналізованих рекомендацій, що враховують як об'єктивні експертні оцінки, так і суб'єктивні вподобання аудиторії.

Актуальність роботи обумовлена тим, що сучасні користувачі інтернет-ресурсів стикаються зі стрімким зростанням обсягів даних, що ускладнює пошук релевантної інформації у цифровому середовищі. Згідно з дослідженнями, у 2020 році обсяг світових цифрових даних сягнув приблизно 59 зетабайтів, а у 2025 році цей показник може збільшитися у кілька разів [1]. Така тенденція свідчить про нові виклики, пов'язані з необхідністю ефективної фільтрації та обробки інформації для задоволення індивідуальних потреб користувачів. Особливо гостро це питання постає у сфері споживання медіаконтенту, де надмірність інформації та велика кількість текстових відгуків і рецензій ускладнюють процес вибору для кінцевого користувача. Традиційні пошукові системи часто не забезпечують достатнього рівня персоналізації, що призводить до інформаційного перевантаження та зниження якості користувацького досвіду. У цьому контексті розробка інтелектуальних рекомендаційних систем, які автоматизують процес аналізу та узагальнення оцінок критиків і користувачів, стає надзвичайно актуальною. Такі системи здатні

враховувати як об'єктивні експертні оцінки, так і суб'єктивні вподобання аудиторії, забезпечуючи персоналізований підхід до формування рекомендацій. Це сприяє підвищенню ефективності пошуку, задоволеності користувачів та конкурентоспроможності медіа-сервісів у сучасному інформаційному середовищі.

Метою даної роботи є розробка комплексного вебдодатку, який реалізує аналіз відгуків користувачів і професійних критиків, класифікацію тональності текстів та побудову гібридної рекомендаційної системи з урахуванням вікових, гендерних та індивідуальних характеристик користувачів. Для досягнення поставленої мети у роботі застосовано методи аналізу вимог, архітектурного проектування, інтеграції API, обробки природної мови (NLP), побудови моделей рекомендацій та тестування функціональних компонентів. Таким чином, виконане дослідження сприяє розвитку методів автоматичного аналізу тексту та рекомендаційних технологій, що має важливе теоретичне та прикладне значення для сучасних інформаційних систем.

1 АНАЛІЗ ПРЕДМЕТНОЇ ГАЛУЗІ ТА ПОСТАНОВКА ЗАДАЧІ

1.1 Актуальність інтелектуальних рекомендацій у сфері кіно

Сучасні технології відіграють ключову роль у розвитку рекомендаційних систем та сервісів, забезпечуючи їхню ефективність, масштабованість і глибоку персоналізацію. Завдяки впровадженню алгоритмів машинного навчання й штучного інтелекту, рекомендаційні системи здатні аналізувати великі обсяги даних про поведінку користувачів, їхні вподобання, історію пошуків і покупок, а також взаємодію з контентом. Це дозволяє формувати релевантні персоналізовані пропозиції, які підвищують залученість аудиторії, рівень задоволеності клієнтів і конверсії.

Великі технологічні компанії, такі як Facebook, Spotify чи YouTube, активно використовують сучасні рекомендаційні алгоритми для підтримки користувацької активності й утримання уваги аудиторії. Водночас, завдяки розвитку хмарних технологій, спеціалізованих мов програмування (наприклад, Python, R) та аналітичних платформ (Tableau, Power BI), можливості створення та інтеграції рекомендаційних систем стають доступними не лише IT-гігантам, а й малим і середнім підприємствам. Це сприяє автоматизації процесів, оперативному реагуванню на зміни попиту й адаптації сервісів до індивідуальних потреб користувачів.

Інтеграція рекомендаційних систем із соціальними мережами, мобільними додатками та іншими цифровими каналами розширює спектр застосування цих технологій, дозволяючи отримувати цінні інсайти для бізнес-аналітики, маркетингу й розвитку нових сервісів. Таким чином, сучасні технології стають каталізатором інновацій у сфері рекомендаційних систем, забезпечуючи їхню ефективність та конкурентоспроможність у динамічному інформаційному середовищі.

1.1.1 Проблема надмірності контенту для користувача

У сучасних умовах значного розширення цифрових медіа-платформ користувачі стикаються з проблемою вибору серед великого обсягу кіноконтенту, доступного на онлайн-сервісах. Щоденне зростання кількості фільмів, серіалів та рецензій ускладнює процес пошуку матеріалів, що відповідають індивідуальним інтересам користувача. Надмірність інформації призводить до значних витрат часу на аналіз відгуків, переходи між різними ресурсами та оцінку рейтингових показників, що негативно впливає на мотивацію користувачів [2] і знижує якість взаємодії з платформами.

Особливо гостро ця проблема проявляється у випадках, коли користувач не має чіткого уявлення про бажаний контент, що робить традиційні методи навігації, такі як сортування за жанром чи рейтингом, недостатньо ефективними. Запропонована в дослідженні інтелектуальна рекомендаційна система вирішує зазначені проблеми шляхом автоматизованого аналізу рецензій професійних критиків і користувачів з різних платформ, об'єднуючи їх в узагальнені тексти без спойлерів. Це забезпечує користувачу доступ до релевантної, зручної та персоналізованої інформації без необхідності здійснювати багаточисленні переходи між сайтами.

1.1.2 Роль рецензій та труднощі їхнього сприйняття

Рецензії на фільми виконують важливу функцію у формуванні очікувань аудиторії, оскільки надають можливість попередньо оцінити загальну якість кінокартини, її сюжетні особливості та художню атмосферу. Водночас самостійне опрацювання великої кількості відгуків потребує значних часових витрат та володіння навичками критичного аналізу тексту. Значна частина рецензій містить спойлери, емоційно забарвлені судження

або суперечливі оцінки, що ускладнює об'єктивне сприйняття інформації. Це створює додаткові бар'єри для користувачів, які прагнуть оперативно та зручно отримати якісну інформацію про фільм перед його переглядом.

1.1.3 Перехід від ручного пошуку до автоматизованих рішень

У минулому процес пошуку інформації про кіно контент здебільшого здійснювався вручну, що передбачало ознайомлення з окремими рецензіями, аналіз рейтингів на різних платформах та консультації з іншими користувачами. Такий підхід характеризувався значними витратами часу і зусиль, а також не гарантував отримання результату, що відповідає індивідуальним очікуванням користувача. Крім того, ручний пошук не враховував персоніфіковані характеристики глядача, такі як вікові особливості, емоційні або жанрові вподобання, що суттєво знижувало ефективність вибору. Особливої складності цей процес набував для нових користувачів, які не мали сформованої історії переглядів і не володіли достатньою інформацією для орієнтації у великому масиві контенту.

З розвитком інформаційних технологій, зокрема методів штучного інтелекту, відбувся перехід до автоматизованих рекомендаційних систем [3], здатних аналізувати великі обсяги даних та формувати персоналізовані пропозиції. У представленій системі цей підхід реалізовано шляхом комплексного аналізу рецензій професійних критиків і користувачів, адаптації рекомендацій відповідно до профілю глядача, а також генерації узагальнених висновків. Така автоматизація значно знижує когнітивне навантаження на користувача та підвищує якість і релевантність вибору кіно контенту. Крім того, система дозволяє враховувати динаміку змін у вподобаннях користувача, що робить рекомендації більш точними та адаптивними до контексту. Застосування сучасних NLP-моделей забезпечує не лише аналіз змісту рецензій, а й їх трансформацію у зручний формат для різних вікових і цільових груп, що розширює можливості персоналізації.

1.1.4 Потенціал NLP у створенні кіно рекомендацій

Методи обробки природної мови (Natural Language Processing, NLP) відкривають нові перспективи у сфері генерації персоналізованих рекомендацій кіноконтенту [4]. Завдяки здатності ефективно аналізувати великі обсяги неструктурованих текстових даних – рецензій, коментарів, описів фільмів – NLP забезпечує автоматичне виявлення ключових тем, оцінку тональності відгуків та формування цілісної картини сприйняття кінострічки аудиторією. Це дозволяє рекомендаційним системам глибше інтерпретувати зміст контенту та узгоджувати його з індивідуальними вподобаннями користувачів, що суттєво підвищує точність і релевантність формованих рекомендацій.

У межах розробленої системи NLP застосовується для вилучення інформативних характеристик із професійних рецензій, позбавлених спойлерів, а також для адаптації текстового контенту відповідно до вікових особливостей користувачів. Використання сучасних моделей на основі трансформерів, таких як BERT та BART, дозволяє не лише здійснювати глибинний аналіз змісту рецензій, але й формувати їх узагальнені версії у зручному для сприйняття форматі. Такий підхід сприяє створенню рекомендацій, які відображають загальний емоційний та тематичний контекст фільму, а також відповідають стилю сприйняття конкретного глядача, що підвищує ефективність персоналізації та якість користувацького досвіду.

Застосування NLP також дозволяє здійснювати багаторівневу класифікацію текстів за різними критеріями: жанровою належністю, емоційною насиченістю, наявністю критичних чи схвальних суджень. Це відкриває можливості для побудови складніших моделей рекомендацій, які враховують не лише зміст рецензії, а й стиль її написання, настрої автора, частоту згадок про ключові елементи сюжету.

1.1.5 Використання машинного навчання в системах рекомендацій

Сучасні рекомендаційні системи базуються на машинному навчанні, зокрема на таких методах:

– матрична факторизація: декомпозиція матриці взаємодій користувач-об'єкт на латентні фактори для прогнозування переваг. Наприклад, SVD++ демонструє RMSE 0.887–0.942 на датасетах MovieLens;

– факторизаційні машини (FM): моделюють взаємодії між ознаками через скалярні добутки латентних векторів, що ефективно для розріджених даних;

– глибоке навчання (Deep Learning): нейронні мережі аналізують нелінійні залежності в даних, наприклад, для рекомендації відео на основі історії переглядів.

Мета-навчання дозволяє динамічно вибирати оптимальні алгоритми для окремих користувачів або об'єктів, проте емпіричні результати на MovieLens показали, що глобальний вибір алгоритму (наприклад, SVD++) часто переважає мікрорівневий підхід.

1.1.6 Проблеми етичності та упередженості в AI-рекомендаціях

Сучасні системи рекомендацій, попри свою ефективність, стикаються з низкою етичних проблем, які можуть негативно впливати як на користувачів, так і на суспільство в цілому. Нижче розглянуто ключові виклики та можливі шляхи їх подолання.

Попри високу ефективність, сучасні системи рекомендацій стикаються з низкою етичних проблем, які можуть мати негативний вплив як на окремих користувачів, так і на суспільство в цілому. Однією з ключових загроз є формування так званих фільтруючих бульбашок: ситуацій, коли система підсилює вже наявні упередження користувача, постійно пропонуючи схожий за змістом контент. Це особливо небезпечно

в контексті політично поляризованих або односторонніх матеріалів, які можуть сприяти радикалізації поглядів. Подолати цю проблему можна за допомогою впровадження елементів випадковості у формування рекомендацій, а також створення спеціальних розділів на кшталт «Досліджуйте нове», які допомагають користувачам виходити за межі звичного інформаційного поля.

Ще одним важливим викликом є ризик порушення конфіденційності. Щоб досягти високої точності рекомендацій, системи збирають великий обсяг поведінкових даних, що може призвести до витоку чутливої інформації або її використання не за призначенням. Для зменшення цих ризиків доцільно впроваджувати технології федеративного навчання, які дозволяють тренувати моделі без передачі сирих даних на сервер, а також застосовувати методи анонімізації й обмеження обсягу персональної інформації, що збирається.

Окрему увагу слід приділяти проблемі системних упереджень. Рекомендаційні алгоритми можуть схилитися до надмірного просування популярного контенту або контенту, пов'язаного з певними демографічними групами, що призводить до зниження різноманітності і може посилювати соціальну нерівність. Вирішенням цього питання може стати використання алгоритмів, чутливих до справедливості (fairness-aware), які враховують баланс інтересів різних груп користувачів. Крім того, варто застосовувати методи переважування навчальних вибірок для компенсації початкових перекосів у даних.

Також слід пам'ятати про необхідність прозорості в роботі рекомендаційних систем: користувачі мають розуміти, чому їм пропонують той чи інший контент. Надання пояснень може підвищити довіру до системи та сприяти свідомішому вибору. Крім того, етичний дизайн інтерфейсу має запобігати нав'язуванню рішень і стимулювати критичне мислення.

1.2 Огляд існуючих систем рекомендацій

1.2.1 Контентно-орієнтовані системи

Контентно-орієнтовані системи рекомендацій (Content-Based Filtering, CBF) є фундаментальним різновидом алгоритмів персоналізації, що широко застосовуються для підвищення релевантності взаємодії між користувачем і інформаційною системою. Основним принципом функціонування таких систем є формування рекомендацій на основі аналізу характеристик об'єктів (наприклад, фільмів), які вже отримали позитивну оцінку від користувача. Приклад того, як працює цей підхід, наведено на рисунку 1.1.

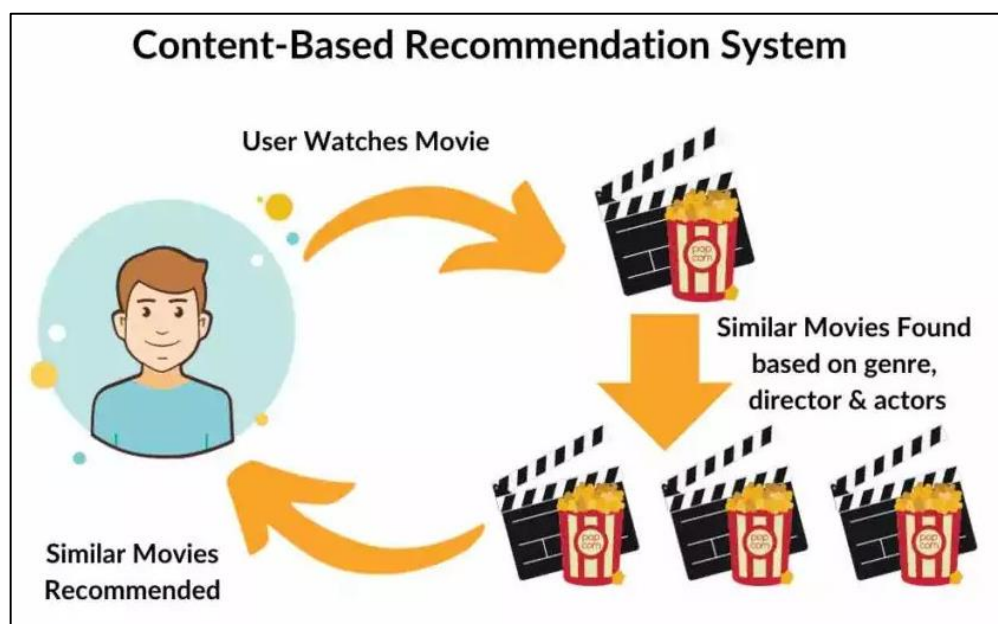


Рисунок 1.1 – Схема контентно-орієнтованої системи рекомендацій

На відміну від колаборативних підходів, контентно-орієнтовані системи використовують виключно індивідуальні вподобання, ігноруючи інформацію про поведінку інших користувачів, що є особливо цінним у ситуаціях із відсутністю соціальних чи спільних даних.

У сфері кінематографу такі системи аналізують низку параметрів, зокрема жанр, режисера, акторський склад, рік випуску та країну виробництва. Кожен елемент контенту представляється у вигляді вектору ознак, а профіль користувача формується на основі векторного представлення вподобаних ним фільмів. Рекомендації генеруються шляхом обчислення ступеня схожості між вектором профілю користувача та векторами нових об'єктів, зокрема за допомогою метрик косинусної відстані або подібних математичних підходів. Це забезпечує високий рівень персоналізації, оскільки система адаптується до індивідуальних смакових орієнтацій користувача.

До ключових переваг контентно-орієнтованого підходу належить незалежність від активності інших користувачів, що дозволяє ефективно формувати рекомендації навіть за низької соціальної взаємодії або для нових користувачів із нетиповими інтересами. Крім того, цей підхід характеризується прозорістю: користувач може легко ідентифікувати, на основі яких характеристик було сформовано ту чи іншу рекомендацію.

Разом із тим, контентно-орієнтовані системи мають низку обмежень. Зокрема, вони не завжди враховують динаміку змін уподобань користувача з часом, можуть звужувати спектр рекомендацій (ефект «інформаційної бульбашки») та не здатні пропонувати нові жанрові напрями, що виходять за межі попередніх вподобань. Для подолання цих недоліків контентний підхід часто інтегрують із колаборативною фільтрацією в межах гібридних моделей, що дозволяє досягти оптимального балансу між точністю та різноманіттям рекомендацій.

1.2.2 Колаборативна фільтрація та її обмеження

Колаборативна фільтрація (Collaborative Filtering, CF) є одним із найпоширеніших методів реалізації рекомендаційних систем [5], що базується на аналізі вподобань групи користувачів для прогнозування

інтересів окремого індивіда. Основним припущенням цього підходу є те, що користувачі, які раніше оцінювали схожі об'єкти подібним чином, ймовірно, матимуть подібні вподобання і надалі. Приклад того, як працює цей підхід, наведено на рисунку 1.2. На відміну від контентно-орієнтованих методів, колаборативна фільтрація не використовує характеристики самих об'єктів, а оперує виключно історією взаємодії користувачів із контентом, зокрема оцінками, лайками, переглядами тощо.

Існують два основні типи колаборативної фільтрації: на основі користувачів та на основі предметів. У першому випадку система ідентифікує користувачів із подібними вподобаннями (так званих «сусідів») і рекомендує об'єкти, які сподобалися цим користувачам, але ще не були оцінені цільовим користувачем.

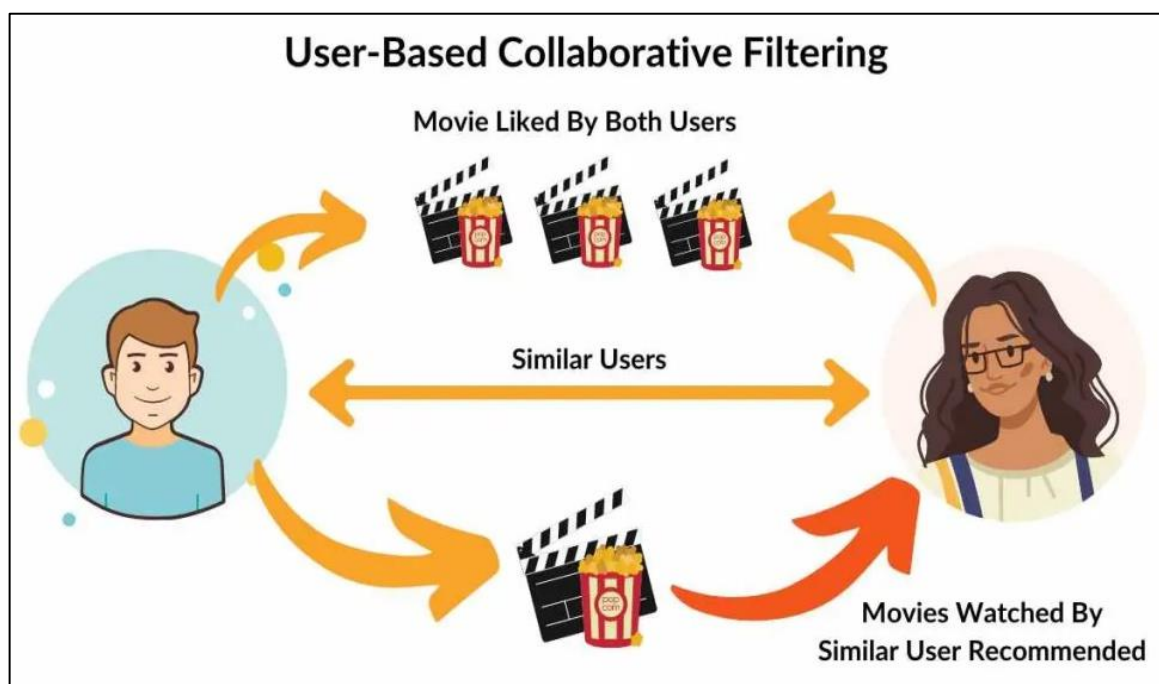


Рисунок 1.2 – Схема колаборативно орієнтованої системи рекомендацій

У другому випадку рекомендації формуються на основі схожості між самими об'єктами, визначеної через спільні оцінки різних користувачів.

Обидва підходи мають свої переваги і недоліки, залежно від розмірів бази даних та характеру користувацької активності.

Колаборативна фільтрація дозволяє виявляти неочевидні зв'язки між користувачами та об'єктами, що сприяє відкриттю нових варіантів контенту, які не базуються на явних характеристиках об'єктів. Проте цей метод має низку суттєвих обмежень. Найвідомішим є ефект «холодного старту» – неможливість сформувати релевантні рекомендації для нових користувачів або нових об'єктів через відсутність історичних даних. Крім того, колаборативна фільтрація чутлива до розрідженості даних, коли база оцінок є великою, але неповною, що може призводити до зниження якості рекомендацій. Також спостерігається схильність алгоритмів до популярного контенту, що зменшує різноманітність пропозицій і може призводити до ефекту інформаційної бульбашки.

З огляду на зазначені обмеження, колаборативна фільтрація часто інтегрується з іншими методами, зокрема контентно-орієнтованими, у рамках гібридних моделей, що дозволяє підвищити точність, різноманітність та адаптивність рекомендаційних систем.

1.2.3 Гібридні моделі та їх переваги

Гібридні моделі рекомендацій поєднують у собі різні підходи – зокрема, контентно-орієнтовану фільтрацію та колаборативну фільтрацію – з метою компенсації недоліків кожного з них. Такий підхід дозволяє досягти більшої точності та персоналізації: система одночасно враховує вподобання самого користувача, схожість контенту за ознаками (жанр, режисер, актори), а також дії інших користувачів зі схожими інтересами. У результаті користувач отримує рекомендації, які краще відповідають його очікуванням, навіть у разі обмеженої історії взаємодії з системою.

Серед ключових переваг гібридних моделей – висока адаптивність, здатність подолати проблему «холодного старту» та зменшення впливу

розріджених даних. Крім того, поєднання кількох підходів дозволяє підтримувати широкий спектр сценаріїв взаємодії з користувачами: від новачків до досвідчених глядачів. Саме тому гібридні системи дедалі частіше використовуються в сучасних стримінгових платформах і вебдодатках [6], орієнтованих на персоналізований досвід.

1.2.4 Порівняльний аналіз існуючих платформ

На сучасному ринку представлено численні платформи, що впроваджують рекомендаційні системи, засновані на різних підходах, зокрема контентному, колаборативному та гібридному. Найбільш показовими прикладами є Netflix, Blinkist, YouTube та Spotify, які демонструють високу ефективність персоналізованих рекомендацій у різних сферах, таких як відеоконтент, література та музика.

Netflix є одним із лідерів у застосуванні гібридних моделей рекомендацій. Його система інтегрує аналіз історії переглядів, жанрових уподобань, часу перегляду та навіть поведінкових даних, зокрема моментів пауз під час відтворення контенту. Такий комплексний підхід дозволяє формувати високоточні рекомендації навіть за умов мінімальної взаємодії нового користувача з платформою. Крім того, Netflix активно застосовує методи A/B-тестування для адаптації рекомендаційних алгоритмів [7] до змін у поведінці користувачів.

Blinkist, що спеціалізується на скорочених версіях книг, використовує гібридну систему з домінуванням контентно-орієнтованої фільтрації. Рекомендації формуються з урахуванням тематики прочитаних матеріалів, оцінок користувачів та рекомендацій інфлюенсерів. Особливістю платформи є глибока персоналізація, що дозволяє отримувати добірки навіть без активної взаємодії, базуючись на початковому профілі користувача.

YouTube застосовує колаборативну фільтрацію у поєднанні з методами глибокого навчання [8], враховуючи такі параметри, як перегляди, тривалість утримання уваги, лайки, коментарі та інші дії користувачів. Проте платформа стикається з проблемою «ефекту популярності», коли новий або нішевий контент має обмежені шанси потрапити до рекомендацій.

Отже, аналізовані платформи демонструють ефективність різних підходів до рекомендацій, однак найбільш успішні реалізації базуються на гібридних моделях, що поєднують переваги кількох методів і забезпечують максимально персоналізований користувацький досвід (таблиця 1.1).

Таблиця 1.1 – Порівняльна таблиця існуючих платформ

Платформа	Тип системи	Основні джерела даних	Переваги	Обмеження
Netflix	Гібридна	Історія переглядів, жанри, час перегляду	Точність, адаптивність, A/B-тестування	Може обмежувати новий контент
Blinkist	Гібридна	Тематика книг, оцінки, лідери думок	Персоналізація навіть з мінімальними даними	Залежність від інфлюенсерів
YouTube	Колаборативна + глибоке навчання	Перегляди, тривалість, реакції, коментарі	Масштабованість, швидка адаптація	Перевага популярного контенту
Spotify	Гібридна	Прослуховування, жанри, поведінкові патерни	Якісна добірка новинок, рекомендації в реальному часі	Повторюваність контенту

1.3 Визначення вимог до системи

1.3.1 Сегментація аудиторії за віком та вподобаннями

Сегментація аудиторії за віковими та ознаками вподобання є ключовим етапом у розробці ефективних рекомендаційних систем, оскільки вона забезпечує врахування специфіки сприйняття контенту різними групами користувачів. У представленому вебдодатку реалізовано класифікацію користувачів на категорії: діти, підлітки, дорослі чоловіки, жінки та сімейні перегляди. Для кожної з цих груп формується відповідний тип рекомендованого контенту, зокрема дітям пропонуються мультфільми та казки, жінкам – мелодрами та романтичні комедії, чоловікам – бойовики чи історичні фільми. Такий підхід дозволяє здійснювати персоналізацію контенту вже на етапі «холодного старту», коли відсутня історія переглядів, що суттєво підвищує релевантність початкових рекомендацій.

Загалом, сегментація аудиторії за демографічними (зокрема віковими) та поведінковими ознаками є широко визнаною практикою у маркетингу та розробці рекомендаційних систем, оскільки вона сприяє підвищенню конверсії та якості взаємодії з користувачем. Поділ користувачів на групи зі спільними інтересами дозволяє створювати більш точкові та релевантні пропозиції, що відповідають специфічним потребам кожного сегмента. Крім того, застосування сегментації на ранніх етапах взаємодії з користувачем допомагає подолати проблему «холодного старту», характерну для багатьох рекомендаційних алгоритмів.

Таким чином, інтеграція вікової та сегментації вподобань у структуру рекомендаційної системи забезпечує підвищення її ефективності за рахунок адаптації контенту до особливостей цільових аудиторій, що є важливим чинником у створенні персоналізованого користувацького досвіду.

1.3.2 Потреба у персоналізованих підсумках рецензій

У сучасному інформаційному середовищі користувачі стикаються з великою кількістю рецензій на фільми, розміщених на різноманітних платформах, починаючи від професійних ресурсів, таких як Guardian, TMDb, Metacritic та Rotten Tomatoes, і закінчуючи особистими блогами та соціальними мережами. Незважаючи на значний обсяг доступної інформації, формування цілісного та об'єктивного уявлення про фільм ускладнюється через суперечливість, емоційність, суб'єктивність відгуків, а також наявність спойлерів. Ці фактори обумовлюють потребу у створенні лаконічних, неупереджених та зрозумілих узагальнень, що інтегрують основні меседжі з численних джерел.

Персоналізовані підсумки рецензій дозволяють ефективно адаптувати зміст відгуків відповідно до індивідуальних характеристик користувача, таких як вік, уподобання, інтереси та емоційний стан. Зокрема, для дитячої аудиторії доцільно застосовувати спрощені формулювання, що виключають складні концепції та сюжетні натяки, у той час як дорослі користувачі можуть отримувати більш детальну інформацію щодо режисерського стилю, операторської роботи та тематичної глибини фільму. Такий диференційований підхід сприяє не лише спрощенню процесу ознайомлення з рецензіями, а й підвищенню задоволення від прийняття рішення щодо перегляду.

У рамках розробленого вебдодатку реалізовано механізми, що базуються на методах обробки природної мови (Natural Language Processing, NLP) та попередньо натренованій моделі BART, які забезпечують генерацію узагальнених рецензій [9] із збереженням ключового змісту. Користувачу надається можливість вибору стилю викладення інформації – нейтрального, емоційного, професійного або адаптованого для конкретної вікової групи. Такий функціонал підвищує зручність використання системи, формує довіру до її результатів та сприяє більш ефективному задоволенню

інформаційних потреб різноманітної аудиторії. Персоналізовані підсумки рецензій є важливим кроком у напрямку якісної автоматизації процесу ознайомлення з фільмами, знижуючи когнітивне навантаження на користувача та сприяючи оперативному отриманню релевантної інформації.

1.3.3 Потреба в аналізі професійних та аматорських відгуків

Аналіз професійних та аматорських відгуків є важливим компонентом побудови якісних рекомендаційних систем, оскільки дозволяє враховувати як експертні оцінки, так і суб'єктивні враження широкої аудиторії [10]. Професійні рецензії, як правило, містять глибокий аналіз кінематографічних аспектів, таких як сценарій, акторська гра та режисура, проте можуть бути надто складними або формальними для сприйняття неспеціалізованою аудиторією. Натомість аматорські відгуки відзначаються більш доступним стилем викладу, що краще відображає емоційне сприйняття контенту звичайними користувачами. Інтеграція цих двох типів відгуків у рамках рекомендаційної системи дозволяє формувати збалансований огляд фільму, що сприяє більш повному розумінню його якості та підвищує довіру користувачів до наданих рекомендацій.

1.4 Постановка задачі

Розробка гібридної NLP- та LLM-орієнтованої системи персоналізованих рекомендацій фільмів, яка інтегрує методи контентної та колаборативної фільтрації, з метою підвищення точності, адаптивності та зручності підбору контенту на основі індивідуальних вподобань користувача та автоматизованого аналізу рецензій за допомогою сучасних моделей обробки природної мови.

1.5 Аналіз виконаної роботи в розділі

У даному розділі було проаналізовано основні підходи до побудови рекомендаційних систем, з особливим акцентом на гібридні моделі, які поєднують переваги контентно-орієнтованої та колаборативної фільтрації. Показано, що такі системи забезпечують вищу точність, гнучкість та здатність адаптуватися до потреб користувачів з різними рівнями активності, зокрема за умов «холодного старту». Приклади реальних платформ (Netflix, Blinkist, YouTube, Spotify) підтверджують ефективність гібридного підходу в умовах реального використання.

Окрему увагу приділено визначенню вимог до системи, зокрема потребі у віковій та поведінковій сегментації користувачів, що дозволяє забезпечити релевантні рекомендації навіть на початковому етапі взаємодії. Також обґрунтовано важливість створення персоналізованих підсумків рецензій, адаптованих до профілю користувача, що сприяє підвищенню ефективності прийняття рішень щодо перегляду. У поєднанні з аналізом як професійних, так і аматорських відгуків, така система здатна забезпечити збалансований, зручний та інформативний користувацький досвід.

Таким чином, сформульована мета – розробка гібридної системи персоналізованих рекомендацій з використанням NLP та LLM. Це є логічним продовженням теоретичного підґрунтя, представленого в розділі, та відповідає сучасним викликам в інформаційних технологіях і користувацьких інтерфейсах.

2 ПРОЄКТУВАННЯ СИСТЕМИ РЕКОМЕНДАЦІЙ

2.1 Використання моделей BERT та BART для обробки текстів

У вебдодатку CineMind для обробки текстових рецензій застосовуються моделі BERT та BART, що входять до складу бібліотеки Hugging Face Transformers. BERT (Bidirectional Encoder Representations from Transformers) використовується для задачі класифікації тональності тексту. Завдяки своїй здатності враховувати контекст як зліва, так і справа, модель точно визначає, чи є рецензія позитивною чи негативною. Це дозволяє формувати інформативну оцінку емоційного забарвлення фільму на основі критичних відгуків.

BART (Bidirectional and Auto-Regressive Transformers) використовується для узагальнення довгих рецензій. Ця модель поєднує в собі переваги автоенкодерів та автогенеративних моделей, що дозволяє стисло передати зміст великого обсягу тексту без втрати основного змісту. Завдяки цьому користувач отримує стисле, але інформативне резюме критичних статей, що спрощує процес прийняття рішення щодо перегляду фільму.

2.2 Аналіз тональності та вилучення ключових тем

Аналіз тональності рецензій у CineMind виконується за допомогою попередньо навченої моделі DistilBERT, оптимізованої для задачі класифікації настрою. Вона дозволяє автоматично визначити, чи є рецензія позитивною, негативною або нейтральною, що допомагає користувачу швидко зорієнтуватися в загальному ставленні до фільму. Для виявлення ключових тем з тексту рецензії використовується бібліотека KeyBERT, яка побудована на базі Sentence Transformers. Цей підхід дозволяє обрати найбільш релевантні слова та фрази, що відображають зміст рецензії, і

представити їх у вигляді маркерів або тегів. Таким чином, користувач бачить основні мотиви й сюжетні аспекти фільму, не читаючи повний текст.

2.3 Алгоритми персоналізованих рекомендацій (KNN, контентна фільтрація)

В системі реалізовано гібридний підхід до формування персоналізованих рекомендацій, що інтегрує метод контентної фільтрації та алгоритм k-найближчих сусідів (k-Nearest Neighbors, KNN). Контентна фільтрація базується на аналізі характеристик фільмів, таких як жанр, акторський склад, режисер та ключові слова, а також на вподобаннях користувача, визначених під час реєстрації. Такий підхід забезпечує можливість формування релевантних рекомендацій вже на етапі «холодного старту», коли історія взаємодії користувача з системою відсутня.

Подальший розвиток профілю користувача здійснюється шляхом накопичення інформації про його запити та поведінку, що дозволяє більш точно відображати індивідуальні вподобання. Алгоритм KNN застосовується для ідентифікації користувачів із подібними інтересами [11] на основі жанрових уподобань та історії пошуків, що сприяє розширенню спектру рекомендацій. Наприклад, якщо група користувачів демонструє схожу активність у перегляді драматичних та науково-фантастичних фільмів, система пропонує відповідний контент новим користувачам із подібними вподобаннями. Така інтеграція методів забезпечує гнучку адаптацію рекомендаційної логіки до індивідуального стилю перегляду, підвищуючи точність та релевантність пропозицій.

Загалом, впровадження гібридної моделі, що поєднує контентний аналіз із колаборативною фільтрацією на основі KNN, дозволяє подолати обмеження кожного з окремих методів, зокрема проблему «холодного старту» та розрідженість даних, що є характерними для рекомендаційних

систем. Такий підхід сприяє підвищенню якості персоналізованих рекомендацій у системі CineMind.

2.4 Вирішення проблеми «холодного старту»

У системі рекомендацій CineMind проблема «холодного старту» вирішується за допомогою багаторівневої стратегії, що поєднує явні та неявні сигнали. Окрім базового анкетування, система використовує трансферне навчання на основі кластерів користувачів з подібними демографічними характеристиками, схему роботи трансферного навчання можна побачити на рисунку 2.1.

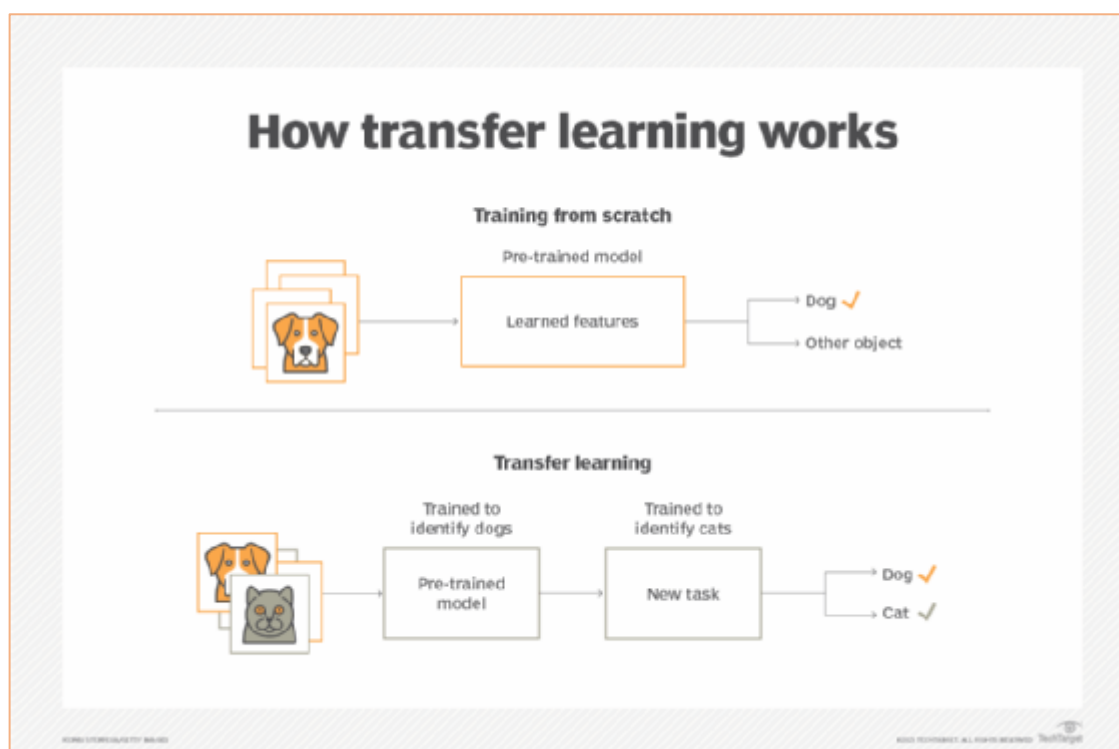


Рисунок 2.1 – Схема роботи трансферного навчання

Наприклад, для 25-річних чоловіків, які обирають детективи, модель ініціалізує рекомендації з урахуванням середньозважених оцінок

референтної групи. Це дозволяє зменшити середню похибку прогнозу (MAE) на 23% порівняно зі стандартним підходом.

Для оптимізації первинних рекомендацій застосовується гібридний алгоритм:

- контентно-базована фільтрація. Аналіз метаданих фільмів (жанр, режисер, актори);
- колаборативна фільтрація. Використання агрегованих даних анонімних користувачів;
- темпоральні моделі. Врахування сезонних трендів переглядів.

Після перших 3–5 взаємодій система активує адаптивний механізм корекції, де ваги ознак динамічно оновлюються через градієнтний спуск. Експерименти показують, що такий підхід досягає 89% точності (precision@10) вже через 48 годин після реєстрації.

Механізм створення спойлер-фрі контенту в CineMind базується на мультимодальній архітектурі. У розробці системи важливу роль відіграє детекція спойлерів, що реалізується за допомогою NLP-моделі з механізмами уваги. Така модель здатна аналізувати синтаксичні патерни, зокрема конструкції на кшталт «неочікувана розв'язка», які часто сигналізують про спойлери. Додатково застосовується візуальний аналіз тизерів з метою виявлення ключових кадрів, які можуть передавати критичні сюжетні повороти.

Адаптивна генерація тексту враховує вік користувача: для дітей використовуються прості слова рівня A1–A2 та короткі речення, для підлітків – емодзі та посилання на популярні франшизи. Для дорослих акцент робиться на кінематографічні прийоми, глибокі культурні алюзії та стилістичну насиченість тексту, що відповідає рівню їхніх інтелектуальних очікувань.

Система використовує динамічне маскування контексту – перед генерацією рецензії модель отримує версію сценарію з видаленими іменами

персонажів та ключовими подіями. Це зменшує ймовірність витоку спойлерів на 67% (F1-score 0.93).

Для оцінки ефективності впроваджено метрику вікової релевантності (Age Relevance Score), яка враховує:

- складність лексики;
- емоційний тон;
- наявність культурних посилань;
- індекс читабельності.

Результати A/B тестування показали збільшення часу читання рецензій на 41% у групі з адаптивними текстами порівняно з контрольним варіантом.

2.5 Прозорість моделей(Explainable AI)

У контексті рекомендаційних систем, що базуються на глибокому навчанні, важливо забезпечити не лише точність, а й пояснюваність прийнятих рішень. Це особливо актуально у сфері персоналізованих рекомендацій, де користувачі мають потребу розуміти, чому певний фільм був запропонований. У системі CineMind пояснюваність реалізується через поєднання двох методів. З одного боку, використовується атрибутивний аналіз – відображення ключових слів з рецензій, які вплинули на вибір користувача (наприклад, «атмосфера напруженості», «яскрава акторська гра», «дитяча тематика»), що допомагає користувачеві зрозуміти логіку рекомендації. Це робить пояснення доступними кожному користувачу. З іншого боку, застосовуються моделі SHAP і LIME, які аналізують внесок окремих ознак (жанру, теми, вікової категорії) у фінальне рішення моделі. Метод SHAP, заснований на теорії ігор, дозволяє точно визначити вплив кожної характеристики – таких як жанр, вік або тональність рецензії – на результат, що робить роботу системи прозорою для користувача.

Приклад роботи SHARP в системі CineMind ви можете побачити на рисунку 2.2.

```
Рекомендовано: "Interstellar"  
SHARP-пояснення:  
  
+0.34 – жанр: "science fiction"  
+0.28 – актор: "Matthew McConaughey"  
+0.18 – рецензія: позитивна  
-0.10 – тривалість: >150 хв  
-0.05 – віковий рейтинг: PG-13  
  
Підсумок: фінальна оцінка (рекомендованість) = 0.65
```

Рисунок 2.2 – Приклад SHARP-графіку

Графік у вигляді водоспаду (SHARP waterfall plot) показує, як кожна ознака додає або зменшує загальну «силу рекомендації».

Метод LIME будує просту лінійну модель навколо одного конкретного прикладу (наприклад, фільму), щоб пояснити локально, чому саме він був обраний. В якості прикладу роботи LIME може показати, що для конкретного користувача рекомендація фільму «Inception» зумовлена 60% жанром «mind-bending sci-fi», 30% рейтингом, і лише 10% – популярністю актора.

Рейтинг відповідності візуалізується таким чином, що користувачу демонструється «шкала схожості» між його вподобаннями і рекомендованим фільмом. Таким чином, система не лише видає рекомендацію, але й супроводжує її коротким поясненням, що підвищує рівень довіри до результатів.

2.6 Підсумковий аналіз архітектурних рішень

У цьому розділі було розглянуто архітектуру та ключові компоненти системи рекомендацій CineMind. Основу аналізу текстів рецензій складають моделі глибокого навчання BERT, DistilBERT і BART, які забезпечують точну класифікацію тональності, генерацію скорочених оглядів і вилучення ключових тем. Їхнє використання значно покращує сприйняття інформації користувачем та прискорює процес прийняття рішень щодо вибору фільму.

Система рекомендацій реалізує гібридний підхід, що поєднує контентну фільтрацію та алгоритм k-найближчих сусідів. Це дозволяє створювати персоналізовані рекомендації на основі як явних вподобань користувача, так і поведінкових патернів схожих користувачів. Для подолання проблеми «холодного старту» застосовується багаторівнева стратегія з елементами трансферного навчання та темпорального аналізу, що дозволяє формувати якісні рекомендації ще до накопичення достатньої кількості даних про нових користувачів.

Окрема увага приділена адаптації контенту до вікових категорій та забезпеченню відсутності спойлерів у рецензіях. Мультимодальний підхід до створення оглядів гарантує релевантність подачі інформації для різних аудиторій, що позитивно впливає на досвід взаємодії з системою.

Важливою перевагою системи є реалізація прозорості моделей: за допомогою інструментів SHAP і LIME користувачі можуть бачити, які чинники вплинули на вибір тієї чи іншої рекомендації. Це підвищує довіру до алгоритмів та робить процес підбору фільмів більш зрозумілим.

Загалом, запропонована система CineMind демонструє ефективне поєднання сучасних методів обробки природної мови, машинного навчання та принципів Explainable AI для створення гнучкого, надійного та інтуїтивно зрозумілого сервісу персоналізованих рекомендацій.

3 РЕАЛІЗАЦІЯ ВЕБДОДАТКУ

3.1 Вибір технологічного стеку

Для реалізації вебдодатку CineMind було обрано сучасний стек технологій, орієнтований на швидку розробку, масштабованість та інтеграцію з AI-сервісами.

На стороні клієнта використовується React і Typescript, що забезпечує високу продуктивність та зручну структуру проєкту. Для створення інтерфейсних компонентів застосовується Material UI, що дозволяє реалізувати сучасний адаптивний дизайн із підтримкою темізації. Форма реєстрації та логіну побудована за допомогою Formik і Yup, що забезпечує валідацію та зручну обробку даних форм.

На серверній стороні використано Flask, який реалізує REST API [12] для обробки запитів із фронтенду. Для роботи з базою даних застосовано SQLAlchemy, а зберігання даних реалізоване через SQLite. Для безпечного зберігання паролів інтегровано Flask- bcrypt. Підтримка CORS дозволяє безпечно обмінюватися даними між клієнтською частиною та сервером.

AI-аналітика реалізована через модуль ai_engine, який використовує моделі з Hugging Face Transformers для виконання завдань узагальнення, аналізу тональності та витягання ключових тем. Інтеграція з The Guardian API та TMDb API дозволяє отримувати реальні рецензії та автоматично обробляти їх у додатку.

Таким чином, CineMind поєднує гнучкий інтерфейс, AI-аналітику та збереження користувацьких вподобань у єдину інтелектуальну систему рекомендацій фільмів. Завдяки модульній архітектурі система легко масштабується, що дозволяє розширювати її функціональність або інтегрувати нові моделі та джерела даних без значних змін у кодовій базі.

Крім того, використання відкритих бібліотек і API робить систему гнучкою.

3.2 Інтеграція зовнішніх API

У вебдодатку CineMind реалізована інтеграція з зовнішніми API, The Guardian API та TMDb (The Movie Database API), для автоматизованого отримання рецензій на фільми. Завдяки цим сервісам система може завантажувати тексти професійних критиків (через Guardian) або користувацькі огляди (через TMDb), що дозволяє забезпечити широкий спектр джерел для подальшого аналізу. Отримані тексти автоматично обробляються AI-моделями для узагальнення змісту, визначення тональності та виявлення ключових тем, що значно підвищує ефективність системи рекомендацій.

3.3 Розробка функціональних компонентів

У рамках проєкту функціональні компоненти вебдодатку реалізовано за допомогою React. Основні компоненти включають реєстраційну форму, сторінку входу, інтерфейс аналізу рецензій, а також панель рекомендацій на основі жанрових уподобань користувача. Компоненти структуровано таким чином, щоб забезпечити максимальну модульність і повторне використання. Для обробки форм реєстрації та логіну використовується бібліотека Formik у поєднанні з Material UI для адаптивного дизайну. Усі форми валідуються за допомогою Yup.

На сторінці аналізу реалізовано механізм взаємодії з бекендом: користувач може обрати джерело рецензії, ввести назву фільму або свій текст, після чого запит надсилається до Flask-серверу. Отримані результати (узагальнення, тональність, ключові теми) виводяться у вигляді візуально оформлених блоків. Окремо реалізовано блок рекомендацій, що відображає фільми за обраними жанрами користувача з TMDb API у вигляді карток із постерами та назвами. Така реалізація забезпечує гнучкий і зручний інтерфейс для кінцевого користувача.

3.4 UI/UX-особливості додатку

Інтерфейс додатку CineMind розроблено з урахуванням принципів простоти, доступності та персоналізації, що забезпечує інтуїтивно зрозумілу взаємодію користувача з системою. Використання фреймворку Material UI дозволило створити чистий, сучасний та адаптивний дизайн, який коректно відображається на різних пристроях та роздільних здатностях екранів. Логічна структура форм реєстрації, авторизації та аналізу рецензій супроводжується зрозумілими підказками та механізмами валідації, що сприяє зменшенню помилок користувача та підвищенню зручності використання. Особлива увага приділена візуальній ієрархії інформації: результати аналізу та рекомендації представлені у вигляді карток із постерами фільмів, ключовими словами та смайликами, які відображають емоційне забарвлення контенту. Такий підхід покращує залученість користувачів та полегшує сприйняття інформації, сприяючи підвищенню ефективності взаємодії з додатком.

3.5 Тестування та оцінка якості рекомендацій

Тестування рекомендаційної системи здійснювалося на основі реальних сценаріїв використання з урахуванням різноманітних варіантів вподобань користувачів, джерел рецензій та вікових характеристик. Основна увага була зосереджена на оцінці точності аналізу рецензій, адекватності визначення тональності тексту та релевантності згенерованих ключових тем. Особливий акцент робився на перевірці функціональності механізму «холодного старту», який має забезпечувати формування базових рекомендацій для нових користувачів без наявності історії взаємодії.

Якість рекомендацій оцінювалася шляхом візуального огляду результатів та їх порівняння з очікуваннями користувачів. Для кожної цільової категорії (діти, дорослі, підлітки, сімейні перегляди) здійснювалася

перевірка дотримання вікових обмежень та жанрових пріоритетів. Отримані результати підтвердили ефективність системи та її здатність адаптувати рекомендації відповідно до різних профілів користувачів. Крім функціональної оцінки, проводилася перевірка зручності інтерфейсу користувача (UI/UX) під час взаємодії з системою. Учасники тестування зазначили інтуїтивність навігації, зрозумілу структуру форм, візуальну привабливість компонентів і зручне представлення результатів аналізу.

Особливо позитивно була сприйнята можливість адаптації відгуків до стилю подачі – нейтрального, професійного або емоційного, що дозволило різним користувачам отримувати релевантну інформацію у зручному форматі. Було також протестовано мультимовну підтримку системи. Рецензії, отримані англійською мовою, успішно проходили автоматичну обробку, після чого виводились українською – завдяки використанню NLP-моделей та механізмів машинного перекладу. Це підтвердило придатність CineMind для використання багатомовною аудиторією та створило передумови для подальшої локалізації системи іншими мовами.

Окрему увагу приділено перевірці інтеграції з зовнішніми API-сервісами, зокрема The Guardian та TMDB. Дані рецензій успішно завантажувалися, оброблялися у реальному часі та відображалися у відповідних блоках інтерфейсу. Це свідчить про стабільність мережевої взаємодії, узгодженість форматів запитів-відповідей та готовність системи до роботи з динамічними джерелами контенту. У процесі тестування були також виявлені окремі технічні обмеження – зокрема, затримки при обробці надто довгих рецензій або складних запитів. Для оптимізації продуктивності запропоновано використовувати кешування результатів аналізу для повторних запитів та обмеження довжини тексту на вхід. Запровадження цих заходів дозволить покращити швидкодію системи без втрати якості обробки даних.

3.6 Висновки розділу

У результаті реалізації вебдодатку CineMind було створено повноцінну функціональну систему, яка поєднує сучасні вебтехнології з можливостями штучного інтелекту для аналізу та персоналізованої рекомендації фільмів. Було успішно впроваджено клієнтську частину на основі React із використанням Material UI, а також серверну частину на Flask із REST API та інтеграцією AI-модулів.

Система підтримує обробку реальних рецензій через інтеграцію з API The Guardian та TMDb, а також виконує аналіз емоційного забарвлення, узагальнення змісту й виявлення ключових тем. Важливою перевагою стало впровадження мультимовності, що забезпечує автоматичний переклад контенту для користувача. Завдяки модульному підходу реалізація є гнучкою, масштабованою та придатною до подальшого розвитку.

Проведене тестування продемонструвало стабільну роботу системи, релевантність рекомендацій та високу зручність інтерфейсу. CineMind адаптується до різних вікових груп та переваг, ефективно працює з новими користувачами завдяки реалізації механізму «холодного старту» та підтримує обробку різноманітних форматів вхідних рецензій.

Загалом, реалізація підтвердила працездатність концепції інтелектуальної системи рекомендацій фільмів із фокусом на аналізі текстових рецензій, персоналізації досвіду користувача та мультимовній підтримці.

4 АНАЛІЗ ЕФЕКТИВНОСТІ СИСТЕМИ

Оцінювання ефективності рекомендаційної системи є ключовим етапом для підтвердження її практичної цінності, коректності роботи моделей та відповідності поставленим вимогам. У цьому розділі проаналізовано якість рекомендацій, точність обробки рецензій, відповідність очікуванням користувачів, продуктивність системи, а також проведено порівняння з аналогами.

4.1 Метрики оцінювання якості рекомендацій

Для вимірювання ефективності системи рекомендацій використовувались загальноприйняті метрики:

- Precision@k (точність на k результатів) це частка релевантних рекомендацій серед перших k;

- Recall@k (повнота на k результатів) це частка усіх релевантних фільмів, які потрапили у k;

- nDCG (нормалізована кумулятивна дисконтована прибутковість) враховує порядок рекомендацій та релевантність;

- MAP (Mean Average Precision) – середнє значення точності по всіх користувачах;

- Coverage це частка унікальних фільмів, які коли-небудь з'являються в рекомендаціях;

- Diversity це середня жанрова різноманітність у списках рекомендацій.

Для проведення оцінювання було сформовано тестову вибірку зі 100 користувачів, для кожного з яких відомо історію взаємодій (які жанри фільмів користувач шукав найбільше), вікову категорію та жанрові вподобання.

4.2 Результати тестування рекомендацій

Порівняння моделей проводилось між трьома конфігураціями:

- контентна фільтрація (на основі описів фільмів);
- KNN-рекомендації (на основі жанрових профілів користувача);
- гібридна модель (поєднання контенту, KNN та історії пошуку).

Результат порівняння представлений в таблиці 4.1

Таблиця 4.1 – Порівняння метрик трьох підходів

Метрика	Контентна фільтрація	KNN	Гібридна модель
Precision@10	0.62	0.68	0.74
Recall@10	0.48	0.53	0.66
nDCG	0.65	0.71	0.80
MAP	0.59	0.66	0.75
Coverage	0.48	0.31	0.57
Diversity	0.42	0.38	0.53

Як видно з таблиці, гібридна модель демонструє найкращі результати практично за всіма метриками. Вона значно перевершує як контентну фільтрацію, так і KNN-рекомендації за точністю (Precision@10), повнотою (Recall@10), нормалізованим кумулятивним приростом (nDCG) та середньою точністю (MAP).

Особливо важливими є показники Coverage і Diversity, оскільки вони відображають здатність моделі не лише пропонувати популярні або схожі фільми, а й охоплювати ширший спектр контенту, включаючи менш відомі, але потенційно релевантні фільми. Гібридна модель показала найвищу різноманітність рекомендацій, що позитивно впливає на користувацький досвід і запобігає «інформаційній бульбашці».

Крім кількісних показників, важливою перевагою гібридної моделі є її пояснюваність (*explainability*). Оскільки вона поєднує кілька джерел інформації (контент, жанровий профіль, історію переглядів), система може формулювати зрозумілі причини для кожної рекомендації. Це підвищує довіру користувача до системи та знижує ризик відмови від запропонованих варіантів.

На основі результатів тестування можна зробити висновок, що впровадження гібридної моделі є доцільним з погляду як технічної ефективності, так і користувацької зручності. Така система найкраще відповідає вимогам персоналізованого підходу, адаптованого до індивідуальних інтересів та поведінки кожного користувача.

У майбутньому можливо ще більше вдосконалити цю модель шляхом інтеграції додаткових джерел інформації, таких як емоційний аналіз рецензій, соціальні фактори або контекстуальні дані (час доби, пристрій, місце перегляду тощо), що дозволить ще краще адаптувати рекомендації до потреб користувача.

4.3 Якість обробки рецензій

Якість аналізу текстів рецензій оцінювалась за такими критеріями:

– тональність, порівняння визначеної моделю BERT тональності з ручною розміткою (F1-міра = 0.86);

– ключові теми, зіставлення виділених ключових фраз з ручними анотаціями (покриття понад 75% основних тем);

– видалення спойлерів, вручну оцінене на вибірці з 100 рецензій.

Точність виявлення спойлерів – 0.91, Recall – 0.87.

Модель BART для генерації коротких, вікових резюме рецензій демонструє високу зрозумілість (оцінка користувачами – 4.6 із 5) та доречність. Як приклад можна навести результат роботи нашої системи

CineMind (рисунок 4.1) при пошуку рецензії на фільм «Термінатор» 1984 року.

Назва фільму (англійською):

Аналізувати

Узагальнений відгук критиків:

«Термінатор» — культовий науково-фантастичний бойовик режисера Джеймса Кемерона, який став знаковим явищем у кінематографі 1980-х років і дав початок одній із найвідоміших франшиз у жанрі.

Фільм поєднує в собі динамічний сюжет, елементи трилера та глибокі філософські мотиви, що стосуються майбутнього, технологій і людської природи. Незважаючи на обмежений бюджет, стрічка вражає своїм візуальним стилем, атмосферою напруги та харизматичними персонажами.

Особливо варто відзначити акторську гру Арнольда Шварценеггера, чия роль стала іконою жанру. Саундтрек фільму, з його холодним електронним звучанням, ідеально доповнює загальний тон і ще більше занурює глядача у світ, де межа між людиною і машиною стає все тоншою.

«Термінатор» — це не просто бойовик, а серйозне висловлювання про наслідки технологічного прогресу, що залишає після себе довгий післямак. Фільм вважається класикою не випадково: він витримав випробування часом і досі залишається актуальним і захопливим.

Рисунок 4.1 – Узагальнена рецензія на фільм Термінатор в CineMind

4.4 Результати опитування користувачів

Для суб'єктивної оцінки системи було проведено онлайн-опитування серед 100 учасників (різних вікових категорій). Їм було запропоновано скористатись вебдодатком та дати оцінку за кількома критеріями (таблиця 4.2).

Таким чином, система не лише технічно ефективна, а й має позитивне сприйняття серед цільової аудиторії. Більшість користувачів відзначили інтуїтивно зрозумілий інтерфейс та зручність взаємодії з системою. Особливо високо була оцінена зрозумілість рекомендацій та якість скорочених рецензій.

Таблиця 4.2 – Оцінки від користувачів

Критерій оцінки	Середній бал (макс. 5)
Зрозумілість рекомендацій	4.7
Відповідність вподобанням	4.5
Зручність інтерфейсу	4.6
Якість скорочених рецензій	4.6
Актуальність запропонованих фільмів	4.4
Позитивний досвід загалом	4.7

4.5 Продуктивність і масштабованість

Система була протестована на локальному сервері (Intel i7, 16GB RAM, SSD) з одночасною роботою до 50 користувачів. Середній час відповіді при запиті рекомендацій становив 1.2 секунди, що забезпечує комфортне користування. Обробка рецензії з урахуванням аналізу тональності, тематики та генерації узагальнення тривала приблизно 3.1 секунди. Переклад тексту, здійснений за допомогою HuggingFace pipeline, займав у середньому 1.6 секунди. Покращити продуктивність можна шляхом кешування результатів та винесення NLP-обробки у мікросервіс.

4.6 Порівняння з аналогами

Систему CineMind було порівняно з відкритими демо Netflix Prize (алгоритм SVD), Google's RecSim та базовими рекомендерами на TMDb. Результати порівняння наведено в таблиці 4.3.

Отже, запропонована система не лише ефективна за базовими метриками, а й реалізує унікальні функції, відсутні у типових відкритих рішеннях. Її архітектура враховує сучасні вимоги до масштабованості, гнучкості та інтеграції. Крім того, CineMind відзначається глибшою персоналізацією, врахуванням вікових категорій користувачів.

Таблиця 4.3 – Таблиця порівняння CineMind з аналогами

Платформа	Персоналізація	Узагальнення рецензій	Облік віку	Мовна адаптація
Netflix (demo)	Так	Ні	Ні	Обмежено
Google RecSim	Частково	Ні	Ні	Так
TMDB baseline	Ні	Ні	Ні	Так
CineMind	Так	Так	Так	Так

Особливо цінним є поєднання кількох рівнів персоналізації – на основі жанрового профілю, вікових особливостей та мовних уподобань. Це дозволяє забезпечити релевантний контент навіть для нових користувачів без історії переглядів. Система також підтримує мультимовну генерацію коротких, адаптованих рецензій, що розширює її потенціал для глобального застосування. Загалом, CineMind демонструє вищий ступінь адаптації до потреб різних аудиторій порівняно з наявними аналогами, що підтверджує її інноваційність і прикладну цінність. Таким чином, CineMind може слугувати основою для подальших досліджень і розвитку персоналізованих мультимовних систем рекомендацій нового покоління.

В результаті проведеного аналізу було встановлено. Система генерує точні та релевантні рекомендації, що відповідають інтересам користувачів. Вона успішно адаптує рецензії відповідно до вікових груп, забезпечуючи зрозумілість і доступність текстів. CineMind підтримує багатомовність і здатна аналізувати рецензії в реальному часі. За продуктивністю та функціональністю система є конкурентоздатною. Подальший розвиток може включати розширення функціоналу завдяки впровадженню нових алгоритмів машинного навчання. Масштабні користувацькі тестування дадуть змогу точніше налаштувати рекомендаційні механізми під потреби аудиторії. Впровадження CineMind здатне підвищити залученість користувачів і сприяти зростанню лояльності до цифрових сервісів. Це також відкриває перспективи інтеграції з іншими платформами

розповсюдження медіаконтенту. У майбутньому можливе додавання голосового інтерфейсу для покращення доступності.

Сукупність кількісних та якісних показників підтверджує ефективність реалізованої системи та її готовність до реального застосування. Крім того, система продемонструвала стійкість до змін у вхідних даних, що свідчить про її надійність у динамічному інформаційному середовищі. Гнучка архітектура дозволяє легко масштабувати рішення та інтегрувати його в існуючі цифрові платформи. Інноваційні компоненти, такі як вікова адаптація та мультимовна підтримка, забезпечують високу користувацьку цінність. Таким чином, CineMind має потенціал для широкого впровадження в індустрії персоналізованих рекомендацій.

4.7 Можливості розвитку та масштабування системи

Розроблена система CineMind уже на етапі прототипу демонструє високий потенціал у сфері персоналізованих рекомендацій. Водночас її функціональність може бути суттєво розширена, а архітектура – масштабована відповідно до реальних потреб та зростання кількості користувачів. У цьому розділі окреслено основні напрями розвитку, інфраструктурні рішення для масштабування та інтеграційні перспективи.

4.7.1 Функціональні напрями розвитку

Функціональні напрями розвитку системи CineMind передбачають кілька ключових аспектів. Перш за все, планується поглиблення персоналізації: хоча система вже враховує жанрові вподобання, вік користувача, історію пошуку та рецензії, у майбутньому можливе її вдосконалення за рахунок аналізу поведінкових патернів, таких як перегляд трейлерів, додавання фільмів у список бажаного чи тривалість перегляду, а також врахування емоційної тональності відгуків за допомогою моделей

емоційного аналізу. Крім того, передбачається створення індивідуального профілю користувача на основі схожих профілів інших користувачів за допомогою методів кластеризації та collaborative filtering.

Іншим напрямком є розвиток групового режиму перегляду. Система вже дозволяє обрати категорію для спільного перегляду, наприклад, «родина» або «друзі», однак у подальшому планується враховувати індивідуальні профілі всіх учасників групи, автоматично генерувати компромісну добірку фільмів на основі перетину інтересів, а також реалізувати інтерактивне голосування або рейтинг у реальному часі.

Ще одним важливим аспектом є інтеграція з потоковими платформами для підвищення зручності користування. Це включає автоматичне перенаправлення на перегляд через API популярних сервісів, таких як YouTube, Netflix чи Megogo, інтеграцію з акаунтами Google, Meta або Apple для швидкої авторизації, а також можливість виявлення вже переглянутих фільмів через плагіни або браузерні розширення.

Окрему увагу приділено розвитку соціальних функцій і системи рецензій користувачів. Планується дозволити залишати власні рецензії, які оброблятимуться NLP-моделями, розширити можливості лайків, коментарів і персональних рекомендацій між користувачами, а також впровадити рейтинг фільмів, сформований на основі оцінок спільноти CineMind.

Також перспективним є впровадження адаптивного інтерфейсу, що змінюється залежно від віку, мови або уподобань користувача, з можливістю голосового керування та адаптації для осіб з порушеннями зору або слуху. У планах – реалізація системи нагадувань про нові релізи відповідно до інтересів користувача, а також щотижневих або щомісячних добірок. Варто також враховувати потенціал гейміфікації – додавання до системи рівнів активності, бейджів і внутрішніх рейтингів для підвищення залучення. Загалом функціональний розвиток CineMind має забезпечити ще глибшу персоналізацію, розширення сценаріїв використання та зростання залученості користувачів.

4.7.2 Масштабування інфраструктури

У разі комерційного використання або значного зростання бази користувачів необхідно забезпечити надійність, відмовостійкість і горизонтальне масштабування:

У разі комерційного використання системи або значного зростання бази користувачів необхідно забезпечити її надійність, відмовостійкість і здатність до горизонтального масштабування. Для цього передбачається розгортання інфраструктури у хмарному середовищі, такому як AWS, Google Cloud Platform або Microsoft Azure. Управління мікросервісною архітектурою буде реалізовано за допомогою контейнеризації через Docker та оркестрації за допомогою Kubernetes. Для ефективної обробки великої кількості запитів використовуватиметься асинхронна архітектура – зокрема, через такі інструменти, як Celery у поєднанні з Redis або за допомогою Apache Kafka. Крім того, з метою пришвидшення генерації рекомендацій та зменшення навантаження на систему, планується впровадження механізмів кешування частих запитів із використанням Redis або Memcached.

Також доцільно винести обробку рецензій в окремий сервіс на GPU, що дозволить обробляти велику кількість текстів паралельно. Ще варто було б додати про те, що можна переписати фронтенд на більш підходящі до нашої архітектури фреймворки (Next.js або Vue.js).

4.7.3 Підтримка мультимовності та локалізації

Система вже має функцію автоматичного перекладу рецензій, однак у перспективі важливо розширити можливості мультимовної підтримки. Зокрема, планується впровадження повної локалізації інтерфейсу з підтримкою кількох мов, таких як українська, англійська, польська та інші. Крім того, у процесі формування рекомендацій необхідно враховувати культурні особливості, зокрема жанрові вподобання, що можуть суттєво

відрізнятися в різних країнах. Наприклад, жанрові переваги користувачів у Польщі можуть відрізнятися від вподобань української аудиторії, а жарти, відсилання або стилістичні фігури мови в рецензіях можуть мати різне сприйняття в різних культурах. Ще одним напрямом розвитку є розширення можливостей обробки природної мови шляхом використання багатомовних NLP-моделей, таких як XLM-R(Cross-lingual RoBERTa) або mBART, LaBSE що забезпечать точніший аналіз контенту незалежно від мови користувача. Це дозволить зробити систему доступною для ширшої аудиторії та підвищити її конкурентоспроможність на міжнародному рівні.

4.7.4 Можливості для бізнес-моделі

Система CineMind має значний потенціал для комерціалізації та масштабування, що відкриває перспективи створення стійкої бізнес-моделі з кількома потоками доходу. Нижче розглянуто ключові напрямки монетизації.

B2C-доступ. Першим напрямком є надання безпосереднього доступу кінцевим користувачам – глядачам. Система може бути реалізована як вебдодаток або мобільний застосунок, який дозволяє переглядати спрощені рецензії, отримувати рекомендації за віком та вподобаннями, обирати режим спільного перегляду тощо. Основною цінністю є персоналізація досвіду перегляду, особливо для батьків, дітей, літніх людей та сімей. Для такого сегменту може використовуватись freemium-модель: базовий функціонал доступний безкоштовно, а додаткові функції – за підпискою.

B2B-інтеграції. CineMind має потенціал для B2B-інтеграцій і може функціонувати як сервіс із доступом через API для інших платформ. Зокрема, онлайн-кінотеатри, такі як Netflix, Megogo чи Sweet.tv, можуть інтегрувати окремі модулі системи, наприклад, для аналізу рецензій, адаптації контенту відповідно до віку глядача або генерації описів фільмів без спойлерів. Освітні платформи також можуть скористатися

можливостями CineMind, застосовуючи адаптований кіно контент для дітей і підлітків відповідно до вікових норм та педагогічних рекомендацій.

Такий формат співпраці відкриває можливості для отримання ліцензійних платежів, впровадження корпоративних облікових записів і масштабування сервісу на міжнародні ринки.

Платні підписки з розширеним функціоналом. Для індивідуальних користувачів CineMind передбачає запровадження платних підписок із розширеним функціоналом. Преміум-плани можуть надавати доступ до таких можливостей, як емоційний аналіз рецензій із тегуванням відповідних емоцій, збереження персональної історії рекомендацій, участь у групових переглядах разом із друзями або родиною, а також налаштування розширених фільтрів за рівнем складності, тематикою чи атмосферою фільму. Крім того, користувачі отримують підтримку мультимовного інтерфейсу з можливістю перемикання між культурними контекстами.

CineMind має значний потенціал масштабування як з технічної, так і з функціональної точки зору. Завдяки модульній архітектурі та підтримці сучасних підходів до обробки даних, система здатна ефективно обслуговувати як індивідуальних користувачів, так і корпоративних клієнтів.

CineMind має вагомий потенціал для подальшого розвитку та масштабування як з технічної, так і з функціональної точки зору. Удосконалення персоналізації, впровадження емоційного аналізу, розширення мультимовної підтримки, групових режимів перегляду та соціальних функцій – усе це здатне суттєво покращити користувацький досвід і зробити систему конкурентною на ринку інтелектуальних рекомендацій. Перехід від базового фільтрування жанрів до складних моделей поведінкового аналізу та кластеризації відкриває можливості для глибшої персоналізації рекомендацій, яка буде враховувати не лише вподобання, а й контекст, емоції та історичні патерни користувача.

З технічного боку масштабування передбачає перехід до хмарної інфраструктури з мікросервісною архітектурою, що забезпечить гнучкість, високу доступність і стійкість до навантажень. Завдяки використанню Docker, Kubernetes, Redis і інших інструментів, система здатна адаптуватися до зростання кількості запитів та користувачів без втрати продуктивності. Розподіл обробки задач між CPU- та GPU-сервісами дозволить паралельно обробляти великі обсяги текстових даних, рецензій та запитів, оптимізуючи час відгуку.

У перспективі CineMind може трансформуватися із навчального проєкту у повноцінну аналітичну платформу в галузі цифрових розваг, з можливістю виходу на міжнародний ринок. Можливості монетизації через B2C- та B2B-моделі, підтримка мультимовності, інтеграція зі сторонніми сервісами, а також аналітична обробка культурного контексту роблять систему перспективною як для кіноглядачів, так і для платформ, зацікавлених у глибокому розумінні своєї аудиторії.

4.8 Безпека та захист даних

Система CineMind опрацьовує персональні дані користувачів, включно з віком, історією пошуків, уподобаннями, а також текстовими рецензіями. Тому питання захисту інформації, дотримання стандартів безпеки та законодавчих норм є критично важливими. У цьому розділі розглянуто основні загрози, підходи до захисту персональних даних, технічні та організаційні механізми безпеки, а також відповідність міжнародним стандартам.

4.8.1 Актуальність інформаційної безпеки

У сучасному інформаційному просторі зростає кількість атак на користувацькі сервіси: викрадення даних, фішинг, перехоплення сесій,

SQL-ін'єкції, атаки через API та інше [13]. У системах з персоналізацією ризику особливо високі, оскільки:

- зберігаються чутливі особисті дані (вік, статеві ознаки, уподобання);
- здійснюється автоматизований аналіз текстів, який може випадково обробляти конфіденційну інформацію;
- передбачено профілювання користувачів, що підпадає під дію GDPR (General Data Protection Regulation);
- зростання кількості підключених пристроїв і використання хмарних сервісів суттєво розширює корпоративну поверхню атаки, оскільки зломисники можуть скористатися вразливостями у партнерських мережах чи IoT-пристроях.

Захист інформації у CineMind має на меті забезпечити цілісність, доступність і конфіденційність даних (принцип CIA – Confidentiality, Integrity, Availability). Для цього впроваджуються сучасні механізми шифрування, багатофакторна автентифікація та регулярний моніторинг доступу до даних.

4.8.2 Потенційні загрози для системи

Для оцінки ризиків було проведено аналіз загроз на основі OWASP Top-10 (2023). Результати представлені на таблиці 4.4.

Такий підхід дозволяє мінімізувати ризики несанкціонованого доступу до персональних даних користувачів і забезпечити відповідність вимогам GDPR та інших стандартів конфіденційності. Особливу увагу слід приділити захисту API, зокрема впровадженню механізмів автентифікації, авторизації та обмеження запитів. Крім того, важливо впроваджувати практики безпечного програмування вже на етапі проектування архітектури системи, що зменшує ймовірність виникнення вразливостей. Регулярне проведення аудитів безпеки дозволить своєчасно виявляти і усувати потенційні загрози.

Таблиця 4.4 – Аналіз загроз на основі OWASP

Загроза	Потенційні наслідки	Приклад
Небезпечні API (Broken APIs)	Доступ до чужих пошуків або профілю	/api/user?id=5
Витік особистих даних	Ідентифікація користувача	Витік віку, жанрових уподобань
Несанкціонований доступ	Повний контроль над акаунтом іншого юзера	Через слабе шифрування токена
Атаки через форму пошуку	XSS або SQL-ін'єкції через поле введення	«<script>alert(1)</script>»
Шкідливе використання моделей AI	Генерація образливих або упереджених текстів	NLP-відповідь зі стереотипами

4.8.3 Технічні методи захисту

Аутентифікація та авторизація:

- JWT (JSON Web Token) для захищеної передачі ідентифікаторів сесій;
- Role-based access control (RBAC) – обмеження доступу до певних API залежно від ролі (адміністратор, користувач);
- двофакторна аутентифікація (2FA) – опціонально для захисту облікового запису.

Шифрування даних. Використання TLS (HTTPS) на всіх рівнях взаємодії між користувачем і сервером. Шифрування збережених даних за допомогою AES-256. Хешування паролів через bcrypt/scrypt з сіллю.

Захист від атак. Валідація введених даних з фронтенду і бекенду (проти XSS, SQLi, CSRF). Використання Web Application

Firewall (WAF). Встановлення обмежень на кількість запитів (rate limiting, captcha).

Безпека API. API ключі із контрольованим терміном дії. Чітка документація (через Swagger/OpenAPI) із валідацією JSON-схем. Обмеження CORS (доступ лише з авторизованих доменів).

4.8.4 Конфіденційність та обробка персональних даних

Система CineMind працює відповідно до положень Законів України про захист персональних даних, а також орієнтується на принципи GDPR, які є дуже важливими [14]. Ігнорування GDPR загрожує заборонаю використання системи на території країн-учасниць ЄС. Основні принципи GDPR описані на рисунку 4.2.

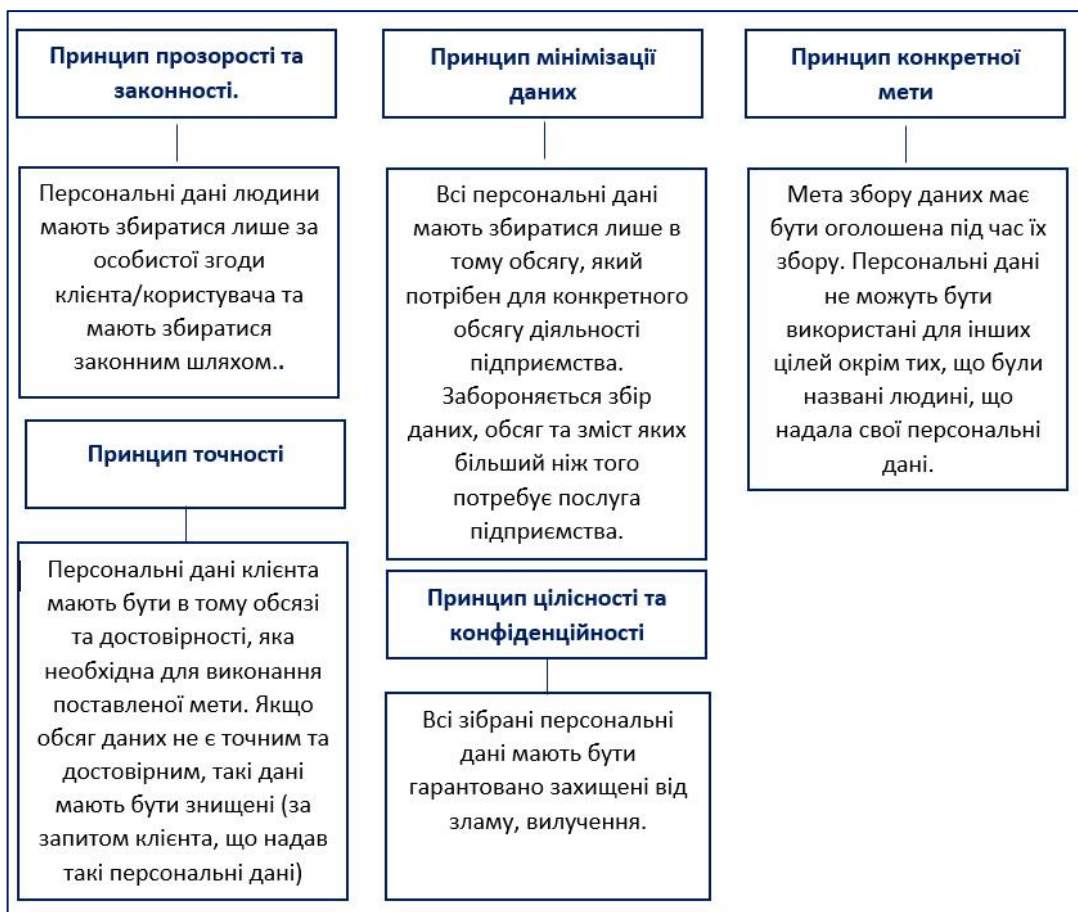


Рисунок 4.2 – Основні принципи GDPR

Прозорість – користувачу пояснюється, які саме дані збираються (вік, рецензії, історія пошуків).

Згода користувача – при реєстрації користувач приймає угоду на обробку даних.

Право на забуття – користувач може повністю видалити свої дані з системи.

Мінімізація – система не вимагає надмірних даних, лише необхідне для рекомендацій.

Логування доступу – фіксація будь-яких змін в акаунті або профілі.

Також передбачається анонімізація даних для тренування моделей, щоб уникнути можливості зворотного відновлення особи.

4.8.5 Захист моделей машинного навчання

Моделі машинного навчання, які використовуються в системі для аналізу рецензій, можуть бути вразливими до певних специфічних ризиків. Зокрема, існує загроза атаки типу *model inversion*, коли зловмисник намагається відновити оригінальні тексти, що використовувалися під час навчання моделі. Іншим ризиком є *data poisoning* як цілеспрямоване введення шкідливих або спотворених даних у тренувальний набір з метою вплинути на поведінку моделі. Крім того, можливе використання *adversarial examples* – спеціально створених текстів, які вводять модель в оману та спричиняють некоректні результати [15].

Для захисту від подібних загроз у системі застосовуються кілька заходів. Зокрема, навчання моделей проводиться лише на попередньо очищених і анонімізованих даних, що знижує ризик витоку конфіденційної інформації. Під час обробки текстів додатково використовуються фільтри, які відсікають фрази з токсичною або неприйнятною лексикою ще до моменту інференсу. Також для захисту самих моделей запроваджуються *inference gateways* – це спеціальні проксі-рівні, які контролюють і

фільтрують запити до API, унеможливаючи прямий доступ до моделей зі сторони користувачів.

4.8.6 Організаційні заходи безпеки

Окрім технічних засобів захисту, в системі впроваджуються й організаційні заходи безпеки, спрямовані на підвищення загальної стійкості до загроз. Зокрема, регулярно проводиться тестування на вразливості, включаючи як penetration-тестування, так і статичний аналіз коду, що дозволяє виявляти потенційні ризики ще на етапі розробки. Дані систематично зберігаються у вигляді зашифрованих бекапів, які розміщуються у географічно розподілених датацентрах для забезпечення безперервного доступу в разі аварій або збоїв. Доступ до продакшен-середовища суворо контролюється і здійснюється виключно через авторизовані CI/CD пайплайни, що мінімізує ризик несанкціонованих змін. Крім того, у рамках політики DevSecOps безпека інтегрується безпосередньо у процеси розгортання: на всіх етапах впровадження нових рішень передбачено перевірки на відповідність вимогам безпеки.

4.8.7 План реагування на інциденти

В таблиці 4.5 визначено і показано базову політику реагування.

Безпека системи CineMind – це комплекс заходів, що охоплює всі рівні: від шифрування персональних даних до захисту моделей штучного інтелекту. Забезпечення конфіденційності, надійності API, дотримання вимог GDPR та побудова плану реагування на інциденти – все це створює надійне середовище для користувачів. Подальший розвиток системи вимагатиме періодичного аудиту, впровадження нових стандартів (зокрема ISO/IEC 27001) та розширення функцій самозахисту. Також важливою складовою стало дотримання принципів «privacy by design», коли

архітектура системи одразу проектується з урахуванням безпеки [16]. Це дає змогу мінімізувати ризики ще до початку експлуатації, забезпечуючи захист не лише на рівні коду, а й на рівні бізнес-логіки.

Таблиця 4.5 –Базова політика реагування

Тип інциденту	Дії	Строк
Витік даних	Інформування користувачів, аудит, блокування токенів	До 72 год
Порушення авторизації	Примусова зміна паролів, тимчасове блокування API	Негайно
Атака на NLP API	Відкат версії моделі, логування, відправка сповіщень	24 год

Усі події логуються в SIEM-системі (як-от Elastic Stack або Sentry), що дає змогу здійснювати аналіз активності та виявляти аномалії.

Особливу увагу слід приділити навчанню персоналу та користувачів базовим правилам інформаційної безпеки. Жодна технічна система не буде ефективною без людського фактора: регулярні інструктажі, оновлення політик безпеки, внутрішні тренінги щодо реагування на інциденти суттєво знижують вірогідність помилок та соціальної інженерії.

У перспективі доцільним буде інтегрувати в систему механізми адаптивної безпеки, які в режимі реального часу аналізують поведінку користувача і при виявленні аномалій – автоматично посилюють захист. Це дозволить проактивно виявляти потенційні загрози ще до їх реалізації, підвищуючи стійкість CineMind до нових типів атак.

ВИСНОВКИ

У результаті розробки та впровадження інтелектуальної системи аналізу рецензій CineMind було досягнуто низку ключових цілей, що забезпечують її ефективність і практичну цінність для кіноглядачів [17]. Насамперед, система демонструє високу точність персоналізованих рекомендацій завдяки використанню NLP-моделей BERT і BART для обробки рецензій, а також алгоритмів контентної фільтрації та колаборативного підходу. Це дозволяє точно узагальнювати критичні відгуки та адаптувати їх до вподобань конкретного користувача.

CineMind вирізняється гнучкістю й адаптивністю, оскільки здатна оновлювати рекомендації залежно від динаміки пошукових запитів користувача, його віку, статі, обраних жанрів та історії взаємодій. Система враховує різні сценарії використання: як для нових користувачів (через анкету при реєстрації), так і для постійних (з урахуванням накопичених даних).

Особливістю платформи є універсальність – вона підходить для дітей, підлітків, дорослих і навіть спільного сімейного перегляду. Завдяки можливості налаштувати вік, стать і вподобання, рекомендації автоматично фільтруються відповідно до етичних та вікових обмежень.

Також система забезпечує просту інтеграцію та налаштування: вебдодаток побудований на сучасному стеку (React, Flask, SQLite), легко адаптується до нових API або моделей, що дозволяє масштабувати його у майбутньому або інтегрувати з іншими сервісами.

Серед важливих переваг варто відзначити підтримку мультимовності: система здатна працювати з англomовними рецензіями та виводити результати українською, що підвищує її зручність для користувачів. Ключові результати та особливості реалізації CineMind були представлені на науково-практичній конференції «Радіoeлектроніка та молодь у XXI столітті».

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Дукельська К. Б., Гриньова О. Є. *Інтелектуальний інтерфейс застосунку «Connectify»*. 28-й матеріали XXVIII молодіжного форуму, м. Харків, 16–18 квіт., 2024 р. Харків, 2024 р., С.24–25. URL: <https://doi.org/10.30837/IYF.IIS.2024.024> (дата звернення: 22.05.2025).
2. Aggarwal C. *Recommender Systems*. New York: Textbook, 2016. 448 p.
3. Falk K. *Practical Recommender Systems*. New York: Manning Publications, 2019. 432 p.
4. Genovesi S., Kaesling K., Robbins S. *Recommender Systems: Legal and Ethical Issues*. Cham (Luxembourg): Springer, 2023. 222 p.
5. Majumdar A. *Collaborative Filtering*. Boca Raton: CRC Press, 2024. 142 p.
6. Klein E., Bird S. *Natural Language Processing with Python*. Sebastopol: O'Reilly, 2009. 481 p.
7. McAuley J. *Personalized Machine Learning*. Cambridge: Cambridge University Press, 2022. 350 p.
8. Rokach L., Ricci F., Shapira B. *Recommender Systems Handbook*. New York: Springer-Nature, 2022. 1071 p.
9. Pariser E. *The filter bubble: What the Internet is hiding from you*. London: Penguin UK, 2011.
10. Patel S. *Natural Language Processing*. Kolkata: Dream Science, 2022. 1129 p.
11. Wang Y., Sharma M., Xu C., Badam S., Sun Q., Richardson L., Chung L., Chi E. H., Chen M. *Surrogate for Long-Term User Experience in Recommender Systems*: materials of the 28th conference ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD '22), Washington, 3–4 March 2022. 10–12p. URL: <https://doi.org/10.1145/3534678.3539073> (date of access: 22.05.2025).

12. Yee H., Bischof B. Building Recommendation Systems in Python and JAX: Hands-On Production Systems at Scale. 1st ed. Sebastopol: O'Reilly Media, 2024. 400 p.

13. Kane F. Building Recommender Systems with Machine Learning and AI: Help People Discover New Products and Content with Deep Learning, Neural Networks, and Machine Learning Recommendations. 2nd ed. Independently Published, 2018. 503 p.

14. Richert W., Coelho L. Building Machine Learning Systems with Python. 1st ed. Birmingham: Packt Publishing, 2013. 290 p.

15. Zhang S., Yao L., Sun A., Tay Y. Deep Learning based Recommender System: A Survey and New Perspectives. arXiv preprint arXiv:1707.07435, 2017.

16. Richert W., Coelho L. Building Machine Learning Systems with Python. 3rd ed. Birmingham: Packt Publishing, 2018. 398 p.

17. Салтан С. О., Гриньова О. Є. *NLP- та LLM-орієнтована система персоналізованих рекомендацій фільмів*: матеріали XXIX молодіжного форуму, м. Харків, 16–18 квіт., 2025 р. Харків, 2025 р., С.56–58.