

ОСОБЛИВОСТІ СИНЕРГІЇ СУЧАСНИХ ТЕХНОЛОГІЙ У СТВОРЕННІ ІНТЕЛЕКТУАЛЬНИХ СИСТЕМ ЕЛЕКТРОННОЇ КОМЕРЦІЇ

Наумов А.В., Сердюк Н.М.

Харківський національний університет
радіоелектроніки
Кафедра Комп'ютерних Інтелектуальних Технологій
та Систем, Україна.

E-mail: artem.naumov@nure.ua

Abstract

This research analyzes the architectural principles and synergistic interaction of modern technologies for building effective, intelligent e-commerce systems. It proposes a specific hybrid architecture combining Kotlin with Jetpack Compose for the client-side, Supabase as a reliable SQL backend, and LiteRT for on-device AI integration.

The scientific novelty consists in a unique hybrid deployment architecture for a Two-Tower model. Unlike traditional server-side approaches, the User Tower is moved to the client device for local inference. Consequently, the backend role shifts from a compute node to a high-speed vector search database storing pre-calculated Item Tower vectors. This design creates two distinct data flows: a real-time, privacy-preserving recommendation flow using anonymized vectors, and a separate asynchronous flow for model retraining. The results are applicable for prototyping e-commerce services and developing modern AI-driven recommendation systems.

У сучасному світі важко уявити життя без електронних пристроїв і онлайн-сервісів, що пронизують майже всі сфери діяльності людини – від купівлі товарів до бронювання квитків і підписок на цифрові платформи. Електронна комерція, яка бере початок у 1960-х роках і активно розвивається з 1990-х, сьогодні стала невіддільною частиною глобальної економіки [1]. З появою генеративного штучного інтелекту та персоналізованих рекомендаційних систем e-commerce переживає черговий етап еволюції. Це створює потребу в ефективних архітектурних рішеннях і розумних технологічних платформах нового покоління, здатних забезпечити швидкість, гнучкість і персоналізацію цифрових сервісів.

Мета дослідження полягає в аналізі архітектурних принципів, переваг та синергетичної взаємодії сучасних технологій для побудови ефективних, масштабованих та інтелектуальних цифрових сервісів у сфері e-commerce. Дослідження зосереджується на обґрунтуванні гібридної архітектури, у межах якої Kotlin з Jetpack Compose формують клієнтську частину, Supabase виступає надійним SQL-бекендом з REST API, а LiteRT використовується для інтеграції on-device AI-компонентів як ключової складової рекомендаційної системи.

Об'єктом дослідження є сучасні технології створення веб- та мобільних застосунків. Предметом дослідження виступають методи інтеграції мов програмування, фреймворків і платформ у єдину архітектуру інтелектуальних систем.

Для глибшого розуміння принципів такої інтеграції доцільно розглянути набір технологій, які у цьому дослідженні використовуються для побудови інтелектуальної архітектури. Нижче описано роль кожного компонента та особливості їхньої взаємодії:

Клієнтська частина формує ядро взаємодії з користувачем. Kotlin виступає як основна мова розробки, пропонуючи потужний набір сучасних можливостей. Ключовими перевагами є його повна сумісність з екосистемою Java, виразний та лаконічний синтаксис і сучасна система типів, що включає вбудовану безпеку щодо null (null-safety). Це суттєво підвищує надійність коду та мінімізує помилки. Окрім того, для ефективної обробки асинхронних операцій можна використовувати Kotlin Coroutines, що дозволяє писати складний асинхронний код у значно простішому та читабельному стилі.

Для побудови інтерфейсу використовується Jetpack Compose – сучасний декларативний інструментарій, який дозволяє описувати UI як функції які автоматично реагують на зміну даних. Це ідеально підходить для архітектури, оскільки інтерфейс, наприклад, список рекомендацій, автоматично оновлюється через механізм recomposition щоразу, коли з'являється новий персоніфікований список.

В якості серверної частини пропонується Supabase – відкрита платформа, що позиціонується як альтернатива Firebase. На відміну від Firebase, що використовує – документо-орієнтовану NoSQL-модель, Supabase побудований на базі PostgreSQL. Це надає критичні переваги реляційної SQL-парадигми, яка є більш придатною для структурованих e-commerce даних, забезпечуючи надійність через транзакції та підтримку складних зв'язків. Ключовою перевагою для інтеграції є те, що Supabase автоматично генерує REST API поверх таблиць. Це усуває необхідність у розробці окремого серверного прошарку для базових операцій, дозволяючи Kotlin-застосунку напряду надсилати вектори та отримувати списки рекомендованих товарів.

Інтелектуальна складова базується на гібридній рекомендаційній системі, робота якої розділена на два етапи:

Перший етап це – тренування на сервері. В основі архітектури лежить модель Two-Tower, що тренується заздалегідь та реалізує гібридний підхід, поєднуючи колаборативну фільтрацію з контентними ознаками [2-3]. Для цього використовується Python, який, наразі, є стандартом для Data Science та Machine Learning завдяки бібліотекам, таким як TensorFlow та PyTorch. Натренована Item Tower використовується для створення векторів товарів, які завантажуються у Supabase, тоді як User Tower готується до розгортання на пристрої.

Другий етап – виконання на пристрої (On-Device Inference). Натренована модель User Tower конвертується у ефективний портативний формат .tflite і запускається через LiteRT безпосередньо на пристрої користувача. Таке рішення, знаходить ефективний компроміс із ключовими обмеженнями традиційних підходів: воно вирішує проблеми конфіденційності та розміру, оптимізує затримку, але вимагає підключення до мережі для фінального пошуку [4].

Отже, новизна запропонованого рішення полягає у розробленій гібридній архітектурі, що інтегрує локальне виконання та хмарний векторний пошук [5]. На відміну від традиційних серверних підходів, де обчислення вектора користувача відбувається на сервері, дане рішення пропонує перенести його безпосередньо на пристрій. При цьому серверна частина змінює свою роль: замість того, щоб бути «обчислювальним» вузлом, вона перетворюється на високошвидкісну «пошукову» базу даних, яка зберігає заздалегідь обчислені вектори Item Tower.

Такий розподіл створює архітектуру з двома чіткими та незалежними потоками даних. Перший – це миттєвий потік рекомендацій, що реалізується через обмін знеособленими векторами. Другий – це асинхронний потік для навчання, в якому «сира» історія кліків окремо збирається, анонімізується і відправляється на сервер виключно для майбутнього перенавчання моделі.

Результати дослідження можуть бути використані для розробки навчальних курсів з сучасної full-stack та мобільної розробки. Бути основою для швидкого створення прототипів (MVP) e-commerce-сервісів з інтелектуальною архітектурою. Надані в роботі рекомендації можуть бути впроваджені для створення гібридних та контентно-орієнтованих AI-рекомендаційних систем у комерційних продуктах.

Література

1. Kenneth L., Carol T. E-Commerce 2021-2022: Business. Technology. Society. 17th ed. Pearson Education, 2021. 912 p.
2. Lee W.-M., Cho Y.-S. A Flexible Two-Tower Model for Item Cold-Start Recommendation. IEEE Access. 2023. Vol. 11. P. 146194–146207. doi.org/10.1109/access.2023.3346918
3. Francesco R., Lior R., Bracha. S. Recommender Systems Handbook 3rd ed. Springer US, 2022. 1071 p.
4. LiteRT overview | Google AI Edge | Google AI for Developers. Google AI for Developers. URL: <https://ai.google.dev/edge/litert>
5. Mark R., Neal F. Fundamentals of Software Architecture: An Engineering Approach. O'Reilly Media, 2020. 432 p.