

РЕЦЕНЗІЯ

рецензента – доктора технічних наук, професора, директора навчально-наукового інституту інформаційних технологій Державного університету інформаційно-комунікаційних технологій Нестеренко Катерини Сергіївни на дисертаційну роботу Коротіна Дениса Сергійовича на тему: «Моделі та методи удосконалення DSP-платформи для персоналізованої реклами», подану на здобуття ступеня доктора філософії в галузі знань 12 - Інформаційні технології за спеціальністю 123 - Комп'ютерна інженерія

Актуальність обраної теми.

Функціонування сучасних DSP-платформ відбувається у жорстких умовах мілісекундних затримок протоколів Real-Time Bidding. Необхідність обробки десятків тисяч запитів за секунду робить інтеграцію ресурсоемних мультимодальних генеративних моделей вкрай складною інженерною задачею. Відсутність інтероперабельності між аналітичним ядром системи торгів та модулями генерації контенту призводить до неефективного витрачання обчислювальних ресурсів та рекламних бюджетів. З огляду на це, дисертаційне дослідження Коротіна Д.С., що пропонує цілісну мікросервісну архітектуру та методи зниження латентності у генеративних DSP-системах, є надзвичайно актуальним завданням для сучасної комп'ютерної інженерії.

Обґрунтованість наукових положень, висновків і рекомендацій дисертації

Теоретичні висновки та розроблені моделі ґрунтуються на фундаментальних засадах теорії систем, проєктування баз даних, машинного навчання та архітектури розподілених високонавантажених платформ. Обґрунтованість підтверджується створенням реального програмного прототипу з використанням Kafka, Redis, Docker та Kubernetes, що довело життєздатність розробленої архітектури. Експериментальні дослідження на масивах реальних логів транзакцій дозволили достовірно підтвердити зниження латентності інференсу та обчислювальної вартості.

Оцінка новизни наукових результатів дисертаційного дослідження

Наукова новизна результатів дисертаційного дослідження полягає у наступному:

- вперше розроблено метод інтеграції мультимодальних генеративних моделей, який, завдяки впровадженню кешування, потокової обробки та керуванню ітераціями генерації, вирішив проблему високої латентності у RTB-середовищі;
- вперше запропоновано адаптивну модель таргетингу на базі RLHF, де формалізована багатокритеріальна оптимізаційна задача враховує не лише поведінкові реакції, а й вартість інференсу генеративної моделі;
- отримала подальший розвиток математична модель системної ефективності DSP, яка об'єднала когнітивні параметри, економічні

показники та обчислювальну складність генерації у вигляді єдиної інтегральної цільової функції.

Практична цінність отриманих результатів

Створено архітектуру інтегрованої DSP-платформи, яка забезпечує асинхронну мікросервісну взаємодію між модулями оцінки та генерації контенту. Впровадження запропонованих рішень на практиці дозволило знизити середню затримку відповіді на 45% (з 210 мс до 115 мс) порівняно з базовими моделями, скоротити час генерації креативів на 56% та зменшити обчислювальні витрати GPU без втрати якості цільового рекламного контенту.

Зв'язок роботи з науковими програмами, планами, темами

Дисертаційна робота тісно пов'язана з практичними запитами ІТ-ринку та виконана відповідно до плану держбюджетних і госпдоговірних досліджень Державного університету інформаційно-комунікаційних технологій, зокрема в рамках НДР «Комплексна розробка прикладних ІТ-рішень для підвищення продуктивності комп'ютерних систем у комерційному та соціальному секторі» (№ ДР 0125U003178).

Повнота викладу основних результатів дисертації в публікаціях

Результати дослідження повноцінно висвітлені у 9 фахових публікаціях автора. З них 6 статей надруковано у наукових фахових виданнях України та базах Scopus, а ключові архітектурні концепції оприлюднено на 3 всеукраїнських та міжнародних конференціях.

Оцінка змісту дисертації, відповідність встановленим вимогам щодо оформлення

Дисертація Коротіна Д.С. обсягом 175 сторінок вирізняється чіткою інженерною логікою викладення матеріалу. Анотація, зміст, рисунки та список використаних джерел оформлені бездоганно і цілком відповідають Порядку присудження ступеня доктора філософії (Постанова КМУ від 12 січня 2022 р. № 44).

Зауваження до проведеного дисертаційного дослідження

При загальній високій оцінці роботи є деякі зауваження дискусійного характеру:

1. В архітектурі системи (розділ 3.1) описано механізм відмовостійкості та використання кешу Redis, проте не наведено детального аналізу масштабування пам'яті при експоненційному зростанні кількості унікальних згенерованих промптів для персоналізованих користувачів.
2. Під час моделювання багатокритеріальної задачі прийняття рішень (формула 2.6) штраф за порушення політик безпеки використовується як лінійний доданок. У реальних високонавантажених системах блокування контенту з високим ризиком часто реалізується як жорсткий фільтр-відсікач, а не вагова компонента оптимізації.

3. Було б доцільно ширше описати апаратні характеристики GPU-серверів, на яких виконувався інференс дифузійних моделей, для точнішого розуміння показників енергоефективності.

Висновок

Дисертаційна робота Коротіна Д.С. є завершеним самостійним науковим дослідженням, яке вирішує актуальне науково-прикладне завдання підвищення ефективності архітектури DSP-платформ за рахунок інтеграції генеративних моделей.

За рівнем наукової новизни, достовірністю отриманих результатів та якістю викладення матеріалу дисертація Коротіна Д.С. повністю відповідає паспорту спеціальності 123 - Комп'ютерна інженерія та вимогам п. 6-9 «Порядку присудження ступеня доктора філософії...», затвердженого постановою Кабінету Міністрів України від 12 січня 2022 р. № 44. Автор роботи, Коротін Денис Сергійович, безумовно заслуговує на присудження наукового ступеня доктора філософії.

Рецензент

Доктор технічних наук, професор,
директор навчально-наукового
інституту інформаційних технологій
Державного університету
інформаційно-комунікаційних технологій

Гресу

Катерина НЕСТЕРЕНКО

Підпис

К. Нестеренко

ЗАСВІДЧУЮ

Учений секретар

Державного університету

інформаційно-комунікаційних технологій

