

ВІДГУК

офіційного опонента,

кандидата технічних наук, доцента,

завідувача кафедри математичного моделювання та аналізу даних

Харківського національного університету імені В. Н. Каразіна

Струкова Володимира Михайловича

на дисертаційну роботу Гребенюка Микити Олександровича

на тему «Дослідження та розробка методів рекомендаційної системи

з використанням алгоритмів нечіткої логіки та генетичного алгоритму»,

подану на здобуття наукового ступеня доктора філософії

в галузі знань 12 – «Інформаційні технології»

за спеціальністю 122 – «Комп'ютерні науки»

Актуальність теми дослідження.

На сьогодні персоналізовані рекомендаційні системи є ключовим інструментом подолання інформаційного перевантаження у медіасервісах, електронній комерції та освітніх платформах. Домінуючий клас методів — колаборативна фільтрація — має фундаментальне структурне обмеження: розрідженість матриці оцінок, за якої обчислення стандартних мір подібності для більшості пар користувачів стає неможливим або статистично ненадійним. Зазначена проблема загострюється в умовах холодного старту, коли для нового користувача наявні лише поодинокі оцінки, а також внаслідок статичності параметрів моделей, що задаються експертно і не адаптуються до індивідуальних когнітивних особливостей користувачів. Вирішення цих задач потребує розробки гібридних підходів, які поєднують нечітке подання різномірних атрибутів профілю користувача з еволюційною оптимізацією структурних та параметричних характеристик моделі. Виходячи з зазначеного, можна зробити висновок про актуальність теми дисертаційної роботи Гребенюка М. О.

Зв'язок роботи з науковими програмами, планами, темами.

Дисертаційну роботу виконано на кафедрі комп'ютерного моделювання та інтелектуальних технологій Харківського національного університету радіоелектроніки відповідно до плану науково-дослідних робіт університету у період з 2022 р. по 2026 р.

Обґрунтованість і достовірність наукових результатів, висновків і рекомендацій.

Аналіз змісту дисертаційної роботи дає змогу зробити висновок про належну обґрунтованість та достовірність наукових положень дослідження. Наукові положення та висновки, представлені у дисертації, обґрунтовуються коректною постановкою і подальшим вирішенням завдань дослідження, використанням відомих наукових методів теорії нечітких множин, еволюційних обчислень та колаборативної фільтрації, застосуванням процедури багаторазової валідації Монте-Карло з ізольованою тестовою вибіркою та статистичною верифікацією отриманих переваг непараметричним критерієм знакових рангів Вілкоксона. Достовірність підтверджується також практичним впровадженням результатів у виробничу практику. Результати дослідження аналізувались та доповнювались під час рецензування наукових публікацій, а також під час апробації на міжнародних та всеукраїнських наукових конференціях.

Основні наукові результати дисертаційної роботи.

Нові наукові результати дисертаційного дослідження опубліковано в 10 наукових працях, у тому числі у 3 статтях у наукових фахових виданнях України категорії «Б» та 2 статтях у збірниках матеріалів міжнародних конференцій, що індексуються наукометричними базами Scopus / Web of Science.

У дисертаційній роботі отримано такі нові результати.

1. Вперше запропоновано метод коеволюційного налаштування гетерогенної хромосоми, який, на відміну від послідовного налаштування та від відомих коеволюційних схем, що поєднують функції належності з нечіткими правилами,

забезпечує одночасну структурно-параметричну адаптацію вагових коефіцієнтів ознак і параметрів функцій належності нечітких множин у компактній гібридній моделі користувача зі збереженням семантичної цілісності, що підвищує якість ранжування рекомендацій (NDCG@10, Precision@10, Recall@10) та стійкість в умовах холодного старту.

2. Удосконалено показник зацікавленості категорією у компактній гібридній моделі користувача шляхом введення штрафного коефіцієнта негативного досвіду, який, на відміну від базової формули, враховує кількість негативно оцінених об'єктів категорії і знижує значення індикатора пропорційно частці негативного досвіду, що усуває семантичну неоднозначність між станами «байдужість» та «антипатія» до категорії та забезпечує точніше формування категорійного профілю в умовах розрідженості матриці оцінок.

3. Удосконалено метрику нечіткої відстані між гібридними профілями користувачів шляхом заміни масштабно-залежного множника на нормалізуючий коефіцієнт, що, на відміну від базової реалізації, забезпечує масштабну інваріантність метрики та коректну агрегацію різнорідних атрибутів у єдиному діапазоні значень, що дозволяє формувати множину сусідів навіть за повної відсутності спільно оцінених об'єктів та суттєво розширює покриття каталогу порівняно з класичними мірами подібності.

4. Удосконалено метод генетичної оптимізації вагових коефіцієнтів гібридної моделі шляхом розроблення композитної функції придатності з регуляризацією, що, на відміну від оптимізації виключно за середньою абсолютною похибкою, включає метрику ранжування безпосередньо до цільового функціоналу та містить штраф за відхилення ваг від рівноважних значень, що дозволяє одночасно знизити похибку прогнозу та підвищити якість ранжування рекомендацій порівняно з рівнозваженою базовою моделлю.

Практичне значення отриманих результатів.

Отримані в дисертаційній роботі нові наукові результати використано при розробці програмного модуля мовою Python, який реалізує запропоновані методи та моделі з використанням відкритих бібліотек і допускає інтеграцію в розподілені архітектури рекомендаційних сервісів. Результати дисертаційного дослідження впроваджено у виробничу практику компанії EFISCO Sp. z o.o. (Республіка Польща) у складі перевикористовного програмного компонента рекомендаційної підсистеми, що підтверджується актом впровадження. Таким чином, практичне значення отриманих результатів підтверджується вирішенням задачі побудови рекомендаційних систем, стійких до розрідженості матриці оцінок та проблеми холодного старту.

Оцінка змісту дисертаційної роботи.

Дисертація складається з анотації, вступу, чотирьох розділів, висновків, списку використаних джерел (136 найменувань) та двох додатків. Загальний обсяг роботи становить 196 сторінок, у тому числі 156 сторінок основного тексту, 44 рисунки, 33 таблиці.

У вступі обґрунтовано актуальність теми роботи, наведено мету, об'єкт, предмет і завдання дослідження, наукову новизну отриманих результатів, повноту їх опублікування й апробацію, а також відображено зв'язок дисертаційної роботи з науковими програмами та темами.

Перший розділ дослідження присвячено аналізу сучасного стану та напрямків досліджень у сфері персоналізованих рекомендаційних систем. Виділено три центральні виклики: розрідженість матриці оцінок, холодний старт та статичність моделі користувача. Виконано порівняльний аналіз методів колаборативної фільтрації, гібридних архітектур на основі компактної гібридної моделі користувача, а також апарату нечіткої логіки та еволюційної оптимізації. Сформульовано постановку задачі дослідження.

Другий розділ присвячено формалізації гібридної моделі користувача. Розроблено удосконалений індикатор зацікавленості категорією зі штрафним

коефіцієнтом негативного досвіду, описано процедуру фазифікації демографічних та категорійних атрибутів, а також удосконалену масштабно-інваріантну метрику нечіткої відстані між гібридними профілями, для якої доведено виконання критичних аксіом коректності.

Третій розділ присвячено розробці методів еволюційної оптимізації параметрів гібридної моделі. Запропоновано метод оптимізації вагових коефіцієнтів з композитною функцією придатності та регуляризацією, обґрунтовано обмеження часткової оптимізації і розроблено коеволюційний генетичний алгоритм з гетерогенною хромосомою, що поєднує вагові коефіцієнти та параметри функцій належності в єдиному еволюційному циклі.

Четвертий розділ дисертації присвячено експериментальній перевірці отриманих у попередніх розділах теоретичних результатів на наборі даних MovieLens 100K. Виконано порівняння нечіткої відстані з класичними мірами подібності у режимі п'ятикратної крос-валідації Монте-Карло, оцінено індивідуальний внесок кожного з розроблених алгоритмів, проведено багатокористувацьке порівняння з індустріальними алгоритмами та перевірку стійкості в умовах симульованого холодного старту. Статистичну значущість переваг підтверджено критерієм Вілкоксона.

Висновки містять узагальнення отриманих в дисертаційній роботі теоретичних та практичних результатів.

У додатках наведено список публікацій здобувача за темою дисертаційної роботи, а також документи, що підтверджують впровадження результатів дисертаційного дослідження.

Дотримання вимог академічної доброчесності.

Аналіз тексту дисертації та наукових праць здобувача дає можливість зробити висновок, що дисертаційна робота є виконаною самостійно завершеною науковою працею, та не містить елементів фальсифікації, плагіату та запозичень. Використані ідеї інших авторів мають посилання на відповідне джерело; зокрема, базова модель

компактного гібридного профілю та первинна нечітко-генетична схема, що слугували відправною точкою дослідження, послідовно цитуються, а власні удосконалення чітко відмежовано від запозичених положень. Отже, дисертаційна робота відповідає вимогам щодо академічної доброчесності.

Зауваження до дисертаційної роботи.

Зауваження до дисертаційної роботи полягають у наступному.

1. У другому розділі використано трикутні функції належності; доцільно було б коротко пояснити, чому не розглянуто інші форми, наприклад гаусові чи трапецієподібні.

2. На с. 103. Ви пишете «Формально, це вимагає розширення задачі оптимізації (3.10) до двокритеріальної:» і далі йде формула з номером (3.9). В наступному розділі наведена формула (3.10), яка не є задачею оптимізації, а описує структуру хромосоми Н. Очевидно, тут помилка з номером формули.

3. На с. 106 (після рисунку 3.5) написано «Сегмент структурної значущості (*Sweight*).» Вище на малюнку скоріш за все він позначений як «Сегмент ваг». Не зрозуміло, про який саме сегмент йде мова. Те ж саме стосується фрази «Сегмент семантичної інтерпретації». Має місце невідповідність позначень на малюнку і в тексті.

4. С. 117. 2-й абзац. Ви пишете «Валідаційна вибірка... слугує незалежним критерієм прийняття знайденого рішення». Сама вибірка не може слугувати критерієм, вона може слугувати основою для обчислення значення критерію.

5. С. 119 (розділ 4.1). Написано – «Для об'єктивного порівняння результати *FD* зіставляються з двома загальновизнаними методами колаборативної фільтрації: косинусною мірою подібності (*Cosine*) та кореляцією Пірсона (*Pearson*)». Косинусна міра подібності (*Cosine*) та кореляція Пірсона (*Pearson*) не є методами колаборативної фільтрації – вони є лише метриками, які можуть використовуватися в методах колаборативної фільтрації.

6. С. 122. Написано – «...в умовах критичної розрідженості (> 93 порожніх клітинок) є...». Скоріш за все там повинно бути $> 93\%$.

Наведені зауваження не знижують загальну наукову та прикладну цінність дисертаційного дослідження і не впливають на загальну позитивну оцінку дисертаційної роботи.

Висновок.

Дисертаційна робота Гребенюка Микити Олександровича на тему «Дослідження та розробка методів рекомендаційної системи з використанням алгоритмів нечіткої логіки та генетичного алгоритму» за своїм змістом відповідає спеціальності 122 Комп'ютерні науки. Дисертація є завершеною науково-дослідною роботою, яка розв'язує важливе наукове завдання підвищення якості персоналізованих рекомендацій в умовах розрідженості матриці оцінок шляхом розроблення методів і моделей гібридної рекомендаційної системи на основі нечіткої логіки та коеволюційної оптимізації структурно-параметричних характеристик профілю користувача.

Дисертаційна робота Гребенюка Микити Олександровича за змістом, структурою, обсягом та оформленням відповідає вимогам Наказу МОН України № 40 від 12.01.2017 «Про затвердження вимог щодо оформлення дисертації» (зі змінами) та «Порядку присудження ступеня доктора філософії та скасування рішення разової спеціалізованої вченої ради закладу вищої освіти, наукової установи про присудження ступеня доктора філософії», затвердженого Постановою Кабінету Міністрів України від 12.01.2022 № 44. Таким чином, здобувач Гребенюк Микита Олександрович заслуговує на присудження наукового ступеня доктора філософії у галузі знань 12 Інформаційні технології за спеціальністю 122 Комп'ютерні науки.

Офіційний опонент:

завідувач кафедри математичного моделювання

та аналізу даних

Харківського національного університету

імені В. Н. Каразіна,

кандидат технічних наук, доцент



Володимир СТРУКОВ