

ВІДГУК

офіційного опонента,
доктора технічних наук, професора,
професора кафедри інформаційних систем та технологій
навчально-наукового інституту комп'ютерних наук та інформаційних технологій
Національного технічного університету «Харківський політехнічний інститут»
Москаленко Валентини Володимирівни
на дисертаційну роботу Гребенюка Микити Олександровича
«Дослідження та розробка методів рекомендаційної системи з використанням
алгоритмів нечіткої логіки та генетичного алгоритму»,
представлену на здобуття наукового ступеня доктора філософії
в галузі знань 12 – «Інформаційні технології»
за спеціальністю 122 – «Комп'ютерні науки»

1. Обґрунтування вибору теми дослідження

Дисертаційна робота присвячена розв'язанню задачі розробки моделей та методів гібридної рекомендаційної системи на основі нечіткої логіки та коеволюційної оптимізації для підвищення якості персоналізованих рекомендацій в умовах розрідженості матриці оцінок. На сьогодні провідні індустріальні алгоритми колаборативної фільтрації функціонують у режимі «чорної скриньки» і поступаються прозорістю пояснюваним нечітким моделям. Статичні нечіткі моделі з експертно заданими параметрами нерідко поступаються їм за точністю.

Наявність потреби у рекомендаційних системах, які одночасно забезпечують інтерпретованість, стійкість до розрідженості та ефективність в умовах холодного старту, підкреслює важливість розробки підходів до структурно-параметричної адаптації гібридних моделей користувача.

Зазначене свідчить про актуальність теми дисертаційного дослідження.

2. Оцінка змісту дисертації, її завершеності в цілому і оформлення

За своєю структурою дисертаційна робота містить анотацію, вступ, чотири змістовні розділи, загальні висновки, бібліографічний список із 136 найменувань та 2 додатки. Загальний обсяг рукопису становить 196 сторінок (із них 156 сторінок складає основний текст). Матеріал належно ілюстровано: у тексті наведено 44 рисунки та 33 таблиці.

У вступі виконано обґрунтування вибору теми дослідження, сформульовано мету, задачі, об'єкт і предмет дисертаційного дослідження, наукову новизну роботи, а також наведено інформацію про апробацію дисертаційної роботи та публікації із зазначенням особистого внеску здобувача.

У першому розділі виконано аналіз сучасного стану досліджень у сфері персоналізованих рекомендаційних систем. Виконано аналіз процесу персоналізованої рекомендації як об'єкта дослідження, порівняльний аналіз методів колаборативної фільтрації, аналіз гібридних архітектур на основі компактної гібридної моделі користувача, а також аналіз можливостей апарату нечіткої логіки та еволюційних обчислень у задачах рекомендації. Сформульовано проблему дослідження, яка є наслідком протиріччя між практичною потребою в точних та інтерпретованих рекомендаціях в умовах розрідженості даних і обмеженнями існуючих статичних гібридних моделей. Відповідно до визначеної проблеми сформульовано мету та завдання дослідження.

У другому розділі виконано формалізацію задачі рекомендації та розроблено компоненти гібридної моделі користувача: удосконалений індикатор зацікавленості категорією зі штрафним коефіцієнтом негативного досвіду, процедуру фазифікації різнорідних атрибутів профілю та удосконалену масштабно-інваріантну метрику нечіткої відстані, для якої доведено виконання аксіом коректності. Удосконалений індикатор відрізняється від базового врахуванням негативно оцінених об'єктів категорії, що усуває семантичну неоднозначність між відсутністю взаємодії та негативним досвідом.

У третьому розділі запропоновано метод оптимізації вагових коефіцієнтів гібридної моделі на основі генетичного алгоритму з композитною функцією придатності та метод коеволюційного налаштування гетерогенної хромосоми, який об'єднує вагові коефіцієнти ознак і параметри функцій належності в нерозривному генотипі з механізмами топологічного контролю мутації для збереження семантичної цілісності нечітких множин.

У четвертому розділі виконано експериментальну перевірку отриманих наукових результатів на наборі даних MovieLens 100K у режимі п'ятикратної крос-валідації Монте-Карло. Результати експериментальної перевірки підтверджують ефективність розроблених моделей та методів; статистичну значущість переваг верифіковано непараметричним критерієм знакових рангів Вілкоксона.

Висновки за результатами виконання дисертаційного дослідження у повній мірі відображають наукову новизну та практичну значимість проведених досліджень.

Список використаних джерел містить переважно публікації за останні 4–5 років. Використання актуальних публікацій підтверджує, що отримані в дисертаційній роботі результати базуються на результатах аналізу сучасних наукових досліджень у даному напрямку.

Дисертація є завершеною науковою працею, а її оформлення відповідає вимогам «Порядку присудження ступеня доктора філософії та скасування рішення разової спеціалізованої вченої ради закладу вищої освіти, наукової установи про присудження ступеня доктора філософії» (постанова Кабінету Міністрів України від 12.01.2022 р. № 44), та наказу Міністерства освіти і науки України від 12.01.2017 р. № 40 «Про затвердження Вимог до оформлення дисертації».

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

Тематика дисертації відповідає освітньо-науковій програмі підготовки докторів філософії зі спеціальності 122 «Комп'ютерні науки» та науковому напрямку кафедри комп'ютерного моделювання та інтелектуальних технологій

ХНУРЕ. Роботу виконано у межах плану науково-дослідних робіт університету на 2022–2026 рр.

4. Ступінь обґрунтованості наукових положень, висновків і рекомендацій, сформульованих у дисертації

Достовірність та обґрунтованість наукових результатів дослідження підтверджується коректним використанням відомих підходів та моделей колаборативної фільтрації, теорії нечітких множин та еволюційних обчислень, результатами теоретичного аналізу, експериментальної перевірки з ізольованою тестовою вибіркою та статистичною верифікацією, а також практичним застосуванням розроблених моделей та методів. Отримані в роботі результати були представлені, прорецензовані та обговорені на міжнародних наукових конференціях.

Отже, аналіз змісту дисертаційної роботи дає можливість зробити висновок щодо обґрунтованості отриманих наукових положень, висновків та рекомендацій.

5. Основні наукові результати, одержані автором, та їх новизна

Основні положення та практичні результати дослідження відображено у 10 наукових працях. Здобувачем опубліковано 3 статті у фахових виданнях України (категорія «Б»), 2 праці у матеріалах міжнародних конференцій, що входять до наукометричних баз Scopus та Web of Science, а також 5 тез доповідей на науково-практичних конференціях і форумі.

У дисертаційній роботі отримано такі наукові результати.

1. Вперше розроблено метод коеволюційного налаштування гетерогенної хромосоми, який здійснює одночасну структурно-параметричну адаптацію вагових коефіцієнтів та параметрів функцій належності нечітких множин у складі компактного гібридного профілю. На відміну від традиційних підходів роздільного або послідовного налаштування, запропоноване рішення забезпечує збереження семантичної цілісності нечітких термів, що дозволяє досягти вищої якості

сортування рекомендаційних списків (NDCG@10, Precision@10, Recall@10) та гарантує стійкість системи в умовах дефіциту оцінок (холодного старту).

2. Удосконалено формалізацію індикатора зацікавленості категорією в межах гібридної моделі користувача. Завдяки впровадженню штрафного коефіцієнта негативного досвіду, який пропорційно знижує підсумкову оцінку категорії залежно від кількості негативно оцінених об'єктів, вдалося розмежувати стани «антипатії» та відсутності інформації про об'єкт («байдужості»). Це дозволило усунути семантичну неоднозначність профілювання та підвищити точність категорійного представлення користувача при високій розрідженості даних.

3. Удосконалено метрику нечіткої відстані, яка застосовувана для порівняння гібридних профілів. Модифікація полягає у заміні масштабно-залежного множника на нормалізуючий коефіцієнт, що дало змогу забезпечити масштабну інваріантність і коректне агрегування різнотипних атрибутів у єдиному нормованому діапазоні. Отримане рішення уможливлює знаходження «найближчих сусідів» навіть за відсутності спільних оцінок, що суттєво розширює охоплення каталогу порівняно з класичними мірами подібності.

4. Удосконалено метод генетичної оптимізації вагових коефіцієнтів гібридної моделі користувача. Запропоновано композитну функцію придатності з вбудованим механізмом L2-регуляризації, яка, на відміну від стандартного орієнтування виключно на мінімізацію похибки MAE, інтегрує метрику ранжування безпосередньо у цільовий функціонал. Це забезпечує одночасне підвищення точності прогнозування та якості сортування видачі порівняно з базовою рівнозваженою моделлю.

6. Практичне значення одержаних результатів

Прикладна цінність одержаних результатів полягає у розробці програмного модуля мовою Python (із використанням сучасних відкритих бібліотек), який характеризується відкритістю інтерфейсів, забезпечує повну відтворюваність експериментів та може легко інтегруватися у розподілені сервіси рекомендацій із

підтримкою паралельних обчислень. Запропоновані інструментальні рішення дозволяють проєктувати алгоритмічно прозорі рекомендаційні системи, які зберігають високу точність у режимах високої розрідженості даних та «холодного старту».

Практичну спрямованість роботи підтверджено впровадженням її результатів у виробничу діяльність компанії EFISCO Sp. z o.o. (Республіка Польща), де розроблені методи та програмні компоненти використовуються у складі рекомендаційних підсистем для e-commerce та сервісів дистрибуції цифрового контенту, що засвідчено відповідним актом впровадження..

7. Дотримання академічної доброчесності

Порушень академічної доброчесності в дисертації та наукових публікаціях, у яких висвітлені основні наукові результати роботи, не виявлено. Отримані наукові результати є оригінальними. Елементи фальсифікації чи фабрикації тексту, використання ідей інших науковців без посилання на їхні публікації в роботі відсутні; базові моделі, обрані за основу для вдосконалення, коректно цитуються.

8. Дискусійні положення та зауваження до змісту дисертації

Відзначаючи новизну, важливість та повноту одержаних наукових та практичних результатів, слід відзначити й наступні недоліки.

1. У першому розділі серед основних проблем названо зміну вподобань користувача з часом, однак експериментально це питання у роботі не досліджувалося; можливо доцільно було б зазначити його як напрям подальшої роботи.

2. У третьому розділі не описано, як саме було підібрано параметри генетичних алгоритмів (розмір популяції, ймовірності схрещування та мутації); доцільно було б стисло це пояснити.

3. Складність алгоритму SEGA оцінено лише теоретично (підрозділ 4.4); доцільно було б навести фактичний час його навчання.

4. У порівняльних експериментах четвертого розділу використано лише класичні алгоритми колаборативної фільтрації, що використовуються в рекомендаційних системах для прогнозування оцінок користувачів (SVD, KNN Baseline, SlopeOne); доцільно було б обґрунтувати такий вибір та пояснити, чому не розглядалися неймережеві методи рекомендацій.

5. У підрозділі 4.5 наведено результати тестування рекомендаційних систем, побудованих за декількома алгоритмами, зроблено висновок щодо їх ефективності за параметрами MAE, Coverage (%), NDCG@10, Precision@10 та Recall@10; доцільно було б провести порівняльний аналіз цих методів більш детально з уточненням умов, за якими запропонований метод є найкращим.

6. У запропонованому алгоритмі CEGA використовується інструментарій нечіткої логіки для семантичної адаптації, зокрема розглянуто трикутну функцію належності у третьому розділі, але не зрозуміло, як вибір типу функції приналежності впливає на ефективність запропонованого алгоритму.

Вказані недоліки мають переважно дискусійний і рекомендаційний характер, вони не впливають на загальну позитивну оцінку дисертаційної роботи, не зменшують рівень значущості наукових результатів та практичної цінності.

9. Загальні висновки щодо дисертаційної роботи

Дисертаційна робота Гребенюка Микити Олександровича «Дослідження та розробка методів рекомендаційної системи з використанням алгоритмів нечіткої логіки та генетичного алгоритму», представлена на здобуття наукового ступеня доктора філософії з галузі знань 12 – «Інформаційні технології» за спеціальністю 122 – «Комп'ютерні науки», є актуальною, завершеною науковою працею, яка виконана на високому науково-теоретичному рівні.

У роботі розв'язана актуальна наукова задача, яка полягає у підвищенні якості персоналізованих рекомендацій в умовах розрідженості матриці оцінок і холодного старту шляхом розроблення моделей та методів гібридної рекомендаційної системи

на основі нечіткої логіки та коеволюційної оптимізації структурно-параметричних характеристик гібридного профілю користувача.

Дисертаційна робота повністю відповідає вимогам до дисертацій на здобуття наукового ступеня доктора філософії, а саме вимогам пунктів 6 – 9 «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 р. № 44 із змінами, а здобувач Гребенюк Микита Олександрович заслуговує на присудження наукового ступеня доктора філософії з галузі знань 12 – «Інформаційні технології» за спеціальністю 122 – «Комп'ютерні науки».

Офіційний опонент,

доктор технічних наук, професор,
професор кафедри інформаційних систем та технологій
навчально-наукового інституту комп'ютерних наук
та інформаційних технологій
Національного технічного університету
«Харківський політехнічний інститут»

Ваш

Валентина МОСКАЛЕНКО

*підпис засвідчено
першим проректором
з науково-
педагогічної роботи*

М



*Русман
ГУЩЕНКО*