

ВІДГУК

офіційного опонента на дисертаційну роботу

Шварца Михайла Євгенійовича

*«Гібридні моделі і методи прогнозування рекомендацій для
інтернет-магазину»,*

подану на здобуття наукового ступеня кандидата технічних наук
за спеціальністю 01.05.03 – математичне та програмне забезпечення
обчислювальних машин і систем

1. Актуальність теми

Стрімкий розвиток та розроблення інформаційних технологій для побудови і впровадження комп'ютерних систем електронного бізнесу вимагає вирішення нових задач. З ростом обсягів даних, що зберігаються у мережі і пропонуються користувачам, зростає актуальність проблеми випередження запиту користувача шляхом пропонування потенційно корисної для нього інформації. Цю проблему розв'язують системи надання рекомендацій, котрі здатні пропонувати відповіді без явного запиту з боку користувача і, як результат, підвищити корисність споживання інформації в сучасному глобальному інформаційному просторі.

Тому розробки методів та засобів для побудови і впровадження рекомендаційних систем електронного бізнесу взагалі та інтернет-магазинів зокрема, як основних суб'єктів електронної комерції, є актуальною науковою проблемою.

Актуальність теми дисертаційної роботи підтверджується тим, що результати дослідження використані при виконанні завдань науково-дослідних робіт Національного університету «Львівська політехніка», а саме: міжнародний науково-дослідний проект – TEMPUS-JPCR – «Розробка програми для нової спеціальності: «Магістр з інженерії проектування мікросистем»», термін виконання проекту 15.10.2012 – 14.10.2016, реєстраційний номер: № 530785-«TEMPUS-1-2012-1-PL-TEMPUS-JPCR»; грант молодих учених – ДБ/Нанокompозит "Моделювання і створення нового класу кристалічних нанокompозитів із контрольованою кристалізацією та їх дослідження в оптичному та субтерагерцовому діапазонах хвиль", термін виконання проекту 01.01.2016 – 31.12.2018, номер державної реєстрації № 0116U004412.

2. Ступінь обґрунтованості наукових положень, висновків і рекомендацій

Для верифікації розроблених наукових положень, висновків та рекомендацій автором використані дані, одержані з результатів аналізу сучасного стану та перспектив розвитку моделей і методів для побудови і впровадження рекомендаційних систем електронного бізнесу.

У дисертаційній роботі виконано повний та всебічний аналіз поставленої задачі, комплексне теоретичне та практичне пророблення усіх аспектів її розв'язання,

показано, що найбільш перспективними є гібридні методи прогнозування рекомендацій.

Автор дисертації розробив метод розрахунку коефіцієнтів подібності, який враховує демографічні характеристики користувачів, гібридний метод прогнозування рекомендацій для груп користувачів на основі кластеризації, який адаптується до розрідженості матриці користувач-предмет, а також метод прогнозування рекомендацій для користувачів інтернет-магазину, який використовує алгоритм пошуку асоціативних правил за допомогою адаптивної зміни підтримки асоціативних правил.

На основі запропонованих методів та алгоритмів розроблено програмне забезпечення, яке включає методи прогнозування рекомендацій, методи розрахунку подібності векторів профілів користувачів і предметів та методи тестування точності.

Для підтвердження правомірності теоретичної складової роботи автором проведені дослідження точності і складності розроблених методів прогнозування рекомендацій. Це додатково підтверджує високий ступінь обґрунтованості положень та висновків роботи.

3. Новизна наукових положень, висновків і рекомендацій, їх достовірність

Науковий доробок автора полягає в вирішенні науково-практичної проблеми розроблення моделей та методів для побудови і впровадження рекомендаційних систем електронного бізнесу. Цей доробок складається з таких компонентів:

1. На основі концепції застосування в одному методі категоріальної, мішаної і числової кластеризації вперше розроблено метод пошуку груп користувачів, який адаптується до розрідженості матриці "користувач-предмет".

2. Отримав подальший розвиток метод розрахунку коефіцієнтів подібності векторів профілів користувачів і векторів профілів предметів, який, на відміну від існуючих, використовує демографічні характеристики користувачів, що дозволяє підвищити точність прогнозування рекомендацій і визначати коефіцієнти подібності для нового користувача і нового предмета.

3. Отримав подальший розвиток метод мішаної кластеризації, який використовується для кластеризації категоріально-числових векторів профілів користувачів і, на відміну від існуючих, автоматично вибирає центри кластерів і дозволяє зменшити час пошуку груп користувачів при високій точності виділення груп.

4. Отримав подальший розвиток метод збільшення різноманітності рекомендованих предметів, який дозволяє врахувати оцінки подібних предметів в околі предметів активного користувача і вирішує проблему "довгого хвоста".

5. Удосконалено метод прогнозування рекомендацій для користувачів інтернет-магазину, який на відміну від існуючих методів, використовує алгоритм пошуку асоціативних правил за допомогою адаптивної зміни підтримки асоціативних правил.

Достовірність наукових положень, висновків і рекомендацій, одержаних в дисертації, підтверджується результатами експериментальних досліджень, коректним

застосуванням математичного апарату, сучасних методів моделювання, а також впровадженням розробок в практику людської діяльності.

4. Практичне значення результатів дисертаційної роботи

Практичне значення одержаних результатів полягає в тому, що:

1. Розроблено метод розрахунку коефіцієнтів подібності векторів характеристик користувачів і предметів, який, враховує демографічні і контентні характеристики користувачів і предметів та дозволяє розраховувати коефіцієнти подібності для нового користувача.

2. Розроблено метод прогнозування рекомендацій для користувачів інтернет-магазину на основі концепції асоціативних правил, який використовує алгоритм пошуку асоціативних правил за допомогою адаптивної зміни підтримки асоціативних правил.

3. Розроблено метод пошуку груп користувачів, який адаптується до коефіцієнту розрідженості матриці "користувач-предмет" і використовує методи категоріальної, мішаної і числової кластеризації.

Результати дисертаційних досліджень використано у навчальному процесі Національного університету «Львівська політехніка», зокрема на кафедрі систем автоматизованого проектування при викладанні дисциплін: «Інноваційні інформаційні технології» для підготовки магістрів за спеціальністю «Інформаційні технології проектування» і «Методи і системи штучного інтелекту» для підготовки бакалаврів за спеціальністю «Комп'ютерні науки».

5. Повнота викладу наукових положень, висновків і рекомендацій в опублікованих працях

Усі наведені у дисертації положення та результати повністю викладені в опублікованих наукових працях. За результатами досліджень надруковано 14 публікацій автора, серед яких 4 статті у фахових наукових виданнях України, 1 публікація – в зарубіжному науковому періодичному виданні, яке входить в міжнародні наукометричні бази, 9 публікацій в матеріалах науково-технічних конференцій, 6 з яких міжнародні, а 3 – проіндексовані у базі SCOPUS.

В авторефераті в необхідній мірі містяться основні положення, висновки та рекомендації, наведені в дисертації. Зміст автореферату є ідентичним до змісту дисертації. Автореферат дисертації відображає персональний вклад автора в спільних публікаціях, що дозволяє оцінити самостійність роботи і конкретні здобутки дисертанта.

6. Аналіз змісту дисертації

У вступі обґрунтовано актуальність теми дисертаційної роботи, сформульовано мету і завдання роботи, наукову новизну та практичну цінність отриманих результатів,

показано зв'язок роботи з науковими програмами кафедри, планами і темами. Наведено дані про впровадження результатів роботи, їх апробацію, публікації та особистий внесок здобувача.

У **першому розділі** виконано аналіз сучасного стану моделей і методів електронної комерції, показано, що інтернет-магазин є однією із основних систем електронної комерції, показано область застосування рекомендаційних систем у структурі роботи інтернет-магазину та їх вплив на числові значення метрик ефективності роботи інтернет-магазинів. Здійснено класифікацію існуючих підходів до побудови рекомендаційних систем та можливість застосування кожного з них у роботі інтернет-магазину.

У **другому розділі** дисертаційної роботи наведено формальну постановку задачі прогнозування рекомендацій для інтернет-магазину методом колаборативної фільтрації, продемонстровано особливості формальної моделі для прогнозування рекомендацій з урахуванням груп користувачів, виконано порівняння існуючих методів розрахунку подібності векторів профілів користувачів і предметів, удосконалено метод розрахунку коефіцієнтів подібності, який, на відміну від існуючих, використовує обернену евклідову відстань між векторами профілів користувачів і демографічні характеристики користувачів, показано можливість застосування цього методу для рішення задачі вироблення рекомендацій для нового користувача (т.зв. "холодного старту").

У **третьому розділі** дисертаційної роботи розроблено гібридний метод пошуку груп користувачів, який адаптується до розрідженості матриці "користувач-предмет"; розроблено новий метод мішаної кластеризації, який враховує категоріальні і числові складові вектора профілю користувача і адаптується до розрідженості матриці "користувач-предмет"; показано застосування методів прогнозування рекомендацій для груп користувачів; розроблено метод прогнозування рекомендацій на основі пошуку асоціативних правил; розроблено теоретико-множинні моделі прогнозування рекомендацій для супутніх продаж і додаткових продаж; розроблено метод прогнозування рекомендацій для супутніх продаж, додаткових продаж і режиму післяпродажної роботи, розроблено метод збільшення різноманітності товарів, які пропонує рекомендаційна система.

У **четвертому розділі** дисертаційної роботи розроблено інформаційне забезпечення для тестування моделей і методів прогнозування рекомендацій для інтернет-магазину, розроблено структуру математичного та програмного забезпечення, яка дозволяє вибрати метод прогнозування рекомендацій, метод пошуку груп користувачів, метод прогнозування рекомендацій у групі користувачів та метод прогнозування рекомендацій для формування додаткових продаж, метод розрахунку точності прогнозування, величину поділу тестової матриці "користувач-предмет" на прогнозовану і тестову.

У **додатках** наведено акт впровадження результатів дисертаційної роботи та список робіт автора.

7. Зауваження по дисертації

1. В першому розділі виконано аналіз сучасного стану розроблення математичних та алгоритмічних основ для рекомендаційних систем, орієнтованих на системи B2C, незрозумілим залишається можливість застосування розроблених алгоритмів у інших типах систем електронної комерції: B2B, C2B, C2C (сторінка 33).

2. Пункт 1.2 та підпункти 1.4.1, 1.4.2, 1.5.1 дисертаційної роботи недоцільно виділяти, як окремі самостійні одиниці тексту, оскільки через мізерний обсяг поданого в них матеріалу вони не становлять цінності, як завершені самостійні фрагменти результатів дослідження.

3. На сторінці 71 дисертації при означенні асоціативного правила вводяться позначення X та Y, які відразу використовуються у виразах без чіткого зазначення, що вони є множинами подій (транзакцій). Про це можна здогадатись лише з подальшого тексту.

4. В третьому розділі на сторінці 74 дисертаційної роботи для пошуку асоціативних правил безальтернативно пропонується метод Apriori без обґрунтування доцільності його використання в контексті даної роботи та без аналізу аналогічних за призначенням алгоритмів.

5. Для програмної реалізації рекомендаційної системи в четвертому розділі запропоновано платформу .NET та мову C# з ORM Entity Framework, а в якості інструментарію розробки – Microsoft Visual Studio 2010 та Microsoft SQL Server 2008, як такі, що забезпечують найкращу продуктивність розробленої рекомендаційної системи (сторінка 102). Однак посилання на джерела, що підтверджують це, чи власне доведення такого ствердження відсутні. Крім того варто обґрунтувати застосування далеко не найновіших версій даних інструментів розробки.

6. В тексті дисертації наявні орфографічні та стилістичні помилки, а також недоліки технічного характеру: описки, неправильні посилання на формули та позначення.

8. Загальні висновки

Перелічені зауваження не знижують цінності роботи та позитивного враження від її досягнень. Дисертаційна робота є закінченою кваліфікаційною науковою роботою, вона містить раніше незахищені наукові положення і одержані автором нові науково обґрунтовані результати в області розроблення моделей і методів для побудови і впровадження рекомендаційних систем для електронного бізнесу.

В підсумку, з огляду на актуальність проблеми, високий рівень виконаних досліджень та результати практичного впровадження, вважаю, що дисертаційна робота «Гібридні моделі і методи прогнозування рекомендацій для інтернет-магазину» відповідає вимогам, які висуваються Міністерством освіти і науки України до кандидатських дисертацій згідно з п.п. 9, 11, 12 "Порядку присудження наукових ступенів і присвоєння вченого звання старшого наукового співробітника", затвердженого постановою Кабінету Міністрів України від 24 липня 2013 р. № 567.



Тема, задачі і зміст дисертаційної роботи відповідають профілю спеціалізованої вченої ради Д 35.052.05, паспорту спеціальності 01.05.03 – “математичне та програмне забезпечення обчислювальних машин і систем” в частині його формули та окремим напрямкам досліджень. Автор роботи, Шварц Михайло Євгенійович, заслуговує присудження наукового ступеня кандидата технічних наук за спеціальністю 01.05.03 – математичне та програмне забезпечення обчислювальних машин і систем.

ОФІЦІЙНИЙ ОПОНЕНТ:

кандидат технічних наук,
завідувач кафедри комп'ютерних наук
Тернопільського національного технічного
університету імені Івана Пулюя

І.О. Боднарчук

Підпис І.О. Боднарчука засвідчую:

Вчений секретар
Тернопільського національного технічного
університету імені Івана Пулюя
к.т.н., доцент



Г.М. Крамар