

ЕКОНОМІЧНІ НАУКИ

УДК 330.16--028.68:005.953.2-027.552

DOI: 10.15587/2313-8416.2018.152949

АНАЛІЗ МЕТОДІВ, МОДЕЛЕЙ ТА АЛГОРИТМІВ ПЕРСОНАЛІЗАЦІЇ ДЛЯ РОЗРОБЛЕННЯ РЕКОМЕНДАЦІЙНИХ СИСТЕМ

© Ю. О. Котлярова

В статті досліджуються методи, моделі та алгоритми персоналізації в цифровому середовищі. Наведена загальна характеристика типів рекомендаційних систем, їх методи, моделі та алгоритми, переваги та недоліки застосування. В роботі запропоновано застосування таксономії знань для створення інтелектуального додатку персоналізації для підтримки прийняття маркетингових рішень для підприємств у цифровому середовищі. Виділено додаткові джерела даних для створення рекомендацій

Ключові слова: персоналізація, методи, моделі, алгоритми та таксономія знань для рекомендаційних систем

1. Вступ

В еру цифрової глобалізації синергетика наук є невід'ємною частиною формування нового інформаційного простору. Серед таких наук можна виділити економіку, маркетинг та психологію.

В сучасній цифровій економіці значну увагу потрібно приділити вивченню саме біхевіористичної економіки, яка розглядає поведінку людини під впливом множини факторів зовнішнього середовища, за принципом «стимул-реакція», де під реакцією можна розглянути кінцеві дії покупців-споживачів, наприклад, придбання товарів, а під стимулом – маркетингові заходи, що викликають у споживачів потреби та бажання придбати товари. Отже, вивчаючи причини між стимулом та реакцією, можна встановити закономірний зв'язок, на певний стимул викликати ті чи інші реакції, а тим самим досягти потрібної поведінки людини.

2. Літературний огляд

Людина як споживач має значну кількість різноманітних потреб, задоволення яких і становить зміст її поведінки.

Однією з авторитетних біхевіористичних моделей маркетингу прийнято вважати модель вибіркової поведінки споживача, що розроблена Ф. Котлером [1]. Він характеризує вибір споживача як послідовність трьох блоків: стимулюючі фактори маркетингу, «чорний ящик» свідомості споживача та реакція. «Чорний ящик» розглядається як набір певних стимулів і складається з двох блоків: перший – це особистісні характеристики покупця, які вплива-

ють на сприйняття стимулюючих факторів маркетингу та інших подразників і викликають певні реакції на них; другий – сам процес прийняття споживачем рішення про покупку товару чи послуги, який впливає на споживчу поведінку загалом.

В роботі [2], що стала класикою в дослідженні та моделюванні поведінки споживачів виділено основні фактори (соціальні, культурні, особистісні, психологічні), які впливають на поведінку споживачів і як ці фактори впливу можуть пояснити, чому споживач приймає те або інше рішення. Авторами розроблена модель, що показує взаємозв'язок конкретних факторів зовнішнього середовища, внутрішніх факторів споживача та процесу прийняття рішення про покупку в єдину систему, що дозволяє описати споживача та його потреби.

Значну увагу персоналізації приділено в дослідженні [3], де наведено актуальність застосування веб-технологій, а саме відслідковування поведінки споживача на веб-сайті. Кожний клік на банер чи посилення надає сукупність інформації про те, чим цікавився споживач, в який час, які сторінки відвідував, як довго перебував на веб-сайті. Такі дані значно збагачують і доповнюють фактори, що описували [2]. Процедура реєстрації відвідувачів сайту надає можливість зв'язати їх демографічні дані з поведінкою, цим самим виявити мотивацію відвідувачів. Аналіз купівельної поведінки в режимі реального часу дозволяє персоналізувати зміст веб-сайтів в залежності від споживчих потреб.

Для формування вдалих рекламних кампаній з просування товарів безпосередній вплив на по-

ведінку споживачів справляють за допомогою маркетингових комунікацій. Запропоновано комплекс просування, що поєднує у собі засоби реклами, стимулювання збуту, персонального продажу, засобів комунікації тощо [4]. Дієвість впливу різних видів комунікації на поведінку споживачів залежить від якості та складових елементів комунікації.

З точки зору психології вибору та прийняття рішень споживачем в роботах [5, 6] запропоновано концепцію прийняття рішень в умовах невизначеності, що отримала назву «теорія перспектив», яка доводить що поведінка людини нерациональна, і що ця нерациональність не випадкова і не безглузда – навпаки, цілком систематична і передбачувана. Авторами запропонована формальна модель, що дозволяє спростити створення механізму підвищення залученості споживача у взаємодію з маркетингом.

Серед моделей, що присвячені вивченню поведінки споживача та прийняття економічних рішень [7, 8] виділяють наступні: теорія раціонального вибору, теорія ігор, теорія прийняття рішень, теорія очікуваної корисності, теорія перспектив, теорія обмеженої раціональності, теорія масового обслуговування.

В умовах розвитку цифрової економіки та діджиталізації усіх сфер людської діяльності актуальним питанням є дослідження поведінки саме в цифровому середовищі із застосуванням математичних методів і алгоритмів.

3. Мета та задачі дослідження

Мета дослідження – провести аналіз методів, моделей та алгоритмів персоналізації для розроблення рекомендаційних систем та задоволення потреб споживачів.

Для досягнення мети були поставлені наступні задачі:

1. Дослідити потреби сучасного суспільства в цифровому середовищі.
2. Дослідити процес персоналізації в цифровому середовищі та його основні етапи.
3. Визначити перспективи застосування рекомендаційних систем в цифровому середовищі.
4. Визначити галузі наук для дослідження поведінки та потреб споживача.
5. Проаналізувати існуючі сервіси, з впровадженими рекомендаційними системами.
6. Охарактеризувати типи рекомендаційних систем, їх методи, моделі, алгоритми. Навести переваги та недоліки застосування різних типів рекомендаційних систем. Класифікувати рекомендаційні системи.
7. Розробити таксономію знань для створення інтелектуального додатку персоналізації контенту. Описати фактори знань для розроблення рекомендаційних систем.

8. Визначити додаткові ресурси для збагачення даними рекомендаційну систему.

4. Сучасні потреби цифрового суспільства

В цифровому світі неможливо обійтися без використання цифрових пристроїв, що дають змогу знайти потрібну інформацію про продукт, але й містять достатню інформацію про самого власника гаджету, що значно доповнює аналітичну картину. Часто людина стикається з надмірністю отримання нерелевантних пропозицій від представників електронного бізнесу, що формує не завжди позитивне враження. Значною проблемою є й великий різновид варіантів пропозицій, що ускладнює вибір того чи іншого товару. Підтримка прийняття рішення в цифровому інформаційному просторі є серйозною проблемою. Тому, щоб допомогти споживачеві знайти і надати необхідну інформацію, digital-маркетологи розробляють і впроваджують рекомендаційні системи (recommender systems), які дозволяють максимально налаштувати, персоналізувати та спрогнозувати, що забажає окремий споживач, і надати йому відповідне комерційне повідомлення під його конкретну потребу. Рішенням визначених проблем є розроблення рекомендаційних систем на основі методів, моделей та алгоритмів персоналізації. Використання таких систем дозволить e-commerce прискорити та збільшити прибуток, а споживачам задовольнити власні потреби та побажання.

Технології персоналізації виводять цифровий маркетинг на новий інтелектуальний рівень, що часто називають превентивний маркетинг або маркетинг без попередження.

5. Результати досліджень

Персоналізація – одне з найактуальніших питань сучасного інтелектуального маркетингу, що знайшло своє місце в рекомендаційних системах, які формують рекомендації незалежно для кожного конкретного споживача на основі його попередніх дій (покупок, пошуків, вподобань, оцінок, відгуків та коментарів), соціально-демографічних даних та враховуючи поведінку інших користувачів цифрового простору. Ці дані потім застосовують для прогнозування того, як клієнти оцінюють набори суміжних товарів або наскільки ймовірно, що клієнт повинен купувати додатковий товар. Рекомендаційні системи базуються на двох принципах: перший – це потреби людини, що визначають напрям дій (поведінка), другий – реакція на стимул (задоволення чи незадоволення потреб), яка на основі даних про користувача і товар (предмет) дає рекомендації. Зазвичай такі системи включають в себе весь процес – від отримання інформації до її представлення користувачу, застосовуючи програмні інструменти та методи (рис. 1).



Рис. 1. Цикл персоналізації в цифровому середовищі в контексті рекомендаційних систем

Метою персоналізації в цифровому середовищі є створення для покупців-споживачів персонального комфортного простору в мережі Інтернет взагалі, і позитивного досвіду взаємодії з e-commerce-ресурсами, зокрема Інтернет-магазинами, електронними торговими майданчиками і прайс-агрегаторами, які також відомі як маркетплейси. Як наслідок, пошукові сервіси та власники e-commerce отримують інструментарій для оптимізації використання своїх ресурсів.

Впровадження таких рекомендаційних систем (технологій персоналізації) дозволить вирішити цілу низку проблем:

- вивчення сегменту та потреб споживачів, враховуючи психологічні та фізичні особливості;
- інформування споживачів про цікаві пропозиції;
- перевантаження нецільовим контентом;
- калібрування вподобань споживачів та утримання їх зацікавленості;
- підвищення вартості кошика;
- заохочення та зацікавленість до продукту чи платформи e-commerce;
- персоналізація контенту для споживачів;
- вивчення поведінки споживача на e-commerce-ресурсах;
- продаж супутніх товарів;
- швидка адаптація до нових інтересів споживача;
- формування актуальних комерційних пропозицій для споживачів;
- заощадження часу на здійсненні пошуку потрібних товарів;
- формування цільових повідомлень через e-mail розсилку;
- підвищення лояльності та прив'язаності до бренду.

Описані проблеми дозволить вирішити синергетика таких галузей наук, як психологія, інформаційні технології, математика та статистика (рис. 2). В даному випадку психологія повинна відповісти на питання, які саме елементи контенту варто враховувати, щоб результат рекомендації можна було б назвати персоналізованим, тобто персональне задоволення потреб окремого споживача. Соціологія виступає як інструмент для спостереження, реєстрації фактів (опитування), узагальнення та виведення закономірностей в деякому соціальному середовищі та соціальній реальності. Інформаційні технології мають надати необхідні програмно-технічні засоби, що інтегровані з метою збирання, опрацювання, зберігання, розповсюдження, відображення та використання інформації, а математична статистика, в свою чергу, повинна забезпечити інструментарій (алгоритми, методи та моделі) для виділення та інтерпретації елементів рекомендації. В даному підході розробляється модель психологічного портрету користувача, яка дозволить з високою імовірністю прогнозувати реакцію споживача на запропоновану йому одиницю контенту.

Досвідом застосування рекомендаційних систем діляться такі великі цифрові сервіси, як Amazon, Netflix [9, 10], Facebook, LinkedIn, YouTube, Spotify, Google, Best Buy, Pinterest, Last.fm, Pandora, та Le Boutique [11, 12]. У кожному з них існує своя специфіка, яка безпосередньо залежить від типу даних, які потрібно рекомендувати. Їх основна мета – збирати величезну кількість даних, що стосується поведінки споживачів та їхніх дій у межах середовища, та прогнозувати, які об'єкти можуть бути цікаві споживачу. Цю інформацію можна систематично зберігати в профілях користувачів і застосовувати для майбутніх взаємодій.

Розглянемо основні методи, моделі та алгоритми рекомендаційних систем (табл. 1).



Рис. 2. Галузі наук персоналізації в контексті рекомендаційних систем

Таблиця 1

Загальна характеристика типів рекомендаційних систем, їх методи, моделі, алгоритми, переваги та недоліки застосування

Тип рекомендаційної системи	Характеристика рекомендаційної системи	Методи, моделі та алгоритми, що застосовують	Переваги	Недоліки
1	2	3	4	5
Колаборативна фільтрація на основі пам'яті (Memory-based collaborative filtering)	У [13–15] описано, що даний вид фільтрації намагається знайти користувачів, які схожі на активного користувача, і використовує їх вподобання для прогнозування одиниць контенту для активного користувача, що формуються на основі обчислення міри схожості по накопиченим даним.	метод найближчих сусідів, кореляція Пірсона, векторна модель (подібність на основі векторного косинуса)	– масштабування для великого об'єму даних; – простий в реалізації; – легке оновлення бази даних контенту.	– повільний в обчисленні, так як використовує всю базу даних кожного разу коли здійснює прогнозування; – неточні та ненадійні рекомендації.
Колаборативна фільтрація на основі аналізу моделі даних (Model-based collaborative filtering)	Використовує машинне навчання і інтелектуальний аналіз даних, що знаходить складні шаблони оцінки в навчальних даних, а потім робить інтелектуальні висновки або рекомендації для користувача на основі вивченої моделі. Такий підхід у [16] спрямований на пошук закономірностей, що дозволяє розкрити приховані чинники, що пояснюють оцінку користувачів.	методи Байєсових мереж, методи кластерного аналізу, методи на основі моделей Маркова, мережі залежностей, моделі регресії, методи розкладання сингулярних значень (SVD), матрична факторизація	– підвищення продуктивності прогнозування; – масштабування до фактичного набору даних; – легко уникає використання фітінгів.	– розрідженість даних, що пояснюється неможливістю надати рекомендації користувачам що не мають оцінок або рейтингу на одиницю контенту; – не підтримують інкрементного навчання, менша точність прогнозу.

Продовження таблиці 1

1	2	3	4	5
Колаборативна фільтрація на основі подібності користувачів (User-based collaborative filtering)	Фільтрація цього типу проходить в два етапи [17]: на першому визначається скільки користувачів у базі даних схожі з поточним користувачем, на другому – оцінюються схожі користувачі, щоб передбачити, який тип контенту надати користувачеві, враховуючи загальну вагу користувачів, що більш схожі на нього.	відстань Евкліда, Хеммінга, відстань міських кварталів, кореляція Пірсона, рангова кореляція Спірмена, алгоритм Weighted Slope One, наївний класифікатор Байеса, кластерна модель, матриця факторизації, косинусна подібність, коефіцієнт Жаккарда.		–ресурсомісткість (вимоги до пам'яті) так як для розрахунку рекомендацій необхідно зберігати всю матрицю даних про користувачів та їх вподобання; – складність (велика кількість обчислень, що необхідні для отримання рекомендацій); – обчислення схожості доцільно для відносно невеликих баз даних; – складність прогнозування для нових користувачів, так як у них ще немає оцінок; – при обробленні великої кількості даних та обчисленні схожості виникають проблеми отримання результатів рекомендацій в реальному часі; – складність прогнозування при появі нових елементів, так як для них ще немає оцінок.
Колаборативна фільтрація на основі подібності елементів (Item-based collaborative filtering)	Для формування рекомендацій використовується подібність між оцінками моделей елементів. Якщо два елементи, як правило, мають однакові оцінки користувачів, то вони схожі, і користувачі повинні мати аналогічні рекомендації для подібних елементів, доведено у [17].		– інтуїтивно зрозуміла і проста в реалізації; – висока точність; – ефективна з точки зору часу формування рекомендацій;	
Колаборативна фільтрація на основі транзакцій (transaction-based)	Для видачі рекомендацій, використовують дані щодо здійснених раніше транзакцій купівлі товарів [18], що представляють собою деякий набір елементів контенту (товарів), що купують разом.	асоціативні правила, алгоритм Apriori	– візуалізація здійснених транзакцій; – висока швидкість аналізу даних; – висока точність і стабільність.	– надмірна кількість товару, що є у правилах, може ускладнити аналіз закономірностей; – надмірна кількість правил може ускладнювати розуміння.

Продовження таблиці 1

1	2	3	4	5
Колаборативна фільтрація на основі тегів (tag-based)	Рекомендаційні системи на основі тегів аналізують терміни, які використовує користувач, виявляючи залежність їх семантичного значення з контекстом де він згадується, хоча саме значення елемента в контексті не відображається [16]. Такі системи можуть «розуміти» інформацію про самого користувача і контекст в якому зазначений тег.	матрична факторизація, імовірнісний факторний аналіз, косинусна подібність, метод найближчих сусідів, методи кореляції, латентно-семантичний аналіз, наївний класифікатор Байєса, алгоритм Argiori, алгоритм Портера, алгоритм Пейса	<ul style="list-style-type: none"> – фіксують важливі поняття щодо контенту; – виявлення інтересів користувачів; – кластеризація контенту; – покращує виявлення подібності користувачів; – легке оновлення профілю користувача (додавання тегів) 	<ul style="list-style-type: none"> – низька продуктивність; – складність побудови; – залучення сторонніх сервісів для отримання додаткових даних про теги та їх використання.
Колаборативна фільтрація на основі латентних моделей (Latent models for collaborative filtering)	На моделі з прихованими факторами великий вплив мають саме набір факторів, які дуже специфічні для предметної області (наприклад, складність персонажів фільму, кількість дій у фільмі). Такі фактори в цілому неочевидні і тому важко оцінити їх вплив на рейтинг. Основною метою даних моделей є виявити приховані фактори з даних рейтингу за допомогою математичних методів, зазначено у [19, 20].	семплювання за Гіббсом, латентне розміщення Діріхле, байєсовські мережі, гаусовське розподілення, метод Ланцоца, стохастичний градієнтний спуск, матрична факторизація, алгоритм Expectation Maximization, алгоритм SVD, метод максимальної правдоподібності, алгоритми multi-armed bandit.	<ul style="list-style-type: none"> – висока точність; – дозволяє врахувати не тільки парні лінійні кореляції, але й множинні взаємодії змінних. – можливість оперувати прихованими змінними; – можливість моделювати складні явища; – можливість перевірки моделей на відповідність вихідним даним. 	<ul style="list-style-type: none"> – складність сумісного використання; – зниження швидкості обчислення при збільшенні обсягу вхідних даних.
Колаборативна фільтрація на основі графів (Graph-based collaborative ranking)	Алгоритми рекомендацій на основі графіків являють собою відносини між користувачами та елементами як двопарний графік, де відображається попарне порівняння вподобань користувача до елемента, який він оцінив [21].	метод найближчих сусідів, індекс Кадца, косинусна подібність, алгоритм Random Walk with Restarts, класифікація за методом Роккіо	– моделює рангові дані у формі структури графів, що дозволяє оцінювати близькість користувачів та елементів, на основі яких можна зробити рекомендацію.	<ul style="list-style-type: none"> – дані залежать від контексту; – висока вартість збирання даних для побудови графу; – трудомісткий процес складання залежностей.

Продовження таблиці 1

1	2	3	4	5
Рекомендаційні системи на основі контенту (Content-based filtering)	Рекомендаційні системи контентної фільтрації працюють з профілями користувачів, що містять вподобання та смаки. Вподобання засновані на рейтингу користувача для різних предметів. В процесі рекомендацій порівнюються елементи, які вже позитивно оцінив користувач, з елементами, які він не оцінив, і знаходяться збіги. Елементи, схожі на позитивно оцінені, будуть рекомендовані користувачеві [22].	матрична факторизація, матриць, косинусна подібність, асоціативні правила, латентне розміщення Діріхле, прихована марковська модель, алгоритм прямого-оберненого ходу, векторна модель, індекс Соренсена	– рекомендації на основі вподобань конкретного користувача; – найбільш ефективний для персональних рекомендацій.	– неможливість виведення цільових рекомендацій в разі незаповненого профілю; – рекомендації аналогічні тим, які користувач надавав перевагу в минулому; – неоцінені користувачем елементи не можуть бути рекомендовані.
Рекомендаційні системи основані на знаннях (knowledge-based filtering)	Це системи, які для формування рекомендації використовують отримані знання з предметної області. Дана система формування рекомендацій не враховує оцінки інших користувачів, а лише профіль конкретного користувача і об'єкти. В зв'язку з цим системи можуть враховувати користувачські побажання, демографічні особливості та ін. [23].	метод найближчих сусідів, наївний класифікатор Байеса, асоціативні правила, методи нечіткої імплікації, асоціативні мережі, карти подібності, байєсові мережі.	– більш глибокий аналіз зв'язки між об'єктами не за критеріям схожості, а враховуючи взаємозв'язок груп об'єктів; – враховують ряд додаткових опцій, які відносяться до індивідуальних ознак конкретного користувача; – простота реалізації при наявному наборі правил.	– труднощі в побудові знань; – складно застосовувати, де товари часто купуються; – залежність від предметної області; –
Рекомендаційні системи основані на випадках (case-based filtering)	Case-based підхід враховує наявність додаткової сутності [24]: вимоги користувача (user requirements). Задача даного підходу заключається в тому, щоб знайти рекомендовані об'єкти, згідно вимогам.	метод найближчих сусідів, асоціативні правила, кореляція Пірсона, кластерна модель, матриця факторизації, коефіцієнт Жаккарда	– здійснюється підбір рекомендацій на основі подібності; – точна відповідність запиту користувача.	– вимагає уточнення від користувача; – використовує дані на основі минулих рекомендацій.

Продовження таблиці 1

1	2	3	4	5
Рекомендаційні системи на основі соціально-демографічних даних (demographic / social-based filtering)	Даний тип рекомендацій категоризує користувачів на основі демографічних груп, враховуючи вік, стать, на скільки він забезпечений, де проживає, скільки заробляє та ін. Системи соціальних рекомендацій вирішують проблеми традиційної рекомендації, використовуючи соціальні дані про користувача, поведінку, його оцінки, відносини, членство в спільнотах, вподобання, коментарі, голоси, закладки та ін. [25].	методи розкладання сингулярних значень (SVD), Обмежена машина Больцмана, методи регресійного аналізу, методи спектрального розкладу, тензорна факторизація, алгоритми баготорівневої кластеризації, матрична кластеризація, генетичні алгоритми	– не потребує історії оцінок користувачів; – точне сегментування та кластеризація користувачів	– великий обсяг даних про користувача; – ресурсомісткість обчислень; – обмеженість даних про користувачів.
Рекомендаційні системи на основі корисності інформації (utility-based filtering)	Utility-based розраховує відносну користь кожного елемента для користувача [26]. В даних рекомендаційних системах можна створювати багато власних областей знань і ранжувати по ним, досягаючи максимальних результатів	регресійний аналіз, нейронні мережі, генетичні алгоритми	– надає рекомендації щодо невідповідних атрибутів продукту; – висока ефективність застосування.	– побудова функції корисності для кожного окремого користувача; – вимагає від користувачів оцінювання елементів для функції корисності, що навантажує користувача.

Описані рекомендаційні системи прийнято класифікувати за факторами персоналізації й виділяють системи колаборативної (спільної) фільтрації, контентно-орієнтовані системи, системи основані на знаннях, на ступені корисності, на соціально-демографічних показниках і гібридні системи (рис. 3).

Гібридні системи містять методи, які поєднують фільтрацію на основі спільного використання та контенту, а також інші методи рекомендації. У [27] наведено основні стратегії гібридних рекомендацій:

- перемикання між методами;
- одночасне використання різних рекомендацій (міксування);
- об'єднання декількох методів;

– уточнення іншими рекомендаційними системами;

– характеристики різних джерел даних рекомендацій виводяться разом у єдиний алгоритм рекомендації;

– одна модель рекомендацій використовується як інша;

– вихідні дані однієї системи застосовують як вхідні для іншої.

Загалом рекомендаційні системи застосовують знання, які можна отримати з профілю користувача, що складається з уподобань, інтересів і потреб та з самого маркетплейсу – дані про контент. Виходячи з цього можна розробити таксономію знань для створення інтелектуального додатку персоналізації (рис. 4).

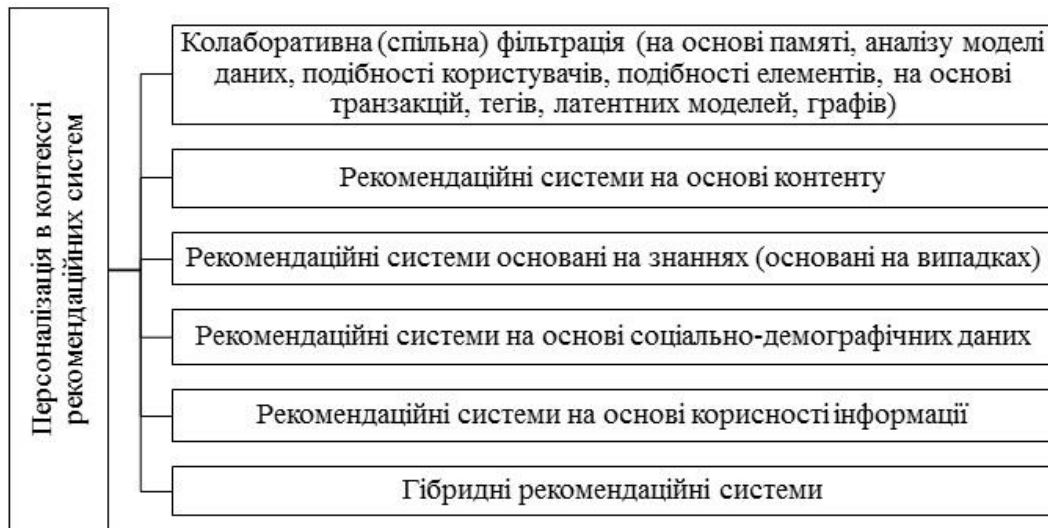


Рис. 3. Класифікація рекомендаційних систем



Рис. 4. Таксономія знань для створення інтелектуального додатку персоналізації

Таким чином розроблену таксономію знань можна представити як набір елементів:

$$TZ_{personal} = \langle Sk, Ik, Dk \rangle,$$

де Sk – множина соціальних знань, що включає підмножину $\{Si, Sb, Sdi, Sc\}$, де Si – соціальні інтереси дослідження предметної області; Sb – соціальна поведінка споживачів товарів чи послуг предметної області; Sdi – соціально-демографічні показники споживачів предметної області; підмножина $Sc = \{Scs, Scr, Scc\}$, де Sc – контекст предметної області, що використовується для соціологічних досліджень, Scs – оцінка контенту предметної області на основі соціологічних досліджень, Scr – рейтинг контенту предметної області на основі соціологічних досліджень, Scc – відгуки про контент досліджуваної предметної області;

Ik – множина індивідуальних знань про користувача, що включає в себе підмножину $\{Ia, Idi, Ir, Iiw\}$, де Ia – дії, які здійснював конкретний користувач; Idi – соціально-демографічні показники окремого користувача; Ir – вимоги та побажання конкретного користувача; Iiw – підмножина, що містить параметри інтересів конкретного користувача і складається з $\{Iiws, Iiwc, Iiwp, Iiwcl\}$, де $Iiws$ – оцінка контенту користувачем; $Iiwc$ – відгуки, що залишив користувач про контент; $Iiwp$ – веб-сторінка, з якої користувач прийшов та з якої завершив своє перебування в цифровому об’єкті; $Iiwcl$ – дії користувача відносно цифрового об’єкта (теплова карта, трекер подій);

Dk – множина знань про предметну область, що включає підмножину знань $\{Dc, Dd, Dck\}$, де Dc – контент деякої предметної області; Dd – характеристика (опис) контенту предметної області за атрибутами $\{Ddv, Ddp, Ddc\}$, де Ddv – числові значення, що

може приймати одиниця контенту, *Ddp* – параметри, які характеризують одиницю контенту і *Ddc* – обмеження для деякої одиниці контенту; *Dck* – знання, щодо специфічних особливостей одиниці контенту заданої предметної області.

Єдиним суттєвим недоліком більшості рекомендаційних систем можна виділити проблему «холодного старту», як для нових користувачів, так і для нових елементів контенту. «Холодний старт» – розрідженість даних, тобто відсутність початкових даних для надання рекомендацій, у випадку з користувачами – це відсутність реакції на контент (оцінка, мітка, рейтинг) та відсутність даних про самого користувача, а з елементами контенту – відсутність тих самих оцінок від користувачів, отже нікому їх рекомендувати. Але дану проблему можна вирішити, здійснивши кореляційно-регресійний аналіз для елементів предметної області або із застосуванням однокласової колаборативної фільтрації (one-class collaborative filtering), де є достатнім сам факт придбання одиниці товару.

Для збагачення даними застосовуються такі категорії даних:

– first party data – власні дані, які накопичив маркетплейс про своїх клієнтів, що зібрані в процесі їх взаємодії з маркетплейсом (чим цікавилися, які сторінки переглядали), а також дані з CRM-системи;

– second party data – це сторонні дані маркетплейсу, які можна придбати для збагачення своїх даних і отримати новий погляд на власних клієнтів. До уваги беруть кліки, соціальну активність (натискання кнопки «поділитись», лайки), інші поведінкові фактори;

– third party data – дані третіх сторін; їх джерело – сервіси email-розсилок, платіжні системи, сторонні веб-ресурси, дані лічильників з веб-ресурсів, облікові записи соціальних мереж, дані з DSP-платформ (Demand Side Platform) і RTB-системи (Real Time Bidding), сервіси обробки і зберігання даних про цільову аудиторію – DMP-платформи (Data Management Platform, платформа керування даними) [28].

Однією з новітніх технологій збору інформації про вибір, поведінку, звички та навіть емоції користувачів, з їх цифрових пристроїв є (фітнес-трекери, мобільні пристрої) – intelligent user interfaces [29], що потім обробляється методами штучного інтелекту (розпізнавання образів і ситуацій, мови, тексту, жестів, класифікація зображень) для формування персоналізованої рекламної пропозиції.

Таким чином проведений аналіз методів, моделей та алгоритмів рекомендаційних систем свідчить про наявність широкого арсеналу інструментів для створення рекомендаційних систем, що може стати підґрунтям для вдосконалення інформаційного забезпечення систем підтримки прийняття рішень та вдосконалення самого механізму формування рекомендацій користувачу в цифровому середовищі. Така система забезпечить стійкий розвиток та функціонування цифрових підприємств та відповідатиме вимогам сьогодення.

6. Висновки

1. Для задоволення потреб сучасного споживача в цифровому середовищі необхідно застосовувати новітні інформаційно-комунікаційні технології, рекомендаційні системи зокрема. Такі системи дозволяють персоналізувати контент та підвищити релевантність комерційних повідомлень.

2. Для формування персональних рекомендацій, необхідно пройти такі етапи:

– зібрати початкові дані про споживача та його дії відносно контенту та маркетплейсу;

– обробити отримані результати методами інтелектуального аналізу даних;

– здійснити прогнозування отриманих результатів та представити рекомендації окремому споживачу.

3. Рекомендаційні системи вирішують низку проблем, що пов'язані з задоволенням потреб споживачів відповідно до їх вподобань. До основних переваг можна віднести:

– формування актуальних комерційних пропозицій для споживачів, тим самим підвищується вартість кошика та обсяги продажів супутніх товарів;

– підвищення лояльності та прив'язаності до бренду;

– заохочення та зацікавленість до продукту та маркетплейсу.

4. Для вирішення проблеми персоналізації варто залучити додаткові галузі наук, такі як психологія, інформаційні технології, математика та статистика, що дозволять дослідити поведінку споживачів, виявити закономірності та потреби, а також прогнозувати реакцію споживача відносно сформованої рекомендації.

5. Світові гіганти цифрового бізнесу визнають високий потенціал у тому, що системи рекомендацій дозволяють збирати величезну кількість інформації щодо поведінки користувачів та їхніх дій в цифровому середовищі. Як показує практика такі інтелектуальні додатки задовольняють потреби споживачів та підвищують їх лояльність.

6. Аналіз різних типів рекомендаційних систем, їх методів та алгоритмів показав їх високу практичну значимість виходячи зі специфіки предметної області. Одні алгоритми виграють за часом, інші за точністю, але для підтримки маркетингових рішень саме в цифровому середовищі було вирішено обрати такі методи: методи розкладання сингулярних значень (SVD), метод найближчих сусідів, наївний класифікатор Байєса, байєсові мережі, асоціативні правила. Комплекс таких методів дозволяє створити інтелектуальні рекомендаційні системи, що пояснюють, переконують та прогнозують, тим самим розширюючи класичні механізми цифрового середовища.

До основних типів рекомендаційних систем прийнято відносити колаборативну фільтрацію, рекомендаційні системи на основі контенту та гібридні. В залежності від специфіки предметної області та вирішення проблеми відсутності початкових даних пропонується використовувати і інші типи рекомендаційних систем, що ґрунтуються на знаннях, на ступені корисності, на соціально-демографічних показниках, які можуть доповнювати, уточнювати, замінювати та збагачувати даними рекомендаційні

системи для надання більш вдалих, персоналізованих рекомендацій.

7. Для подальшого проектування інтелектуального додатку з персоналізації контенту було розроблено таксономію (онтологію) знань, де виділено такі фактори: множину соціальних знань, множину

індивідуальних знань про споживача та множину знань про предметну область.

8. Для розширення знань, щодо персональних рекомендацій доцільно використовувати спеціальні джерела знань цифрового простору, – RTB-систему, DSP- та DMP-платформи.

Література

1. Котлер Ф. Основы маркетинга. Москва: Прогресс, 1990. 672 с.
2. Блэкуэлл Р., Минард П., Энджел Дж. Поведение потребителей: уч. Санкт-Петербург: Питер, 2007. 944 с.
3. Алешина И. В. Поведение потребителей: уч. Москва: Экономистъ, 2006. 525 с.
4. Прокопенко О. В., Троян М. Ю. Поведінка споживачів: навч. пос. Київ: Центр учбової літератури, 2008. 176 с.
5. Kahneman D., Tversky A. Prospect Theory: An Analysis of Decision under Risk // *Econometrica*. 1979. Vol. 47, Issue 2. P. 263–292. doi: <http://doi.org/10.2307/1914185>
6. Tversky A., Kahneman D. Advances in prospect theory: Cumulative representation of uncertainty // *Journal of Risk and Uncertainty*. 1992. Vol. 5, Issue 4. P. 297–323. doi: <http://doi.org/10.1007/bf00122574>
7. Поведенческая экономика: современная парадигма экономического развития: монография / ред. Журавлёва Г. П., Манохина Н. В., Смагина В. В. Москва-Тамбов: Издательский дом ТГУ им. Г.Р. Державина, 2016. 340 с.
8. Поведение потребителей: уч. пос. Тамбов: Изд-во ФГБОУ ВПО «ТГТУ», 2012. 84 с.
9. Linden G., Smith B., York J. Amazon.com recommendations: item-to-item collaborative filtering // *IEEE Internet Computing*. 2003. Vol. 7, Issue 1. P. 76–80. doi: <http://doi.org/10.1109/mic.2003.1167344>
10. Amatriain X., Basilico J. Netflix Recommendations: Beyond the 5 stars (Part 1). Netflix Technology Blog. 2012. URL: <https://medium.com/netflix-techblog/netflix-recommendations-beyond-the-5-stars-part-1-55838468f429>
11. Corinna U. Use Cases of Recommendation Systems in Business – Current Applications and Methods. 2018. URL: <https://emerj.com/ai-sector-overviews/use-cases-recommendation-systems/>
12. Златьева Д. Опыт персонализации контента в Le Boutique: как интуиция побеждает machine learning. 2016. URL: <https://retailers.ua/news/tehnologii/5102-opyit-personalizatsii-kontenta-v-le-boutique-kak-intuitsiya-pobejdaet-machine-learning>
13. Джонс М. Т. Рекомендательные системы: Часть 1. Введение в подходы и алгоритмы. Принципы работы рекомендательных механизмов Интернета // IBM Corporation. 2014. URL: <https://www.ibm.com/developerworks/ru/library/os-recommender1/index.html>
14. Tilahun B., Awono C., Batchakui B. A Survey of State-of-the-art: Deep Learning Methods on Recommender System // *International Journal of Computer Applications*. 2017. Vol. 162, Issue 10. P. 17–22. doi: <http://doi.org/10.5120/ijca2017913361>
15. Korb D. Recommendation System Algorithms. Main existing recommendation engines and how they work. 2017. URL: <https://statsbot.co/blog/recommendation-system-algorithms/>
16. Tag-based collaborative filtering recommendation in personal learning environments / Chatti M. A. et. al. // *IEEE Transactions on Learning Technologies*. 2013. Vol. 6, Issue 4. P. 337–349. doi: <http://doi.org/10.1109/tlt.2013.23>
17. Пятикоп Е. Е. Исследование метода коллаборативной фильтрации на основе сходства элементов // *Информатика, кибернетика та обчислювальна техніка*. 2013. С. 109–114.
18. Machine Learning (8) – Recommender Engine: Collaborative Filtering. 2015. URL: oliviaklose.azurewebsites.net
19. Hofmann T., Puzicha J. Latent Class Models for Collaborative Filtering // *IJCAI '99 Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence*. San Francisco: Morgan Kaufmann Publishers Inc., 1999. P. 688–693.
20. Langseth H., Nielsen T. D. A latent model for collaborative filtering // *International Journal of Approximate Reasoning*. 2012. Vol. 53, Issue 4. P. 447–466. doi: <http://doi.org/10.1016/j.ijar.2011.11.002>
21. Shams B., Haratizadeh S. Graph-based collaborative ranking // *Expert Systems with Applications*. 2017. Vol. 67. P. 59–70. doi: <http://doi.org/10.1016/j.eswa.2016.09.013>
22. Various Methods of Using Content-Based Filtering Algorithm for Recommender Systems / Swathi S. et. al. // *International Journal of Innovative Research in Science, Engineering and Technology*. 2017. Vol. 6, Issue 2. P. 1606–1613.
23. Asanov D. Algorithms and Methods in Recommender Systems. Berlin: Institute of Technology, 2004. URL: https://www.snet.tu-berlin.de/fileadmin/fg220/courses/SS11/snet-project/recommender-systems_asanov.pdf
24. Lorenzi F., Ricci F. Case-based recommender systems: a unifying view // *ITWP'03 Proceedings of the 2003 international conference on Intelligent Techniques for Web Personalization*. Acapulco: Springer-Verlag, Berlin: Heidelberg, 2005. P. 89–113. doi: http://doi.org/10.1007/11577935_5
25. Marung U., Theera-Umporn N., Auephanwiriyakul S. Top-N Recommender Systems Using Genetic Algorithm-Based Visual-Clustering Methods // *MDPI Reference List and Citations Style Guide*. Basel, 2016.
26. Feng D. Utility-based Recommender Systems Using Implicit Utility and Genetic Algorithm // *International Conference on Mechatronics, Electronic, Industrial and Control Engineering*. 2015. P. 860–864. doi: <http://doi.org/10.2991/meic-15.2015.197>
27. Brusilovsky P., Kobsa A., Nejdl W. The adaptive web: methods and strategies of web personalization. Heidelberg: Springer-Verlag Berlin, 2007. 762 p.
28. Dent A. An In-Depth Look At Second-Party Data For Digital Marketers. 2015. URL: <https://marketingland.com/second-party-data-digital-marketers-128254>
29. Sonntag D. ISMAR Tutorial on Intelligent User Interfaces, Proceedings of ISMAR 2015. Related DFKI Technical Report. 2015.

Рекомендовано до публікації д-р екон. наук Степаненко О. П.

Дата надходження рукопису 09.10.2018

Котлярова Юлія Олександрівна, аспірант, асистент, кафедра інформаційних систем в економіці, Навчально-науковий інститут «Інститут інформаційних технологій в економіці» Державного вищого навчального закладу «Київський національний економічний університет імені Вадима Гетьмана», пл. Львівська, 14, м. Київ, Україна, 04053
E-mail: julkot@ukr.net