

УДК 004.89

## МАТЕМАТИЧНА МОДЕЛЬ РЕКОМЕНДАЦІЙНОГО СЕРВІСУ НА ОСНОВІ МЕТОДУ КОЛАБОРАТИВНОЇ ФІЛЬТРАЦІЇ

Кюльян Артур Геннадійович, Польгуль Тетяна Дмитрівна, Хазін Марат Борисович

Вінницький національний технічний університет, Україна

### Анотація

У даній роботі розроблено підхід до побудови сервісів рекомендації товарів для підвищення ефективності збуту в інтернет-магазинах на основі методу колаборативної фільтрації, представлено структуру рекомендаційної системи та описано критерії визначення схожості між користувачами для видачі найбільш оптимального рекомендаційного набору.

### Abstract

In this paper, the approach of building recommendation services based on collaborative filtering method for improving the effectiveness of sales in online stores was developed, the structure of recommendation system is presented and criterias of similarity between users for the puprose of most optimal recommendation set are described.

### Вступ

Компанії-гіганти створюють програмні комплекси для заохочування клієнтів. Прикладами таких систем є сайти знайомств, які допомагають людям швидше знаходити собі пару; комплекси, що прогнозують зміну цін на авіаквитки; і буквально всі жадають краще розуміти своїх клієнтів, щоб проводити цілеспрямовані рекламні кампанії. Компанія Amazon збільшила обсяги продажу на 3% завдяки таким системам (обіг товарів компанії 30 мільярдів доларів, 3% - 900 мільйонів). Засновники Google придумали новий підхід до ранжування результатів пошуку, який базується на використанні посилань з мільйонів сайтів. Це збільшило кількість обслуговуваних ним пошукових запитів до 85%. Компанії Netflix використання таких систем дозволило збільшити кількість користувачів і прокатів на 4% [1]. З кожним днем об'єми продажу інтернет-площадок збільшується. За прогнозами експертів за наступні роки вони збільшаться ще на сорок відсотків, витіснивши звичайні магазини. На відміну від магазинів, до яких звикли старші покоління, в інтернет-магазинах немає продавця, який допомагає покупцю підібрати товар. Таку задачу розв'язують рекомендаційні системи, значно підвищуючи збут продукції в інтернеті. Здатність збирати інформацію та наявність обчислювальних потужностей для її інтерпретації відкрили нові можливості співпраці і дозволили краще зрозуміти потреби користувачів і замовників. У зв'язку з вищевказаним, задача розробки програмних комплексів для надання рекомендацій у сфері споживання є актуальною і має важливе практичне значення.

У даній роботі запропоновано математичну модель сервісу для рекомендацій покупцям споживчих товарів. Структурна схема такої рекомендаційної системи приведена на рисунку 1.

Розроблена модель базується на модифікованому методі колаборативної фільтрації [2] і розв'язує багатокритеріальну задачу визначення рекомендацій для сервісу продажу.

Система складається з чотирьох модулів. Модуль збору інформації обробляє такі вхідні дані системи: геолокація, демографія, освіта, інтереси, персональні вподобання, статистика попередніх продажів тощо. Задачі категоризації та класифікації вхідних даних розв'язано завдяки модулю визначення схожості даних. Для вирішення цих задач використовується комбінована метрика схожості, яка формується на основі коефіцієнта косинусної схожості між двома векторами (див. формулу 1) та коефіцієнта Танімото (див. формулу 2) [3].

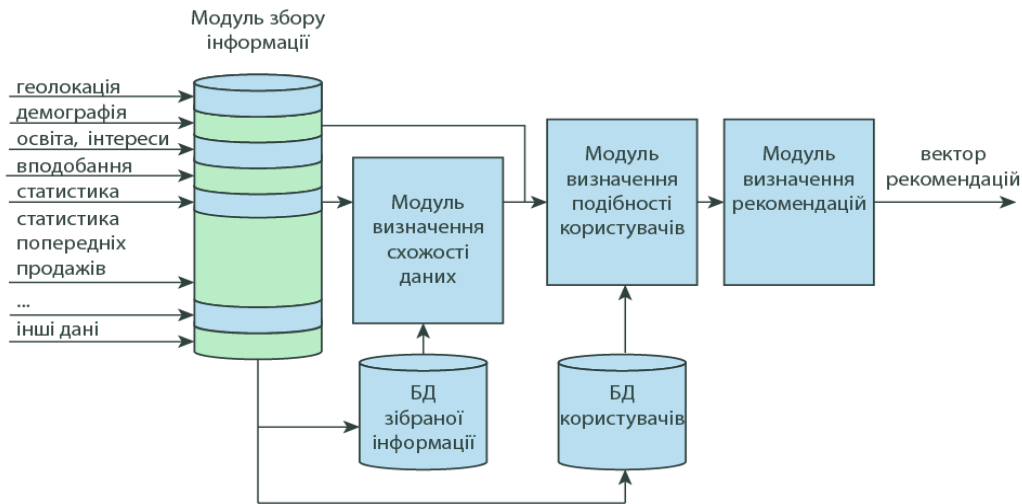


Рисунок 1 - Структурна схема рекомендаційної системи

Подібність смаків користувачів визначається в модулі визначення подібності користувачів за допомогою коефіцієнта кореляції Пірсона [4], використовуючи дані, які сформувалися у модулі визначення схожості товарів.

Оперуючи коефіцієнтами подібностей між клієнтами та коефіцієнтами подібності даних, модель рекомендацій обчислює зважену суму оцінок, на основі якої отримано вектор рекомендацій.

$$k = \cos(A, B) = \frac{A \cdot B}{AB} \quad (1)$$

де  $A, B$  - вектори, елементами яких є частоти появи окремих слів у заданому наборі інформації.

$$k = \frac{N_c}{(Na + Nb - N_c)} \quad (2)$$

де  $Na$  – кількість елементів в певному наборі даних користувача  $A$ ,  $Nb$  – кількість елементів у наборі даних користувача  $B$ ,  $N_c$  – кількість елементів в їх перетині.

На основі описаної моделі розроблено програмний комплекс сервісу рекомендацій. Розроблену систему використано на сайті [nadiupo.com.ua](http://nadiupo.com.ua). Застосування сервісу дало можливість збільшити обсяг продажу за другий квартал 2012 року на 5% порівняно з обсягами продажу у другому кварталі 2011 року.

Отже, у даній роботі розроблено модель рекомендаційної системи на основі методу колаборативної фільтрації. Результати роботи впроваджено у комерційному проекті. Також проведено аналіз моделі та доведено її адекватність, показано переваги над існуючими методами.

#### Список використаних джерел:

1. Сегаран Т. Программируем коллективный разум. / Т. Сегаран; пер. с англ. А. Слинкина – СПб: Символ-Плюс, 2008. – 368 с., ил. – ISBN 5-93286-119-3.
2. Alexander T. Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions / T. Alexander – IEEE Trans. On Knowledge and Data Engineering, vol. 17, no. 6, June 2005, pp. 734-749.
3. Linden G.D. Collaborative Recommendations Using Item-to-Item Similarity Mappings / G.D.Linden, J.A.Jacobi, E.A Benson – US Patent 6,266,649 (to Amazon.com), Patent and Trade-mark Office, Washington, D.C., 2001.
4. Sarwar B.M. Item-Based Collaborative Filtering Recommendation Algorithms / B.M. Sarwar, – 10th Int'l World Wide Web Conference, ACM Press, 2001, pp. 285-295.