

**Худік Б.О.***Державний університет інформаційно-комунікаційних технологій, Київ*

## МЕТОД АВТОМАТИЗОВАНОГО ФОРМУВАННЯ РЕКОМЕНДАЦІЙ ДЛЯ ЗДОБУВАЧІВ ОСВІТИ НА ОСНОВІ ГІБРИДНОГО ПІДХОДУ

**Анотація:** Найбільш популярними сферами використання рекомендаційних систем є платформи, що містять фільми, музику, книги та інші об'єкти, що є доволі однотипними за своїми характеристиками. Для таких систем розроблено доволі багато як загальних так і специфічних методів побудови рекомендацій. В статті розглянуто метод автоматизованого формування рекомендацій для здобувачів освіти на основі гібридного підходу, який передбачає врахування контенту об'єктів інтересу здобувача. Проведено аналіз існуючих підходів до побудови рекомендацій та виявлено ключові проблеми. Визначено, що врахування контенту об'єктів інтересу та користувачів дозволяє підвищити точність рекомендацій. В статті наведено типові види рекомендацій, які потребують студенти в освітньому процесі. Описано джерела для формування рекомендацій, зокрема, явні оцінки об'єктів інтересу, отримані на основі обов'язкових опитувань студентів; неявні оцінки, отримані на основі аналізу про поведінку користувачів в системі та контент об'єктів інтересу і користувачів. Досліджено характеристики об'єктів інтересу та визначено, що вони обов'язково мають короткі або довгі текстові описи. Визначено ключові етапи методу автоматизованого формування рекомендацій з урахуванням специфіки домену предметної галузі. Для підвищення точності рекомендацій запропоновано перед застосуванням колаборативної фільтрації виконувати попередній відбір потенційно цікавих для користувача об'єктів на основі методів обробки текстової інформації (Natural Language Processing). Це дозволить врахувати, в тому числі, семантичну близькість інтересів користувача та об'єктів в системі навіть при невеликих обсягах текстової інформації про об'єкти інтересу та користувачів. Додатково точність рекомендацій може бути підвищена за рахунок використання явних оцінок об'єктів інтересу, отримання яких на рівні закладу освіти забезпечується регулярними заходами з моніторингу якості освіти.

**Ключові слова:** рекомендаційна система, колаборативна фільтрація, гібридний підхід, Natural Language Processing, інформаційна технологія, база даних, автоматизація, підтримка прийняття рішень.

**Khudik B.O.***State University of Information and Communication Technologies, Kyiv*

## A METHOD OF AUTOMATED GENERATION OF RECOMMENDATIONS FOR EDUCATION APPLICANTS BASED ON A HYBRID APPROACH

**Abstract:** The most popular areas of use of recommender systems are platforms containing movies, music, books, and other objects that are fairly similar in their characteristics. Quite a lot of both general and specific methods of building recommendations have been developed for such systems. The article discusses the method of automated generation of recommendations for education applicants based on a hybrid approach, which involves taking into account the content of the objects of interest of the student. An analysis of the existing approaches to the construction of the recommendation was carried out and key problems were identified. It was determined that taking into account the content of objects of interest and users allows to increase the accuracy of the recommendation. The article provides typical types of recommendations that students need in the educational process. Sources for the formation of recommendations are described, in particular, explicit evaluations of objects of interest obtained on the basis of mandatory student surveys; implicit evaluations obtained on the basis of analysis of the behavior of users in the system and the content of

*objects of interest and users. The characteristics of the objects of interest were studied and it was determined that they necessarily have short or long text descriptions. The key stages of the method of automated formation of the recommendation are determined, taking into account the specifics of the domain of the subject area. In order to increase the accuracy of the recommendations, it is proposed to perform a preliminary selection of potentially interesting objects for the user based on Natural Language Processing methods before applying collaborative filtering. This will allow taking into account, among other things, the semantic proximity of the user's interests and objects in the system even with small amounts of textual information about objects of interest and users. In addition, the accuracy of the recommendation can be increased due to the use of explicit assessments of objects of interest, which are obtained at the level of the educational institution by regular measures to monitor the quality of education.*

**Keywords:** *recommendation system, collaborative filtering, hybrid approach, Natural Language Processing, information technology, database, automation, decision support.*

## **1. Вступ**

Підтримка освітнього процесу засобами сучасних інформаційних технологій є невід'ємною частиною організації взаємодії з усіма стейкхолдерами вищої освіти. Суттєвий поштовх до використання дистанційних технологій в Україні в рамках формальної освіти надали карантинні заходи, зумовлені епідеміологічною ситуацією (пандемією коронавірусу SARS-CoV-2), що дало змогу закладам вищої освіти використовувати цей досвід при організації дистанційного навчання в період воєнного стану. В більшості випадків сферою застосування інформаційних технологій залишаються процеси, які безпосередньо пов'язані з проведенням навчальних занять в асинхронному чи синхронному форматі та оцінюванням студентів [1]. Практика перенесення організаційних питань в онлайн також має успішні приклади реалізації починаючи від звичайних опитувань студентів і закінчуючи реалізацією значної частини функцій деканату засобами спеціалізованих інформаційних систем [2].

Однак, відсутність чи суттєве зменшення «живого» спілкування та перенесення більшості процесів в онлайн створює іншу проблему – велика кількість інформаційних ресурсів для аналізу та необхідність самостійно приймати рішення щодо оцінки релевантності цих ресурсів, викликають у користувачів таких систем інформаційне перенавантаження. Усунення цієї проблеми потребує створення методів та засобів для автоматизованого формування рекомендацій, які допомогли б кінцевому користувачеві орієнтуватися в інформаційних ресурсах, які необхідні йому для успішної реалізації своїх задач в межах освітнього процесу.

Гібридний підхід до формування рекомендації передбачає використання як колаборативного підходу, що ґрунтується на аналізі взаємодії користувачів з об'єктами інтересу, так і на методах, що дозволяють врахувати характеристики (контент) об'єктів інтересу та самих користувачів.

## **2. Аналіз останніх досліджень і постановка проблеми**

Рекомендаційні системи знайшли широке розповсюдження в сфері електронної комерції, соціальних мережах, що обумовлено їх доволі високою ефективністю при формуванні пропозицій для потенційних покупців/клієнтів компаній чи побудові рекламних кампаній або вибудові цифрових зв'язків між користувачами.

Робота [3] присвячена огляду рекомендаційних систем, їх типів, особливостей застосування та типових проблем побудови і використання. Розглядаються методи, які використовуються в рекомендаційних системах, включаючи фільтрацію на основі вмісту, колаборативну (соціальну) фільтрацію, послідовну, сеансову фільтрацію та інші. Наведена загальна таксономія методів, що використовуються в рекомендаційних системах, а також метрики оцінки їх ефективності. Автори проводять аналіз наборів даних та особливостей доменів декількох найбільш популярних предметних галузей використання рекомендаційних

систем, серед яких в якості ТОП-5 галузей виділено фільми, книги, їжа, музика та деякі продукти чи послуги. Доводиться користувачів різноманітних Інтернет-систем. Розглядаються різноманітні проблеми рекомендаційних систем та пропонуються деякі рішення для їх усунення з метою підвищення точності надання рекомендації.

В роботі [4] досліджуються проблеми сучасних рекомендаційних систем, в тому числі, пов'язані з холодним стартом – ситуацією, в якій не накопичено достатньої кількості даних для коректної роботи системи. Для усунення цієї проблеми рекомендується використовувати гібридний підхід до побудови рекомендації, який враховує контекстну інформацію.

Конкретні методи та підходи до формування рекомендацій наведено в роботах [5-8].

Інтелектуальна система рекомендацій, яка розуміє уподобання користувача, виявляє зміни в уподобаннях користувача з часом, передбачає невідомі уподобання користувача та досліджує адаптивні механізми для забезпечення інтелектуальних дій у складному та змінному середовищі описується в роботі [5]. В результаті аналізу впровадження рекомендаційної системи в банківській сфері автори доходять до висновку, що існуючі рекомендаційні системи мало або зовсім не використовують знання експертів галузі в процесі адаптації до нової ситуації, часто не в змозі адекватно передбачити переваги того чи іншого продукту, який обрав користувач/клієнт, не підтримують збирання даних та аналітику діяльності користувачів і не використовують ці дані для надання оперативних рекомендацій. Запропонована в статті система когнітивних рекомендацій характеризується тим, що вона керується знаннями, дозволяє імітувати роботу експертів на основі методів краудсорсингу, використовує технології штучного інтелекту та машинного навчання для розуміння великих даних та дозволяє зрозуміти особистість кінцевих користувачів і проаналізувати їхню поведінку та ставлення з протягом деякого часу.

В роботі [6] розглянуто один із найбільш популярних методів реалізації рекомендаційних систем – метод колаборативної фільтрації та його реалізації: фільтрація на основі класів, на основі пам'яті та на основі моделі, наводяться приклади реалізації різновидів методів та їх порівняльна характеристика. Суттєвою проблемою рекомендаційних систем автори вважають різноманітність предметних галузей, що ускладнює або унеможливорює точний прогноз з використанням однієї й тієї ж техніки побудови рекомендації для різних програм.

В статті [7] пропонується гібридна система рекомендацій фільмів і метод оптимізації, заснований на зваженій класифікації та алгоритмі колаборативної фільтрації користувачів. Розріджена лінійна модель використовується як базова модель рекомендацій, локальна модель рекомендацій навчається на основі кластеризації користувачів, а персоніфікована рекомендація фільмів із топ-N реалізується шляхом об'єднання з моделлю зваженої класифікації. Згідно з уподобаннями категорії предметів, оціночна матриця перетворюється на низькорозмірну, щільну матрицю переваг категорії предметів, отримується кілька центрів кластерів, обчислюється відстань між цільовим користувачем і кожним центром кластеру, а цільовий користувач класифікується на найближчий кластер. Після цього алгоритм колаборативної фільтрації використовується для прогнозування балів товарів без оцінки цільового користувача для формування списку рекомендацій. Елементи кластеризуються за допомогою параметра категорії елемента, а багатовимірний рейтинговий матрицю перетворюється на низьковимірну матрицю переваг категорії елемента, що додатково зменшує розрідженість даних.

Питанням використання рекомендаційних систем в галузі освіти присвячено роботу [8]. Розглядається задача підтримки користувача у виборі навчального курсу в системі електронного навчання. Для формування рекомендації пропонується використовувати три

основні методи: рекомендації на основі вмісту, на основі колаборативної фільтрації та на основі знань.

### 3. Мета і задачі дослідження

Метою даної публікації є дослідження питань автоматизованого формування рекомендацій для здобувачів освіти засобами інформаційних технологій, зокрема, з використанням рекомендаційних систем на основі гібридного підходу. Для досягнення поставленої мети вирішуються такі завдання:

- визначення видів рекомендацій та джерел для їх формування;
- розробка методу автоматизованого формування рекомендації на основі гібридного підходу.

### 4. Результати дослідження

На відміну від більшості існуючих рекомендаційних систем, де об'єктом інтересу користувачів виступають ресурси якогось одного виду (наприклад, фільми, музика), об'єкти інтересу здобувачів освіти, відносно яких можна побудувати рекомендацію, утворюють декілька окремих груп, кожна з яких має свої специфічні характеристики. До таких груп можна віднести:

- 1) дисципліни вільного вибору здобувача;
- 2) літературні та інші інформаційні джерела, які студент може використовувати в процесі навчання та самоосвіти;
- 3) керівники курсових, дипломних робіт, практик, гуртків тощо;
- 4) тематика курсових, дипломних робіт, творчих проєктів тощо.

Для формування рекомендацій відносно зазначених об'єктів можуть використовуватись:

- результати цільових опитувань/оцінювань відносно об'єктів інтересу (отримання явної оцінки користувачів);
- результати аналізу інформації про поведінку користувачів відносно об'єкту інтересу та їх взаємодії (отримання неявної оцінки користувачів);
- результати аналізу контенту об'єкту інтересу з метою визначення рівня релевантності інтересів користувача та характеристик об'єкту інтересу.

Етапи методу автоматизованого формування рекомендації на основі гібридного підходу можуть бути описані алгоритмом на рис.1. Для побудови рекомендації на першому кроці визначається об'єкт інтересу користувача. Інформація про кожен з об'єктів інтересу повинна бути представлена такими параметрами, які дозволили б реалізацію наступних кроків методу.

Особливістю об'єктів інтересу здобувача є те, що для кожної групи наявний короткий чи повний текстовий опис, що характеризує об'єкт. Для побудови рекомендації вимоги до такого текстового опису визначаються наборами даних, які можна сформувати в умовах реального закладу освіти:

- для дисциплін вільного вибору здобувача мінімальний контент повинен складатись з назви та анотації дисципліни, в ідеальному варіанті – повного опису вмісту дисципліни у формі робочої програми чи силабусу;
- інформаційні джерела можуть бути представлені безпосередньо текстовим контентом або метайнформацією про них (автори, ключові слова тощо);
- дані про керівників повинні містити текстові описи наукових та професійних інтересів викладача та/або список публікацій, проєктів тощо;
- курсові, дипломні роботи, творчі проєкти можуть бути представлені короткими описами, що включають лише назву роботи, але бажаним варіантом представлення, що

дозволить отримати більш точну рекомендацію, є наявність анотації чи повного текстового опису роботи/проекту.

Критерії відбору та порогові значення для подальшого відбору можуть коригуватись в залежності від групи об'єктів або користувацьких налаштувань. Оскільки кожен з визначених об'єктів інтересу містить деякий текстовий опис, на першому етапі формування рекомендації застосовуються методи обробки текстової інформації (Natural Language Processing) для вилучення даних про об'єкти інтересу. Перелік даних, що вилучаються, залежить від типу об'єкту. Слід зазначити, що для зменшення часу формування рекомендації, ключова інформація про об'єкти інтересу, що використовується в подальших етапах методу, може бути сформована заздалегідь та оновлюватись тільки у випадках зміни контенту об'єктів інтересу чи додаванні/видаленні об'єкту.

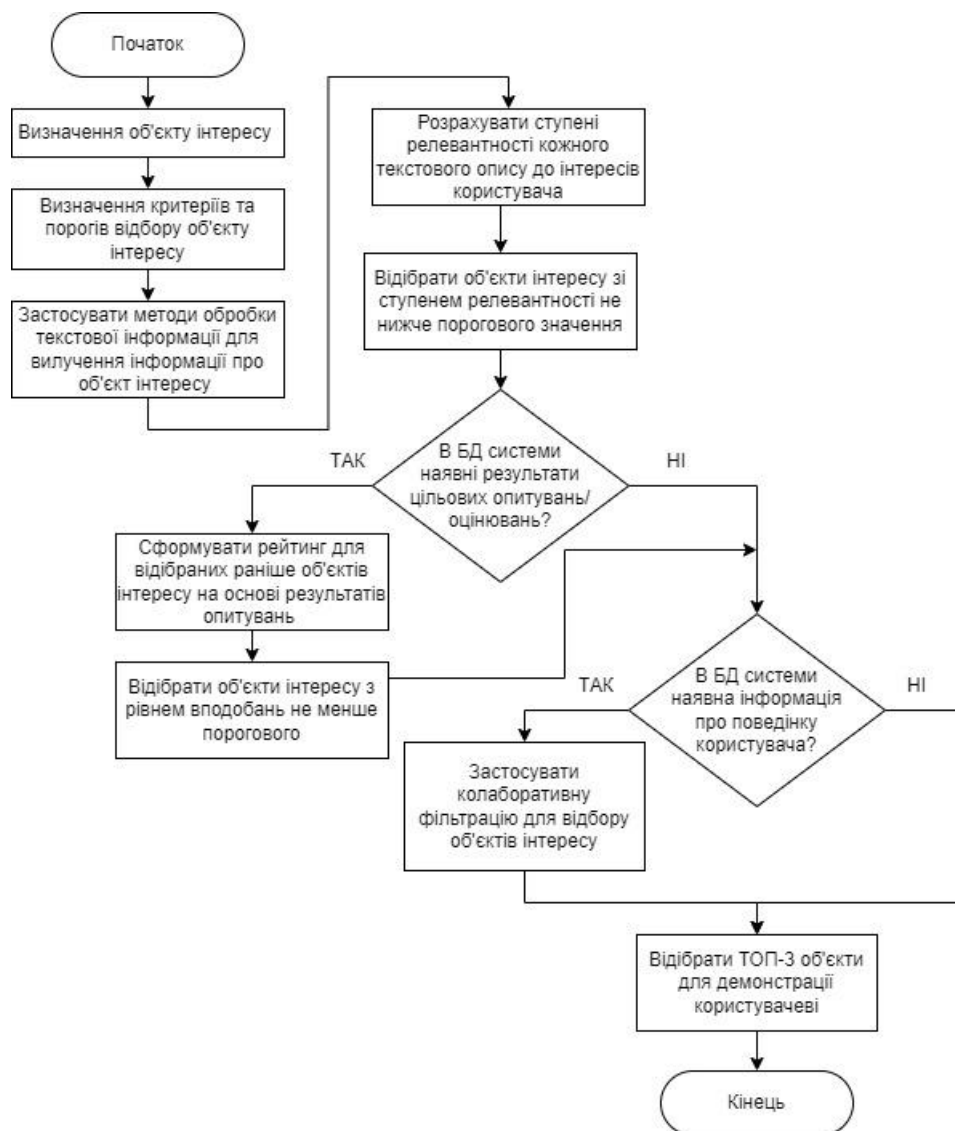


Рис.1. Етапи методу автоматизованого формування рекомендації на основі гібридного підходу

Для розрахунку ступеня релевантності об'єкту до інтересів користувача можуть використовуватись різні метрики подібності. Також, сама релевантність може трактуватись різними способами: це може бути обчислення семантичної подібності (наприклад, з використанням методів машинного навчання) або відстань між векторизованими

представленнями об'єкта та інформації про інтереси користувача (наприклад, косинусна відстань між вектором пошукового запиту користувача та вектором об'єкту інтересу).

Після етапу попереднього відбору об'єктів інтересу на основі їх контенту система повинна використовувати дані взаємодії користувачів з об'єктами. Для освітньої сфери це можуть бути взаємодії двох видів:

– явні у вигляді цільових опитувань користувачів про об'єкти інтересу (оцінка студентами якості викладання дисциплін, рейтинги викладачів, оцінки підручників, посібників тощо);

– неявні, які отримуються за результатами моніторингу дій користувачів в системі, що містить об'єкти інтересу (наприклад, відвідування сторінки дисципліни в системі дистанційного навчання, відвідування сторінки підручника на сайті бібліотеки і т.п.).

Використання даних цільових опитувань дозволяє сформувавши попередню користувацьку оцінку для переважної більшості об'єктів, оскільки такі цільові опитування є обов'язковим елементом моніторингу якості освітнього процесу. Виключенням можуть бути такі об'єкти інтересу, які з'являються в освітньому процесі вперше (новий викладач, нова дисципліна, новий підручник, новий напрямок досліджень тощо). Однак, завдяки етапу відбору на основі аналізу вмісту текстового контенту для цих об'єктів можна отримати деяку рейтингову оцінку, співставивши їх зі схожими об'єктами відповідно до обраної метрики схожості.

## 5. Висновки

Запропонований метод автоматизованого формування рекомендації на основі гібридного підходу дозволяє врахувати особливості характеристик об'єктів інтересу відповідно до предметної галузі формування рекомендації для здобувачів освіти. Застосування попередньої фільтрації об'єктів інтересу з використанням методів обробки текстової інформації дозволяє скоротити розмір вектору потенційно корисних об'єктів. Специфічні для освітнього процесу цільові опитування використовуються з тою ж метою, що дозволяє в подальшому суттєво зменшити розміри матриць колаборативної фільтрації. Також, рейтингування об'єктів інтересу на основі цільових опитувань дозволяє зменшити проблему «холодного старту», оскільки кількість об'єктів, про які немає рейтингової інформації в системі, є дуже низькою, і навіть для таких об'єктів можна легко отримати попередню оцінку на основі аналогічних оцінок схожих з ними об'єктів.

## Список використаної літератури

1. Ткаченко, Л. В., Хмельницька, О. С. (2021). Особливості впровадження дистанційного навчання в освітній процес закладу вищої освіти. Педагогіка формування творчої особистості у вищій і загальноосвітній школах. № 75, Т. 3. С.91-96. Doi: <https://doi.org/10.32840/1992-5786.2021.75-3.18>
2. Нікіфорова, Т. Д., Капшук, О. А., & Нечепуренко, Д. (2022). Навчальна та організаційна робота деканату в умовах воєнного стану. Матеріали II Форуму академічної спільноти «Освіта в умовах війни: реалії, виклики та шляхи подолання». С.41-43.
3. Patel, Dhruval & Patel, Foram & Chauhan, Uttam. (2023). Recommendation Systems: Types, Applications, and Challenges. 2210-142. 10.12785/ijcds/130168.
4. Мелешко Є. В. Проблеми сучасних рекомендаційних систем та методи їх рішення / Є. В. Мелешко // Системи управління, навігації та зв'язку. - 2018. - Вип. 4. - С. 120-124. - Режим доступу: [http://nbuv.gov.ua/UJRN/suntz\\_2018\\_4\\_25](http://nbuv.gov.ua/UJRN/suntz_2018_4_25)
5. Beheshti, A., Yakhchi, S., Mousaeirad, S., Ghafari, S. M., Goluguri, S. R., & Edrisi, M. A. (2020). Towards cognitive recommender systems. Algorithms, 13(8), 176.

6. Raghuwanshi, Sandeep & Pateriya, R.. (2019). Collaborative Filtering Techniques in Recommendation Systems. 10.1007/978-981-13-6347-4\_2.
7. Yuan, Zhenning & Lee, Jong & Zhang, Sai. (2021). Optimization of the Hybrid Movie Recommendation System Based on Weighted Classification and User Collaborative Filtering Algorithm. Complexity. 2021. 1-13. 10.1155/2021/4476560.
8. Zhang, Q., Lu, J., & Zhang, G. (2021). Recommender Systems in E-learning. Journal of Smart Environments and Green Computing, 1(2), 76-89.

### References

1. Tkachenko, L. V., & Khmelnytska, O. S. (2021). Peculiarities of the implementation of distance learning in the educational process of a higher education institution. Pedagogy of creative personality formation in higher and secondary schools. No. 75, Vol. 3. P.91-96. Doi: <https://doi.org/10.32840/1992-5786.2021.75-3.18>
2. Nikiforova, T. D., Kapshuk, O. A., & Nechepurenko, D. (2022). Educational and organizational work of the dean's office under martial law. Materials of the II Forum of the academic community "Education in conditions of war: realities, challenges and ways of overcoming". P.41-43.
3. Patel, Dhruval & Patel, Foram & Chauhan, Uttam. (2023). Recommendation Systems: Types, Applications, and Challenges. 2210-142. 10.12785/ijcids/130168.
4. Meleshko E. V. Problems of modern recommendation systems and methods of their solution / E. V. Meleshko // Management, navigation and communication systems. - 2018. - Issue 4. - P. 120-124. - Access mode: [http://nbuv.gov.ua/UJRN/suntz\\_2018\\_4\\_25](http://nbuv.gov.ua/UJRN/suntz_2018_4_25)
5. Beheshti, A., Yakhchi, S., Mousaeirad, S., Ghafari, S. M., Goluguri, S. R., & Edrisi, M. A. (2020). Towards cognitive recommender systems. Algorithms, 13(8), 176.
6. Raghuwanshi, Sandeep & Pateriya, R.. (2019). Collaborative Filtering Techniques in Recommendation Systems. 10.1007/978-981-13-6347-4\_2.
7. Yuan, Zhenning & Lee, Jong & Zhang, Sai. (2021). Optimization of the Hybrid Movie Recommendation System Based on Weighted Classification and User Collaborative Filtering Algorithm. Complexity. 2021. 1-13. 10.1155/2021/4476560.
8. Zhang, Q., Lu, J., & Zhang, G. (2021). Recommender Systems in E-learning. Journal of Smart Environments and Green Computing, 1(2), 76-89.