УДК 004.738.5

ДОСЛІДЖЕННЯ МЕТОДІВ ПЕРСОНАЛІЗАЦІЇ ОГОЛОШЕНЬ У МАРКЕТПЛЕЙСІ НЕРУХОМОСТІ ЗА ІСТОРІЄЮ ПЕРЕГЛЯДІВ

Юшко К.В. [2]; Книрік Н.Р.,к.т.н.,доц [2]
[1,2]Національний університет кораблебудування імені адмірала Макарова
[1,2]Україна, Миколаїв
[1]kostya.dank47@gmail.com; [2]natalya.knyrik@nuos.edu.ua

Анотація. *Розглянуто персоналізацію видачі оголошень у централізованому маркетплейсі нерухомості як задачу ранжування за історією переглядів. Запропоновано модель відповідності оголошення профілю інтересів, у якій профіль числових ознак описано ядерною оцінкою густини, а ваги критеріїв підбирає генетичний алгоритм. Розроблено програмний модуль із трьома режимами видачі, зокрема для користувача без реєстрації.*
Ключові слова: *нерухомість; маркетплейс; персоналізація; рекомендаційна система; генетичний алгоритм; історія переглядів.*

Вступна частина. Вибір об’єкта нерухомості потребує порівняння значної кількості альтернатив за кількома критеріями одночасно. У централізованих маркетплейсах зростання кількості оголошень призводить до інформаційного перевантаження користувача [1], а стандартний каталог із сортуванням і фільтрами не враховує його індивідуальних вподобань [2]. Для нерухомості персоналізацію ускладнено: кожне оголошення описує один унікальний об’єкт, який після продажу залишає каталог, історія окремого користувача коротка, а частина відвідувачів не має облікового запису [2, 3]. Водночас історія переглядів є неявним сигналом, що не потребує від користувача явного оцінювання об’єктів [4].

Публічно доступні матеріали провідних українських платформ не розкривають відкритої моделі персоналізації. OLX Group заявила про платформу Otodim, яка за допомогою штучного інтелекту формує персоналізовані результати пошуку, проте принципи її роботи не описано [5]; в офіційному описі DOM.RIA акцент зроблено на перевірених об’єктах, фільтрах і тривимірній карті [6], а Rieltor.ua подано як інформаційний портал про нерухомість [7]. Оскільки внутрішні алгоритми платформ недоступні, актуальність задачі полягає не у відсутності персоналізації як такої, а у відсутності відкритої відтворюваної моделі, придатної для вбудовування в існуючу платформу та експериментальної оцінки.

Мета роботи. Дослідити методи персоналізації оголошень у маркетплейсі нерухомості та розробити математичну модель і програмний модуль, що формують персональну видачу за історією переглядів користувача з адаптивним підбором ваг критеріїв.

Основна частина. Об’єктом дослідження є процес формування персоналізованої видачі оголошень, предметом - методи та моделі оцінювання релевантності оголошень за історією переглядів. Порівняно три підходи. Контентна фільтрація прозора й незалежна від інших користувачів, але має статичні ваги, а усереднення ознак спотворює інтерес при перегляді різнобюджетних об’єктів. Колаборативна фільтрація чутлива до холодного старту [2], що критично для ринку унікальних і швидко зникаючих об’єктів. Третій підхід застосовує генетичний алгоритм (ГА) [8] не як окрему рекомендаційну модель, а для адаптивного підбору ваг критеріїв у контентній моделі. Він працює з короткою історією одного користувача, не потребує даних інших користувачів і зберігає інтерпретованість оцінки, тому його обрано за основу.

Нехай C - множина кандидатів (оголошення зі статусом AVAILABLE), кожне з яких описується ціною x1, площею x2, рейтингом агента x3 і населеним пунктом ℓ, а H = (v1, …, vm) - історія переглядів від найновішого до найстарішого, m ≤ 20. Оцінку відповідності оголошення p профілю інтересів задано зваженою сумою:

$$S(\mathrm{p}|\mathrm{H},\mathrm{w}) = \sum_{k=1}^{4} w_{\mathrm{k}} s_{\mathrm{k}}(p) \quad (1)$$

де w = (w1, …, w4) - вектор ваг критеріїв (ціна, площа, рейтинг, локація) на одиничному симплексі (wk ≥ 0, Σwk = 1); sk(p) ∈ [0, 1] - оцінка за критерієм k, що залежить від H. Оцінка є лінійною комбінацією, тому внесок кожного критерію відомий. Профіль числових критеріїв (k = 1, 2, 3) описано ядерною оцінкою густини з гаусовим ядром [9] за мінімаксно нормалізованими значеннями x̃k, шириною ядра h = 0,15 та вагою давності ρj-1 (ρ = 0,9) для j-го за новизною перегляду:

$$D_{\mathrm{k}}(p) = \sum_{j=1}^{m} p^{j-1} \exp\left(-\frac{1}{2}\left(\frac{x_{\mathrm{k(p)}} - x_{\mathrm{k}(v_{\mathrm{j}})}}{\mathrm{h}}\right)^2\right) \quad (2)$$

$$S_{\mathrm{k}}(p) = \frac{D_{\mathrm{k}}(p)}{max D_{\mathrm{k}}(q)}, q \epsilon \mathrm{C} \quad (3)$$

На відміну від середнього, така оцінка зберігає кілька піків: для історії зі значеннями 10 000 і 100 000 високі оцінки отримують оголошення поблизу кожного з них. Для локації частку населеного пункту в історії (з вагами давності) порівняно з часткою в каталозі зі згладжуванням Лапласа (α = 1); відношення цих часток liftℓ перетворюється на оцінку (для порожньої історії всі sk = 0,5):

$$a_{\ell} = \frac{lift_{\ell}}{1 + lift_{\ell}}, \quad S_{(4)}(\mathrm{p}) = \frac{a_{\ell(\mathrm{p})}}{\max a_{\ell(\mathrm{q})}}, \mathrm{q} \epsilon \mathrm{C} \quad (4)$$

Хромосомою є вектор ваг w. Початкова популяція (P = 30) містить базові ваги w0 = (0,30; 0,20; 0,20; 0,30), решту особин згенеровано рівномірно на симплексі. Якість ваг оцінюється за схемою leave-one-out: кожне переглянуте оголошення vj по черзі приховується, профіль будується за рештою історії, а vj ранжується серед кандидатів. Нехай J – множина переглядів, для яких приховане оголошення входить до кандидатів і не повторюється в решті історії; тоді функція пристосованості:

$$F(\mathrm{w};\mathrm{H}) = \frac{1}{|J|} \sum_{j \epsilon J} \frac{1}{rank_{\mathrm{j}}(w)} - \lambda ||\mathrm{w} - w^0||^2 \quad (5)$$

де $rank_{\mathrm{j}}$ (w) - місце оголошення vj. Перший доданок - середній зворотний ранг (MRR), другий - регуляризація, що обмежує відхилення від w0 і зменшує ризик перенавчання на короткій історії (λ = 0,15). Рангова метрика є кусково-сталою функцією ваг, тому градієнтні методи для неї неприйнятні, а еволюційний пошук виправданий [8]. Застосовано турнірну селекцію (розмір турніру 3), арифметичне схрещування з ймовірністю 0,9, гаусову мутацію (σ = 0,08) з нормуванням ваг до суми одиниці та елітизм (дві найкращі особини); алгоритм виконує G = 25 поколінь. Адаптація вмикається за m ≥ 4 і |J| ≥ 2, інакше використовуються w0. Генератор випадкових чисел ініціалізується послідовністю ідентифікаторів історії, тому однакова історія завжди дає однакову видачу. Модуль реалізовано як самостійний компонент, який можна вбудувати в існуючу платформу через REST API: клієнт на Vue.js, сервер на Java 21 зі Spring Boot та JPA/Hibernate, база даних PostgreSQL у Docker-контейнері. Рекомендаційне ядро виділено в клас без залежностей від фреймворків, тому його можна переносити між платформами. Джерело історії визначає один із трьох режимів видачі: для

ідентифікованого користувача використовується до 20 останніх переглядів із бази даних; для гостя - до 10 переглядів, що зберігаються в локальному сховищі браузера й передаються в запиті, тож персоналізація працює без реєстрації; для користувача без історії - останні перегляди всіх користувачів, тобто глобальний тренд. Клієнту повертаються до 12 оголошень із числовою оцінкою відповідності.

Висновки. Обґрунтовано, що специфіка нерухомості (унікальні об'єкти, коротка й часто анонімна історія) обмежує застосування колаборативної фільтрації та робить доцільним поєднання контентної моделі з адаптивним підбором ваг. Запропоновано модель, у якій профіль числових ознак описано ядерною оцінкою густини з урахуванням давності переглядів, локацію - згладженим показником lift, а ваги чотирьох критеріїв підбирає генетичний алгоритм за ранговою функцією пристосованості leave-one-out. Розроблено програмний модуль із REST API та трьома режимами видачі, зокрема для користувача без реєстрації. Обмеження моделі: параметри (h, ρ, λ, σ, P, G) задано апріорі, сигналом є лише факт перегляду, а ціна й площа корельовані. Гіпотезу про перевагу адаптивних ваг над фіксованими ще не перевірено; подальші дослідження стосуватимуться експериментального порівняння якості ранжування (MRR) за фіксованих і підібраних ваг, оцінки чутливості до параметрів та перевірки на реальних даних.

Використання інструментів ШІ. Інструменти штучного інтелекту не використовувалися.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

[1] Eppler, M. J., & Mengis, J. (2004). The concept of information overload: A review of literature from organization science, accounting, marketing, MIS, and related disciplines. *The Information Society, 20*(5), 325–344. https://doi.org/10.1080/01972240490507974

[2] Ricci, F., Rokach, L., & Shapira, B. (Eds.). (2015). *Recommender systems handbook* (2nd ed.). Springer. https://doi.org/10.1007/978-1-4899-7637-6

[3] Yuan, X., Lee, J.-H., Kim, S.-J., & Kim, Y.-H. (2013). Toward a user-oriented recommendation system for real estate websites. *Information Systems, 38*(2), 231–243. https://doi.org/10.1016/j.is.2012.08.004

[4] Hu, Y., Koren, Y., & Volinsky, C. (2008). Collaborative filtering for implicit feedback datasets. In *Proceedings of the 8th IEEE International Conference on Data Mining* (pp. 263–272). IEEE. https://doi.org/10.1109/ICDM.2008.22

[5] OLX запустив платформу для пошуку житла в Україні із ШІ-інструментами. (2026, 12 червня). *Mezha.ua*. https://mezha.ua/news/olx-zapustiv-platformu-dlya-poshuku-zhitla-v-ukrajini-312220/

[6] DIM.RIA. (б. д.). *Про нас*. Дата звернення: 1 вересня 2026, https://dom.ria.com/uk/about-us/

[7] Rieltor.ua. (б. д.). *Про проєкт*. Дата звернення: 1 вересня 2026, https://rieltor.ua/pages/about/

[8] Goldberg, D. E. (1989). *Genetic algorithms in search, optimization, and machine learning*. Addison-Wesley002E

[9] Silverman, B. W. (1986). *Density estimation for statistics and data analysis*. Chapman & Hall.

Yushko K.V., Knyrik N.R.

A study of listing personalisation methods in a real estate marketplace based on browsing history

Abstract. *Listing personalisation in a centralised real estate marketplace is considered as a ranking problem based on browsing history. A model of listing relevance to a user interest profile is proposed, in which the profile of numeric attributes is described by kernel density estimation and the criteria weights are selected by a genetic algorithm. A software module with three feed modes, including one for users without registration, is developed.*

Keywords: *real estate; marketplace; personalisation; recommender system; genetic algorithm; browsing history*