

Олег Масюк, Олександр Овсяк*

¹ НЛТУ України, Львів, Україна, ORCID: 0009-0007-3894-0929,
olh.masiuk@gmail.com

² НЛТУ України, Львів, Україна. ORCID: 0000-0003-2620-1938,
ovsjak@nltu.edu.ua

Персоналізовані гібридні системи: методології оцінювання та сфери застосування

***Анотація.** Розглянуто сучасні підходи до оцінювання персоналізованих гібридних систем та їх сфер застосування. Описано методи інтеграції моделей, включаючи ранню, пізню та змішану гібридизацію, а також механізм позиційного голосування (метод Борда). Систематизовано критерії оцінювання гібридних алгоритмів та наведено приклади їх використання в реальних задачах. На основі аналізу сформовано узагальнені рекомендації щодо вибору підходів до побудови гібридних систем і оцінювання їх ефективності.*

***Ключові слова** – базова модель, рекомендаційні системи, гібридні алгоритми.*

Вступ. У наш час, коли складність задач постійно зростає і необхідно опрацьовувати все більше і більше даних, майже неможливо знайти одну модель на всі випадки, яка буде працювати стабільно, швидко та точно. Натомість є альтернатива – гібридні системи. Вони поєднують різні сильні сторони багатьох підходів: ансамблі алгоритмів, каскадні моделі, інтеграцію даних та дозволяють керувати компромісами точність-швидкодія-вартість. На практиці це означає, що замість однієї найкращої моделі ми застосовуємо суміш різних підходів саме під задачу яка нам необхідна.

Математична модель. Гібридизація може виконуватися трьома підходами [1]:

- Рання інтеграція (early fusion) – з'єднує показники різних каналів на вході моделі. Це найкращий варіант для задач, де важливо мати спільний контекст.

- Пізня інтеграція (late fusion) поєднує виходи моделей, це може бути як проста зважена сума виходів, так і мета-модель і позиційне голосування.
- Змішані схеми комбінують як ранню, так і пізню інтеграцію. У таких випадках зважене середнє можна обчислити за формулою:

$$S_{hyb}(x) = \sum W_m S_m(x)_{m=1}^M, \sum W_m = 1, W_m \geq 0 \quad (1)$$

де $S_{hyb}(x)$ – підсумковий бал об’єкта, $S_m(x)$ – бал m -тої базової моделі, W_m – вага цієї моделі, M – кількість моделей, x – вектор ознак об’єкта.

Є ще окремий вид, а саме позиційне голосування. Принцип роботи наступний: кожна модель сортує дані за позиціями і об’єкту надається певний бал, який додається з вагами самої моделі. Метод Борда є стійким до викидів окремих моделей, тому використовується там, де прогнози можуть мати різні оцінки та значення. Схему узагальненої гібридної архітектури подано на рис. 1, а приклад роботи позиційного голосування – на рис. 2.



Рисунок 1. Узагальнена архітектура гібридної системи

Програмна реалізація. Практична реалізація гібридних методів передбачає такі етапи як збір даних та видачу результатів через єдиний API. Дані користувачів надходять у єдине сховище ознак, де виконуються трансформації. З цим сховищем працює сервісний застосунок (наприклад ASP.NET Core). Щоб оцінити та налаштувати гібридні методи потрібно визначитись з метриками. Для класифікації беруть до уваги точність та повноту, для регресії MAE,

MAPE, RMSE, а для ранжування це параметри MAP@K, NDCG@K і Recall@K. Обчислення метрик реалізують у кодї вручну або із залученням сторонніх бібліотек для.NET (наприклад, Accord.NET або ML.NET), залежно від вибраної архітектури.

Об'єкт	Бали від А	Бали від В	Підсумок
item1	5	1	6
item2	4	4	8
item3	3	5	8
item4	2	0	2
item5	1	2	3
item6	0	3	3

Рисунок 2. Приклад роботи метода Борда

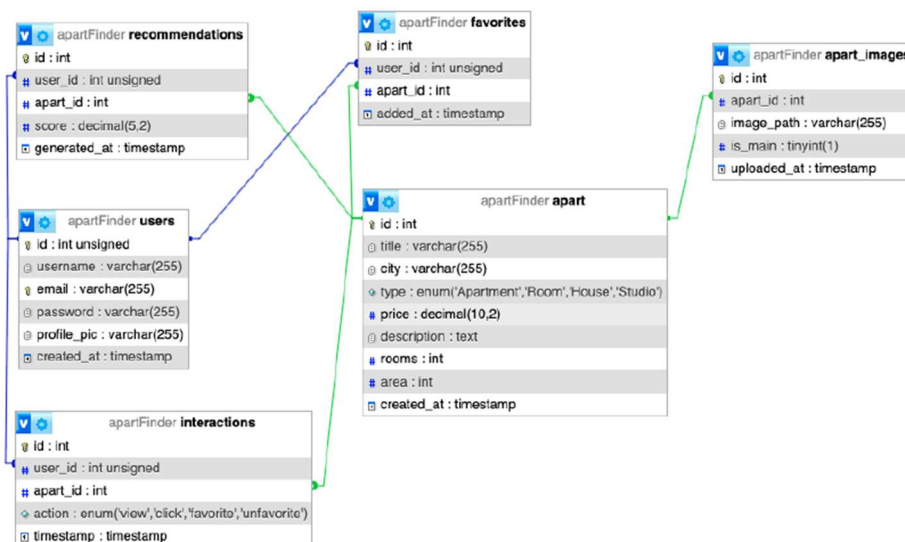


Рисунок 3. Приклад інтеграційної схеми даних

Для уникнення проблем з дублюванням логіки та розсинхронізацією між тренувальними даними та реальними бізнес-даними, підтримують уніфіковану модель даних [2]. Виділяють об'єкти та їх типи, події, зв'язки між ними, формують похідні ознаки. Схему даних доцільно подавати у табличному вигляді, що полегшує відтворюваність експериментів (рис. 3).

Висновок. Отже, гібридні системи надають змогу поєднати сильні сторони різних підходів та алгоритмів, що дозволяє досягти значного приросту якості без критичного зростання часу роботи. Водночас навчання моделі може бути

дуже дорогим, адже потребує збір та обробку даних, оцінку та вибір критеріїв і також вміння обробляти шум. Подальша робота передбачає розширення системи та реалізацію власного гібридного метода, автоматизацію та формування стабільних, економічно доцільних сценаріїв використання гібридного алгоритму.

Список використаних літературних джерел

- [1] Martinez, M. J. Barranco, L. G. Perez, and M. Espinilla. A knowledge based recommender system with multigranular linguistic information. International Journal of Computational Intelligence Systems. – 2008. – Т. 1, № 3. – С. 225–236.
- [2] Roy D.; Dutta M. A systematic review and research perspective on recommender systems Journal of Big Data. 2022. Vol. 9, Art. 59

Personalized hybrid systems: evaluation methodologies and areas of application

Oleh Masiuk, Oleksandr Ovsyak

Modern approaches to evaluating personalized hybrid systems and their areas of application are examined. Methods of model integration are described, including early, late, and mixed hybridization, as well as a mechanism of positional voting (the Borda method). Evaluation criteria for hybrid algorithms are systematized, and examples of their use in real-world tasks are provided. Based on the analysis, generalized recommendations are formulated for selecting approaches to building hybrid systems and assessing their effectiveness.