

А. І. Кулягін<sup>1</sup>, В.В. Нарожний<sup>1</sup>, В. М. Ткачов<sup>2</sup>, Г. А. Кучук<sup>3</sup>

<sup>1</sup> Національний аерокосмічний університет імені М. Є. Жуковського «ХАІ», Харків, Україна

<sup>2</sup> Харківський національний університет радіоелектроніки, Харків, Україна

<sup>3</sup> Національний технічний університет «Харківський політехнічний інститут», Харків, Україна

## ДОСЛІДЖЕННЯ МЕТОДІВ ПОБУДОВИ РЕКОМЕНДАЦІЙНИХ СИСТЕМ ДЛЯ РОЗВ'ЯЗАННЯ ЗАДАЧІ ВИБОРУ НАЙБІЛЬШ РЕЛЕВАНТНОГО ВІДЕО ПРИ СТВОРЕННІ ВІРТУАЛЬНИХ АРТ-КОМПОЗИЦІЙ

**Анотація. Актуальність.** Через зростання цифровізації мистецтва постають задачі покращення імерсивності під час взаємодії користувача з системами extended reality мистецтвом. **Методи дослідження.** Колаборативна фільтрація матричною факторизацією, метод рекомендацій, заснований на знаннях, глибока нейронна мережа з шарами занурення. **Мета статті:** використовуючи моделі рекомендаційних систем побудованих на різних принципах провести низку обчислювальних експериментів на модельних даних і, порівнюючи результати, перевірити, який із існуючих підходів до побудови рекомендаційних систем покаже найліпші результати при розв'язанні нашої задачі - побудови системи вибору віртуальної арт-композиції. **Отримані результати.** Проаналізовано ефективність різних методів побудови рекомендаційних систем для розв'язування задачі вибору відео у віртуальних арт-композиціях з урахуванням явного та неявного фідбеку користувачів. Перевірено, що найбільш ефективним виявився підхід з використанням гібридної моделі, яка поєднує метод колаборативної фільтрації, метод рекомендацій, заснований на знаннях і глибоку нейронну мережу з шарами занурення. Доведено, що завдяки математичному апарату глибоких нейронних мереж можна ефективно розв'язати задачу вибору відео у віртуальних арт-композиціях з урахуванням вподобань користувача. **Висновок.** Розроблений у роботі підхід може бути використаний для покращення імерсивності під час взаємодії користувача з системами extended reality мистецтвом.

**Ключові слова:** колаборативна фільтрація, метод рекомендацій, заснований на знаннях, глибока нейронна мережа з шарами занурення, extended reality, імерсивність.

### Вступ

Цифровізація мистецтва, зокрема, способів взаємодії користувачів з витворами мистецтва, в останні роки активно розвивається. Перспективним напрямом є поєднання живопису та технологій VR/AR/MR, загалом XR (extended reality) [1].

Припустимо, що є мобільний додаток, який дозволяє за допомогою камери сканувати зображення-маркери, робити їх AR якорями та показувати на їх місці відео, завантажене з серверу. В якості зображення-маркера може виступати, наприклад, витвір живопису у музейній композиції. Сукупність зображення та відео, яке прив'язане до зображення, називається віртуальною арт-композицією.

Розпізнавання зображення-маркера відбувається на сервері. Після розпізнавання система повинна обрати відео, яке найбільш відповідає вподобанням користувача (за жанрами, кольоровою гамою, композицією тощо). Далі воно коригується з урахуванням явного (сподобалась композиція чи ні, рейтингова оцінка, додаткове анкетування) та неявного (додивився до кінця, як довго фокусувався на композиції тощо) фідбеку. Початкове визначення вподобань користувача може здійснюватися за допомогою анкетування, яке кожен користувач має пройти на початку користування додатком. За його результатами до профілю користувача будуть додані дані, які можуть бути вказувати на його вподобання, наприклад: гендер, вік, улюблений колір, улюблені жанри живопису та ін.

На наш погляд, очевидним рішенням розв'язання задачі визначення відео, яке необхідно обрати згідно з вподобаннями користувача, є використання рекомендаційних систем.

Система вибору відео для показу користувачеві повинна відповідати таким вимогам:

- 1) вибір має здійснюватися на основі профілю користувача та/або його минулих фідбеків;
- 2) у кожного зображення-маркера та відео будуть різні характеристики, які потрібно врахувати при виборі;
- 3) при наданні рекомендації система повинна враховувати відгуки на віртуальну композицію від інших користувачів.

Таким чином, система вибору відео може мати на вході дві робочі ситуації:

*Ситуація 1.* На вході системи віртуальна арт-композиція  $K: A, \{V_1 \dots V_n\}$  ( $A$  - якор +  $V_i$  конкретне відео), яка не має рейтингових оцінок жодного з користувачів (ситуація холодного старту). В такому випадку система має обрати композицію  $K$ , в якій відео має ознаки, що найбільше відповідають профілю користувача (наприклад, по кольоровій гамі, жанрам живопису тощо).

*Ситуація 2.* На вході системи віртуальна арт-композиція  $K: A, \{V_1 \dots V_n\}$  ( $A$  - якор +  $V_i$  конкретне відео), яка вже має деяку кількість рейтингових оцінок від інших користувачів. Користувач, який потребує рекомендації для цієї композиції, має рейтингові оцінки інших віртуальних композицій. В такому випадку система має обрати композицію  $K$  на основі спільних оцінок даного користувача та інших користувачів, які оцінювали дану віртуальну композицію.

**Мета статті:** ідея даного дослідження полягає в наступному: на модельних даних провести низку обчислювальних експериментів з метою порівняння ефективності використання різних моделей рекомендаційних систем для розв'язання задачі побудови системи вибору віртуальної арт-композиції.

## Виклад основного матеріалу

Почнемо з підсумку того, які дані пропонує нам реальний світ:

Цільова змінна — рейтинги можуть бути явними (тобто користувач залишає відгук) або неявними (тобто передбачають позитивний відгук, якщо користувач додивився до кінця відео частину віртуальної арт композиції). У будь-якому випадку вони необхідні. #Review

Властивості продукту — опис характеристик продукту (наприклад, жанрів живопису, в яких виконано картини у віртуальних арт композиціях), які в основному використовуються в методах на основі вмісту.

Профіль користувача — інформація, яка описує характеристики користувача: може бути демографічною (тобто стать і вік) або поведінкою (тобто вподобання користувача у живопису або у відеотворах). Здебільшого використовується для рекомендацій на основі знань.

Контекст — додаткова інформація щодо контексту, в якому було отримано конкретний рейтинг (тобто коли, де, історія пошуку). Також часто використовується у рекомендаційних системах на основі знань.

Сучасні рекомендаційні системи всі перелічені види інформації для побудови прогнозу щодо нашого смаку.

Наприклад, YouTube рекомендує нам наступне відео для перегляду, використовуючи все, що Google знає про нас, зокрема наші вподобання.

Існує 4 основні підходи до створення рекомендаційних систем [2, 3]. Це рекомендаційні системи колаборативної фільтрації, на основі контенту, на основі знань та гібридні моделі.

*Колаборативна фільтрація (collaborative filtration)*. Суть першого підходу полягає в тому, що у користувачів, які зацікавилися однаковими елементами одного разу, можливо будуть однакові вподобання наступного разу. Процес пошуку співпадінь у спрощеному вигляді можна описати наступним чином: пошук користувачів, котрі мають ті ж шаблони оцінок, що і активний користувач (для якого призначений прогноз); використання рейтингів користувачів зі схожими вподобаннями, що знайдені на попередньому кроці для прогнозу майбутніх вподобань користувача; корекція результатів за фідбеком користувача.

*Рекомендації на основі контенту (content-based recommendations)*. Рекомендаційні системи на основі контенту засновані на створених вручну або автоматично витягнутих характеристиках елементів та профілю користувача, в якому надається вага значущості цих характеристик. Загалом, такі системи мають на меті знайти подібні елементи до тих, що користувач вподобав раніше. Порівняно з колаборативною фільтрацією, рекомендації на основі контенту мають дві переваги: по-перше, для досягнення розумної точності рекомендацій не потрібні великі групи користувачів. Крім того, нові елементи можна рекомендувати відразу після того, як з'явиться доступ до характеристик елементів. Однак, в багатьох областях, більш

суб'єктивні характеристики елемента, які можуть бути корисними для генерації рекомендацій, важко отримати автоматично [4]. Це означає, що така інформація повинна вводиться в систему вручну, що є потенційно дорогим і схильним до помилок процесом [5].

*Рекомендації на основі знань (knowledge-based recommendations)*. Коли в якомусь нішевому домені немає можливості покладатися на історію покупок або переглядів, з'являється передумова для спільного використання колаборативного підходу та фільтрації на основі контенту. Іноді недостатньо просто уявити елементи, які відповідають заданому набору функцій, що запитуються, оскільки відсутній аспект персоналізації. Так, кожен користувач з однаковими функціями, що запитуються, отримає один і той же набір рекомендацій. У цьому випадку потрібна система, яка використовує додаткові знання для вироблення рекомендацій. У таких підходах система зазвичай використовує додаткову інформацію, що часто надається вручну, як про поточного користувача, так і про доступні елементи. Інший аспект — це «взаємодія з користувачем», оскільки у багатьох рекомендаційних системах, що ґрунтуються на знаннях, вимоги користувача повинні виявлятися в інтерактивному режимі. Отже, для визначення потреб та переваг користувача потрібні складніші типи взаємодії, в основному через відсутність історії покупок, яку можна було б використати. Простий підхід — безпосередньо запитати користувача про його чи її вимоги. Проте, такий підхід потребує не тільки детального технічного розуміння властивостей предмета, а й створює додаткове когнітивне навантаження у сценаріях із великою кількістю характеристик предмета. Тому при більш складних підходах намагаються реалізувати більше стилів розмовної взаємодії, у яких система намагається поступово з'ясувати переваги в інтерактивному та персоналізованому діалозі.

*Гібридний підхід (hybrid approaches)*. Різні підходи, які були розглянуті вище, залежно від постановки задачі мають певні переваги та недоліки. Одним із очевидних рішень є об'єднання різних методів для вироблення кращих або більш точних рекомендацій. Наприклад, якщо існують знання спільноти та доступна подібна інформація про окремі елементи, система рекомендацій може бути покращена шляхом гібридизації колаборативної фільтрації з методами на основі контенту [6]. Така зв'язка може бути використана для подолання описаних проблем у чисто колаборативних підходах та використання контенту для нових елементів або нових користувачів [7]. На рис. 1 наведено приклад загальної схеми гібридної рекомендаційної системи.

Можна підсумувати, що для вирішення задачі рекомендацій при різних наявних вхідних даних є необхідність у використанні різних типів рекомендаційних систем. Натомість, різні підходи рекомендаційних систем мають різні вимоги до вхідних даних, які представлені у табл. 1. Ідея полягає в тому, щоб порівняти ефективність використання описаних вище методів побудови рекомендаційної системи віртуальних арт-композицій.

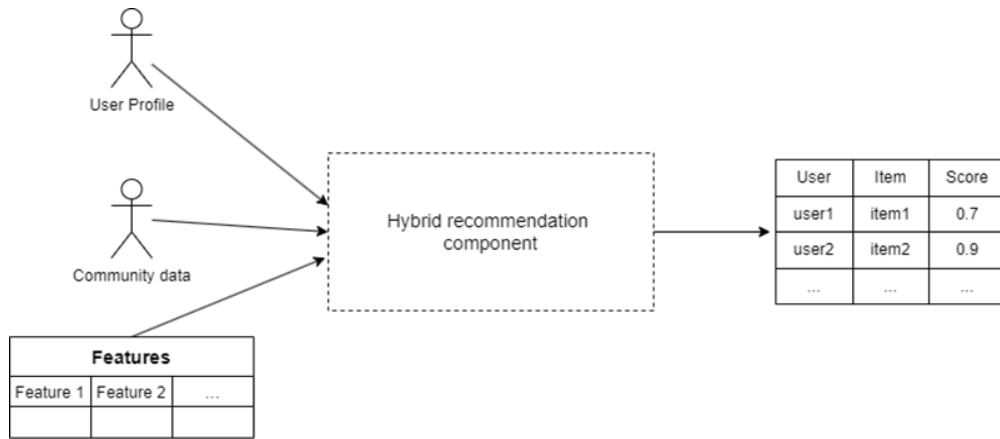


Рис. 1. Загальна схема гібридної рекомендаційної системи

Таблиця 1 – Вимоги до вхідних даних різних рекомендаційних підходів

Підхід	Профіль користувача та параметри контексту	Дані спільноти	Характеристики елемента	Моделі знань
Колаборативний	Так	Так	Ні	Ні
На основі контенту	Так	Ні	Так	Ні
На основі знань	Так	Ні	Так	Так
Гібридний	Так	Можливо	Можливо	Можливо

В основі першої моделі лежать матриця користувачів Users і матриця продуктів Products, створені за допомогою шарів вбудовування (Embeddings Layer) нейронної мережі. Тут вхідні дані — пари "користувач-продукт", а вихідні дані — рейтинг,

присуджений користувачем відповідному продукту. Для прогнозування нової пари "користувач-продукт" модель виконує пошук користувача у просторі занурення Users і продукт у просторі Products. Схема слів моделі показана на рис. 2.

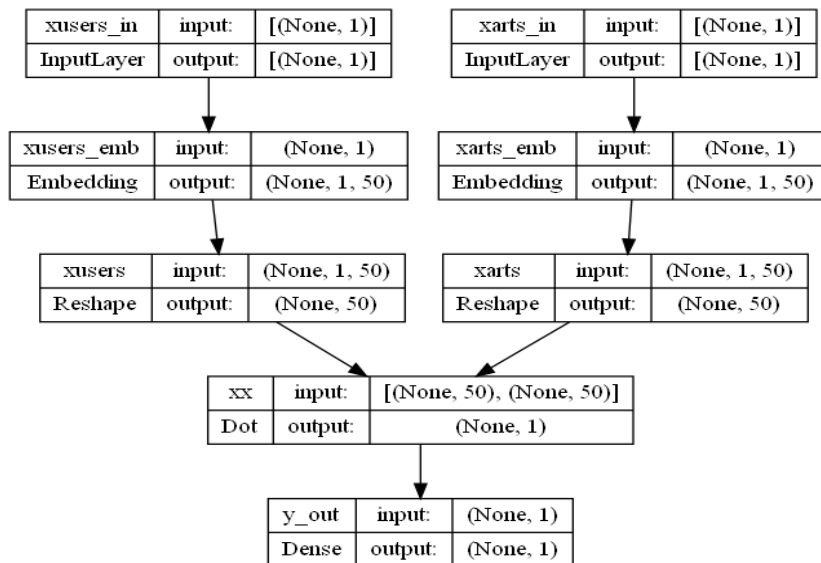


Рис. 2. Слої моделі на основі матричної факторизації

Наразі найсучасніші рекомендаційні системи використовують глибоке навчання. Зокрема, поєднують нелінійність нейронних мереж з матричною факторизацією. Ця модель створена для максимального використання простору вбудовування (embedding space), використовуючи його не лише для традиційної колаборативної фільтрації, але й для повнозв'язної глибокої нейронної мережі. Глибока нейронна мережа має фіксувати шаблони та властивості, які може пропустити матрична факторизація. Схема слів моделі зображена на рис. 3.

Третя модель, яку ми використали для проведення експериментів, є гібридною. Завдяки гнучкості сучасного апарату нейронних мереж стало можливим об'єднати в рамках однієї моделі різні підходи до створення рекомендаційних систем. До моделі Neural Collaborative Filtering ми додали дві окремі глибокі нейронні мережі, які для вибору відео використовують інформацію про властивості віртуальних арт-композицій та інформацію із профіля користувачів. Це, на нашу думку, дозволить підвищити якість рекомендацій, які ми отримуємо від системи. Слої гібридної моделі на рис. 4.

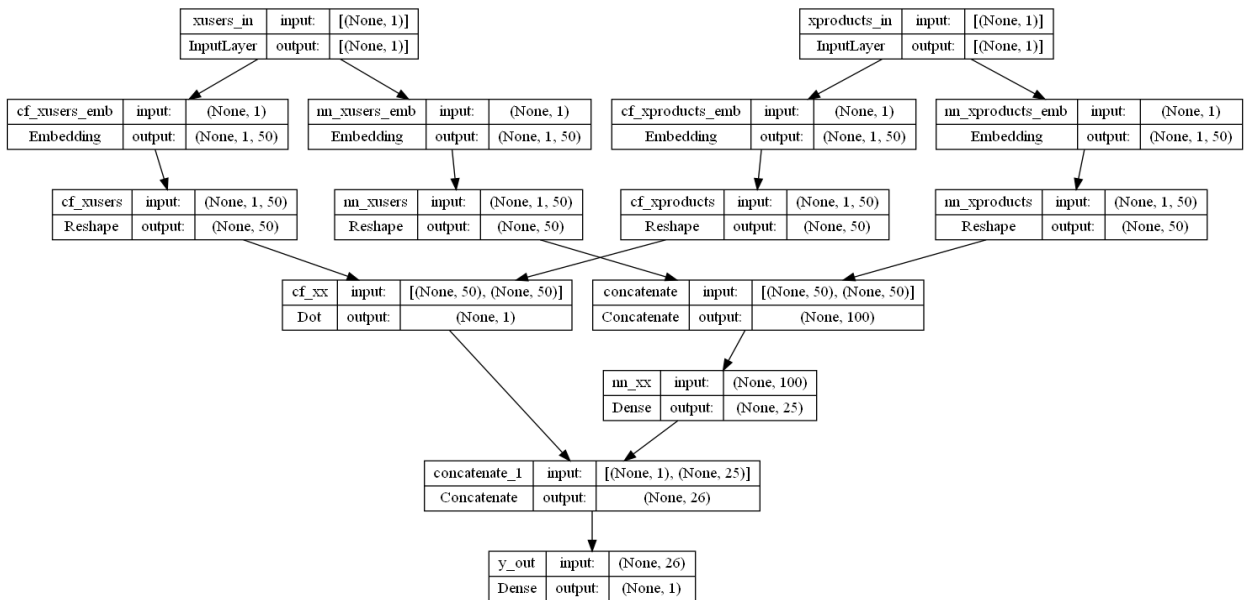


Рис. 3. Слої моделі на основі нейронної колаборативної фільтрації

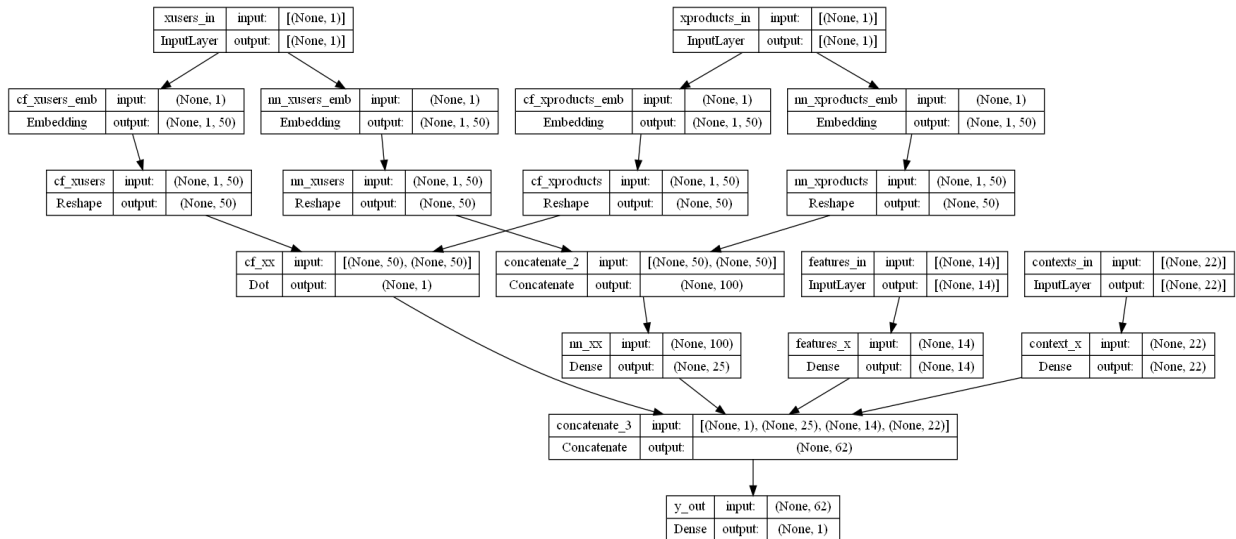


Рис. 4. Слої гібридної моделі

Набір даних для навчання моделі. Для проведення обчислювальних експериментів для порівняння ефективності описаних моделей обрано дані з датасету WikiArt. В якості датасету, що представляє користувачів та їх рейтинг було застосовано псевдо-рандомно згенеровані дані.

Ці дані складаються з трьох наборів, які містять інформацію про користувачів та їх вподобання, а також арт-композиції.

Virtual Art Composition Dataset. Набір даних віртуальних арт композицій містить інформацію про композиції та їх властивості. Його структура зображена на рис. 5. Набір даних містить перелік віртуальних арт-композицій з визначенням ID-номера кожної композиції (image\_path) та жанру живопису, до якого відноситься композиція (class\_name).

Users Dataset. Набір даних користувачів, який містить інформацію про користувачів, рис. 6.

image_path	class_name	rating
/kaggle/input/wikiart-gangogh-creating-art-gan/genre-painting/3d2f9a915a61136f314313ee	genre-painting	6
/kaggle/input/wikiart-gangogh-creating-art-gan/genre-painting/d57e51532362214de671dd3	genre-painting	6
/kaggle/input/wikiart-gangogh-creating-art-gan/genre-painting/7baac30940e875b1eddd04	genre-painting	

Рис. 5. Структура набору даних віртуальних арт-композицій.

id	first_name	last_name	gender	favorite_color_hex	age	genre_1	genre_2	genre_3
1	Inglebert	McKerton	Male	#6d2c6f	16	still-life	animal-painting	landscape
2	Spike	Hanhardt	Male	#d49fd1	53	landscape	abstract	cityscape
3	Muhammad	Hasely	Male	#f6de77	60	nude-painting-nu	mythological-painting	portrait
4	Kaleb	Verine	Male	#4b3552	50	abstract	mythological-painting	still-life

Рис. 6. Структура набору даних користувачів

Кожен рядок в цьому наборі містить характеристики окремого користувача, а саме унікальний ID-номер користувача (id), його прізвище та ім'я (first\_name, last\_name), вік (age), перелік улюблених жанрів живопису (genre\_1, genre\_2, genre\_3) тощо.

*Ratings Dataset.* Набір даних рейтингового оцінювання користувачами віртуальних композицій із Virtual Art Composition Dataset, рис. 7.

user_id	vcomposition_id	y
0	0	4.0
0	2	4.0
0	5	4.0
0	43	5.0

Рис. 7. Структура набору даних рейтингового оцінювання віртуальних арт-композицій.

Цей набір містить дані про кожний випадок при своєму користувачем user\_id рейтингу у віртуальній композиції vcomposition\_id.

*Результати.* Нижче наведені результати обчислювальних експериментів, проведених на описаних моделях з використанням даних, описаних вище. Хочемо звернути увагу, що у використаних моделях задача прогнозування рейтингу відео розглядається як задача регресії.

Для проведення експериментів набір даних було розділено на навчальний та тестовий набори у співвідношенні 80 до 20. Тридцять відсотків навчального набору використовувалось для валідації моделі в процесі навчання.

На рис. 8–10 наведено графіки зміни значень функції втрат і метрики при навчанні відповідних моделей на навчальних даних і перевірки на валідаційних даних.

У всіх трьох моделях було використано в якості функції втрат середню абсолютну помилку (mean absolute error) і середню абсолютну процентну помилку (mean absolute percentage error) в якості метрики, показника, який оцінює модель під час навчання та тестування для моніторингу процесу

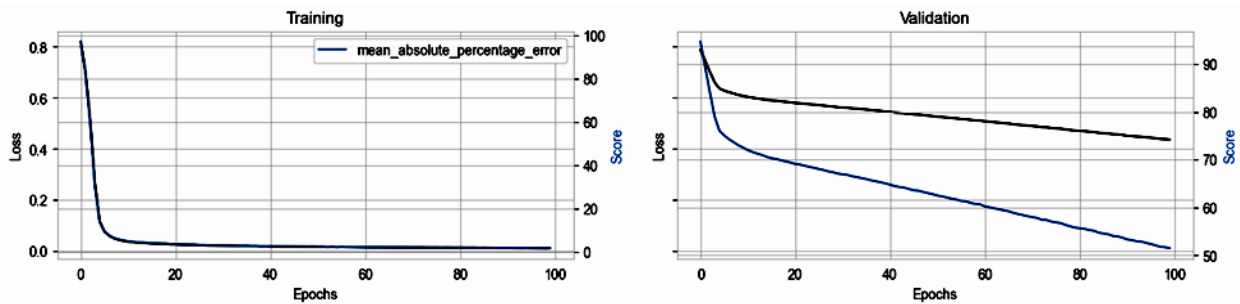


Рис. 8. Графік змін значень функції втрат і метрики під час навчання моделі на основі колабораційної фільтрації матричною факторизацією

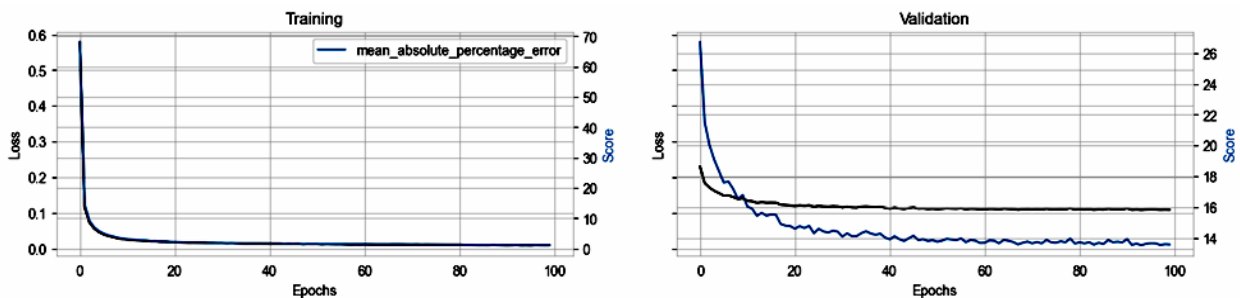


Рис. 9. Графік змін значень функції втрат і метрики під час навчання моделі на основі знань про елементи арт-композицій

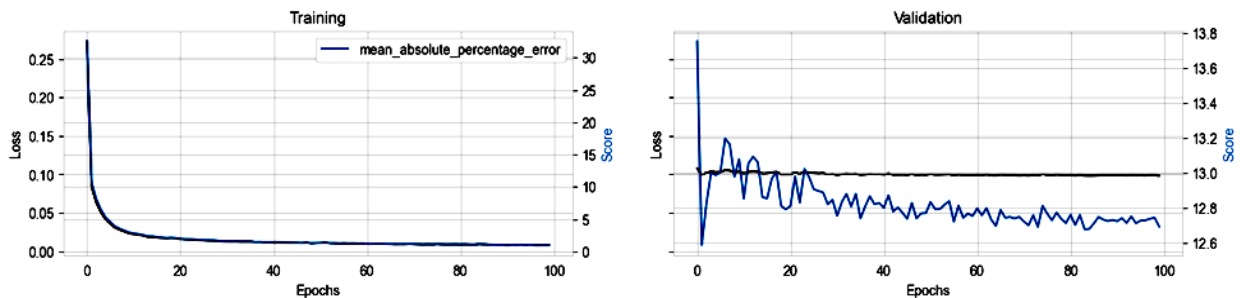


Рис. 10. Графік змін значень функції втрат і метрики під час навчання глибокої нейронної мережі з шарами занурення

Після навчання для перевірки моделі було використано тестовий набір даних і було отримано такі

значення середньої абсолютної помилки (mean absolute error) і середньої абсолютної процентної

помилки (mean absolute percentage error):

Mean Absolute Error ( $\Sigma|y-\text{pred}|/n$ ): 0.4302.

Mean Absolute Perc Error ( $\Sigma(|y-\text{pred}|/y)/n$ ): 50.91%.

В результаті тестування моделі 2 було отримано наступні показники:

Mean Absolute Error ( $\Sigma|y-\text{pred}|/n$ ): 0.1058.

Mean Absolute Perc Error ( $\Sigma(|y-\text{pred}|/y)/n$ ): 13.26%.

Значення показників функції втрат і метрики, отримані під час тестування моделі 3:

Mean Absolute Error ( $\Sigma|y-\text{pred}|/n$ ): 0.0922.

Mean Absolute Perc Error ( $\Sigma(|y-\text{pred}|/y)/n$ ): 12.07%.

Як видно із результатів експериментів, найгірший результат показала перша модель, а найкращий - третя.

На нашу думку, це пояснюється тим, що фактично кожна наступна модель відрізняється від попередньої тим, що в неї додано новий елемент гібридизації. Перша модель реалізує колаборативну фільтрацію методом матричної факторизації. Друга модель є гібридом методів матричної факторизації і колаборативної фільтрації глибокою нейронною мережею. Третя модель додатково містить третій елемент, ще одну глибоку нейронну мережу, яка враховує характеристики користувачів при визначенні рейтингової оцінки відео у віртуальній арт-композиції.

## Висновки

В цій статті ми ставили собі за ціль порівняти ефективність різних методів побудови рекомендаційних систем для розв'язання задачі вибору відео у віртуальних арт-композиціях з урахуванням вподобань користувача. Для проведення дослідження було обрано три методи: метод колаборативної фільтрації матричною факторизацією, метод рекомендацій, заснований на знаннях, і глибоку нейронну мережу з шарами занурення. Як показали результати обчислювальних експериментів, найбільш ефективним виявився підхід з використанням гібридної моделі, яка поєднує в собі метод колаборативної фільтрації, метод рекомендацій, заснований на знаннях, і глибоку нейронну мережу з шарами занурення. Завдяки гнучкості сучасних нейронних мереж можна побудувати таку архітектуру глибокої нейронної мережі, яка буде включати в себе різні підходи, що підвищить ефективність за рахунок виникнення ефекту синергії. Це доводить, що завдяки математичному апарату глибоких нейронних мереж можна ефективно розв'язати задачу вибору відео у віртуальних арт-композиціях з урахуванням вподобань користувача.

**Подальші дослідження** доцільно проводити у напрямі поєднання існуючої системи з генерацією віртуальних арт-композицій з метою збільшення вибірки рекомендацій та нівелювання обмеженої кількості мануально створених віртуальних арт-композицій.

## СПИСОК ЛІТЕРАТУРИ

1. F. Ye. Image Art Innovation based on Extended Reality Technology // 2022 7th IEEE International Conference on Data Science in Cyberspace (DSC), 2022, С. 584-587.
2. Michael D. Ekstrand. Collaborative Filtering Recommender Systems [Текст] / Michael D. Ekstrand, John T. Riedl, Joseph A. Konstan // University of Minnesota. – 2010. – С. 82–112.
3. Joonseok Lee. A Comparative Study of Collaborative Filtering Algorithms [Текст] / Joonseok Lee, Mingxuan Sun, Guy Lebanon // Cornell University. – 2012. – С. 2-3.
4. Badrul Sarwar. Item-Based Collaborative Filtering Recommendation Algorithms [Текст] / Badrul Sarwar, George Karypis, Joseph Konstan, John Riedl // University of Minnesota, Minneapolis. – 2001. – С. 2-3.
5. Francesco Ricci. Recommender Systems Handbook [Текст] / Francesco Ricci, Lior Rokach, Bracha Shapira, Paul B. Kantor // New York Dordrecht Heidelberg London, 2011. – С. 76-85.
6. Dietmar Jannac. Recommender Systems: An Introduction [Текст] / Dietmar Jannac, Markus Zanker, Alexander Felfernig, Gerhard Friedrich // Technische Universität Dortmund, Germany, 2010. – С. 49-59.
7. Peter Bostron. Comparison of User Based and Item Based Collaborative Filtering Recommendation Services [Текст] / Peter Bostron, Melker Philipson // SKOLAN FÖR DATAVETENSKAP OCH KOMMUNIKATION. – 2017. – С. 6-7.

Received (Надійшла) 30.08.2022

Accepted for publication (Прийнята до друку) 16.11.2022

### Study of methods of building recommendation system for solving the problem of selecting the most relevant video when creating virtual art compositions

A. Kuliakin, V. Narozhnyi, V. Tkachov, H. Kuchuk

**Abstract. Topicality.** Due to the growing digitization of art, the tasks of improving immersiveness during user interaction with extended reality art systems arise. **Research methods.** Collaborative Filtering by Matrix Factorization, a Knowledge-Based Method, Deep Neural Network with Immersion Layers. **The purpose of the article:** using models of recommendation systems built on different principles, conduct a number of computational experiments on model data and, comparing the results, check which of the existing approaches to building recommendation systems will show the best results when solving our problem - building a system for choosing a virtual art composition. **The results obtained.** The effectiveness of various methods of building recommendation systems for solving the problem of video selection in virtual art compositions is analyzed, taking into account explicit and implicit user feedback. It has been verified that the most effective approach using a hybrid model, which combines the method of collaborative filtering, a method based on knowledge and a deep neural network with immersion layers. It is proven that thanks to the mathematical apparatus of deep neural networks, it is possible to effectively solve the problem of video selection in virtual art compositions, taking into account the user's preferences. **Conclusion.** The approach developed in the work can be used to improve immersiveness during user interaction with extended reality art systems.

**Keywords:** collaborative filtering, knowledge-based method, deep neural network with immersion layers, extended reality, immersiveness.