УДК 004-032-26
https://www.doi.org/10.47813/rosnio.5.2026.19.2015
EDN YUZFWL

Гибридная двухбашенная нейронная модель для персонализированных рекомендаций в цифровой библиотеке

О.А. Митина*, В.В. Зверев

МИРЭА – Российский технологический университет, проспект Вернадского, 78, Москва, Россия

*E-mail: alogmi@yandex.ru

Аннотация. Рассмотрены подходы к персонализированному подбору книг в цифровых библиотеках на основе нейросетевых методов. Разработана информационная система «Умная библиотека», в основе которой лежит двухбашенная архитектура Two-Tower с интеграцией BERT-эмбеддингов текстовых описаний изданий. Система обучена на открытом датасете GoodBooks-10k-extended и оценена по точечным и ранжирующим метрикам до и после двухфазной процедуры обучения. Рост объёмов цифровых библиотечных фондов формирует запрос на автоматизированную персонализацию подборки литературы. Задача осложняется высокой разреженностью матриц взаимодействий (97-99 %) и проблемой «холодного старта» для новых пользователей, существенно ограничивающими применимость классических методов коллаборативной фильтрации. Большинство коммерческих сервисов формируют рекомендации непрозрачно: читатель получает готовую подборку, не видя, какие именно его оценки, жанровые предпочтения или история чтения определили результат. Профессиональные библиотечно-информационные системы ориентированы на каталогизацию и лишены полноценного рекомендательного модуля, а аналитический инструментарий для библиотекаря в обоих случаях фактически отсутствует. Решением является информационная система, объединяющая гибридную нейронную модель, средства визуализации факторов персонализации и аналитический дашборд для библиотечного персонала. Данное исследование способствует распространению применения двухбашенных нейронных архитектур в задачах библиотечной персонализации.

Ключевые слова: рекомендательная система, двухбашенная архитектура, Two-Tower, BERT-эмбеддинги, коллаборативная фильтрация, цифровая библиотека, персонализация.

Hybrid two-tower neural model for personalized book recommendations in a digital library

O.A. Mitina*, V.V. Zverev

MIREA – Russian Technological University, Vernadsky Avenue, 78, Moscow, Russia

*E-mail: alogmi@yandex.ru

Abstract. Approaches to personalised book selection in digital libraries using neural network methods are considered. An information system has been developed on the basis of the Two-Tower architecture with BERT embeddings of textual book descriptions. The model has been trained on the GoodBooks-10k-extended dataset and evaluated using pointwise and ranking quality metrics. The growth of digital library collections creates demand for automated personalisation of literature selection. The task is complicated by high interaction-matrix sparsity (97-99 %) and the cold-start problem for new readers. Most commercial services generate recommendations opaquely, without revealing personalisation rationales or providing analytical tools for library staff. The proposed system integrates a hybrid neural recommendation module, visualisation of personalisation factors, and an analytics dashboard for library personnel.

Keywords: recommendation system, two-tower architecture, BERT embeddings, collaborative filtering, digital library, personalisation.

1. Введение

Рекомендательные системы превратились в ключевой инструмент удержания читательской аудитории цифровых библиотек, конкурирующих с крупными потоковыми сервисами по качеству персонализации. С переносом читательской активности в онлайн-пространство стал возможен систематический анализ данных о взаимодействии пользователей с книжным фондом: выставленных оценок, истории чтения и жанровых предпочтений.

Применение нейросетевых методов к задаче подбора литературы сопряжено с рядом структурных ограничений. Матрицы взаимодействий «пользователь-книга» характеризуются высокой разреженностью (97-99 %), что существенно затрудняет обучение коллаборативных моделей. Новые пользователи не имеют истории оценок, необходимой для расчёта персональных векторов. Существующие коммерческие решения – ЛитРес, Яндекс Книги, LiveLib – реализуют сложные алгоритмы, однако формируют рекомендации непрозрачно для пользователя, не раскрывая, какие именно его предпочтения определили выдачу. Профессиональные АБИС (ИРБИС64) ориентированы на каталогизацию и лишены рекомендательного модуля как такового. Аналитический контур для библиотекаря отсутствует в обоих классах систем.

Результатом настоящей работы является информационная система «Умная библиотека», обеспечивающая персонализированный подбор литературы на основе двухбашенной нейронной архитектуры Two-Tower. Система визуализирует факторы формирования рекомендаций непосредственно в интерфейсе читателя и предоставляет библиотечному персоналу аналитический дашборд для мониторинга эффективности рекомендаций и тенденций читательского спроса. Данная работа способствует распространению применения гибридных нейронных архитектур в задачах библиотечной персонализации.

2. Постановка задачи (Цель исследования)

Цель исследования – разработать информационную систему для цифровой библиотеки, способную одновременно учитывать коллаборативные сигналы (историю пользовательских взаимодействий с книжным фондом) и контентные признаки изданий (семантику текстовых описаний, жанровую принадлежность, библиографические метаданные), и оценить качество ранжирования рекомендаций на репрезентативном открытом датасете.

Практическая ценность разработки состоит в создании комплексного решения для библиотечных организаций, объединяющего прозрачный рекомендательный модуль, средства

визуализации факторов персонализации и панель аналитики, доступную сотрудникам библиотеки без специальных технических знаний. Результаты применимы разработчиками цифровых библиотечных платформ, специалистами по управлению читательским опытом, а также способствуют популяризации двухбашенных нейронных архитектур в задачах, требующих обработки сильно разреженных матриц взаимодействий.

3. Методы и материалы исследования

Персонализация контента в библиотечных сервисах предполагает учёт индивидуальных предпочтений читателя на основе явных сигналов (оценки, выбранные жанры) и неявных данных (история чтения, порядок взаимодействия с фондом). В отличие от традиционного универсального каталога, персонализированная система подстраивает порядок и состав предлагаемых изданий под конкретного пользователя [1]. Современные исследования показывают, что гибридные подходы, совмещающие коллаборативную фильтрацию с контентным анализом на основе глубоких нейронных сетей, эффективно справляются с задачей ранжирования рекомендаций даже в условиях высокой разреженности данных [2].

Глубокое обучение на разреженных матрицах взаимодействий требует архитектур, обеспечивающих совместное представление пользователей и объектов рекомендации. Двухбашенная нейронная модель (Two-Tower Neural Collaborative Filtering) кодирует признаки пользователя и книги независимыми подсетями в общем латентном пространстве; близость представлений вычисляется через скалярное произведение [3]. Интеграция языковых моделей типа BERT для извлечения семантических признаков из текстовых описаний книг существенно обогащает векторное представление объектов, особенно для изданий с малым числом пользовательских оценок [4]. С принципами работы двухбашенных сетей в задачах рекомендаций можно ознакомиться в работе [5].

Реализованный рекомендательный модуль построен на нейронной двухбашенной архитектуре. Пользователь кодируется вектором, объединяющим его обучаемый эмбеддинг (размерность 384), средний выставленный рейтинг и суммарное количество оценок; башня пользователя преобразует этот вектор полносвязным блоком с Layer Normalization и Dropout(0,25).

Книга представлена конкатенацией трёх типов признаков: BERT-эмбеддинга текстового описания, multi-hot вектора жанровой принадлежности (топ-30 наиболее частотных жанров датасета) и нормализованных числовых метаданных – количества страниц, года публикации, среднего рейтинга издания, числа оценок и числа рецензий. Башня книги

обрабатывает этот вектор блоком с промежуточным слоем размерности 768 и аналогичной нормализацией.

Предсказанный рейтинг пользователя для книги вычисляется как скалярное произведение латентных представлений:

$$\hat{r}_{u,i} = u^T \cdot i, \quad (1)$$

где u – рейтинг пользователя, i – номер книги.

Ранжирование рекомендаций формируется сортировкой всех непросмотренных книг по убыванию предсказанного рейтинга, обучение – минимизацией среднеквадратичной ошибки (MSE) на известных взаимодействиях. Такая архитектура позволяет эффективно учитывать как коллаборативные сигналы через эмбеддинги пользователей, так и контентные признаки через семантику BERT и жанровую принадлежность книг.

Исходный набор данных сформирован на основе расширенной версии датасета GoodBooks-10k-extended [6]. Данный датасет содержит метаданные книг (название, авторы, год издания, ISBN, жанры, описание, ссылки на обложки и другие атрибуты) и историю пользовательских оценок. Для обогащения библиографических метаданных привлекался открытый API Open Library. В результате начальной подготовки сформированы три ключевых файла данных: books_enriched.csv с расширенными метаданными изданий, ratings.csv со всеми пользовательскими оценками и to_read.csv со списком книг, отмеченных пользователями к прочтению. Все данные хранятся в реляционной базе данных MySQL с таблицами books, ratings и users.

В ходе предобработки выполнены следующие операции: обогащение данных через внешний API с нормализацией рейтингов к исходным шкалам; заполнение пропусков в числовых полях (количество страниц, год издания, средний рейтинг) медианными по выборке значениями; замена нераскрываемых текстовых пропусков заглушками; дедупликация оценок с сохранением последней по времени записи; проверка уникальности ISBN. Далее для каждой книги формировался расширенный вектор признаков: выполнялось multi-hot кодирование жанров по топ-30 категориям, Z-нормализация всех числовых атрибутов и генерация BERT-эмбеддингов описаний. Полнота данных по ключевым атрибутам таблицы books составила 98,7 %, для критически важных полей таблицы ratings – 100 %.

Итоговый датасет включает 10000 книг, 53424 активных пользователя и 5976479 оценок; разреженность матрицы взаимодействий – 98,88 %, что соответствует типичному диапазону 97-99 % для задач коллаборативной фильтрации.

Обучение модели реализовывалось в два этапа. На первой фазе (12 эпох) применялся

оптимизатор Adam со скоростью обучения 7×10^{-4} и планировщиком ReduceLROnPlateau. На второй фазе тонкой настройки (18 эпох) использовался оптимизатор AdamW с пониженной скоростью обучения, ограничением нормы градиентов (max_norm = 0,5) и стабилизацией весов методом Exponential Moving Average (коэффициент затухания $\beta = 0{,}999$).

Параметры EMA обновлялись по формуле (2) и применялись при финальном предсказании:

$$\theta_E^{MA}(t) = \beta \cdot \theta_E^{MA}(t-1) + (1-\beta)\cdot\theta_t, \quad (2)$$

где β – коэффициент затухания.

Наблюдаемая динамика монотонного снижения MSE-loss по ходу обучения подтверждает стабильную сходимость модели на данных с высокой разреженностью.

4. Полученные результаты

В таблице 1 представлены метрики качества модели, полученные на тестовой выборке (20 % датасета) по точечным и ранжирующим показателям.

Таблица 1. Результаты оценки гибридной Two-Tower модели на тестовой выборке.

Метрика	До аугментации признаков	После двухфазного обучения
RMSE	1,0241	0,6917
MAE	0,7893	0,5167
R^2	0,2104	0,4579
Precision@5	—	0,3271
NDCG@5	—	0,3322
MAP@5	—	0,4774
MRR	—	0,5213

Ошибки RMSE = 0,69 и MAE = 0,52 на пятибалльной шкале свидетельствуют о приемлемой точности предсказания рейтингов. Коэффициент детерминации $R^2 \approx 0{,}46$ показывает, что модель объясняет около 46 % вариации пользовательских оценок. Для данных с разреженностью 98,88 % подобный результат является типичным и закономерным: модель явно превосходит базовый уровень предсказания по глобальному среднему, достигая при этом относительно высокой объяснительной силы [7].

Ключевым результатом по ранжирующим метрикам является MRR = 0,5213: первая релевантная книга занимает в среднем вторую позицию рекомендательного списка, что

представляет хороший практический результат. Значение Precision@5 = 0,33 означает, что каждая третья рекомендованная книга соответствует предпочтениям пользователя. Метрика NDCG@5 = 0,33 оценивает не только факт попадания релевантных книг в топ, но и корректность их ранжирования – модель достаточно уверенно расставляет книги по степени важности для читателя. MAP@5 = 0,48 подтверждает стабильное размещение релевантных позиций ближе к началу выдачи. Относительно низкий Recall объясняется большим объёмом каталога: модель не всегда находит все потенциально интересные книги, однако уверенно выводит наиболее релевантные варианты в первые позиции.

Для пользователей с числом оценок до 10 включительно система автоматически активирует режим «холодного старта»: рекомендации формируются на основе популярных жанров и авторов, что подтверждает способность системы корректно работать с новыми читателями. Регрессионный анализ поведения пользователей выявил устойчивую обратную зависимость: уравнение $y = -0{,}0014x + 4{,}09$ показывает, что по мере накопления оценок средний балл снижается, отражая естественный рост требовательности опытных читателей. Жанровый анализ подтвердил доминирование категории fiction (свыше 5 млн. оценок), за которой следуют fantasy, classics, young-adult и contemporary. Эти данные образуют базу для управленческих решений библиотекаря по комплектованию фонда.

5. Выводы

Разработанная информационная система «Умная библиотека» реализует гибридную двухбашенную нейронную модель, интегрирующую коллаборативные и контентные сигналы. Система обеспечивает стабильное ранжирование рекомендаций, успешно обрабатывает сценарии «холодного старта» и формирует интерпретируемые объяснения подбора литературы непосредственно в пользовательском интерфейсе. Аналитический дашборд предоставляет библиотечному персоналу инструменты мониторинга метрик точности модели и тенденций читательского спроса без необходимости выгрузки данных из сторонних источников.

В целом, полученные результаты демонстрируют применимость архитектуры Two-Tower для задач персонализации в условиях сильно разреженных библиотечных данных и способствуют популяризации использования двухбашенных нейронных моделей с BERT-эмбеддингами в прикладных задачах информационного поиска. Дальнейшее повышение качества возможно при увеличении объёма взаимодействий, более детальной предобработке жанровых метаданных и применении техник негативного сэмплирования.

Список литературы

1. Ушакова, О.Б. Инструменты персонализации библиотечно-информационного обслуживания пользователей в цифровой среде / О.Б. Ушакова // Менеджмент вузовских библиотек. – Минск: БГУ, 2025. – С. 167-171. – EDN: QQBRAS.
2. Фальк, К. Рекомендательные системы на практике: практическое руководство / К. Фальк; пер. с англ. – М.: ДМК Пресс, 2021. – 412 с.
3. Сергеев, А.В. Разработка рекомендательной системы на основе двухбашенной архитектуры / А.В. Сергеев // 81-я научная конференция студентов и аспирантов Белорусского государственного университета: материалы конференции. – Минск, 15-25 мая 2024 г. – Минск: БГУ, 2024. – С. 382-386. – EDN: AEAMHF. – URL: https://elib.bsu.by/bitstream/123456789/328337/1/382-386.pdf.
4. Raschka, S. Machine Learning with PyTorch and Scikit-Learn / S. Raschka, Y. Liu, V. Mirjalili. – Birmingham: Packt Publishing, 2022. – 770 p.
5. Recommender Systems Handbook. 3rd Edition / F. Ricci, L. Rokach, B. Shapira. – Cham: Springer, 2022. – 1050 p.
6. malcolmosh. goodbooks-10k-extended: Расширенная версия датасета GoodBooks-10k [Сайт] // GitHub repository. – URL: https://github.com/malcolmosh/goodbooks-10k-extended.
7. Кульшин, Р.С. Обобщенная метрика оценки эффективности алгоритмов рекомендательных систем на основе этропийного метода / Р.С. Кульшин, А.А. Сидоров // Проблемы управления. – 2024. – № 4. – С. 44-60. – DOI: 10.25728/pu.2024.4.4 – EDN: DPLIXL. URL: http://pu.mtas.ru/archive/Kulshin_pu_424.pdf.