Research on Content-Based Generative Collaborative Book Recommendation Methods

Authors: Cheng Chuansheng, Xin Xin, Wang Yidan, Li Xiaojie, Li Xiaojie

Date: 2024-09-01T00:00:00+00:00

Abstract

To alleviate information overload in library resources and enrich readers' online experience through generative book recommendation methods. The method is applied to the book recommendation service of Shandong University Library, encompassing 111,479 book entries purchased after 2017, 31,840 reader profiles, and 390,300 borrowing records. For books with borrowing frequency > 1, a generative recommendation algorithm enhanced by fusing collaborative and content feature information is employed. For books with borrowing frequency ≤ 1, a content-based recommendation method utilizing pre-trained language models is adopted to address the cold-start problem. Experimental results demonstrate that for books with borrowing frequency exceeding 1, the proposed method achieves a Recall@100 improvement of 18.9% and an NDCG@100 improvement of 16.8% compared to the LightGCN algorithm; for books without borrowing history, the proposed method attains a Recall@100 of 0.3053 and an NDCG@100 of 0.088 compared to LightGCN; for books with borrowing frequency not exceeding 1, the proposed method achieves a Recall@100 of 0.0145 and an NDCG@100 improvement of 132% compared to LightGCN. The generative book recommendation method that integrates content features and collaborative information demonstrates superior performance and can significantly contribute to alleviating information overload and enriching readers' online experience.

Full Text

Research on Content-Based Generative Book Collaborative Recommendation Methods

Cheng Chuansheng[1] **Xin Xin**[2] **Wang Yidan**[2] **Li Xiaojie**[1]
[1]Shandong University Library, Jinan 250101
[2]School of Computer Science and Technology, Shandong University, Qingdao 266000

Abstract

[**Objective**] This study proposes a generative book recommendation method to alleviate information overload in library resources and enrich readers' online experience.

[**Application Background**] The method is applied to the book recommendation service of Shandong University Library, covering 111,479 book entries purchased after 2017, 31,840 reader profiles, and 390,300 borrowing records.

[**Method**] For books with more than one borrowing record, we employ a generative recommendation algorithm enhanced by fusing collaborative information with content features. For books with at most one borrowing record, we adopt a content-based recommendation method using pre-trained language models to address the cold-start problem.

[**Results**] Experimental results demonstrate that for books borrowed more than once, our proposed method achieves an 18.9% improvement in Recall@100 and a 16.8% improvement in NDCG@100 compared to the LightGCN algorithm. For books without any borrowing history, our method attains Recall@100 of 0.3053 and NDCG@100 of 0.088 compared to LightGCN. For books with no more than one borrowing record, our method achieves Recall@100 of 0.0145 and improves NDCG@100 by 132% over LightGCN.

[**Conclusion**] The generative book recommendation method that unifies content features and collaborative information demonstrates superior performance and can play a significant role in reducing information overload and enriching readers' online experience.

Keywords: book recommendation; generative model; collaborative filtering
Classification Number: G252

"Every reader has their book" and "Every book has its reader" are fundamental tenets of library science. Precisely meeting readers' personalized needs to achieve efficient knowledge dissemination and match books with their most suitable audiences has remained a critical research focus. In recent years, the continuous growth of print resources and the increasing abundance of electronic resources in libraries have made this issue particularly prominent and urgent. However, books are subjective and emotional products; readers' preferences vary significantly, making effective recommendation through simple content matching difficult. Numerous scholars have actively explored higher-quality, more efficient, and personalized book recommendation services through sustained research and practice. As an emerging machine learning technology, generative artificial intelligence has demonstrated tremendous potential across many domains, and its transformative impact on human-machine relationships can similarly revolutionize library reading promotion through collection resource dissemination, reading activities, and reader development[1]. Yet practical explorations of generative models in book recommendation remain relatively scarce. Therefore, this paper proposes a generative book recommendation framework that unifies content features and collaborative information to serve readers' personalized and diversified information needs.

This work is supported by the Key Project of Education and Teaching Reform Research at Shandong University titled "Future Learning Ecosystem Construction in the Era of Digital Intelligence" (Grant No. 2023Z21).

Authors: Cheng Chuansheng, Deputy Director of the Library, Associate Research Librarian, Master's degree; Xin Xin, Lecturer, PhD; Wang Yidan, Graduate Student; Li Xiaojie, Director of Network Information Technology Center, Librarian, Master's degree, lixiaojie@sdu.edu.cn, Corresponding Author.

1.1 Book Recommendation Systems

Over the past two decades, book recommendation systems have received extensive attention and in-depth research from both academia and industry as a crucial means of addressing information overload. From the perspective of recommendation methods, collaborative filtering, content-based recommendation, and hybrid recommendation are the most widely adopted approaches. Among these, neighbor-based or model-based collaborative filtering algorithms represent the mainstream methodology, followed by content-based methods that recommend similar books using feature information, and hybrid systems that combine multiple recommendation techniques. Other approaches such as association rule-based and knowledge-based recommendation methods have also been continuously explored. Regarding data characteristics, relevant studies comprehensively consider book features, reader information, and interaction behaviors to enrich the information dimensions of recommendation systems. On the reader side, research has explored personalized book recommendation scenarios for special user groups including online community users[2], OPAC non-logged-in users[3], external library users[4], and children[5], integrating readers' natural attributes, interest attributes, social attributes, and micro-behavior sequences[6] to construct user profiles[7]. On the book side, beyond collecting content features such as titles, keywords, authors, and publishers, studies have also considered table of contents information[8], rating information[9], review information[10], and bestseller effects[11]. For interaction information, historical borrowing records constitute the most critical interaction data, while subscription records[12], co-borrowing relationships[13], and implicit feedback data including favorites, shares, and clicks[14] have also been incorporated into research. From a system performance perspective, common issues such as dataset scale, sparsity, and real-time requirements have been addressed with targeted optimizations in different research outcomes. For instance, Chen Linghong et al.[15] proposed integrating knowledge graph and reader profiling techniques into book retrieval to effectively alleviate data sparsity and cold-start problems. Liu Yuanyuan[16] designed a deep learning recommendation model based on reader interest mining to improve the weak generalization capability of traditional recommendation algorithms.

1.2 Generative Recommendation Systems

Traditional book recommendation algorithms are predominantly based on discriminative thinking, completing model training by continuously optimizing the decision boundary between positive and negative samples. However, the discriminative recommendation framework struggles to ensure the authenticity and representativeness of negative samples in experimental data, and its discrimination process is limited to known candidate item sets. This restriction severely constrains the system's ability to explore and capture user interests, resulting in inherent limitations that are difficult to circumvent[17]. In recent years, AI-Generated Content (AIGC) has achieved remarkable success across multiple domains and has demonstrated tremendous potential in the information recommendation field to break through the limitations of traditional recommendation paradigms[18]. Generative models learn data distributions to generate new data similar to real data distributions, thereby providing more accurate recommendations and effectively meeting users' diverse information needs. Common generative models in personalized recommendation include Generative Adversarial Networks (GAN)[19] and Variational Autoencoders (VAE)[20]. With continuous technological development, generative recommendation algorithms based on large language models[21], diffusion models[22], and multi-task learning[23] have emerged. These algorithms possess distinct characteristics and are suitable for different scenarios and datasets, better satisfying diverse user needs by generating personalized, innovative, and proactive recommendation results. Although research on generative recommendation systems is deepening, several challenges and issues remain. These include data sparsity and cold-start problems, computational complexity and scalability, and the interpretability of recommendation results. Data sparsity refers to the limited borrowing data between readers and numerous books, making it difficult for models to accurately capture reader preferences. The cold-start problem involves new readers or books lacking sufficient historical data in the recommendation system, hindering effective recommendations. To address these issues, existing research has proposed various solutions. For example, leveraging knowledge graphs combined with feature learning[24] can mine potential relationships between readers to solve data sparsity problems and improve the accuracy and diversity of recommendation decisions. Regarding computational complexity and scalability, researchers have optimized model structures and algorithm designs, such as employing efficient attention mechanisms and downsampling techniques, to reduce computational costs and enhance the ability to process large-scale data. For recommendation result interpretability, methods based on knowledge graph reasoning and natural language generation can be explored to produce fluent and personalized recommendation explanations, thereby improving reader experience.

1.3 Summary

In summary, book recommendation systems build upon classical recommendation methods in computer science, continuously optimizing and improving algorithm performance by deeply mining information such as users' historical behaviors, interest preferences, and book content features, while integrating emerging technologies like knowledge graphs[25], deep learning[26], and neural networks[27]. The remarkable achievements and potential of AIGC have also brought innovative solutions for the development of next-generation recommendation systems. Therefore, modeling the interaction relationships between books and readers through generative models and observing the potential generative distribution of reader behavior data can inspire traditional book recommendation systems and actively contribute to realizing a new paradigm of generative book recommendation. However, existing research on generative model-based book recommendation remains limited, and how to effectively fuse book feature information with user historical interaction information in generative models represents an urgent academic problem requiring in-depth investigation. Consequently, this study analyzes and explores book recommendation systems from the perspective of generative recommendation.

2 A Generative Book Recommendation Method Framework with Unified Content Features and Collaborative Information

This paper proposes a generative book recommendation method framework that unifies content features and collaborative information. The research addresses two primary recommendation scenarios.

Scenario one focuses on intelligent recommendation for books with borrowing records (more than one borrowing record), employing a generative recommendation algorithm enhanced by fusing collaborative and content information. This scenario comprises two tasks: reader-book recommendation and book-book indexing. The input consists of relevant data from books that have interacted with readers, including reader-book borrowing behavior data and book text description content, while the output is the recommended books. Book text descriptions are fed into language models to encapsulate content information. The MT5 model employed retains the Transformer-based encoder-decoder architecture.

Scenario two primarily targets newly added books (with no more than one borrowing record). The experimental data used is divided into two categories, applying a content-based recommendation method using pre-trained language models to datasets with zero borrowing records and those with at most one borrowing record respectively. The BERT model is used to obtain book embedding representations, followed by cosine similarity calculation to determine inter-book similarity, and finally recommending to readers books similar to those they have previously borrowed.

2.1 Problem Definition

Let u and i denote a specific reader and book, respectively. The set of books borrowed by reader u is denoted as I_u^+, and the set of readers who have borrowed book i is denoted as U_i^+. The textual description content of book i is denoted as c_i. The randomly assigned unique identifier representing book i is expressed as the atomic identifier ad_i for the book. Additionally, where l represents the length of GID_i.

The generative recommendation task takes input describing UI^+ information and generates a GID list as the recommendation result. GID is generated through an autoregressive approach. The probability of recommending book i to reader u is estimated as: MATH_3-1 . The recommender selects the top-N ranked books as the recommendation list for reader U.

This paper primarily addresses two recommendation scenarios. The first scenario involves intelligent recommendation for books with behavioral data (reader and book borrowing records), employing a generative recommendation algorithm enhanced by fusing collaborative and content information. The second scenario concerns newly added books (without corresponding borrowing records), which uses a content-based recommendation method with pre-trained language models. The following sections will introduce the recommendation algorithms used in these two scenarios respectively.

2.2 Generative Recommendation Algorithm Enhanced by Collaborative and Content Information Fusion

[Figure 3: see original paper]-1 illustrates that the proposed collaborative and content information fusion-enhanced generative recommendation algorithm is a generative recommendation framework with an encoder-decoder architecture. It can simultaneously consider collaborative signals and book content information, where collaborative signals refer to borrowing records between readers and books, and book content information corresponds to book textual descriptions. Specifically, the algorithm first constructs GID using a graph neural network-based collaborative filtering model. Second, the training of this generative recommendation algorithm includes two tasks: the reader-book recommendation task and the book-book indexing task. The reader-book recommendation task aims to map the content information of books historically borrowed by readers to the GID of recommended books. The book-book indexing task targets mapping from book-side information to book GID. Both tasks are implemented through a language model based on a shared encoder-decoder to better capture textual content information. To this end, the recommendation task unifies collaborative signals and book content information to achieve better recommendations, while the indexing task performs alignment between collaborative signals and content information. It is worth emphasizing that the parameters of the involved language model are also fine-tuned to better adapt the language model for recommendation.

2.2.1 Construction of Generative Identifiers

GID construction plays a crucial role in generative recommendation, primarily utilizing hierarchical clustering methods to build GID from graph neural network-based collaborative filtering models. Specifically, book representations are first extracted from a LightGCN model trained on borrowing records. Then, a constrained K-means algorithm is employed for hierarchical clustering of book representations. Clustering at each current level uses books from the parent cluster as the entire instance set. For leaf nodes at the final layer, indices from 1 to K are randomly assigned to these books. This establishes a K-ary tree to organize the book collection. Each book corresponds to a leaf node, and the path from root to leaf node constitutes the book's GID. Since LightGCN is trained on the reader-book borrowing graph, GID can model collaborative signals. Simultaneously, each position in GID has a codebook embedding matrix to help integrate content information into the book's GID.

2.2.2 Reader-Book Recommendation

The input for a reader is an unordered tuple, with each tuple describing the content information of a book borrowed by that reader. For book i, its textual description c_i is formulated as a book "sentence" using a universal data format. Specifically, the textual description of book i originates from a flattened attribute dictionary composed of key-value attribute pairs v_k, with ad_i added to increase model fidelity. The borrowing item set of reader u is denoted as I_u^+. Therefore, the input for the reader-book recommendation task is: MATH_3-1 and MATH_3-2 .

Subsequently, an encoder-decoder-based model is employed to process the text. The Encoder produces hidden states. Afterward, given the previously generated tokens $z_{<t}$ before step t, this process can be formulated as: the Decoder computes the generation probability at step t using the Encoder's hidden states and the codebook embedding matrix E_t at position t, expressed as: MATH_3-3 , MATH_3-4 , and MATH_3-5 .

In this work, the pre-trained language model MT5 is utilized as the model in the generative framework. MT5 is a multilingual version developed by Google based on its T5 model, trained on datasets covering 101 languages, demonstrating significant multilingual processing capabilities. Through its cross-lingual understanding ability, strong performance, optimized datasets and training methods, and broad application prospects, MT5 brings revolutionary changes to the NLP field. In this paper, MT5 parameters are fine-tuned via backpropagation to better adapt the language model for recommendation tasks.

Building upon this, a ranking task is further introduced to enhance its ranking capability. Here, a book i that reader u has never borrowed and whose GID does not overlap with the positive sample is randomly sampled as a negative sample i^-. BPR loss is employed to optimize ranking, as shown in formula MATH_3-6 .

2.2.3 Book-Book Indexing Task

To align collaborative signals with book content information, a book-book indexing task is introduced to map the content-based language space to the borrowing-based collaborative space. The input sequence for the book indexing task contains book textual information. Additionally, information about readers who have borrowed the book is introduced to further encode collaborative signals. Therefore, the input for the indexing task is represented as: MATH_3-7 .

The indexing task is conducted through the same language model and codebook embeddings as the recommendation task. The generation probability of the indexing task is similar to Eq.(4) and Eq.(5), except the model input is X_i rather than X_u. Cross-entropy loss is subsequently employed for parameter tuning. The book indexing loss is defined as: MATH_3-8 .

To achieve more effective alignment between collaborative signals and content information, a contrastive learning task is further introduced. The core idea is that books with similar GIDs should also be similar in the content-based language space. To this end, a book i^+ with overlapping subsequences in GID is randomly sampled as a positive sample, while another book i^- with non-overlapping GID is randomly sampled as a negative sample. The contrastive learning task loss is defined as: MATH_3-9 , where $h(\cdot)$ denotes the final hidden state of the Encoder$(\cdot)$, and σ represents the sigmoid function. Such contrastive loss helps the encoder learn better representations for book inputs.

2.3 Content-Based Recommendation Method Using Pre-trained Language Models

[Figure 3: see original paper]-2 Cold-start Flowchart

For newly added books, it is difficult to use collaborative filtering-based recommendation models to recommend new books to readers due to insufficient reader and book borrowing records. Therefore, we explore personalized recommendation for new books using the BERT pre-trained language model, recommending to readers new books that are content-wise similar to those they have previously borrowed.

The specific approach is illustrated in [Figure 3: see original paper]-2. For library collection books, book embedding representations are obtained through the BERT pre-trained language model based on book content descriptions, including both newly added books and previously added books with sufficient borrowing records. The BERT model used is a Natural Language Processing (NLP) pre-training model proposed by Google in 2018. Book content descriptions include book titles, authors, and classification category information corresponding to the Chinese Library Classification numbers.

BERT can process multilingual data, and its emergence has greatly advanced NLP development, with applications in various domains including text classification (e.g., sentiment analysis, topic classification), named entity recognition, question-answering systems, text summarization, and machine reading comprehension. Simultaneously, BERT still has considerable room for algorithmic optimization.

Subsequently, a content-based book similarity score matrix is obtained using cosine similarity based on the acquired book embedding representations. Since recommendations are based on similarity, the next step updates readers' scoring status for unborrowed books, jointly estimating a reader's preference for a particular unborrowed book based on inter-book similarity and the reader's borrowing history. The specific formula is as follows: MATH_3-10 , where S_{ni} is the similarity between books i and n, and r_{un} represents the borrowing status between reader u and book n, with a value of 1 if the reader has borrowing records for the book and

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv — Machine translation. Verify with original.