Proceedings of the 2nd International Conference on Machine Learning and Automation
DOI: 10.54254/2755-2721/97/20241175

Optimizing E-commerce Recommender Systems: A Comprehensive Review of Techniques and Future Directions

Zhijing Wu

School of Engineering, University of Connecticut, CT, USA

zhijing.2.wu@uconn.edu

Abstract. This paper examines recommender systems in e-commerce by reviewing technologies and real-world applications and identifying the importance of big data analytics in recommender systems. In the study, three kinds of recommendation algorithms are discussed: collaborative filtering, content-based recommendation, and hybrid models. Collaborative filtering methods did well in the case of large-scale user data, but had cold-start and sparsity problems. Based on the content, recommended methods have strong personalized suggestion functions, but the "information cocoon" phenomenon is a risk, which decreases the contents' diversity. Hybrid models are a combination approach between the two techniques, providing a flexible and robustness solution, but only a more complex computationally. This article also looks at the technology trends that have emerged lately, for example, the use of deep learning models, as well as the privacy-preserving techniques utilized in recommender systems. By analyzing and summarizing the existing research, this paper provides a reference basis for future optimization and application of recommender systems and points out potential research directions.

Keywords: Data Analytics, E-Commerce, Recommender Systems, Collaborative Filtering.

1. Introduction

With the increasing use of e-commerce, platforms have discovered a pressing need to improve the user interface and the conversion rate. To eliminate this issue, platform designers deploy personalization tools, otherwise, referred to as recommender systems. Such systems can give product recommendations to the consumers after analyzing their historical behaviors and personal preferences. This results in the increased user-level satisfaction which further leads to the increased revenue of the platform. The growing use of big data technology creates opportunities for the continuous optimization of the recommendation scheme aimed at understanding in a more precise manner user needs and improving recommendations efficiency through the analysis of the huge amount of users' behavior data [1]. The research revealed that significant business value exists in the field of e-commerce as a result of big data analysis, offering capabilities that can help organizations to identify customer preferences, better decision making, as well as promotion of personalized services innovation and improvement [2].

In the modern economy, where environmental threats and opportunities fluctuate occasionally, dynamic capabilities enable companies to be agile in real-time adjustments of resources and processes. This way, the acquisition of competitive advantage is sustainable. E-commerce aggregator sites that involve hybrid recommender systems aim to offer two perceivers for shoppers: the products and brands [3]. Referral research in recommender systems is currently concentrated on on collaborative filtering

DOI: 10.54254/2755-2721/97/20241175

techniques focused on examining the purchase history, content-based recommendation algorithms that in general generate tailor-made suggestions by examining the attributes of products and preferences of users, and hybrid approaches that combine a mixture of techniques to be more efficient in case of data scarcity and lack of information for new customers [4]. G3sr and GraphSAGE are some of the deep learning models that have achieved significant improvement of session-based recommender systems, focusing on capturing a manifold of complex relationships and interactions by employing graph neural networks [5]. Multilayer perceptrons (MLPs) combined with Long Short-Term Memory (LSTM) recurrent neural networks in predictive analytics are particularly effective in providing real-time data to suggest a course of action for online systems improvement and overall user experience [6].

Through this research study, optimization of the e-commerce recommender systems using big data analytics will be achieved. This kind of work concerns the use of public datasets for e-commerce data obtaining, step-by-step data processing involving cleaning and relevant tools, application of collaborative and content-based recommendation algorithms for construction and optimization of the recommender system, as well as a test of algorithms' efficiency during the experiments aimed at improvement of the recommendation effect. To carry out this research paper, the recommendation system will be strengthened by big data analysis technology to make users satisfied as well as platform revenue [7]. For practical recommendation enhancements, the goal is to integrate big data processing technology and a detailed analysis of the recommendation algorithms. Afterward, the accuracy of methods will be compared and analyzed to find the optimization solutions for e-commerce platforms.

2. Research Methodology

2.1. Literature search and screening

The literature review of this study was searched through academic databases such as Google Scholar, Web of Science, IEEE Xplore, and SpringerLink. In order to ensure the comprehensiveness and relevance of the literature, the search terms include "big data analytics in e-commerce," "recommender system," "collaborative filtering," "hybrid recommendation model," and "deep learning recommendation system." The selection criteria included peer-reviewed literature published within the last five years related to e-commerce recommender systems and their optimization. The initial search yielded more than 100 papers, and after de-duplication and quality screening, 10 core papers were finally selected for detailed analysis.

2.2. Literature classification and analytical framework

In order to systematically analyze existing studies, this paper classifies the literature into the following categories:

1. Collaborative filtering algorithms: these algorithms make recommendations by analyzing the similarity between users and are more suitable when there is user behavioral data with complex information. According to Nguyen et al., an alternative method may involve using the extended KNN algorithm with cognitive similarity, which in turn may enhance the accuracy of predictions [8].
2. Content-based recommendation algorithms: this class of algorithms achieves personalized recommendation by computationally analyzing product features and user preferences, thus is appropriate for the initial-phase when a user has no history with the recommender system. chakraborty et al. demonstrated the application of image content in personalized recommendation in fashion recommendation [9].
3. Hybrid recommendation models: combining collaborative filtering and content-based recommendation models, hybrid recommender systems are better able to address the problems with data sparsity and also increase personalization. In the model suggested by Wang et al., GraphSAGE is an improvement of capturing complex relationships and interactions through Graph Neural Networks, based on deep learning methods [3].

To enhance the ability of different techniques to be compared in a general way, this paper will be examining the characteristics of various key performance indicators like algorithm accuracy,

DOI: 10.54254/2755-2721/97/20241175

computational complexity, and user experience in detail. What is classified as the best model and the worst model is illustrated by the following figures about the pluses and minuses of classifying differently and their performance in a real industry setting.

3. Applications

3.1. Collaborative Filtering

Among the recommendation systems in e-commerce websites such as Amazon and Netflix, collaborative filtering models are the most used recommendation algorithm. Measuring users' actions and tastes, the algorithm finds specific comparability among the users and gives personal recommendations. Its main advantage is effective handling of mass-scale user data that leads to exact and relevant suggestions. The collaborative filtering model has issues to solve like a cold-start problem and data sparsity, that are of the extreme when dealing with new users or the occasional goods. This method, however, does not have these problems, while recently solving the cold-start problem has gained [2, 3]. For example, Nguyen et al. first proposed an enhanced KNN-based Collaborative Filtering method in that credit users, cognitive similarity is used, and this is proved to have higher accuracy and at the same time more effective for practical applications [8].

3.2. Content-Based Recommendation Systems

Content-based recommendation systems are particularly handy when it comes to the circumstance that user tastes can just pair with the feature information in an item. Platforms like Netflix use content-based algorithms that recommend movies and TV shows based on the user's watching history and preferences by personalized recommend. These systems work well in chilly scenarios where there is a bit less data of user's behavior as they focus much on the description of products instead of the way users interact with the items. Nevertheless, the document-based method has bad sides as well, which are like the "filter bubble" when a user is always stumbling on the same literature, and the personalization is just not enough to be closed to him all necessary information [6, 9]. Chakraborty et al. carried out the examination of content-based systems in fashion design suggestions and showed the utility of image processing to boost the modern approach to buying wholeheartedly [9].

3.3. Hybrid Recommender Systems

Hybrid recommendation systems unify the positives of collaborative filtering and content-based methods to provide a holistic answer that tackles the challenges of each. Module combination is realized by hybrid systems, and they are successfully overcoming such problems of data sparsity and cold start. Multi-domain systems in common, i.e., YouTube and Google-oriented recommendation engines use hybrid systems to come up with comprehensive and diverse content suggestions. Wang et al. evidenced the role of deep learning in hybrid systems where GraphSAGE neural networks were applied to analyze complex interactions between users and items [3]. This work takes into account that the combination of TSs is particularly real time, flexible, and self-organized for a scenario at which TS operation strategies required to be tuned based on environmental conditions.

3.4. Comprehensive analysis and advantages and disadvantages

Based on this review, there are different types of recommendation systems that each have their strengths and weaknesses, which can be suitable in specific application scenarios. While collaborative filtering systems excel in understanding user data in large amounts, they encounter issues such as cold-start and sparsity; content-based recommenders can make recommendations with high personalization, albeit through data quality, which may easily lead to the phenomenon known as “information cocoon;” and hybrid recommenders can produce both flexible and diverse recommendations by joining the advantages of different methods, but have a high computational complexity degree. However, hybrid recommender systems enable further flexibility and versatility in terms of recommendation offerings as they combine the advantages of various techniques. Nevertheless, the algorithm becomes more complex. Table 1

DOI: 10.54254/2755-2721/97/20241175

shows that different types of recommender systems have advantages and disadvantages that affect their performance in various application scenarios.

Table 1. Comparison of Recommender System Types, Advantages, Disadvantages, and Typical Applications

Recommender System Type	Advantages	Disadvantages	Typical Applications	References
Collaborative filtering	Efficient handling of large-scale user data	Cold start problem, data sparsity	E-commerce platforms such as Amazon	[8], [3]
Content-based	Strong personalization	Highly dependent on data	Media recommendations such as Netflix	[9], [6]
Hybrid model	Flexible and diverse	Computationally complex	Multi-scenario applications	[3], [7]

4. Recommendations

4.1. Algorithmic optimization

Since the incorporation of more advanced techniques is foundational to the efficiency of e-commerce recommendation systems, there is a need to involve them, that is, deep learning and graph neural networks into pre-established algorithms. For instance, including the concession of the GraphSAGE model into the conventional models of collaborative filtering can easily grasp the deeper and critical interactions between users and products hence the accuracy of recommendations dramatically improves [3]. Furthermore, coming up with a multi-hybrid model which could be a combination of collaboration filtering and content-based models can then make it even better in the sense of the interface with data sparsity and cold-start problems [7]. At the same time, ensemble learning offers the possibility to alleviate the weaknesses of a unique model, rather by enhancing recommendation results via uniting the strength of multiple algorithms for more robust recommendation.

4.2. Data processing and management

Data processing is the core of business intelligence systems, and because of this, e-business platforms should always and invariably strive to ensure that data quality and diversity are improved at all costs, and additional data on users' behavior is acquired or collected too i.e., clickstream data and purchase history. A case in hand, transfer learning and active learning, among other things, should be used for dealing with the data sparsity and cold-start issue, where transfer learning facilitates knowledge exchange from related domains, or active learning whereby user feedback can be enriched for; among other techniques [6]. In addition, the system should use real-time data processing techniques and updating facilities to ensure that the relevant system's components are in line with the users' current preferences/circumstances.

4.3. System implementation and user experience

From a practical approach, it might not be difficult to incorporate the recommender system into the user interface of the platform, however, this is of great importance to ensure a better user experience. Recommender systems are not just for showing accurate recommendations but also need to be clear, understandable, and reasonable so that the users can sense the fair and honest aspect of why certain recommendations are delivered to them. AI technology can be another development to rectify the doubts of the users in the system, as it offers the explanations for different actions and decisions [8]. On the contrary, the platform should let users tinker and manage their personalize settings, so they can change and modify the options given to them, making the information more relevant to their needs.

DOI: 10.54254/2755-2721/97/20241175

4.4. Future Research Directions

Further studies could be focused on the field of cross-domain recommendation systems to explore the strategies of how the system can help if the user knowledge can be provided in a way of a recommendation between different locations and domains. This should be complemented by tying recommender systems with social media analytics for the purpose of emphasizing on the way in which recommendation strategies can be improved through the analysis of user sentiment and behavior. The third dominating issue that stands out is the privacy issue. So various researchers must boost research on privacy-preserving recommendation techniques wherein federated learning concept, the technology is enabled to permit cross-device parallelized model training, however, the devices maintain an unshared control of their respective data [5, 9].

5. Conclusion

The present paper discusses the three main types of algorithms used in e-commerce recommendation systems (collaborative filtering, content-based recommending, and hybrid models). The performance of these algorithms in a practical setting is also analyzed. The study reveals that collaborative filtering deals very well with big data of users; limited generalizability of recommendations due to cold and infrequent use of the system is the main drawback here; content-based recommender systems possess a strong quality of personalizing recommendations but lack the data variety as mentioned earlier; finally, hybrid recommender system leverages the merits of several algorithms but further limits the sparseness of data and improves the recommendation accuracy. Especially, the optimization of recommenders should consider other decisions, such as algorithm performance, computational complexity, and user experience.

Through the analysis of the pros and cons of the diverse techniques constructed for recommendation, the paper notes the importance of the deep learning models integration into the recommender systems, builds a new cross-domain recommender system, and innovates the personalized recommendation effectiveness while the users' privacy is well protected. With the persistent artificial intelligence creation and diversification of user needs, recommender systems will have a much greater mission in e-commerce and other fields, and peer research will make recommender systems more adaptive and effective in various application situations.

References

[1] Roy, D., & Dutta, M. (2022). A systematic review and research perspective on recommender systems. Journal of Big Data, 9(1), 1–36. https://doi.org/10.1186/s40537-022-00592-5

[2] Huy, P. Q., & Phuc, V. K. (2023). Big data in relation with business intelligence capabilities and e-commerce during COVID-19 pandemic in accountant's perspective. Future Business Journal, 9(1), 40–21. https://doi.org/10.1186/s43093-023-00221-4

[3] Wang, J., Lee, L. K., Wu, N. I. (2023). A Systematic Review of Recommendation System Based on Deep Learning Methods. In: Nedjah, N., Martínez Pérez, G., Gupta, B.B. (eds) International Conference on Cyber Security, Privacy and Networking (ICSPN 2022). ICSPN 2021. Lecture Notes in Networks and Systems, vol 599. Springer, Cham. https://doi.org/10.1007/978-3-031-22018-0_12

[4] Akter, S., & Wamba, S. F. (2016). Big data analytics in E-commerce: a systematic review and agenda for future research. Electronic Markets, 26(2), 173–194. https://doi.org/10.1007/s12525-016-0219-0

[5] Cirqueira, D., Hofer, M., Nedbal, D., Helfert, M., & Bezbradica, M. (2020). Customer Purchase Behavior Prediction in E-commerce: A Conceptual Framework and Research Agenda. In New Frontiers in Mining Complex Patterns (pp. 119–136). Springer International Publishing. https://doi.org/10.1007/978-3-030-48861-1_8

[6] Alsmadi, A. A., Shuhaiber, A., Al-Okaily, M., Al-Gasaymeh, A., & Alrawashdeh, N. (2023). Big data analytics and innovation in e-commerce: current insights and future directions. Journal of Financial Services Marketing, 1–18. https://doi.org/10.1057/s41264-023-00235-7

DOI: 10.54254/2755-2721/97/20241175

[7] Chen, Y., Li, C., Wang H. Big Data and Predictive Analytics for Business Intelligence: A Bibliographic Study (2000–2021). Forecasting. 2022; 4(4):767-786. https://doi.org/10.3390/forecast4040042

[8] Nguyen, L. V., Vo, Q.-T., & Nguyen, T.-H. (2023). Adaptive KNN-Based Extended Collaborative Filtering Recommendation Services. Big Data and Cognitive Computing, 7(2), 106-. https://doi.org/10.3390/bdcc7020106

[9] Chakraborty, S., Hoque, Md. S., Rahman Jeem, N., Biswas, M. C., Bardhan, D., & Lobaton, E. (2021). Fashion Recommendation Systems, Models and Methods: A Review. Informatics (Basel), 8(3), 49-. https://doi.org/10.3390/informatics8030049