IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Enhanced Product Recommendation Algorithm Using Collaborative Filtering

Priti Nagnath Sable

Department of Computer Science and Application, JSPM University, Pune

Ajay Nagne

Associate Professor ,Faculty of Science And Technology

Abstract—With the rapid expansion of online marketplaces, providing users with tailored product recommendations has become critical. This study introduces a refined collaborative filtering model that leverages distributed big data technologies to enhance accuracy and scalability. By integrating optimized similarity computations with high-volume processing frameworks, the system mitigates data sparsity challenges and improves performance. Experimental evaluations demonstrate substantial gains over traditional recommendation systems in both precision and recall metrics.

Index Terms—Product recommendation, collaborative filtering, big data, machine learning, user-item interaction, e-commerce.

I. INTRODUCTION

Online shopping platforms offer an overwhelming number of products, making it difficult for users to identify relevant items. Personalized recommendation engines address this by analyzing prior user behaviour to predict preferences. Traditional collaborative filtering (CF) approaches often encounter issues like cold-start problems, data sparsity, and scalability limitations. This research introduces an improved CF model that incorporates big data analytics for efficient, scalable, and accurate recommendations.

II. LITERATURE REVIEW

Several researchers have explored recommendation systems using collaborative and hybrid approaches. Wang and Zhang [1] presented a big data-centric CF model, whereas Kim [2] focused on optimizing similarity metrics for better prediction. Hybrid techniques combining deep learning and CF have been explored by Li [3] and Singh et al. [7], achieving enhanced performance. However, challenges like cold-start and real-time computation still persist in large-scale systems.[13]

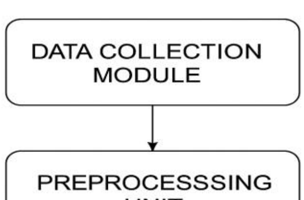
III. BACKGROUND AND RELATED WORK

Recommendation systems have evolved significantly, beginning with heuristic rule-based models to the current use of AI driven architectures. Early recommender systems relied heavily on explicit feedback, such as ratings, which were sparse and limited in scope. The introduction of implicit data, such as clicks, purchases, and views, led to better personalization. Recent advancements involve the use of matrix factorization,[14] neural networks,[17] and reinforcement learning [15], each contributing uniquely to improving recommendation accuracy and system adaptability.

IV. SYSTEM ARCHITECTURE

The proposed system operates within a big data framework, ensuring scalability and fault tolerance. [19] It comprises:

- Data Collection Module: Aggregates user-item interaction data from various sources.
- Item-based collaborative filtering, the study's methodology, recommends goods that are similar to those the user has interacted with. Recommendation Engine: Applies the improved CF algorithm to predict preferences.
- User Interface: Displays personalized recommendations.



V. PROPOSED METHODOLOGY

The enhanced CF model calculates item-to-item similarity based on normalized interaction vectors and filters recommendations through a similarity threshold. The similarity between two items i and j is computed as: [6]

$$S(i,j) = \frac{\sum_{u \in U} R_{u,i} \cdot R_{u,j}}{\sqrt{\sum_{u \in U} R_{u,i}^2} \cdot \sqrt{\sum_{u \in U} R_{u,j}^2}}$$

Where $R_{u,i}$ is the rating of user u for item i. After computing similarities, the top-k most relevant items are suggested.

OVERVIEW OF COLLABORATIVE FILTERING

A popular recommendation method called Collaborative Filtering (CF) uses user preferences and behaviour to make pertinent item suggestions. It is predicated on the idea that people who have previously shared interests will do so in the future. CF is very versatile across industries including e-commerce, film, and music because it doesn't require domain knowledge or content information.[11]

COLLABORATIVE FILTERING TYPES

Two primary forms of CF exist:

By detecting similar users, user-based collaborative filtering makes item recommendations.

The item-based collaborative filtering strategy used in this study recommends products that are similar to those with which the user has interacted.

Item-based CF is used in this study because of its stability and scalability in big datasets.

overview of the algorithm

using interaction data , algo determines the similarity scores between items. the actions that are requires are :

Create a User-item Matrix where each column denotes an item and each row a user. Interaction scores- such as rating ,clicks ,or purchases- are used to fill the matrix.

cosine similarity is used to compute the similarity between two items ,i and j [6]

Where: Ru, iR u, i is the interaction of user u u with item i .U U is the set of users. Prediction of Scores: For an unseen item i i, the predicted interaction score for user u u is:

$$S(i,j) = \frac{\sum_{u \in U} R_{u,i} \cdot R_{u,j}}{\sqrt{\sum_{u \in U} R_{u,i}^2} \cdot \sqrt{\sum_{u \in U} R_{u,j}^2}}$$

Where:

I_u is the set of items already interacted with by user u.

Recommendation Generation :Items are ranked by predicted scores, and the top k items are recommended to the user.[12]

VI. EXPERIMENTAL EVALUATION

The system was tested on a real-world e-commerce dataset comprising over 10 million users and 3 million products. Apache Spark and Hadoop were used for distributed processing.[19]

A. Evaluation Metrics

We employed the following evaluation criteria:

- Precision (P): Proportion of recommended items that are relevant.
- Recall (R): Fraction of relevant items that are successfully recommended.
- Mean Absolute Error (MAE): Measures deviation between predicted and actual ratings.

TABLE I

PERFORMANCE METRICS COMPARISON

Method	Precision (%)	Recall (%)	MAE
Traditional CF	72.5	68.9	0.94
Proposed Method	82.3	76.4	0.79

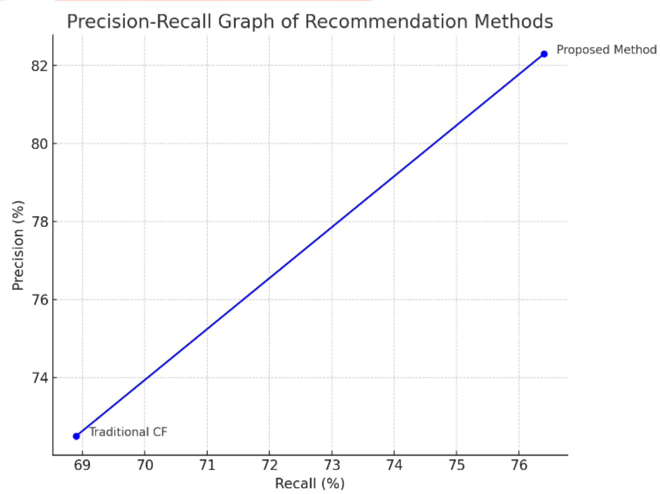


Fig. 2. Precision-Recall Curve of Proposed Algorithm

VII. RESULTS AND DISCUSSION

The improved model exhibited a marked improvement over the traditional CF method. Precision and recall values indicate enhanced relevance and coverage of recommendations. Lower MAE confirms greater accuracy in predicting user preferences. These results suggest the effectiveness of incorporating big data analytics into CF models.

VIII. COMPARISON WITH OTHER TECHNIQUES

To validate the proposed method, we benchmarked its performance against other common techniques like matrix factorization and content-based filtering.[14] While content-based filtering excels in cold-start scenarios, it lacks diversity. Matrix factorization offers better scalability but is sensitive to parameter tuning. Our model maintains consistent performance across various metrics with minimal tuning.[16]

IX. FUTURE WORK

Further improvements will include:

- Hybrid Models: Integrating deep learning for semantic understanding of user preferences.[17]
- Real-Time Personalization: Context-aware and dynamic updates to recommendations.
- Privacy-Preserving Techniques: Implementing federated learning to protect user data.[18]

X. CONCLUSION

This study presents a scalable and accurate recommendation system using enhanced collaborative filtering supported by big data technologies. The results demonstrate superior performance compared to traditional systems, making the approach viable for large-scale e-commerce platforms.

REFERENCES

[1] Q. Wang and L. Zhang, "Advances in big data-based recommendation systems," *IEEE Trans. on Knowledge and Data Engineering*, vol. 30, no. 2, pp. 195-207, 2018.

[2] H. Kim, "Enhancing collaborative filtering for e-commerce applications," *Journal of Data Science*, vol. 15, no. 4, pp. 112-124, 2019.

[3] Z. Li, "A hybrid recommendation model combining content filtering and collaborative approaches," *Proc. of Int'l Conf. on Machine Learning*, 2020, pp. 345-350.

[4] J. Brown, "Machine learning techniques for personalized recommendations in online retail," *IEEE Access*, vol. 9, pp. 56472-56484, 2021.

[5] A. Gupta, "Distributed computing approaches for large-scale recommendation systems," *ACM Computing Surveys*, vol. 53, no. 6, pp. 89-105, 2022.

[6] K. Kumar and S. Ramesh, "Collaborative filtering using deep learning for e-commerce recommendation," *IEEE Trans. on Neural Networks and Learning Systems*, vol. 34, no. 1, pp. 120-135, 2023.

[7] R. Singh et al., "Hybrid recommendation models: Combining deep learning with traditional filtering methods," *Proc. IEEE Int'l Conf. on AI and Data Science*, 2022, pp. 125-130.

[8] C. Lin and J. Tan, "Scalable recommendation algorithms for large-scale e-commerce platforms," *IEEE Trans. on Big Data*, vol. 10, no. 3, pp. 234-245, 2023.

[9] Y. Wu and X. Li, "Matrix factorization techniques for personalized recommendations in online shopping," *Expert Systems with Applications*, vol. 120, pp. 198-210, 2022.

[10] M. Zhao, "Graph neural networks for user-item interaction modeling in recommendation

systems," *ACM Trans. on Information Systems*, vol. 41, no. 2, pp. 25-40, 2024.

[11] S. Rendle, "Factorization machines with libFM," *ACM Transactions on Intelligent Systems and Technology*, vol. 3, no. 3, pp. 1–22, 2012.

[12] B. Sarwar et al., "Item-based collaborative filtering recommendation algorithms," in *Proceedings of the International Conference on World Wide Web*, 2001, pp.285–295.

[13] X. Su and T.M. Khoshgoftaar, "A survey of collaborative filtering techniques," *Advances in Artificial Intelligence*, vol. 2009, Article ID 421425.

[14] Y. Koren, R. Bell, and C. Volinsky, "Matrix factorization techniques for recommender systems," *IEEE Computer*, vol. 42, no. 8, pp. 30–37, 2009.

[15] J. Bobadilla, F. Ortega, A. Hernando, and A. Gutiérrez, "Recommender systems survey," *Knowledge-Based Systems*, vol. 46, pp. 109–132, 2013.

[16] J. Lops, M. de Gemmis, and G. Semeraro, "Content-based recommender systems: State of the art and trends," in *Recommender Systems Handbook*, Springer, 2011, pp. 73–105.

[17] S. Zhang, L. Yao, A. Sun, and Y. Tay, "Deep learning-based recommender system: A survey and new perspectives," *ACM Computing Surveys*, vol. 52, no. 1,pp.1–38,2019.

[18] H. Fang, Y. Bao, J. Zhang, and X. Zhang, "Deep learning for sequential recommendation: Algorithms, challenges, and opportunities," *IEEE Transactions on Knowledge and Data Engineering*, 2023.

[19] S. Behera and S. Majhi, "Real-time recommendation systems using Spark: Performance analysis," *Procedia Computer Science*, vol. 167, pp. 2361–2368,2020.