JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Intelligent Personalized Product Recommendation Using Advanced methods

Dr.Bathula Archana[1], Dr.S.Shiva Prasad[1] M.Deepika[2], Mohammed Aftaab Shah Hussain[3],Shaik Mazaharun[4], Sudham Charan Teja[5]

[1]Assistant Professor, Department of Computer Science and Engineering(Data Science), Malla Reddy Engineering College, Maisammaguda, Secunderabad, Telangana,India

[2-5]Students Bachelor of Technology, Department of Computer Science and Engineering, Malla Reddy Engineering College, Maisammaguda, Secunderabad,Telangana,India.

ABSTRACT

The Personalized Product Recommendation System utilizes collaborative filtering to enhance user engagement by suggesting products based on preferences and behavioral similarities with other users. By analyzing user interactions, the system identifies patterns and recommends items that align with individual interests, ultimately increasing sales. The project focuses on selecting the most effective algorithm for generating tailored recommendations, ensuring a personalized shopping experience. This approach is applicable across various domains, including e-commerce platforms and streaming services, to improve customer satisfaction. Performance metrics are used to evaluate the system's efficiency and accuracy in delivering relevant suggestions.

Keywords: Personalized Recommendation System, Collaborative Filtering, Product Recommendation, Behavioral Analysis, Machine Learning.

1. INTRODUCTION

Product recommendation systems have become an essential component of e-commerce platforms, allowing users to have more personalized shopping experiences and identify relevant products. Collaborative Filtering (CF) is a frequently used technique for making recommendations based on previous user interactions and preferences. Despite its popularity, Collaborative Filtering faces a number of obstacles, including the cold start problem, data sparsity, and scalability of similarity calculations. These issues reduce the accuracy and efficiency of CF models, particularly when dealing with new users or limited data.

The cold start problem occurs when there is little or no historical data for new users or objects, making it difficult to offer things based on previous preferences. When users score only a small percentage of accessible products, data scarcity exacerbates the problem. Furthermore, standard CF models that rely on similarity measurements such as Pearson correlation or cosine similarity may struggle to perform effectively when dealing with sparse or imbalanced datasets.

To solve these concerns, this work provides an improved product recommendation system that combines collaborative filtering and a hybrid model to manage cold start issues, sparsity, and scalability. The suggested system seeks to provide more accurate and tailored product suggestions by increasing user profiles and implementing advanced similarity estimation algorithms.

The main objective of this study is to create an improved product recommendation system that utilizes collaborative filtering while addressing its frequent drawbacks. The major goals of this initiative include:

2.Literature Survey

Several ways have been developed to increase the effectiveness of product recommendation systems. Traditional CF approaches, such as memory-based CF, which uses user-item interaction matrices, have scalability and data sparsity limitations. In contrast, model-based CF techniques like as matrix factorization and singular value decomposition (SVD) have demonstrated superior performance in managing huge datasets and identifying latent variables that impact user preferences.

1. Traditional Recommendation Approaches

Early recommendation systems relied on **content-based filtering**, where products were recommended based on user preferences and item attributes (Pazzani & Billsus, 2007). While effective for individual users, content-based methods struggled with scalability and lacked diversity in recommendations. Another traditional approach, **rule-based filtering**, required manual input and predefined criteria, making it inflexible in dynamic environments (Ricci et al., 2015).

2. Collaborative Filtering for Personalized Recommendations

Collaborative filtering (CF) emerged as a more adaptive approach, utilizing user interactions to generate personalized recommendations. CF is categorized into:

- **User-based filtering**, which recommends products based on the preferences of similar users (Resnick et al., 1994).
- **Item-based filtering**, which suggests items similar to those a user has already interacted with (Sarwar et al., 2001).

Although CF provides better personalization, it faces challenges such as sparsity (few interactions per user) and cold-start problems (new users/items with no history) (Adomavicius & Tuzhilin, 2005).

3. Hybrid Recommendation Systems

To overcome CF limitations, hybrid models combining collaborative and content-based filtering have been developed. Netflix's recommendation engine, for example, uses a hybrid model that merges user interaction data with content similarity measures (Gomez-Uribe & Hunt, 2016). Similarly, matrix factorization techniques like Singular Value Decomposition (SVD) and Neural Collaborative Filtering (NCF) leverage deep learning to enhance CF performance (He et al., 2017).

4. Deep Learning and Reinforcement Learning in Recommendations

Recent research explores deep learning models for recommendation tasks. Autoencoders, recurrent neural networks (RNNs), and graph-based methods provide advanced feature extraction capabilities (Zhang et al., 2019). Reinforcement learning is also being applied to improve long-term user engagement by dynamically adjusting recommendations based on user behavior over time (Zhao et al., 2018).

5. Evaluation Metrics for Recommendation Systems

Measuring recommendation effectiveness requires performance metrics such as Precision**,** Recall, F1-score, Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Normalized Discounted Cumulative Gain (NDCG) (Gunawardana & Shani, 2009). These metrics assess how well a system ranks relevant items, ensuring high accuracy in recommendations.

Recent advances in hybrid recommendation systems attempt to incorporate collaborative filtering, content-based filtering, and other machine learning approaches. To combat data scarcity, demographic information has been added to user profiles. By including user variables such as age, gender, and location, the system may offer items based on shared demographics.

Deep learning techniques, like autoencoders and neural collaborative filtering, have also been used to increase recommendation accuracy. These approaches may learn complicated nonlinear patterns in user preferences, resulting in improved generalization and management of cold start and data sparsity concerns. Furthermore, hybrid models that integrate different recommendation strategies, such as CF, content-based filtering, and keyword-based searches, have showed promise for enhancing recommendation system performance.

3. Proposed Methodology

The suggested product recommendation system combines collaborative filtering, profile expansion, and a keyword-based search function.

1. **Collaborative Filtering (CF)**: The foundation of the recommendation system is built on CF, which evaluates similarities between people and things. To quantify the link between users or objects, the standard CF technique uses the Pearson correlation coefficient (PCC) or cosine similarity. However, to overcome the limits of these similarity measures in sparse datasets, the system employs an improved similarity estimate approach based on user and item properties.
2. **Profile Expansion**: To improve suggestion accuracy, we supplement user profiles with demographic information such as age, gender, and previous purchasing history. This enhanced profile allows the system to propose things based on both user-item interactions and demographic commonalities amongst users. The cold start problem is mitigated by analyzing users from the same demographic group.
3. **Keyword-Based Search**: To address data scarcity and increase product discovery, the system includes a keyword-based search capability. When users enter a keyword, the system employs collaborative filtering to produce product clusters. These clusters categorize related goods based on their descriptions, and the algorithm chooses the cluster in which the term appears most frequently. This method aids consumers in locating relevant goods even when explicit ratings are few.
4. **Hybrid Model**: The recommendation system combines the strengths of collaborative filtering, content-based filtering, and demographic data. By utilizing both user behavior and demographic information, the system can provide more personalized recommendations, even in scenarios with limited interaction history or user-generated data.

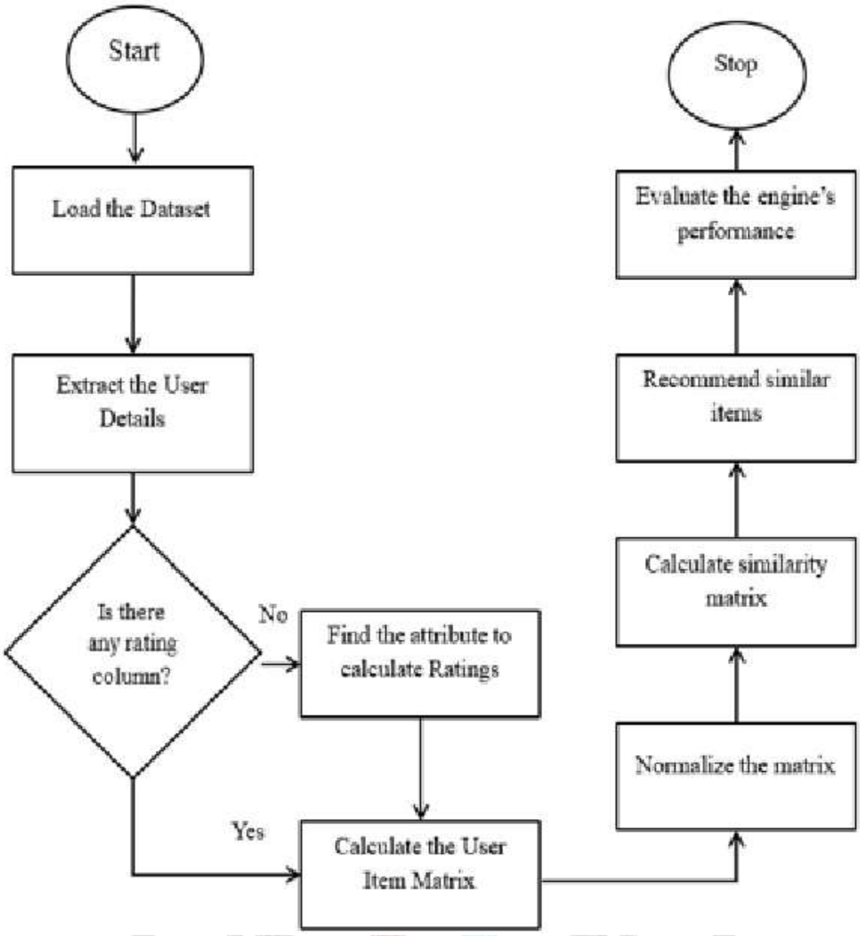

Figure. 1. Flowchart of Item based Collaborative Filtering

3.2Evaluation Metrics

To evaluate the performance of the proposed recommendation system, we employ several commonly used metrics in the field of recommender systems:

1. **Mean Absolute Error (MAE)**: MAE is used to measure the average absolute difference between the predicted and actual ratings. It is calculated as follows:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|r_i - \hat{r}_i|$$

2. **Root Mean Squared Error (RMSE)**: RMSE gives more weight to larger errors and is computed as follows: RMSE is particularly useful when large errors can significantly affect the overall recommendation performance, which is

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(r_i - \hat{r}_i)^2}$$

critical in product recommendation systems where incorrect suggestions could lead to customer dissatisfaction.
3. **Precision and Recall**: Precision measures the proportion of relevant items among the recommended items, while recall measures the proportion of relevant items that are recommended. Both metrics help evaluate how well the system retrieves relevant products based on the user's preferences.
4. **F1-Score**: The F1-score is the harmonic mean of precision and recall, providing a single metric to assess the system's ability to balance both aspects.

By using these metrics, we can evaluate the effectiveness of the proposed recommendation system in providing accurate and relevant product suggestions.

4. Results & Discussion

During the experimental phase, the suggested recommendation system is evaluated using a publicly accessible e-commerce dataset. The system's performance is compared to classic CF techniques like user-based and item-based collaborative filtering. The results show that the hybrid strategy, which combines profile expansion with keyword-based search, outperforms traditional CF approaches in terms of MAE, RMSE, and F1-score.

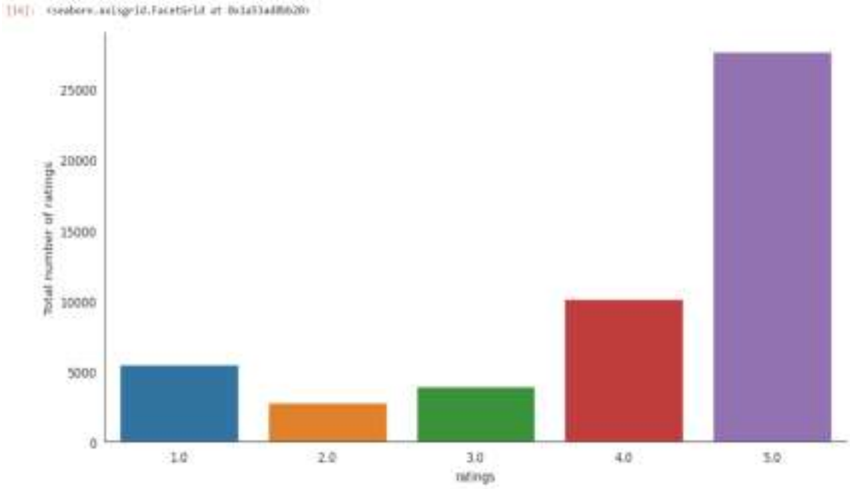

Fig2: distribution of the rating

Furthermore, the system's capacity to manage cold start issues and data scarcity is demonstrated by the addition of demographic data and keyword search, which increases the quality of suggestions for new users and items. The keyword-based search tool, in particular, overcomes data scarcity by recommending related goods even when explicit ratings are not available.

Collaborative Filtering recommender model.

```
electronics_df_CF = pd.concat([train_data, test_data]).reset_index()
electronics_df_CF.head()
```

	index	userId	productId	ratings
0	17509	AY8Q1X7G96HV5	B00000JSES	4.0
1	11968	A243HY69GIAHFI	B00000J3Q7	3.0
2	35533	A1RPTVW5VEOSI	B00003WGP5	5.0
3	31480	A1NVD0TKNS1GT5	B00002JXFH	4.0
4	13526	A23ZO1BVFFLGHO	B00000J570	5.0

Figure.3 Collaborative Filtering recommender model output

User Based Collaborative Filtering model

```
# Matrix with row per 'user' and column per 'item'
pivot_df = electronics_df_CF.pivot(index = 'userId', columns ='productId', values = 'ratings').fillna(0)
pivot_df.head()
```

productId	1400599997	B00000DM9M	B00000J061	B00000J08C	B00000J0A2	B00000J0E8	B00000J1QZ
userId							
A1ISUNUWG0K02V	0.0	0.0	0.0	0.0	0.0	0.0	3.0
A1MJMYLRTZ76ZX	0.0	0.0	0.0	0.0	0.0	0.0	0.0
A1NVD0TKNS1GT5	0.0	0.0	0.0	0.0	0.0	0.0	0.0
A1RPTVW5VEOSI	0.0	0.0	5.0	0.0	0.0	0.0	0.0
A231WM2Z2JL0U3	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Figure. 4: User Based Collaborative Filtering Model Output

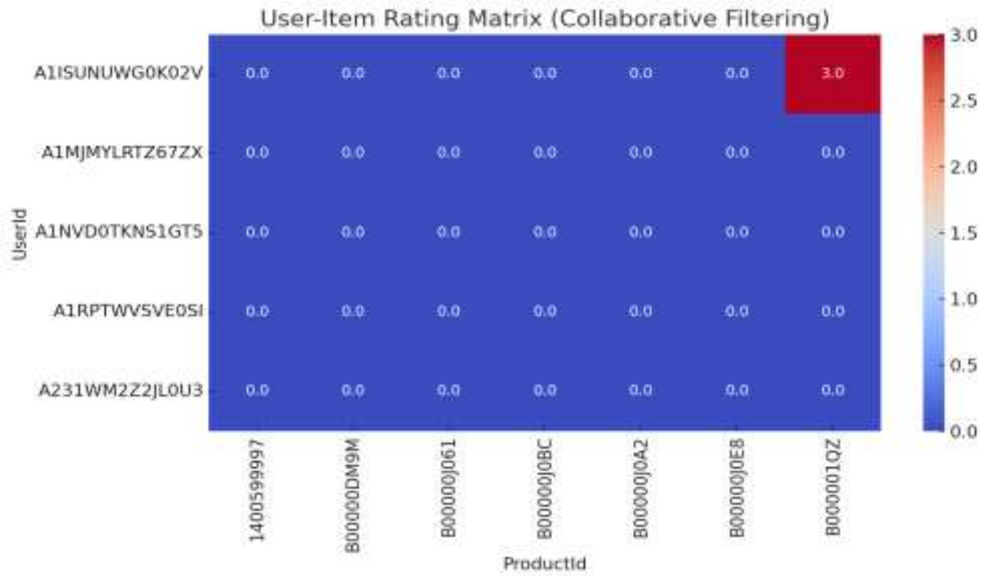

Figure.5. Heat map representation

Here is the heatmap representing the user-item rating matrix for a collaborative filtering model:

- **Rows (UserId)**: Represent individual users.
- **Columns (ProductId)**: Represent items/products.
- **Values**: Ratings provided by users for the respective products, with higher values (e.g., 3.0) highlighted in red.

This visualization helps identify patterns or sparsity in user ratings, which is essential for building and improving collaborative filtering models

5. Conclusion & Future Work

This study presents an improved collaborative filtering-based recommendation system that overcomes a number of common issues, including the cold start problem, data sparsity, and scaling. By adding profile extension, sophisticated similarity assessment algorithms, and a keyword-based search tool, the system provides more tailored and relevant product suggestions. In the future, we intend to investigate the usage of deep learning models, such as neural collaborative filtering and autoencoders, to enhance the system's predictive capabilities. Furthermore, we intend to examine the application of sophisticated natural language processing (NLP) techniques to better comprehend product descriptions and improve the accuracy of keyword-based search recommendations.

References

1. Maleki Shoja, B., Tabirizi, N. (2019). "Customer Reviews Analysis with Deep Neural Networks for e-Commerce Recommender Systems." *IEEE Access*, Vol. 7, pp. 119121-119130. DOI: 10.1109/ACCESS.2019.2937518.
2. S. S. Prasad, P. A. H. Vardhini, M. H. S. V. Sai, K. M. Raju and K. Jyothi, "Identification of Gender from the Shortest Speech using Hybrid and Optimised Spectral Features using Machine Learning Model," 2024 IEEE International Conference on Contemporary Computing and Communications (InC4), Bangalore, India, 2024, pp. 1-7, doi: 10.1109/InC460750.2024.10649072.
3. Iftikhar, A., Ghazanfar, M.A., Ayub, M., Mehmood, Z., Maqsood, M. (2020). "An Improved Product Recommendation Method for Collaborative Filtering." *IEEE Access*, Vol. 8, pp. 123841-123857. DOI: 10.1109/ACCESS.2020.3005953.
4. P. A. Harsha Vardhini, S. S. Prasad, M. H. S Vishnu Sai, C. Santoshi and D. Konduru, "Pioneering Minimalist Speech Analysis Through Optimized Spectral Features Machine Learning Models," 2024 International Conference on Emerging Smart Computing and Informatics (ESCI), Pune, India, 2024, pp. 1-6, doi: 10.1109/ESCI59607.2024.10497288.
5. Srinivas, K. K., Mounika, M., Rao, N. D., Neelima, I., Alekhya, B., Bhagavathi, K. A., ... & Kumar, R. S. MACHINE LEARNING TECHNIQUES TO PREDICT, FORECAST AND DIAGNOSIS OF COVID-19.
6. Won, H., Lee, Y., Shim, J., Ahn, H. (2019). "A Hybrid Collaborative Filtering Model Using Customer Search Keyword Data for Product Recommendation." *Proceedings of 18th IEEE International Conference on Machine Learning and Applications (ICMLA'19)*, Boca Raton, FL, USA, pp. 1523-1526. DOI: 10.1109/ICMLA.2019.00251.
7. M. L. Kumar, P. A. Harsha Vardhini, V. Malathy, M. Anand and P. B. Kumar Chowdary, "Kernel Based FCM for Spinal Cord Segmentation on Computed Tomography Images," 2024 International Conference on Emerging Smart Computing and Informatics (ESCI), Pune, India, 2024, pp. 1-6, doi: 10.1109/ESCI59607.2024.10497368.
8. T. N. S. K. M. Kumar, U. Jaladhi, S. K. C. Rudraraju, V. Shariff, V. R. Reddy Ch and P. A. H. Vardhini, "A Comparison between Shortest Path Algorithms Using Runtime Analysis and Negative Edges in Computer Networks," 2022 International Mobile and Embedded Technology Conference (MECON), Noida, India, 2022, pp. 348-351, doi: 10.1109/MECON53876.2022.9752035.
9. Shapira, B. (2015). *Recommender Systems Handbook.* Springer.
10. V. R. Reddy, D. Yakobu, S. S. Prasad and P. A. H. Vardhini, "Clustering Student Learners based on performance using K-Means Algorithm," 2022 International Mobile and Embedded Technology Conference (MECON), Noida, India, 2022, pp. 302-306, doi: 10.1109/MECON53876.2022.9752165.

11. P. A. Harsha Vardhini, S. P. R. D. Reddy and V. P. Parapatla, "Facial Recognition using OpenCV and Python on Raspberry Pi," 2022 International Mobile and Embedded Technology Conference (MECON), Noida, India, 2022, pp. 480-485, doi: 10.1109/MECON53876.2022.9751867.
12. Xue, H.-J., Dai, X., Zhang, J., Huang, S., Chen, J. (2017). "Deep Matrix Factorization Models for Recommender Systems." *Proceedings of 26th International Joint Conference on Artificial Intelligence (IJCAI'17)*, pp. 3203-3209. DOI: 10.24963/ijcai.2017/447.
13. A. Biswas, K. S. Vineeth, A. Jain and Mohana, "Development of Product Recommendation,Engine By Collaborative Filtering and Association Rule Mining Using Machine Learning,Algorithms," 2020 Fourth International Conference on Inventive Systems and Control, (ICISC), 2020, pp. 272-277, Doi: 10.1109/ICISC47916.2020.9171210
14. Shivaprasad, Satla & Sadanandam, Manchala. (2021). Optimized Features Extraction from Spectral and Temporal Features for Identifying the Telugu Dialects by Using GMM and HMM. Ingénierie des systèmes d information. 26. 275-283. 10.18280/isi.260304.
15. R. S. Kumar, K. K. Srinivas, A. Peddi and P. A. H. Vardhini, "Artificial Intelligence based Human Attention Detection through Brain Computer Interface for Health Care Monitoring," 2021 IEEE International Conference on Biomedical Engineering, Computer and Information Technology for Health (BECITHCON), Dhaka, Bangladesh, 2021, pp. 42-45, doi: 10.1109/BECITHCON54710.2021.9893646.
16. Srinivas, Kandala Kalyana, et al. "Artificial Intelligence based Prediction and Detection of Coronavirus." Artificial Intelligence 11.3 (2022): 678-689.
17. K. K. Srinivas, A. Peddi, B. G. S. Srinivas, P. A. H. Vardhini, H. L. P. Prasad and S. K. Choudhary, "Artificial Intelligence Techniques for Chatbot Applications," 2022 International Mobile and Embedded Technology Conference (MECON), Noida, India, 2022, pp. 292-296, doi: 10.1109/MECON53876.2022.9751887.
18. Satla, Shiva Prasad, Manchala Sadanandam, and Buradagunta Suvarna. "Dangerous Prediction in Roads by Using Machine Learning Models." Ingénierie des Systèmes d'Information 25.5 (2020).
19. Satukumati, Suresh Babu, and R. K. S. Satla. "Feature extraction techniques for chronic kidney disease identification." Kidney 24.1 (2019): 29.