International Journal of Scientific Research in Computer Science, Engineering and Information Technology
ISSN : 2456-3307 (www.ijsrcseit.com)
doi : https://doi.org/10.32628/CSEIT239026

Recommendation System using Data Mining a Review

Mr. Pankaj S. Desai*
Assistant Professor, Department of Computer Engineering, SNJB's Late Sau. K. B. Jain College of Engineering, Chandwad, Maharashtra, India

ARTICLEINFO

Article History:
Accepted: 13 March 2023
Published: 18 March 2023

Publication Issue
Volume 10, Issue 2
March-April-2023
Page Number
82-85

ABSTRACT

Recommendation systems are widely used for suggesting products, social media content, and web content to users. These systems utilize information filtering techniques to predict the preferences or ratings that a user would give to a particular item. This paper presents an overview of various data filtering techniques, algorithms, and application areas utilized in recommendation systems. It also includes a comparison between different algorithms used for recommendation systems.

Keywords : Data Mining, Recommendation System, Content Recommendation, Information Filtering.

I. INTRODUCTION

Recommendation systems, also referred to as recommendation engines, are a type of information system that allows users to discover relevant information from vast amounts of data. The primary function of recommendation systems is to suggest or recommend items or actions to users. Examples of such systems include media recommendations by Netflix, YouTube, and Spotify; online dating suggestions by Tinder; news feeds by Facebook; and product recommendations by Amazon, among others. These systems emphasize personalization, which can lead to data sparsity. In the Media and Entertainment industry, several recommender systems are commonly used, including Collaborative Recommender systems, Content-based recommender systems, Demographic-based recommender systems, Utility-based recommender systems, Database recommender systems, and Hybrid recommender systems.

To identify similarities among items and customer preferences, recommendation systems utilize a variety of techniques, such as information retrieval, machine learning, and data mining. However, as recommendation systems gather voluminous data from multiple sources, the quality of the data collected may be low, which can negatively impact the data mining process. Consequently, various data preprocessing techniques are used to improve the efficiency of the data mining process.

II. RELATED WORK

A recommender system is a decision-making strategy designed to assist users in making choices under complex information environments. It is a means of support that accelerates the social process of using

recommendations from others to make choices when there is insufficient item data or experience of preferences. In recent research, several approaches to building recommendation systems have been developed, including collaborative filtering, content-based filtering, and hybrid filtering. Among these approaches, collaborative filtering is the most widely used and mature.

Collaborative filtering recommends items by linking users with similar item views and using their opinions to recommend items to the active client. Collaborative recommender systems have been implemented in various operational areas. To overcome scalability issues, these systems use cooperative filtering to generate an item-to-item matrix offline, and they recommend other products that are similar online according to the user's search or purchase history.

Content-based filtering uses similarities in features to make recommendations, using algorithms designed to recommend items to users based on data collected about the user. However, the cooperative approach exhibits freezing, sparsity, and scalability problems that can reduce the quality of recommendations. To address these problems, hybrid filtering has been proposed, combining two or more filtering styles in different ways to increase the sensitivity and performance of recommender systems. These approaches combine the strengths of different filtering styles while balancing out their corresponding drawbacks.

Cooperative filtering and content-based filtering approaches are widely used, with some implementations using them separately, while others combine their features. A hybrid recommender algorithm is used by numerous operations due to the new user problem of content-based filtering approaches and the average user problem of cooperative filtering.

III. CLASSIFICATION OF RECOMMENDATION SYSTEM

Recommendation systems predict items for users based on their preferences, using either implicit or explicit information [1]. One of the main goals of the recommendation task is to reduce the time needed to predict what a user would like, and to find a list of items that the user is most likely to enjoy. Explicit information is specific information provided by the user, such as ratings or rankings. Collaborative Filtering (CF) is one of the most successful algorithms that uses explicit information to recommend items, and it has been implemented by online platforms such as TiVo, Amazon, and Netflix [2]. CF-based recommender systems rely solely on ratings provided by a large user community to generate personalized recommendation lists for each individual user.

CF algorithms can be divided into two categories: memory-based algorithms and model-based algorithms [3]. Memory-based (heuristic-based) algorithms use the entire item-user database, and generate prediction based on the ratings in the neighborhood of the active user. Model-based recommendation methods construct a model of user preferences for offline phase before making recommendations.

Content-based approaches, also known as Context-Based Filtering (CBF), build on the assumption that a user likes items with features similar to those of other items they liked in the past [6]. CB recommender systems construct a user profile from rating information, identify like-minded users using a similarity function, and recommend top-N items that like-minded users preferred after their ratings are predicted.

Data mining techniques, mathematical modeling techniques, and software tools are used to find patterns in data and make models. In the context of recommender systems, these methods are used to make recommendation models from large data sets. Recommender systems that incorporate data mining

techniques make their recommendations using data learned from the behavior and attributes of users [7].

DATA PREPROCESSING IN DATA MINING

Data Preprocessing is an essential step in Data Mining, which involves transforming raw data into a format that can be easily understood by algorithms. The quality of data needs to be checked before applying machine learning or data mining algorithms, and this can be determined by checking accuracy, completeness, consistency, timeliness, believability, and interpretability.

The major tasks in data preprocessing are data cleaning, data integration, data reduction, and data transformation. Data cleaning involves removing incorrect, incomplete, or inaccurate data from datasets and replacing missing values. Techniques such as handling missing values, binning, regression, and clustering are used for data cleaning.

Data integration combines multiple sources into one dataset. Schema integration, entity identification problem, and detecting and resolving data value concepts are some of the issues that need to be considered during data integration.

Data reduction helps in the reduction of the degree of the data, making the analysis easier and reducing storage space and computational time. Dimensionality reduction and numerosity reduction are the techniques used for data reduction.

Data compression is used to reduce the size of the data. Lossless compression and lossy compression are the two types of compression. Data transformation involves modifying the format or structure of the data. Techniques such as smoothing, aggregation, and discretization are used for data transformation.

Overall, data preprocessing is crucial for successful data mining, and proper handling of data quality issues and the use of appropriate techniques can significantly impact the quality and relevance of the results obtained from data mining algorithms.

IV. CONCLUSION

The paper aims to provide an overview and critical analysis of recommendation systems that use data mining techniques. The review will likely cover various types of recommendation systems, such as collaborative filtering, content-based filtering, and hybrid models. The paper may also discuss the advantages and limitations of different data mining techniques used in recommendation systems. Overall, this paper can be a useful resource for researchers and practitioners interested in the field of recommendation systems and data mining.

V. REFERENCES

[1]. Nilashi M., Bagherifard M., lbrahim K., Alizadeh O., Arodele H., Nazanin N.L., & Nazanin R. (2013). Collaborative filtering recommendation systems. Research Journal of Applied Sciences, Engineering and Technology, 5, 4168-6182.

[2]. Farokhi N., Vahid M., Nilashi M., & lbrahim O. (2016). A multi-criteria recommendation system for tourism using fuzzy approaches. Journal of soft computing and decision support systems, 3(4), 19-29.

[3]. Sarwar, B., Karypis, G., Konstan, J., & Riedl, J. (2001). Item-based collaborative filtering recommendation algorithm. In Proceedings of the 10th international conference on World Wide Web (pp. 285-295). ACM.

[4]. Nilashi M. (2016). An Overview of Data Mining Techniques in Recommender Systems. Journal of soft Computing and Decision Support Systems, 16-44, October.

[5]. Ininkaye F.O., Folajimi Y.O., & Ojakoh B.A. (2015). Recommendation Systems: Principles, Methods and Evaluation. Egyptian Information Journal, 261-273.

[6]. Modarresi, K. (2016). Recommendation System Based on Complete Personalization.

International Conference on Computational Science, 2190-2204.

[7]. Jooa, J., Bangb, S. W., & Parka, G. (2016). Information Technology and Quantitative Management. Procedia Computer Science, 944-952.

[8]. Vijay Kumar, M., & Pavan Kumar, P. N. V. S. (2017). A Study on Different Phases and Various Recommendation System Techniques.

[9]. Nuñez-Valdez, E. R., Cueva-Lovelle, J. M., Sanjuan, O., & Montenegro-Marin, C. E. (2018). Social Voting Techniques: A Comparison of the Methods Used for Explicit Feedback in Recommendation Systems.

Cite this article as :

Pankaj S. Desai "Recommendation System using Data Mining a Review ", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 9, Issue 2, pp.82-85, March-April-2023.
Available at doi :
https://doi.org/10.32628/CSEIT239026
Journal URL : https://ijsrcseit.com/CSEIT239026