Recommendation Systems: Techniques, Challenges, Applications and Evaluations

Djelloul Bettache and Nassim Dennouni

December 28, 2023

Recommendation Systems: Techniques, Challenges, Applications and Evaluations

1st Djelloul BETTACHE
Computer Science department
Hassiba Benbouali University Of Chlef
Chlef, Algeria
d.bettache@univ-chlef.dz

2nd Nassim DENNOUNI
Ecole superieure de management-Tlemcen
Tlemcen, Algeria
dennouninas@gmail.com

***Abstract*— The rapid expansion of the Internet, mobile technology, and e-commerce has created an overwhelming surplus of information. In response, recommendation systems have been developed to sort and prioritize relevant information for users. These systems empower individuals to discover customized knowledge, products, and services. Since their inception, researchers have diligently worked on improving recommendation systems by utilizing a variety of filtering techniques to enhance both the user experience and system performance. This paper offers an initial examination of recommendation systems that rely on filtering methods, discussing the challenges they encounter and the fields in which they are applied.**

***Keywords*— Recommendation system, Recommendation techniques, Recommendation challenges, Recommendation applications, Evaluation metrics.**

I. INTRODUCTION

When people have many options to choose from, it can be confusing to pick the right one. That's why recommendation systems were created. They help us decide by giving suggestions based on our preferences[1]. People have always sought advice from others for various reasons. These systems help by narrowing down our choices and suggesting what might be best for us[2].The huge amount of information nowadays make these recommendation systems more important. The first computer-based recommendation system was made in 1992, called Tapestry[3]. It helped manage lots of documents like emails and news articles, using human input to be more effective. The main reason for creating recommendation systems is to reduce the load of information and make it easier for users to find what they want. This benefits both users and the companies providing the service[4]. Nowadays, many big companies like Google, Twitter and Netflix use recommendation systems to help them make decisions that increase their profits and reduce risks[5]. Some well-known recommendation systems today include Group Lens, Amazon.com, Netflix, Google News, Youtube, Instagram and Facebook.

II. TECHNIQUES

Numerous approaches have been proposed for creating a recommendation system. Among them, two serve as basic principles techniques: content filtering and collaborative filtering. These techniques are then extended and, in the current RS techniques, they are categorized as follows:

A. Content Filtering

Content-based (CB) techniques use the characteristics of an item to find recommendations for a user based on their past preferences. This method works in two main steps: it creates a user profile using the characteristics generally preferred by the user, and then compares the characteristics of each item to this user profile, recommending items that have a high similarity[1,6]. To use content-based filtering, an article profile is built, which includes the essential characteristics of the article. For example, in the case of a film, the profile might include details such as cast, director, year of release and genre.

CB filtering is a simple approach that does not rely on user feedback. Sometimes a single preference can lead to recommendations for several items. This method works well when information about items is well structured and readily available, as is the case for films, songs, products or books. However, there are limitations, as not all items have a detailed description, making it difficult to measure the similarity among items[7]. These recommendation systems also tend to provide consistent but somewhat static results over time.

B. Collaborative Filtering (CF)

Collaborative filtering (CF) is the most widely used recommendation technique. Its basic principle is that individuals with similar interests tend to share preferences for new and upcoming items[8]. This approach is based on two key principles. Firstly, it identifies a group of users with similar interests, known as 'nearest neighbours', whose opinions are aggregated to form the basis of recommendations. Secondly, it extends this idea by creating a larger group to exert a more significant influence on the recommendation process.

Collaborative filtering techniques are applied in a variety of fields, taking advantage of large datasets. These techniques use ratings or preferences provided by users for various items to predict other items likely to appeal to an active user[1,8].

Collaborative methods can be classified into three categories: memory-based approaches, model-based approaches, and hybrid approaches.

1) Memory-Based Filtering

Memory-based are known for their simplicity and ease of implementation. The most common approach in this category is memory-based neighbourhood filtering[7], which predicts preferences by referring to users with similar tastes to the queried user or items that resemble the queried item. The effectiveness and efficiency of the neighbourhood technique is highly dependent on how the similarity between users or items is calculated. Memory-based techniques fall into two categories: user-based and item-based filtering.

a) User-based: determines the predicted preferences for an active user's items by examining their similarity to other users who have rated the same items [1,10].

b) Item-based: calculates predictions by measuring the similarity between items. This method retrieves all the items rated by the user, evaluates the similarities of these items to the target item and selects the N most similar items to predict the user's preference for the target item[1,10].

2) Model-Based Filtering

Model-based method leverage data mining methodologies to forecast a user's item preferences. They encompass a variety of methods, including association rules, clustering, decision trees, Bayesian classifiers, regression, and latent factor models [1,6]. These models streamline the user-item preference matrix by reducing dimensionality and learning latent variables that aid in predicting user preferences for items during the recommendation process.

3) Hybrid Filtering

Hybrid filtering techniques is created by merging the strengths of two or more filtering methods and overcomes their individual limitations[9]. This approach leads to more effective and improved recommendation results. An example of this is the combination of memory-based and model-based approaches to create a hybrid filtering system, resulting in enhanced prediction accuracy and operational efficiency.

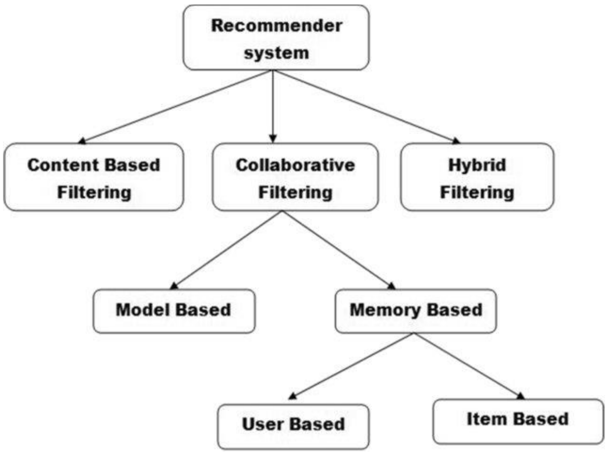

Fig. 1. Recommender System techniques

III. Challenges

Recommender systems (RS) face several challenges, including:

A. Data Sparsity

Numerous e-commerce and online shopping platforms employ recommender systems to evaluate extensive item catalogs. However, as the item sets grow larger, the user-item interaction data becomes increasingly sparse, presenting a challenge for many recommender systems[1,8]. Limited user ratings or preferences can result in less accurate predictions. Additionally, new items are challenging to recommend until they have received sufficient user ratings, and new users may struggle to receive relevant recommendations due to their limited preference history.

B. Scalability

As the number of users and items grows, the computational resources required for recommendation tasks can become a significant challenge[8].

C. Cold Start

New users and items pose a challenge because the system lacks historical data to make accurate suggestions until more interactions occur[8].

D. Evaluation Metrics

Determining the effectiveness of RS poses challenges as traditional evaluation metrics may not reflect the quality of recommendations accurately[8,11].

E. Gray Sheep

Users whose preferences don't align with any particular group are referred to as "gray sheep." These users pose a challenge for collaborative filtering systems as they don't fit neatly into the typical user profiles. To address these issues, a hybrid techniques of content-based and collaborative filtering, can help mitigate the problems associated with "gray sheep" users[8].

F. Synonymy

Synonymy is the phenomenon where several identical or highly similar items are listed under different names. Many recommender systems struggle to identify this underlying connection, resulting in the separate treatment of these products[1,8].

Addressing these challenges is an ongoing effort in the field of recommender systems to enhance the quality and usability of recommendation services[1,8].

IV. Applications

Recommender systems have seen significant growth and use in various service domains. Their applications now include personal, social and business services, all of which are of practical importance in our daily lives and have a significant impact. In general, the applications of recommender systems can be categorised as follows:

A. Social Network

Online social networks like Facebook, Instagram and Twitter are significant digital platforms where users not only share details about their daily lives, hobbies, and interests but also engage and interact with other users[6,9]. The widespread adoption of these social networking services has led to a substantial growth in user-generated data.

B. E-Commerce

This system was created to offer guidance to online shoppers. It's widely used, particularly in the field of e-commerce, and relies on ratings and preferences to generate recommendations. Tagging and user reviews are also utilized to establish connections between users and items[9]. Well-known e-commerce platforms like iTunes, Amazon, and eBay make use of these recommendation systems to enhance the shopping experience for their users.

C. Entertainment

In the realm of entertainment, recommendation systems play a pivotal role in enhancing user experience by delivering personalized and engaging content tailored to individual preferences[1,6]. The application of recommendation systems in the entertainment industry spans various platforms, including streaming services, music platforms, video games, and more.

D. Contents

In recent times, recommender systems have emerged as a vital component of the e-content system, enabling users to discover information and knowledge within digital libraries. They are applied in various domains, including personalized web pages, recommending new articles, and filtering emails[8].

E. Service Oriented

The Internet and mobile devices have created significant opportunities for accessing diverse types of information. This, in turn, has spurred the development of service-based recommendation systems across various domains, including tourist recommendations, travel services, matchmaking services, and consultation services[1,6].

F. Tourism

With the growing interest in travel, the tourism industry has embraced recommendation systems to suggest tourist destinations, travel routes, and transportation options. These recommendation systems rely on situational data, including reviews, location information, user details, time, and weather, which are collected through social networking sites (SNS). This has led to a surge in research focusing on SNS-based recommendation systems in the tourism service sector[12].

G. Education

Traditional classroom education is evolving into a new form known as "Smart Learning" or e-learning, where education takes place online. This shift is driven by the widespread use of smart devices and advancements in wireless networks. Smart education can tap into extensive digital resources and offer personalized learning experiences tailored to individual learners' needs, goals, talents, and interests, all without constraints of time and space[1,8]. This new educational approach aligns with the digital age's learning trends, and education services employing recommendation systems play a crucial role in delivering learning resources that cater to each learner's style and knowledge level. This leads to an effective and efficient learning experience, providing personalized learning content to learners.

V. EVALUATION METRICS

Evaluating the effectiveness and perfecting the algorithms of recommender systems depends on the process of evaluating their performance. To determine the quality of recommendations, a set of metrics is used. These metrics serve as indicators, provide a comprehensive analysis of the performance of recommender systems and provide valuable information for optimisation.

1) Precision: is a pivotal metric in the evaluation of recommendation systems, particularly in assessing the relevance and accuracy of the suggested items[13]. Defined as the ratio of true positive recommendations to the total number of recommended items, precision provides a quantitative measure of the system's ability to accurately identify items that align with the user's preferences. In the context of recommendation systems, a high precision value indicates that a significant proportion of the suggested items are indeed relevant to the user, reflecting a system's capacity to generate precise and valuable recommendations.

2) Recall: in the context of recommendation systems, is a pivotal metric that quantifies the system's ability to capture and recommend all relevant items within a given dataset[13]. Specifically, recall is defined as the ratio of true positive recommendations to the total number of relevant items available in the system. Unlike precision, which focuses on the accuracy of the suggested items, recall emphasizes the system's completeness by evaluating its capacity to identify and recommend all items that would be of interest to the user. In recommendation systems, a high recall value signifies that the system effectively captures a substantial portion of the relevant items, reducing the likelihood of missing items that align with the user's preferences.

3) F1 Score: is a critical evaluation metric in recommendation systems that seeks to strike a balance between precision and recall. It is particularly useful when the goal is to harmonize the trade-off between accuracy and completeness in the recommendation process[13]. The F1 Score is calculated as the harmonic mean of precision and recall, offering a consolidated measure that considers both the precision and recall values simultaneously. This metric is especially valuable in scenarios where precision and recall are of equal importance, and there is a need to assess the overall effectiveness of a recommendation system with a single metric. A high F1 Score indicates a system that not only provides accurate and relevant recommendations but also ensures a comprehensive coverage of all pertinent items.

4) Mean Squared Error (MSE): stands as a fundamental evaluation metric in recommendation systems, particularly in the context of continuous rating predictions[14]. MSE quantifies the average squared difference between the predicted and actual ratings for items in the recommendation system. Mathematically, it is computed as the mean of the squared differences between the predicted and observed ratings. A low MSE value indicates that the recommendation system's predictions align closely with the actual ratings, signifying higher accuracy in predicting user preferences. MSE is valuable in scenarios where the emphasis is on numerical precision, such as in movie or product rating predictions, where the goal is to minimize the squared differences between predicted and actual ratings for a diverse set of items.

5) Root Mean Squared Error (RMSE): is a crucial metric in the realm of recommendation systems, serving as an extension of Mean Squared Error (MSE) for continuous rating predictions[14]. RMSE is calculated as the square root of the average squared differences between predicted and actual ratings for items in the recommendation system. RMSE provides a more interpretable measure than MSE, as it is expressed in the original rating scale. A lower RMSE value indicates a higher degree of accuracy in the recommendation system's predictions, suggesting that the predicted ratings closely match the actual ratings. This metric is particularly valuable in scenarios where the precision of numerical predictions is paramount, such as in movie or product rating predictions.

6) Mean Average Precision (MAP): is a key evaluation metric in the realm of recommendation systems, particularly in information retrieval tasks[15]. It provides a comprehensive assessment of the quality of the ranked recommendations by considering both precision and recall across various levels of cutoffs. MAP is particularly relevant when the focus is on ranking and retrieving relevant items in a ranked list. Mathematically, MAP is calculated as the mean of the Average Precision (AP) values for each user. The Average Precision for a user is determined by computing the precision at each relevant item's position in the recommended list and then averaging these precision values. A higher MAP score indicates a recommendation system that not only provides relevant items but also ranks them effectively, ensuring that highly relevant items appear at the top of the list.

7) Normalized Discounted Cumulative Gain (NDCG): is a vital metric in the evaluation of recommendation systems, specifically in the context of ranked lists[15]. NDCG assesses the relevance and ranking quality of recommended items by considering both the position of each relevant item in the recommendation list and the discounted gain associated with its rank. This metric is particularly valuable when the order of recommendations is crucial, as is often the case in scenarios such as search engine result pages or recommendation lists where users may only view a limited number of items. A higher NDCG score indicates a recommendation system that not only provides relevant items but also ranks them effectively, giving more weight to items at higher positions in the list.

8) Coverage: is a critical metric in the evaluation of recommendation systems, providing insights into the extent to which the system is capable of recommending a diverse set of items from the entire catalog[16]. This metric is particularly relevant in scenarios where the goal is to ensure that the recommendation system effectively spans a broad range of items, avoiding over-reliance on a subset of popular or frequently recommended items. Mathematically, Coverage is calculated as the proportion of unique items in the recommendation list relative to the total number of distinct items available in the system's catalog. A higher Coverage value indicates a recommendation system that suggests a more comprehensive set of items, contributing to a more diverse user experience.

9) Diversity metrics: are crucial components in the evaluation of recommendation systems, designed to quantify the variety and distinctiveness of the recommended items within a given list or across multiple recommendation lists[16]. These metrics play a pivotal role in assessing the system's ability to provide diverse and well-rounded suggestions, enhancing user satisfaction and preventing over-specialization on certain types of items. In the context of recommendation systems, diversity metrics can be categorized into two main types: intra-list diversity and inter-list diversity. Intra-list diversity focuses on the variety of items within a single recommendation list, evaluating how different the recommended items are from each other. This is particularly relevant for ensuring that the user is presented with a heterogeneous set of options within a single interaction. On the other hand, inter-list diversity examines the diversity of recommendations across multiple lists, assessing how distinct the suggestions are when considering a sequence of interactions or recommendation sessions.

This list is not exhaustive, and the choice of metrics often depends on the specific context of application and the particular objectives of the recommendation system (user study, offline evaluation, or online evaluation).

VI. Conclusion

Recommendation systems have become part of our daily lives. With the development of the internet and the amount of information available, it is important to have better recommendation systems that work well. These systems help users find things they like and might not otherwise have discovered. This article describes the different recommendation systems, how they work, the challenges they face and where they are used.

References

[1] Malik, S., Rana, A., & Bansal, M. (2020). A survey of recommendation systems. Information Resources Management Journal (IRMJ), 33(4), 53-73.

[2] Nagarnaik, P., & Thomas, A. (2015, February). Survey on recommendation system methods. In 2015 2nd international conference on electronics and communication systems (ICECS) (pp. 1603-1608). IEEE.

[3] Goldberg, D., Nichols, D., Oki, B.M., Terry, D: Using collaborative filtering to weave an information tapestry. Commu. ACM 35(12), 61–70 (1992)

[4] Patel, K., & Patel, H. B. (2020). A state-of-the-art survey on recommendation system and prospective extensions. Computers and Electronics in Agriculture, 178, 105779.

[5] Bouneffouf, D., Bouzeghoub, A., Ganarski, A.L.: Risk-aware recommender systems. In: Neural Information Processing, pp. 57–65. Springer, Berlin, Heidelberg (2013).

[6] Zuva, T., Ojo, S. O., Ngwira, S., & Zuva, K. (2012). A survey of recommender systems techniques, challenges and evaluation metrics. International Journal of Emerging Technology and Advanced Engineering, 2(11), 382-386.

[7] Afsar, M. M., Crump, T., & Far, B. (2022). Reinforcement learning based recommender systems: A survey. ACM Computing Surveys, 55(7), 1-38.

[8] Shapira, B., Rokach, L., & Ricci, F. (2022). Recommender systems handbook.

[9] Raza, S., & Ding, C. (2022). News recommender system: a review of recent progress, challenges, and opportunities. Artificial Intelligence Review, 1-52.

[10] Papadakis, H., Papagrigoriou, A., Panagiotakis, C., Kosmas, E., & Fragopoulou, P. (2022). Collaborative filtering recommender systems taxonomy. Knowledge and Information Systems, 64(1), 35-74.

[11] Zangerle, E., & Bauer, C. (2022). Evaluating recommender systems: survey and framework. ACM Computing Surveys, 55(8), 1-38.

[12] Evaluating Recommender Systems: Survey and Framework

[13] Samundeeswary, K., & Krishnamurthy, V. (2017, June). Comparative study of recommender systems built using various methods of collaborative filtering algorithm. In 2017 International Conference on Computational Intelligence in Data Science (ICCIDS) (pp. 1-6). IEEE.

[14] Zhang, F., Gong, T., Lee, V. E., Zhao, G., Rong, C., & Qu, G. (2016). Fast algorithms to evaluate collaborative filtering recommender systems. Knowledge-Based Systems, 96, 96-103.

[15] Kaminskas, M., & Bridge, D. (2016). Diversity, serendipity, novelty, and coverage: a survey and empirical analysis of beyond-accuracy objectives in recommender systems. ACM Transactions on Interactive Intelligent Systems (TiiS), 7(1), 1-42.

[16] Kaminskas, M., & Bridge, D. (2016). Diversity, serendipity, novelty, and coverage: a survey and empirical analysis of beyond-accuracy objectives in recommender systems. ACM Transactions on Interactive Intelligent Systems (TiiS), 7(1), 1-42.