Performance Analysis of a Hybrid Recommender System

Uzair Sultan, Hajra Khan, Yasir Saleem Afridi, Mian Ibad Ali Shah, Muniba Ashfaq, Affera Sultan
Department of Computer Systems Engineering, UET, Peshawar, Pakistan.
*Correspondence: uzairsultan10@gmail.com, hajrakh3@gmail.com, yasirsaleem@uetpeshawar.edu.pk, ibad@uetpeshawar.edu.pk, muniba@uetpeshawar.edu.pk, sultan.affera@gmail.com

Citation | Sultan. U, Khan. H, Afridi. Y. S, Shah. M. I. A, Ashfaq. M, Sultan. A, "Performance Analysis of a Hybrid Recommender System", IJIST, Special Issue, pp 257-265, May 2024
Received | May 13, 2024, **Revised |** May 20, 2024, **Accepted |** May 25, 2024, **Published |** May 28, 2024.

In the prevailing information age, human confrontation with extensive information makes it difficult to segregate the relevant content on the basis of choices and priorities. This gives rise to the need for effective recommendation systems that can be incorporated into distinct and diversified domains such as e-commerce, social media, and news media websites and applications. By giving suggestions, these recommender systems efficiently reduce huge information spaces and direct the users toward the items that best match their requirements and preferences. Hence, they play an important role in filtering out the relevant user-specific information. Based on the working principle, recommender systems can be classified into Content-Based Systems, Collaborative Filtering Systems, or Popularity-Based Systems. However, to cope with the problems of cold-start and plasticity that are associated with standalone recommender systems, hybrid recommendation systems are being introduced. This research is therefore focused on the development of a Weighted Hybrid Model that combines the scores of the three standalone recommender models in a linear fashion. The performance of the proposed hybrid model is tested against all three standalone models on an online News dataset. Using a Top-N accuracy metric, it is found that the accuracy of the weighted hybrid model is higher than the standalone Content-Based, Collaborative, and Popularity-Based models against the same dataset. An efficiency of 90% for the Hybrid model was achieved compared to the best-performing standalone model having an efficiency of 53%.

Keywords: Machine Learning, Data Mining, Recommender System, Collaborative Filtering, Hybridization Techniques, Evaluation Metrics, Mean Absolute Error (MAE), Cross-Validation, Training Data.

Introduction:

Data and information have gained unprecedented pace in today's world of information and technology. People have become considerably dependent on technology and living without it seems inconceivable. People are overwhelmed with the extent of data available; hence, it is almost impossible to browse through such vast information space in search of the required and related news stories [1]. Consequently, the interest of users in the consumption of news either gets lowered or is completely lost. Recommendation of news articles faces many challenges because of the dynamic environment, such as new updates in the articles and changes in user preferences. Therefore, an efficient news recommendation system must be able to process and address the rapidly evolving and continuous inflow of news [1]

Content-based and collaborative filtering recommendation systems (discussed in Section 2) can provide some effective recommendations but each of the approaches comes with a few disadvantages. [2] To tackle this issue, this research is therefore focused on the development of a "Hybrid Recommendation Model" which has been built by combining content-based filtering and collaborative filtering systems. This hybrid model will also be combined with the popularity model to increase the accuracy of recommendations and to cope with problems such as cold start.

Literature Review:

Recently, recommending news articles or other documents in the format of web objects has gained more research attention. Several adaptive news recommending systems, such as Google News and Yahoo! News provide personalized news recommendation services for a substantial number of online users [3]. A personalized new recommendation system with the help of the popular microblogging service "Twitter" has been proposed by [4]. News articles are recommended based on the popularity of the article identified from Twitter's public timelines.

Models and memory-based algorithms have been used which utilize the weighted average of past ratings from other users and the weight is proportional to the similarity between the users [5]. Measures that have been used to find similarity are the Pearson Correlation Coefficient and Cosine Similarity. This approach models the user preferences based on prior information. The research carried out by [5] was taken further by using information filtering that filters the relevant information from the unwanted information stream [6]. Firstly, an analysis of user's interests spanned over a period of 14 months was carried out by using click distribution. Then Bayes rule was applied to predict the users' interest for a particular period of time. These predictions made for the particular times were then combined and a final prediction was made for the user against a longer period of the time.

In earlier times, Popularity was one of the only metrics used for news recommendations [7][8] and [9]. The popularity-based news recommendations were based on the number of times a news article has been read by the users. However, this approach had a lot of shortcomings especially in cases where the news gets old, even though it is popular it might not be of interest to the users. On the other hand, the newly published news with low popularity might be of more interest to the users. A notable work in the area of News Recommendation Systems is SCENE (Scalable two-stage Personalized News Recommendation system) [10]. The main focus of this research was on the news selection. Experiments were conducted on how to match the news with the user interests while maintaining the highest accuracy and diversity. The goal was achieved by maintaining a huge metadata of the users, thereby affecting the news selection and improving the news recommendation.

The main application area of the Content-Based Filtering (CBF) technique is semantic-based Recommender and specific news websites. In order to address the shortcomings of both Collaborative Filtering (CF) and CBF approaches, a Hybrid recommendation approach is adopted. It has been observed that the Hybrid approach has been exhaustively used for news Recommender since 2010 (i.e., 60%) because the challenges faced by CBF and CF techniques can be overcome by using a complementary technique such as a hybrid model [11] The collaborative filtering approach in the recommendation system is demonstrated in Scalable Collaborative Filtering with Jointly Derived Neighbourhood Interpolation Weights. This improves the prediction accuracy by improving the interpolation precision and simultaneously deriving the interpolation weights for all nearest neighbors [12]. This technique has been called Singular Value Decomposition (SVD) recommenders in [13].

Other related research includes Europe Media Monitor (EMM) News Explorer [14] and Newsjunkie [15] which uses the news content and named entities for carrying out recommendations of news. However, EMM News Explorer does not provide personalized services and Newsjunkie does not address the news selection, news presentation, and scalability issues. In [16], the recommendation is carried out by using two modules: an offline module that pre-processes the data to build reader and content models, and an online module that uses these models in real-time to recognize the reader's needs and goals and predict a recommendation list, accordingly. The recommended objects are obtained by using a range of recommendation strategies mainly based on content-based filtering and collaborative filtering approaches, each applied separately or in a combination [16].

Novelty of the Research:

The related work from the literature review mostly uses algorithms that involve collaborative filtering or content-based filtering when it comes to news/blog recommendations. To achieve more promising and accurate results, two algorithms (namely, Collaborative and Content-Based Filtering) have been combined to propose a hybrid recommendation system that takes the weighted average of collaborative filtering scores with content-based scores and performs recommendation ranking by resultant scores. Further, to cope with the cold start issue, this developed model has also been combined with a popularity-based model. The accuracy of results achieved by this novel hybrid recommendation model increased exponentially.

Recommendation Systems:

The definition of recommendation systems has evolved over the past 14 years. Recommender systems aim to predict users' interests and recommend items that are most likely of interest to them. These are some of the most powerful machine learning algorithms that are used to increase traffic. [17] Recommendation systems are typically used to speed up the search process and make it easy for users to reach the content that truly interests them. Several Recommendation systems have been deployed over the decade for different domains [15]. The two main types of recommendation systems [18] are discussed in Sections 2.1 and 2.2.

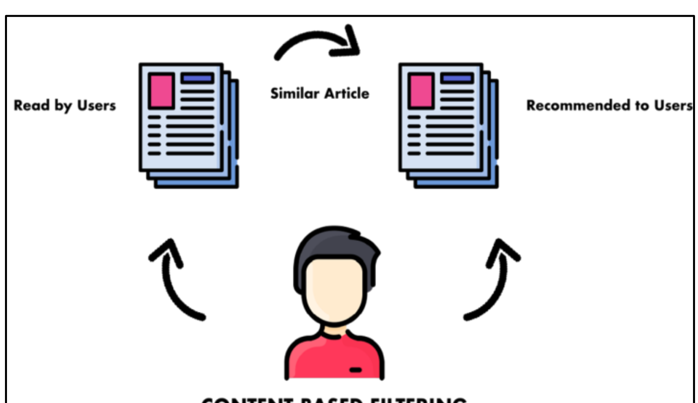

Figure 1: Working of Content-Based Filtering

Content-Based Filtering Approach:

In content-based filtering a user profile is created which is then used to make recommendations to the user [19]. The data is collected, either explicitly (rating) or implicitly (clicking on a link). So, the idea in content-based filtering is to tag products using certain keywords, understand what the user likes, look up those keywords in the database, and recommend different products with similar attributes [19]. The main focus in the content-based approach is on the attributes of the items, where the similarity among items is determined by comparing the attributes of these items. In the content-based filtering approach, users or items are considered atomic units, and better-personalized recommendations are made by gathering more information about the user or item [19].

For content-based filtering, the most popular vector space model Term Frequency–Inverse Document Frequency (TF-IDF) is used to select a set of candidate items. TF-IDF score is the product of term frequency (the number of times a term appears in the document) and Inverse document frequency (a measure of whether a term is rare or common in the collection of documents) [20]. Mathematically, the TF-IDF score for the word **t** in document **d** from the document set **D** is calculated as follows [20]:

$$TF-IDF\ (t,d,D)=tf\ (t,f) \times idf\ (t,D)$$

Where,

$$tf\ (\mathrm{t}, d) = log\ log\ (1 + freq\ (t, d))$$

$$idf\ (t\ , D)=log\ log\ ((N/count)(d \in D : t \in d)$$

TF-IDF algorithms apply to different dimensions of an article by selecting contents and calculating similarity among them.

Challenges Faced by Content-Based Filtering:

As content-based filtering decides to only use the tags, it faces a few challenges in recommending as:

Limited Content Analysis: If the content doesn't contain enough information to discriminate the items precisely, the recommendation itself risks being imprecise [1].

No Diversity: The user will never be recommended different items that he might like. Due to this business might not expand [1].

New Users: User Profile Information collected requires a considerable amount of data and features to define an item. Adequate information is thus required to be able to recommend effectively [1].

Collaborative Filtering Approach:

Collaborative filtering is a technique that can filter out items that a user might like based on reactions by similar users. [1]. It is based upon historical data and the assumption here is that the user who has liked or viewed something in the past is likely to like or view similar content in the future as well. Collaborative filtering can be of two types [1].

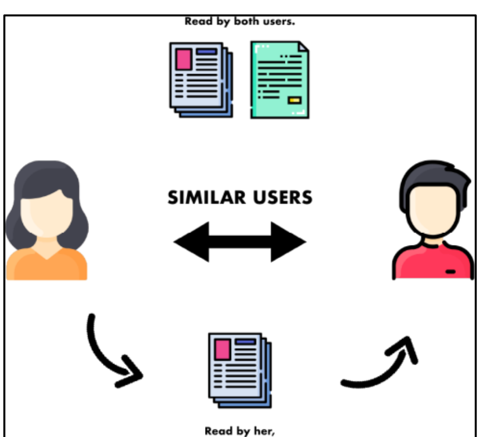

Figure 2: User-Based Collaborative Filtering

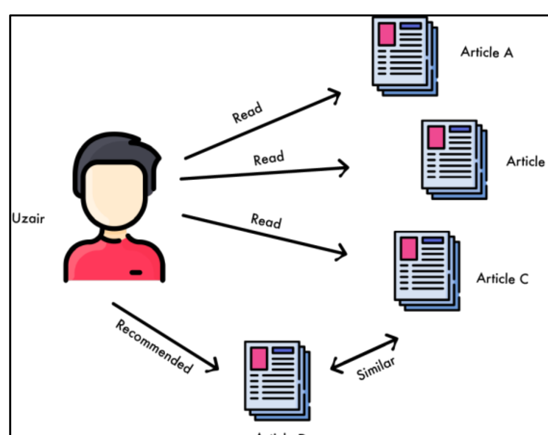

Figure 3: Item-Based Collaborative Filtering

User-Based Collaborative Filtering:

In user-based collaborative filtering, recommendations are made based on similar users who share the same rating pattern as active users.

Item Based Collaborative Filtering:

Item-based collaborative filtering is a type of recommendation method that looks for similar items based on the items users have already liked or positively interacted with.

Single Value Decomposition:

(SVD) is used as a Matrix factorization technique to find the similarities between users as well as similarities between the articles. An advantage of using this approach is that instead of having a high-dimensional matrix containing an abundant number of missing values, a much smaller matrix in lower-dimensional space is required. The more factors there are, the more specific the model will be. However, the model may result in overfitting, if too many factors are taken into consideration.

Challenges Faced by Collaborative Filtering:

Some of the challenges faced by the collaborative filtering are:

Cold Start: Cannot effectively handle the fresh items. Since fresh items or users do not have enough ratings to find a relation [7].

Scalability:

Collaborative filtering essentially depends on the number of users of a system. As the number of users increases, the complexity of analyzing similar users and examining the history of the items used by these users increases exponentially. Scalability is also an issue where online systems need to react immediately to the requirements of the users [7].

Shilling Attack: The abusive use of liking or disliking a product can affect this recommendation model [7].

Data Sparsity: People mostly don't rate the items. Due to this there are missing ratings of a new item. This leads to poor recommendations. [7]

Methodology:

The proposed method was to combine the two most popular recommendation models, i.e., Content-based filtering and Collaborative filtering, and develop a more accurate hybrid model by eliminating drawbacks of each other. Further, to cope with the issue of cold start and also increase the model accuracy, the proposed hybrid recommendation model is combined with the popularity-based recommendation model.

Experimental Setup:

In this research, the recommendation models have been developed in Python. To overcome overfitting and to randomize the dataset, a cross-validation technique called Holdout has been used to validate the results. The data set was split into train data and test data, with 20% of the data for testing purposes and 80% for training the model.

Model Selection Hybrid Recommendation Model

Content-based and collaborative filtering can provide some effective recommendations but each of the approaches has some disadvantages. [2] Thus, in order to improve the results, a hybrid model has been used in which the two algorithms content-based filtering and collaborative filtering were combined. It takes the weighted average of Collaborative filtering - Scores with Content-based scores and performs recommendation ranking by resultant scores.

$$Resultant\ Hybrid\ Score = CBF\ Score \times SVD\ Score$$

This Hybrid model has also been combined with the popular model to cope with cold start and to increase the accuracy of recommendations even more. Popularity-Based Recommendation Model – It takes the weighted average of Collaborative filtering Scores with Content-based scores and performs recommendation ranking by resultant scores. popularity-based recommendation system works with the trend. It basically uses the items that are in trend right now. For example, if any item is usually viewed by every new user, then there are chances that it may suggest that item to the user who just visited. "Popularity-Based Filtering Model sums up users' interactions with each article and recommends the article with the most interactions."

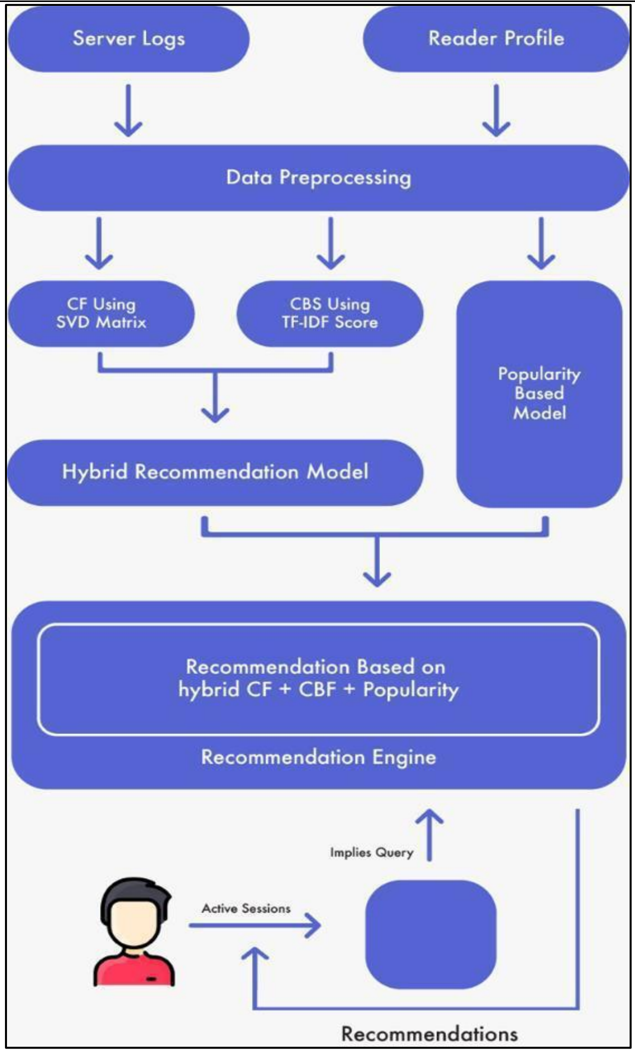

Figure 4: Flowchart of the overall methodology

Dataset:

The dataset used to train and test the hybrid recommendation model has been downloaded from Kaggle [17] which contains the real samples of 12 months of news. The dataset is properly labelled and has two parts. One part contains the **Articles** and their timestamp, content, author, content type, URL, title, and text. The second part contains **User Interactions** such as event Type, content, session ID, user Agent (Browsers, etc.), user Region, user Country, etc. It contains data of about **73000 logged-in users** and more than **3000 user interactions.**

Data Pre-Processing:

There are five types of user interactions in the dataset. **VIEW** (The user has opened the article), **LIKE** (The user has liked the article), **BOOKMARKS** (The user has bookmarked the article for easy return in the future), **FOLLOW** (The user chose to be notified on any new comment in the article), **COMMENT** (The user created a comment in the article). Since there are different interaction types, specific weights should be associated with the interactions. For instance, a "comment" on an article weighs more than a "simple view". Henceforth, in the first step of pre-processing, weights were associated with each interaction that the user performs on the data based on their importance in accurate recommendations such as: View 1.0, Like 2.0, Bookmark 2.5, Follow 3.0, Comment 4.0.

In the second step, the users and interactions having weights less than five were removed. This is because the users and interaction with such low weight cannot be considered as a base for recommendation. The details of the dataset after pre-processing can be seen in the table.

Interactions before removal	72312
Interactions after removal	69869
Users before removal	1897
Users after removal	1140

In the final step, the Unique Interactions were calculated using an already calculated weighted sum of interactions for each news article of every user.

Total Unique Interactions	39106

Evaluation metrics:

To find the efficiency of the proposed model, evaluation has been carried out as it is one of the most important aspects of any machine learning project. Evaluation enables to comparison of different hyper-parameter choices and algorithms for the models. [21]. Evaluation also helps us identify the generic nature of the proposed model, also known as Generalization. Hence, the holdout method of cross-validation has been used to generalize our model by tuning it accordingly.

The evaluation technique that has been used in this research is the Top-N accuracy metrics. It evaluates the accuracy by comparing the recommendations provided to the user with the items the users actually interacted with in the test set [21]. The Top-N accuracy metric that was chosen for the hybrid model is Recall@N which evaluates whether the interacted item is among the top N items in the ranked list. [21]b

Results and Discussions:

The evaluation of the Popularity Model was performed by the method discussed in section 3, and it is surprising that a popularity model can perform this well. It achieved the Recall@5 of 0.241, meaning 24% of interacted items in test sets were ranked by popularity model among the top-5 items. Recall@10 was 37% higher than expected. In the content-based filtering, a Recall@5 of 0.414 was achieved, which means about 42% of interacted items in the test set were tanked by content-based filtering among the top 5 items. Recall@10 was higher at 53%.

In the Collaborative filtering, a Recall@5 of 0.333 was achieved, which means about 33% of interacted items in the test set were tanked by content-based filtering among the top 5 items. Recall@10 was higher at 47%. After these results, the developed Hybrid Recommendation model was also evaluated, and the results were higher than all other recommendation models. A Recall@5 of 0.434 was achieved, which means about 44% of the interacted items in the test set were ranked by the Hybrid Model among the top-5 items. Recall@10 was 54%. After combining the popularity-based model with the Hybrid model, Recall@5 increased up to 65% and Recall@10 reached up to 90%. These results can be thoroughly seen in Table 1.

Table 1: Tabular Comparison of Recommendation Models

Model Name	Recall@5	Recall@10
Popularity Model	24%	37%
Content-Based Filtering	42%	53%
Collaborative Filtering	33%	47%
Hybrid Model	44%	54%
Hybrid + Popularity	65%	90%

Comparison of Different Recommendation Models:

The results of recommendation models working independently were compared with the proposed hybrid model. The results were clearly in favor of the hybrid recommendation model. It has been found that when the proposed Hybrid model was merged with the Popularity model, the accuracy was even increased to 90%. A comparison of the results has been shown in Table 1 and represented graphically in Figure 5.

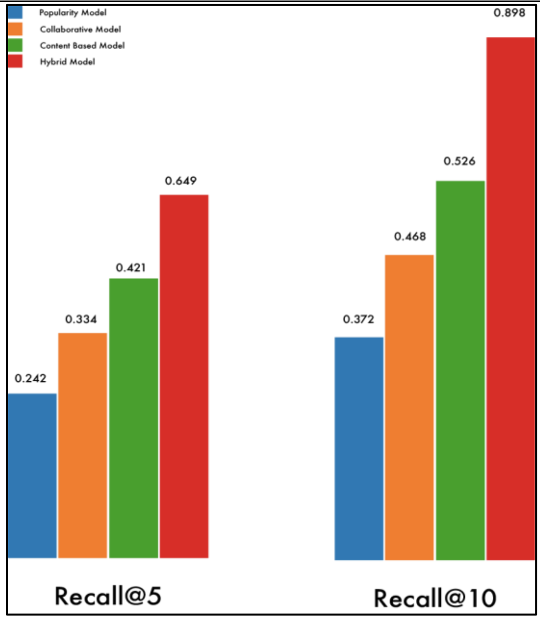

Figure 5: Graphical representation of the accuracy of Different Recommendation Models

Conclusion:

Content-based and collaborative filtering can provide some effective recommendations but each of the approaches has some disadvantages. In this research, the development of a Weighted Hybrid Model was proposed that combines the scores of the three standalone recommender models in a linear fashion. The hybrid recommendation model takes the weighted average of the collaborative filtering score and the content-based filtering scores to carry out the recommendation ranking based on the resultant scores. This Hybrid model has also been combined with the popular model to cope with cold start and to increase the accuracy of recommendations even more. The performance of the proposed hybrid model is tested against all three standalone models on an online News dataset. Using a Top-N accuracy metric, it is found that the accuracy of the weighted hybrid model is higher than the standalone Content-Based, Collaborative, and Popularity-Based models against the same dataset. An efficiency of 90% was achieved for the proposed Hybrid model compared to the best-performing standalone model having an efficiency of 53%. This research is not limited to news websites only, but it can be extended to other recommendations such as in e-commerce.

References:

[1] A. K. Chaturvedi, F. Peleja, and A. Freire, "Recommender System for News Articles using Supervised Learning," Jul. 2017, Accessed: May 18, 2024. [Online]. Available: https://arxiv.org/abs/1707.00506v1

[2] N. Jonnalagedda, S. Gauch, K. Labille, and S. Alfarhood, "Incorporating popularity in a personalized news recommender system," PeerJ Comput. Sci., vol. 2016, no. 6, p. e63, Jun. 2016, doi: 10.7717/PEERJ-CS.63/SUPP-2.

[3] "A HYBRID NEWS RECOMMENDER SYSTEM." Accessed: May 18, 2024. [Online]. Available: https://www.researchgate.net/publication/325550103_A_HYBRID_NEWS_RECOMMENDER_SYSTEM

[4] R. Steinberger, B. Pouliquen, and E. van der Goot, "An introduction to the Europe Media Monitor family of applications," Sep. 2013, Accessed: May 18, 2024. [Online]. Available: http://arxiv.org/abs/1309.5290

[5] Y. Zhou, D. Wilkinson, R. Schreiber, and R. Pan, "Large-Scale Parallel Collaborative Filtering for the Netflix Prize," Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 5034 LNCS, pp. 337–348, 2008, doi: 10.1007/978-3-540-68880-8_32.

[6] R. M. Bell and Y. Koren, "Scalable collaborative filtering with jointly derived neighborhood interpolation weights," Proc. - IEEE Int. Conf. Data Mining, ICDM, pp. 43–52, 2007, doi: 10.1109/ICDM.2007.90.
[7] C. Feng, M. Khan, A. U. Rahman, and A. Ahmad, "News Recommendation Systems-Accomplishments, Challenges Future Directions," IEEE Access, vol. 8, pp. 16702–16725, 2020, doi: 10.1109/ACCESS.2020.2967792.
[8] I. Cantador, P. Castells, and L. Gardens, "Semantic contextualisation in a news recommender system," 2009.
[9] W. Zhou, J. Wen, Q. Qu, J. Zeng, and T. Cheng, "Shilling attack detection for recommender systems based on credibility of group users and rating time series," PLoS One, vol. 13, no. 5, p. e0196533, May 2018, doi: 10.1371/JOURNAL.PONE.0196533.
[10] L. Li, D. Wang, T. Li, D. Knox, and B. Padmanabhan, "SCENE: A scalable two-stage personalized news recommendation system," SIGIR'11 - Proc. 34th Int. ACM SIGIR Conf. Res. Dev. Inf. Retr., pp. 125–134, 2011, doi: 10.1145/2009916.2009937.
[11] P. Viana and M. Soares, "A Hybrid Approach for Personalized News Recommendation in a Mobility Scenario Using Long-Short User Interest," https://doi.org/10.1142/S0218213017600120, vol. 26, no. 2, Apr. 2017, doi: 10.1142/S0218213017600120.
[12] G. Adomavicius and A. Tuzhilin, "Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions," IEEE Trans. Knowl. Data Eng., vol. 17, no. 6, pp. 734–749, Jun. 2005, doi: 10.1109/TKDE.2005.99.
[13] D. Khattar, V. Kumar, M. Gupta, and V. Varma, "Neural Content-Collaborative Filtering for News Recommendation", Accessed: May 18, 2024. [Online]. Available: http://ceur-ws.org
[14] S. Jiang and W. Hong, "A vertical news recommendation system: CCNS - An example from Chinese campus news reading system," Proc. 9th Int. Conf. Comput. Sci. Educ. ICCCSE 2014, pp. 1105–1114, Oct. 2014, doi: 10.1109/ICCSE.2014.6926634.
[15] E. Gabrilovich, S. Dumais, and E. Horvitz, "Newsjunkie," pp. 482–490, May 2004, doi: 10.1145/988672.988738.
[16] J. Liu, P. Dolan, and E. R. Pedersen, "Personalized news recommendation based on click behavior," Int. Conf. Intell. User Interfaces, Proc. IUI, pp. 31–40, 2010, doi: 10.1145/1719970.1719976.
[17] "Articles sharing and reading from CI&T DeskDrop." Accessed: May 18, 2024. [Online]. Available: https://www.kaggle.com/datasets/gspmoreira/articles-sharing-reading-from-cit-deskdrop/data
[18] M. D. Ekstrand, "LensKit for Python: Next-Generation Software for Recommender Systems Experiments," Int. Conf. Inf. Knowl. Manag. Proc., pp. 2999–3006, Oct. 2020, doi: 10.1145/3340531.3412778.
[19] X. Su and T. M. Khoshgoftaar, "A Survey of Collaborative Filtering Techniques," Adv. Artif. Intell., vol. 2009, pp. 1–19, Oct. 2009, doi: 10.1155/2009/421425.
[20] A. Garcia Esparza, G. Esparza, M. P. O, B. Smyth, and S. Garcia Esparza, "A multi-criteria evaluation of a user generated content based recommender system," Freyne, J. al. (eds.). Proc. 3rd ACM RecSys'10 Work. Recomm. Syst. Soc. Web, pp. 23–27, Oct. 2011, Accessed: May 18, 2024. [Online]. Available: http://hdl.handle.net/10197/3509
[21] M. Deshpande and G. Karypis, "Item-based top-N recommendation algorithms," ACM Trans. Inf. Syst., vol. 22, no. 1, pp. 143–177, Jan. 2004, doi: 10.1145/963770.963776.