

# Integration of social media and Hybrid Recommendation System

<sup>[1]</sup>Smitha N, <sup>[2]</sup>Soumyajyoti Sanyal, <sup>[3]</sup>Siddhartha V, <sup>[4]</sup>Ashish M Jacob

<sup>[1]-[4]</sup> CMR Institute of Technology, Department of Computer Science & Engineeringt, Bengaluru, India

**Abstract**—Advancement in technology has one major purpose – to make life easier. A majority of people shop online and a lot of products bought online are based on the technology of recommendation systems. Learning the behaviors and patterns of customers is crucial to the functioning of such systems. Larger data leads to better performance of recommendation systems. This study aims to review work done in the area of recommendation systems and future scope. Insight into the integration of social media content to hybrid recommendation systems and its advantages are highlighted.

**Index Terms** —Collaborative Filtering, Content-based Filtering, Hybrid Recommendation.

## 1. INTRODUCTION

E-commerce applications have expanded quickly as a result of the transformation of consumer behavior and corporate operations, as well as the rapid advancement of technology. The rapid and trouble-free delivery of goods or services to customers is one of the most crucial goals of online commerce. E-commerce is now more common due to factors including lower costs, time savings, and decreased storage expenses [1].

The Recommendation System has emerged as a result of e-commerce sites' efforts to boost the sales potential of their goods and services. Recommendation Systems is a system that uses a variety of machine learning algorithms to provide users with personalized goods and services as well as simple purchase options [2]. The topic of which product will be recommended to whom has arisen due to the growing diversity of products available and the fact that each user has unique preferences and experiences. This issue is resolved by suggestion systems, which identify users' interests and suggest the best product for them.

By incorporating user-provided data as well as their preferences, interests, and experiences, recommendation systems can learn from the decisions made by their users. With this data, it determines whether a user and a product or service are a good fit, determines whether a user and an item are comparable, and provides the best suggestion possible based on the users' requirements.

Collaborative recommendation systems, popularity-based recommendation systems, content-based recommendation systems, and hybrid recommendation systems are the four different types of recommendation systems. This study analyzed the algorithms used by the four distinct recommendation systems and looked at how they were used on e-commerce websites [3].

One of the most popular recommendation systems is the collaborative recommendation system. It is solely dependent on prior performance and the similarity of preferences and tastes between two users [4]. This recommendation system gathers and examines data on user interests and makes predictions about the things users will like based on their affinity for other users. When it discovers users who are similar to them, it assesses how much their preferences match and makes appropriate product recommendations [5].

In content-based recommendation systems, the item information that the user has previously engaged with is used in addition to the user-item interaction [5].

The basis of this recommendation system is the notion of showing people products that are comparable to their previous tastes. For example, The user is given a recommendation by looking at the films with comparable subject matter to the one they are now watching.

Users are directly recommended items using popularity-based recommendation systems, which identify items that are trending or that are popular with other users. It is unaffected by the user's preferences, in contrast to other recommendation systems.

Collaborative recommendation systems and content-based recommendation systems are the foundation of hybrid recommendation systems [6]. A model can be developed and utilized to get the greatest performance

by including content-based features in collaborative suggestion systems or by using cases where the opposite is true. By combining these two methods, it is hoped to prevent the shortcomings of each system and achieve more effective results. Large corporations like Netflix, Spotify, and Amazon greatly profit from this arrangement. For instance, Netflix provides recommendations while also considering user ratings and prior user behavior [7][8].

## 2. LITERATURE REVIEW

A person requires recommendations for movies, web series, books, songs, places, videos, websites, and a lot more. The recommendation system issue is an active platform for academics to work constantly on. The idea of memory-based filtering, content filtering, and model-based filtering techniques constitute the foundation of many recommendation systems.

Online systems' recommendations are based on user input, data collecting, ratings, similarities, preferences, and other elements. The data are gathered based on the input and used to display the recommendations. Without any prior selections or involvement, the recommendation is displayed for those people and is based on ratings and popularity.

A large part of the population today enjoys researching things online [7]. The author of [8] has shown how the number of individuals utilizing the internet is rising rapidly.

Filtering techniques that anticipate user preferences are required in the present day. These approaches, in which user data is examined and the best options are offered to the user, are applied in a variety of contexts, such as watching movies and videos, buying books, reading the news, and making purchases from online shops. The aforementioned goal is achieved using K-means and cooperative strategies [9].

Performance analyses of various recommendation systems have been conducted by Prateek et al. On smaller datasets, they discovered that employing user-based collaborative filtering lowers mean square error. The system Prince Praveen et al. suggested uses a content-based approach and would recommend to the user items with comparable content by taking into account related items. Additionally, it searches for related information using cosine similarity techniques [10]. The collaborative filtering strategy was employed by Furtado et al. This will depend on user reviews, which will limit one's ability to travel farther. This system is built on a teamwork model. Future research is aiming for individual choices, like emotions. The k-means algorithm was employed by Mahmudul Hasan et al. to assess the rating disparity across articles and produce a new cluster. The K-medoid clustering approach was also used by them [11]. The prediction produced by Prateek Sappadla et. al's algorithm was calculated using mean squared error metrics. Both collaborative filtering and content-based filtering are taken into consideration in this approach [12][13].

To provide viewers an indication of the performance from the global average, they have tried out the following fundamental guidelines method. Global average: It computes the average user rating across all films. The rating matrix's missing items are all predicted using this global average. This approach will assess the average rating for each user [14]. Since this method considers users' rating habits, it can be expected to perform slightly better than the average globally. With this method, the product's average rating is provided for each missing rating entry. They employed this technique, called adjusted average, to combine some user and product-related data while predicting the entry.

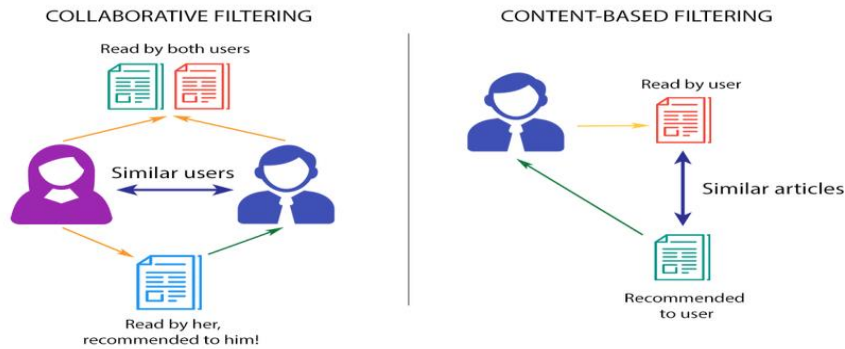
The key benefit of this technique is that it can propose movies even if they are hardly related. It is not necessary to use the other users' data. It suggests products outside the user-rated product category. Additionally, it suggests fresh items that haven't received user ratings. Scalability problems resulting from working with the large data collection caused memory limitations were the most significant difficulties they encountered when creating this system. The data is so large that it cannot be put in a dense matrix. Therefore, sparse matrix representation needs to be used in order for the program to run without running out of memory.

Additionally, evaluation and recording of intermediate outcomes, such as the user-user similarity matrix, is not possible [15][16].

As shown in Fig 1 CBF suggests similar articles to the user based on his past experience, whereas in CF if user A reads an article similar to user B then whichever article user A reads will be recommended to user B.

In the proposed model Fig 2, the dataset will generate 4 major values that are to be passed as parameters to the recommendation system. These values are Ratings Count, Average Ratings, User Ratings, and

Books Data. These are passed as parameters to the standard recommendation systems - Content-Based Recommendation, Popularity Rating, and Collaborative Recommendation. Once the recommendations have been generated here, they are passed to the Hybrid Recommendation system which performs its own calculations to generate a recommendation that is its own.

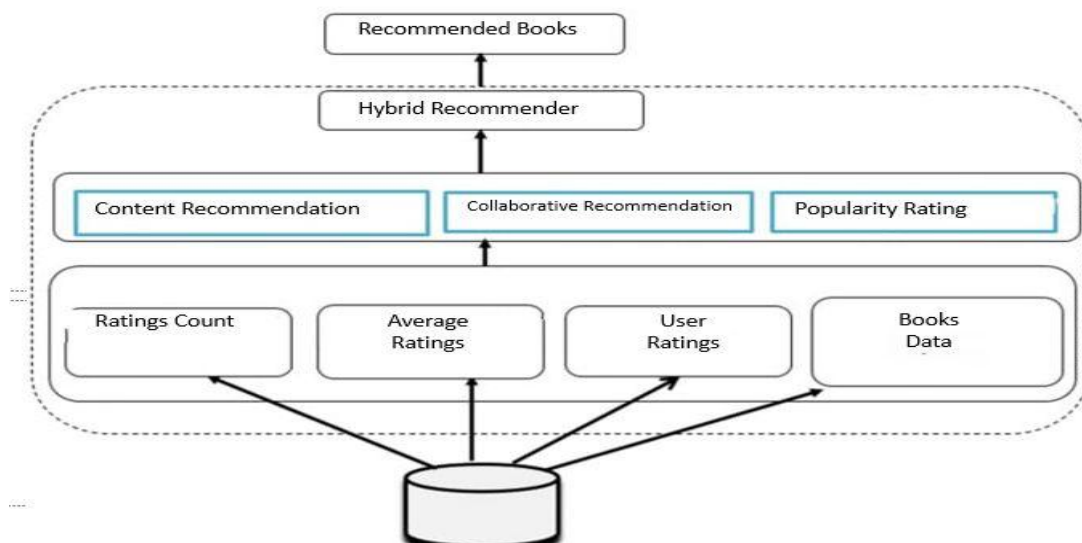


**Fig. 1** Content-Based Filtering VS Collaborative Filtering

### 3. COMPARISON OF CONTENT, COLLABORATIVE, AND HYBRID

In this table, a few papers are compared where Content-based and collaborative filtering is mainly considered for building recommendation models. The key tenet of the collaborative filtering strategy is that users will favor products that similar people value. The new definition of CF is “Having collaborations among the users to assist each other through filtering and recording their reviews about the viewed information”. Issues like “gray sheep” users (users who do not fit in any taste cluster), “cold-start” (new user or new item), etc. have an impact on CF Recommenders [17][18][19]. By considering the user profile and using the items and their attributes, content-based systems are recommended. Additionally, it is based on the product's description. Methods like cosine similarity, etc [20][21]. Hybrid combinations of collaborative filtering and content filtering are used with cosine similarity to obtain high accuracy, however, this method revealed a cold problem as a shortcoming [22].

Hybrid recommendation systems integrate two or more recommendation methods with the goal of achieving high-performance levels with fewer issues on an individual basis as shown in table:02 table:03 The limitations of a few methods and surveys have been specified and the changes that can be made for a better optimal solution is given.



**Fig. 2:** Architecture of Proposed Model

#### 4. RESULT AND ANALYSIS

##### • Popularity model

This model will sort according to an average of all the ratings given by the user for any given book. It recommends the top books of various genres to a new user. The recommendations in this model are user-independent and give the user a wide variety of genres initially. It also solves the cold start problem which is a common issue in recommendation systems.

**Table 1:** Content-based and Collaborative Filtering Recommendation System Papers

| Model                                                                          | Paper                                        | Dataset    | Method                                                                       | Result                                                                                                                               |
|--------------------------------------------------------------------------------|----------------------------------------------|------------|------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------|
| Content-Based Recommender system using the social network for cold-start users | Felipe G.Contrates, Solange N A Souza        | BestBuy    | Naive Bayes, Decision Tree and SVM, Supervised ML algorithm                  | The experiment was conducted for more generic e-commerce than those of movies or music, confirming the efficiency of the proposed RS |
| Collaborative Filtering Recommender Systems                                    | Ben Schafer, Shilad Sen, Dan Frankowski      | Movie Lens | User-Based nearest Neighbor Algorithm, Item-Based nearest Neighbor Algorithm | Authors have attempted to provide a snapshot of the current understanding of collaborative filtering systems and the methods.        |
| Amazon.com Recommendation Item-to-Item Collaborative Filtering                 | Greg Linden, Brent Smith, Jeremy York        | Amazon     | To determine the most similar match for a given item.                        | The algorithm builds a similar-item table by finding items that customers mostly purchase together.                                  |
| Recommender System Based on Collaborative Filtering for Spotify's User         | Javier Perez-Marcos, Vivian F. Lopez Batista | Spotify    | Pearson's correlation coefficient Phi coefficient                            | Phi coefficient is clearly faster than the person's coefficient.                                                                     |

##### • Collaborative Model

Collaborative filtering is a technique used to filter items on the basis of the likes of similar users. In this model, we consider user's reactions, and based on the reactions we classify similar types of users into clusters. We use Single Value Decomposition for the same as it is found to give better recommendations than K-nearest neighbors.

- **Content-based Model**

In this model, the information of the items that the user has engaged with will be considered for filtering. It solves the first rater. We apply Count Vectorizer and Cosine Similarity to generate results for the same.

**Table 2:** Hybrid Recommendation System Papers.

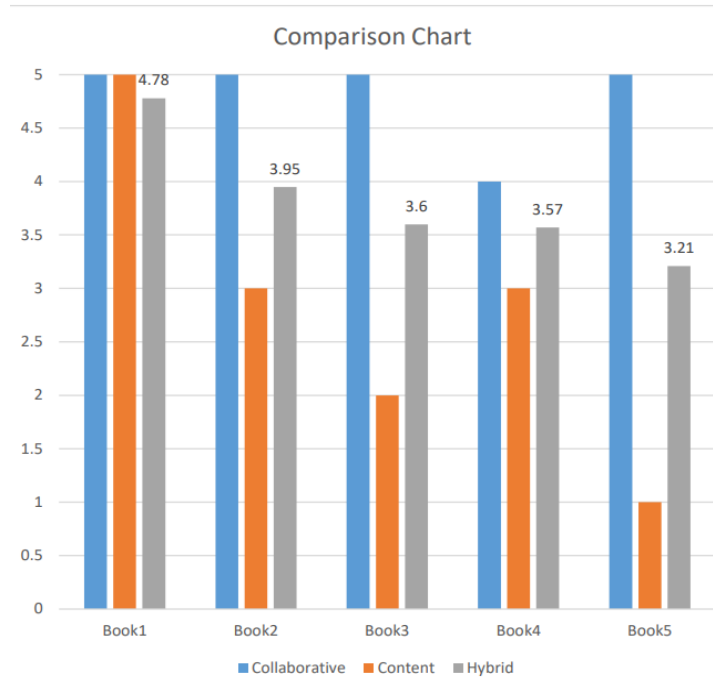
| Model                                                                             | Paper                                                                                       | Dataset                                  | Method                                                                                         | Result                                                                                                                        |
|-----------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------|------------------------------------------|------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------|
| Exploiting Non-content Preference Attributes through Hybrid Recommendation Method | Fernando Mourao, Wagner Meira Jr, Leonardo Rocha & Netflix, LastFm, ML-1M, ML-10M, Millions | Netflix, LastFm, ML-1M, ML-10M, Millions | Matrix Factorization, Latent Feature Log-Linear Model, Biased Matrix Factorization.            | Evaluate the attributes of popularity, similarity, and recency derived for each item from consumption data.                   |
| Cosine Similarity and KNN                                                         | Praveen Goud, Prince Praveen, Sagar Parmar(2020)                                            | Movie lens                               | To find and compare user content and make predictions using KNN                                | Hybrid engine which means combining content and collaborative content and collaborative filtering to show the recommendations |
| Implicative Rating-Based Hybrid Recommendations System                            | Lan Phuong Phan, Hiep Xuan Huyah                                                            | MSWeb, CourseRegistration                | Association Rule Based, User Based.                                                            | We will use the hybrid model based on implicative rating measures to develop a realistic recommendation system.               |
| Building Hybrid Recommendation System Based on Hadoop Framework                   | Mahindra Kumar Gourisaria                                                                   | Each Movie                               | Parallel, K-Mean Algorithm, Hadoop cluster we can know that with the increase of Hadoop nodes. | We can reduce the time consumption of parallel algorithms, and we can increase the computing ability of the system            |

- **Hybrid Model**

In the hybrid model we have combined the above models which give the users more accurate results and enhances user experience. This model increases efficiency and provides the user with better recommendations.

$$R_{\text{hybrid}} = (1-2\alpha) * R_{\text{popularity}} + \alpha * R_{\text{collaborative}} + \alpha * R_{\text{content}} \quad \text{Eq (1)}$$

We have analyzed the ratings given by two different users. Based on the ratings given by the user, 5 books were recommended to the users.



**Fig. 3** Collaborative, Content, Hybrid Rating Comparison for User1

In the Graphical representation of Fig 3, Fig 4 we have shown the results of collaborative, content, and hybrid models. The above formula Eq 1 is used for the hybrid model for assigning the weights for every specific model.  $\alpha$ = The Learning rate = 0.4.

**Table 3:** Limitations

| Model                                                                      | Paper                                  | Limitation                                                                                                                                                                                                                                                                                                                                       |
|----------------------------------------------------------------------------|----------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Content-Based Recommender system using social network for cold-start users | Felipe G.Contrates, Solange N A Souza  | The nature of personalization presents a number of significant obstacles for RS using social networks in the future, including the need for highly computational ways to handle sparse, diverse, and huge datasets in order to identify user preferences.                                                                                        |
| Collaborative Filtering Recommender Systems                                | Ben Schafer Shilad Sen, Dan Frankowski | In some cases, content-based personalization can be successful, but it might probably take decades or longer before the hardware and software technology can start to automatically comprehend the complexities of information that are significant to people.                                                                                   |
| Amazon.com Recommendations Item-to-Item Collaborative Filtering            | Greg Linden, Brent Smith, Jeremy York  | The retail industry should more broadly utilize recommendation algorithms for customized marketing, both offline and online. The technique will be appealing to offline retailers for use in postal mailings, discounts, and other types of client contact due to its higher conversion rates when compared to conventional broad-scale tactics. |

|                                                                                   |                                                                                |                                                                                                                                                                                                                                                                                                                      |
|-----------------------------------------------------------------------------------|--------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Recommender System Based on Collaborative Filtering for Spotify's Users           | Vivian F. Lopez Batista , Javier Perez-Marcos                                  | To make the recommendations more precise, the model could be enhanced. Despite the fact that the Phi coefficient reduces computation time, it would be interesting to investigate alternative approaches that enable better scaling in the relationship between the number of users and the time of recommendations. |
| The YouTube Video Recommendation System                                           | Palash Nandy, James Davidson, Benjamin Liebold, Taylor Van Vleet , Junning Liu | When averaged across the full period, we find that co-visitation-based recommendations perform at 207% of the Most Viewed page's baseline, whereas Top Favorited and Top Rated recommendations perform at levels comparable to or lower than the Most Viewed baseline.                                               |
| Exploiting Non-content Preference Attributes through Hybrid Recommendation Method | Fernando Mourao, Wagner Meira Jr, Leonardo Rocha                               | To better address the actual requirements of recommendation domains, simplistic implementation decisions, like a single global linear combination weight, should be modified.                                                                                                                                        |
| Cosine Similarity and KNN                                                         | Prince Praveen, Praveen Goud, Sagar Parmar(2020)                               | A powerful movie recommendation system can be created by combining advanced deep learning with additional filtering methods like collaborative filtering and hybrid filtering                                                                                                                                        |
| Implicative Rating-Based Hybrid Recommendations System                            | Lan Phuong Phan, Hiep Xuan Huyah                                               | The number of recommendations to be presented to improve the hybrid recommendation model when the number of given known ratings is less than 2.                                                                                                                                                                      |

## 5. CONCLUSION

Due to our fast-paced lifestyles, online buying has dramatically increased globally. Online purchasing is becoming more and more popular. For customers to shop online. Recommendation systems are provided by machine learning approaches to determine client interests. We can identify the things that customers buy frequently and propose them to other customers or users using recommendation techniques. When an e-commerce site offers a wide selection of products, it can be challenging for users to choose the right product from among the many different products. This article illustrates the various recommendation systems available and through the performance recorded, hybrid recommendation systems are found to be the most effective.

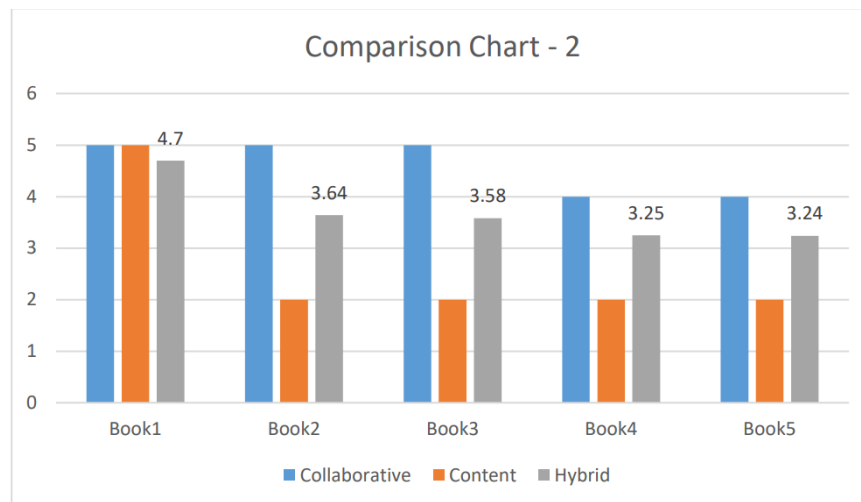


Fig. 4: Collaborative, Content, Hybrid rating Comparison for User2

## 6. FUTURE WORK

All e-commerce businesses will be able to use recommendation systems in the future. Recommendation algorithms are currently being successfully used by big businesses like Netflix and Amazon. Due to the amount of data and the high running costs, medium-sized and small businesses cannot employ the recommendation systems. By managing massive data more cheaply, our updated recommendation algorithms will increase the accessibility of recommendation systems. With the integration of social media in the future, the recommendations will be more accurate to the user based on the various factors that can be considered in social media.

## References

- [1] Coppel, Jonathan. "E-commerce: impacts and policy challenges." (2000).
- [2] Burke, Robin. "Hybrid recommender systems: Survey and experiments." *User modeling and user-adapted interaction* 12 (2002): 331-370.
- [3] Javier Pérez-Marcos, *Recommender System Based on Collaborative Filtering for Spotify's Users*, Conference Paper in *Advances in Intelligent Systems and Computing*, June 2018.
- [4] Filtering-Greg Linden, Brent Smith, and Jeremy York, *Amazon.com Recommendations Item-to-Item Collaborative Filtering*, IEEE Computer Society, 2003.
- [5] T. Keerthana ,T. Bhavani, N. Suma Priya, V. Sai Prathyusha, K.Santhi Sri-Flipkart Product Recommendation System, IV B.Tech, Department of Information Technology, Vignan's Nirula Institute of Technology & Science for Women, 2020
- [6] Yonghong Tian, Bing Zheng, Yanfang Wang", Yue Zhang". Qi Wu, College Library Personalized Recommendation System Based on Hybrid Recommendation Algorithm, 11th CIRP Conference, 2019
- [7] Cilingir I. Recommendation Systems. June 24, 2019. Retrieved from <https://medium.com/@irmcilingir/%C3%B6neri-sistemlerirecommendation-systems-28a3f341c0a9>
- [8] J.B. Schafer, J.A. Konstan, and J. Reidl, "E-Commerce Recommendation Applications," *Data Mining and Knowledge Discovery*, Kluwer Academic, 2001, pp. 115-153.
- [9] Solange N A Souza, *Content-based Recommender System using Social Media for Cold-start Users*, Conference Paper, 2017
- [10] Bateria, J.L: *Spotify-ed-music recommendation and discovery in spotify* (2014)
- [11] Dhruv, Ajay, et al. "Artist Recommendation System Using Hybrid Method: A Novel Approach," *Emerging Research in Computing, Information, Communication and Applications*, Springer, Singapore, 2019. 527-542

- [12] Thorat, Poonam B., Rajeshwari M. Goudar, and Sunita Barve. "Survey on collaborative filtering, content-based filtering and hybrid recommendation system." *International Journal of Computer Applications* 110, no. 4 (2015): 31-3
- [13] Rodriguez G. Introduction to recommender systems. 2018, Retrieved from <https://tryolabs.com/blog/introduction-to-recommender-systems/>
- [14] Dwivedi R. What are recommendation systems in machine learning? 16 April 2020. [Online]. Retrieved from <https://www.analyticssteps.com/blogs/what-are-recommendation-systems-machine-learning>
- [15] Rocca B. Introduction to Recommender Systems. June 3, 2019. Retrieved from <https://towardsdatascience.com/introduction-to-recommendersystems-6c66cf15ada>
- [16] N. Smitha, D. Anusha, C. Chaithanya, J. Sindhu, R. Tanuja and H. S. Hemanth Kumar, "A Review on Movie Recommendation System Using Machine Learning," 2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV), Tirunelveli, India, 2021, pp. 769-773, doi: 10.1109/ICICV50876.2021.9388619.
- [17] Le J. The 4 Recommendation Engines That Can Predict Your Movie Tastes. May 1, 2018. Retrieved from <https://towardsdatascience.com/the4-recommendation-engines-that-can-predict-your-movie-tastes109dc4e10c52>
- [18] Allen RB. User models: Theory, method, and practice. *International Journal of Man-Machine Studies* 1990; 32(5): 511-543.
- [19] Ozkok H. Recommendation Engine. April 17, 2020. Retrieved from <https://www.datascienceearth.com/recommendation-engine-tavsiyeoneri-sistemleri/>
- [20] Guler S. Recommendation Systems and the Usage of Recommendation Systems in E-commerce. MSc Thesis, Sakarya University, 2019.
- [21] Papagelis M, Plexousakis D, Kutsuras T. Alleviating the sparsity problem of collaborative filtering using trust inferences. In *iTrust'05 Proceeding of Third International Conference on Trust Management*, May 2005, pp. 224-239.
- [22] Bobadilla J, Ortega F, Hernando A, Bernal J. A collaborative filtering approach to mitigate the new user cold start problem. *Knowledge-based systems* 2012; 26: 225-238.