

**SCHOOL OF COMPUTER SCIENCE AND ENGINEERING**

WINTER SEMESTER 2023-24

Programme Int M Tech (MIS)

Course Title Data Mining Techniques

Course Code SWE2009

Slot C2

**Title: NOVEL NAVIGATOR: Using recommendation system**

**Team Members:**

A M Harshitha – 22MIS1051

S. Bindu Sri – 22MIS1061

K. Siva Nandini – 22MIS1074

**Faculty:** Dr Pattabiraman V.

**TABLE OF CONTENTS:**

|  |  |
| --- | --- |
| **S. No** | **Content** |
| 1 | Abstract |
| 2 | Scope |
| 3 | Objective |
| 4 | Introduction |
| 5 | Literature Review |
| 6 | Data Description |
| 7 | Architecture |
| 8 | Proposed Methodology |
| 9 | Novelty |
| 10 | Results and Discussion |
| 11 | Results Conclusion |
| 12 | References |

**1.Abstract:**

Novel navigator is a recommendation-based system. It recommends books for the user based on the previous search history (previous database). We use a hybrid recommender system to power our recommendations. Hybrid systems are the combination of two other types of recommender systems: content-based filtering and collaborative filtering. Content-based filtering is a method of recommending items similar to those the user has liked in the past. Collaborative filtering is a method by which user ratings are used in order to determine user or item similarities.

**Keywords:** Recommendation system, Recommender, Content-based filtering, Collaborative filtering, Hybrid

**2.Scope:**

The target of our project "Prediction of Fertilizer with Least Harmful Effects" is to recommend books to a user based on his past likes, ratings history search of the books. User Profiling: To generate user profiles, the system gathers and examines user data. This could contain details on the user's reading preferences, interest genres, favourite authors, and ratings or reviews. Content Analysis: The programme examines a book's text, taking into account metadata like the title, author, genre, and publishing date. Additionally, it might use natural language processing methods to glean details about the plot, topics, and writing style of the book. Collaborative Filtering: This method makes book recommendations based on the tastes of individuals who are similar to one another. The technology finds other users who share the same reading preferences and suggests books that those users have liked but the target user hasn't yet read. Content-Based Filtering: This method recommends books based on how closely the user's tastes and the books' content match. It evaluates book elements like author, genre, and storyline and suggests books that are comparable to ones the user has already loved. Hybrid Approaches: To produce more accurate and varied recommendations, a lot of recommendation systems blend content-based and collaborative filtering strategies. These hybrid strategies take advantage of the advantages of both techniques to get around obstacles and raise the standard of recommendations overall.

**3.Objective:**

Building a machine learning model that can suggest appropriate books to consumers based on their interests and preferences is the primary goal of a book recommendation system. With the use of user ratings, popularity, and past usage, the system seeks to offer customers personalised book recommendations.

With the growth of websites like YouTube, Amazon, and Netflix in recent years, recommender systems—including those for books—have taken on a greater significance. These systems present users with information items (books, films, music, etc.) that they are likely to find interesting using a variety of strategies, including collaborative filtering and content-based filtering.

A book recommendation system's goal can be summed up as follows: - To suggest appropriate books to users based on their tastes and interests. To offer individualised book recommendations by considering user ratings, popularity, and past readings. To improve user experience and engagement by making personalised book recommendations. A book recommendation system's main goal is to assist consumers in finding new books that they might find intriguing and pleasurable to read.

**4.Introduction:**

The main purpose of the Novel Navigator is to recommend/suggest the books in a easy way. This Novel Navigator propose the products to the users which are most relevant to that particular user. In our project we are using hybrid algorithm. A recommendation system is one type of information filtering system, which filters items by user’s interests. Recommender systems have become extremely common and are utilized in a variety of areas like in the applications of movies, articles, diet plan, etc…, Recommendation algorithms are of three types in our project namely: Content based filtering, Collaborative filtering, and the Hybrid recommendation. In the current era the recommendation system have developed for discovering the interesting item according to the user’s interest as discussed in every point of the project. The mainly used applications of this recommendation system are amazon, Netflix, hotstar, YouTube, aha, Facebook, movie lens, etc... also in the e-commerce and the online learning, etc…,

The amount of data, particularly on the Internet, has grown very rapidly today, the search for the necessary information has become more difficult. Recommendation systems aim to solve this type of problem. With the assistance of them one can quickly get to material information without looking through the web physically.

The main purpose of this project is to support those who have an interest in reading and to steer those people who are inculcating the habit of reading. From the massive range of books, it’s extremely tough to decide on a selected book. The

recommendation framework is therefore essential in helping users locate books that satisfy their needs and preferences. Reading is frequently cited as one of the best pastimes or routines. It has many benefits. Reading strengthens our brain, will increase our ability to empathize, builds our vocabulary and lots of more. Furthermore, reading is additionally good for our health each physically and mentally. These days with the ascent of the web and web-based media numerous individuals don’t give a lot of significance to reading. This isn’t simply because they are not intrigued yet additionally on the grounds that because of the tremendous determination of types like collections of memoirs and books, they regularly don’t pick the correct book for their mind-set. Book recommendation has numerous advantages when contrasted with public libraries. It gives recommendation of most appropriate books for the user. By giving the user what they need will help them save their time.

**5.Literature Review:**

## Background of terminologies discussed.

We provide an overview of five main research areas related to the topic: recommendation techniques, data used in RS, machine learning techniques in RS, recommendation system evaluation and architecture of Hybrid recommendation system.

Our research is based on the content-based, collaborative filtering and Hybrid recommendation system.

### Research based on content-based (CB) recommender systems.

Content based recommender system works with the data user has generated like giving ratings on a particular book, their search history, and their preferences from which recommendations are given to that user based on their similar choices. An extensive analysis of content-based recommendation systems is provided in this review study. Numerous topics are covered, including feature extraction methods, similarity measures, and user modelling. The article offers readers a thorough grasp of the subject by addressing various tactics and algorithms. It is an invaluable resource for both scholars and practitioners because it also covers evaluation methodologies for content-based recommendation systems (Pazzani and Billsus' "Content-Based Recommendation Systems"). This paper focuses on the state of content-based recommender systems today, offering insights into recent advancements and new directions. It investigates cutting-edge strategies including integrating collaborative filtering techniques and making suggestions based on context. The report facilitates informed decision-making and future research paths by keeping researchers and industry professionals informed about improvements in content-based recommendation technology by highlighting the state of the art(Melville, Mooney, and Nagarajan's "Content-Based Recommender Systems: State of the Art and Trends).This study investigates how user personality traits can affect the recommendation process, with a focus on improving content-based recommender systems with personality data. In order to enhance user pleasure and personalize recommendations, the authors want to use personality information. The study assesses the effectiveness of using personality data in content-based recommendation systems and highlights both its potential advantages and disadvantages through testing with a dataset of movies (Canamares, Carrasco, and del Jesus' "Improving Content-Based Recommender Systems with Personality Information). This research offers a thorough analysis of multiple approaches and strategies with a focus on content-based recommender systems in the music industry. It discusses the limitations and usefulness of methods like text analysis and audio feature extraction for music recommendation. The study is a useful resource for scholars and practitioners interested in this field since it provides insights into the distinctive qualities of music recommendation systems (Celma and Herrera's "An Overview of Content-Based Recommender Systems in Music Domain). This research paper proposes an audio- and metadata-driven content-based recommendation engine that tackles the difficulties of recommending obscure or specialized music. The method tries to enhance music discovery in the long tail of content by utilizing metadata and audio attributes. This study examines the challenges of making tailored music recommendations and suggests creative ways to improve customer pleasure and experience (McFee and Lanckriet's "Music Recommendation and Discovery in the Long Tail").

### Research based on Collaborative filtering (CF)

Collaborative filtering uses the previous behaviour of the user to recommend the books/items that are liked by the similar users.

This paper focuses on item-based collaborative filtering techniques for recommendation systems. It investigates methods for producing item recommendations based on user-item interactions, including neighbourhood selection, Pearson correlation, and cosine similarity.In item-based collaborative filtering, users are recommended items based on how similar they are to each other, as opposed to s2( "Item-Based Collaborative Filtering Recommendation Algorithms" by Sarwar, Karypis, Konstan, and Riedl).This paper presents matrix factorization techniques for collaborative filtering in recommender systems. It talks about factorising the user-item interaction matrix to find latent components using methods like Singular Value Decomposition (SVD) and Alternating Least Squares (ALS).Enhancing collaborative filtering systems' scalability and accuracy is the aim, especially when working with big datasets(Matrix Factorization Techniques for Recommender Systems" by Koren, Bell, and Volinsky).This paper presents a trust-based collaborative filtering method that integrates user trust relationships with item-based suggestions. It investigates the use of random walk algorithms to include trust information into the recommendation process.The system seeks to offer users recommendations that are more individualised and trustworthy by integrating trust relationships(Jamali and Ester's paper "TrustWalker: A Random Walk Model for Combining Trust-Based and Item-Based Recommendation" ). The integration of implicit feedback data into collaborative filtering recommender systems is the main objective of this research study.Rather than using explicit evaluations, implicit feedback refers to user activities that reveal preferences, including browsing patterns or previous purchases.The study looks into ways to use implicit feedback signals to produce recommendations that are more precise and pertinent(Enhanced Collaborative Filtering Recommender System using Implicit Feedback" by Hu, Koren, and Volinsky).

### Research based on hybrid recommendation.

An integrated suggestion in order to achieve exact performance while minimizing the shortcomings of traditional recommendation techniques, the framework combines content-based recommendation systems (CBRS) and collaborative filtering recommendation systems (CFRS). This survey study provides an overview of hybrid recommender systems and compares various hybrid approaches. It covers the integration of content-based filtering, collaborative filtering, and other techniques. Experimental findings contrasting different hybrid models are presented(Hybrid Recommender Systems: Survey and Experiments" by Burke). According to Oord, Dieleman, and Schrauwen, "Hybrid Music Recommendation Based on Convolutional Neural Networks": In this paper, a hybrid neural network-based collaborative filtering system for music recommendation is presented. It investigates the extraction of information from audio data for personalized music recommendations using deep learning and collaborative filtering.

The hybrid recommendation system model proposed in this study combines collaborative filtering and deep learning techniques. It looks at how deep neural networks can be used to better capture complex interactions between users and items and increase the accuracy of recommendations (Zhang, Yao, and Sun's paper "Hybrid Collaborative Filtering and Deep Learning Model for Recommendation Systems”).

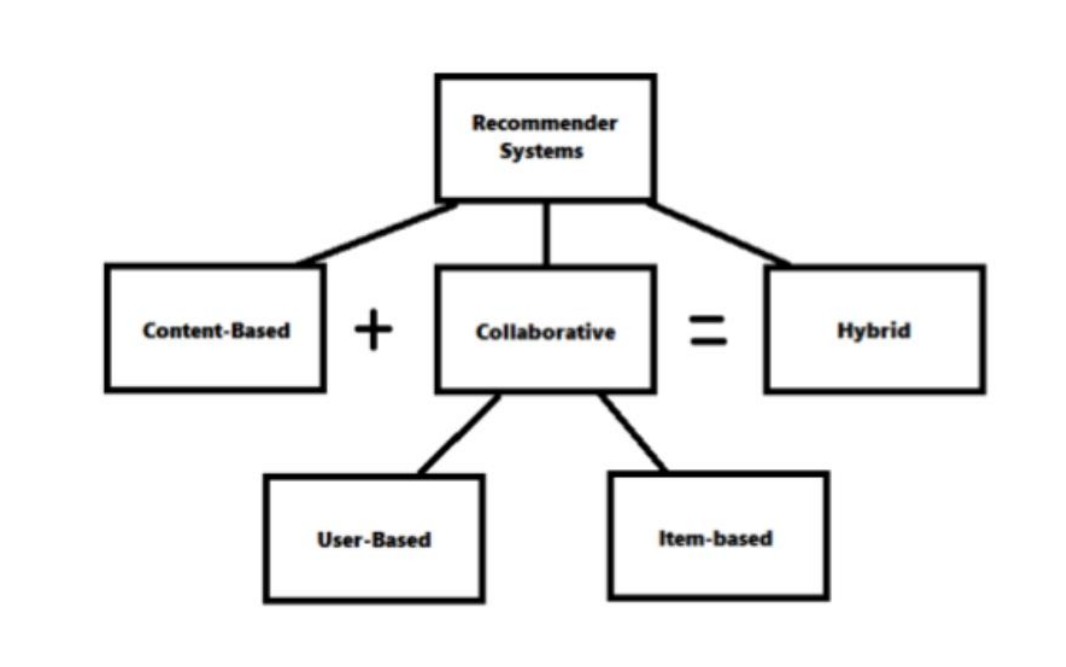
By using social network analysis, Ma, King, and Lyu improve collaborative filtering recommender systems. The main objective of this paper is to incorporate social network analysis into recommender systems for collaborative filtering. It looks at how user interactions and relationships within social networks can help with "cold start" issues and increase the accuracy of suggestions. We introduce a hybrid recommender system designed for online news portals in this paper. Personalized news suggestions are provided based on user preferences, item qualities, and user similarities through a combination of demographic, collaborative, and content-based filtering (Pitsilis, Litos, and Palpanas'"Hybrid Recommender Systems for Online News Portals").

**6.Data Description:**

Book metadata\*: Details about a book, including the title, author, genre, year of publication, synopsis, and image on the cover. This metadata facilitates comprehension of each book's qualities. User Interaction Data: Information about how users have interacted with the site, including what books they have reviewed, read, bought, or put on their wish list. This information sheds light on user preferences and actions. User Profile Data: Extra details on users, including location, age, and gender; reading preferences; reading habits; and previous reading experiences. The creation of customised recommendations is aided by this data. User Reviews and Ratings\*: Clearly expressed opinions expressed by users on books they have read in the form of reviews and ratings. Ratings can be binary (liked/disliked) or numerical (on a scale of 1 to 5 stars, for example). Textual information found in reviews is frequently analysed for emotion and content. Social Network Data: Details about the social ties of users, including groups with shared interests, friends, and followers. Social recommendations, in which users get book recommendations based on what their friends or peers have read or enjoyed, can be incorporated using social network data. Contextual Data: Details on the device being used to access the platform, the season, the day of the week, and the time of day. This information aids in providing prompt and pertinent advice. based on user’s current context. External Data Sources: Information gathered from outside sources, including bestseller lists, literary honours, critics' reviews, and information from other websites devoted to books. The diversity of recommendations and the recommendation process itself can be enhanced by incorporating external data sources.

Book recommendation systems can produce customised recommendations that correspond to each user's interests and preferences by utilising these kinds of data and recommendation algorithms, such as collaborative filtering, content-based filtering, or hybrid techniques.

**7.Architecture:**



**8.Proposed Methodology:**

The proposed methodology for this recommendation system is leveraging Hybrid recommendation and clustering.

**HYBRID RECOMMENDATION:** We use a hybrid recommender system to power our recommendations. Hybrid systems are the combination of two other types of recommender systems: content-based filtering and collaborative filtering. Hybrid approaches could be more effective in some cases. These methods can also be used to overcome some of the common problems in recommender systems such as cold start and the sparsity problem. The sparsity problem occurs when available data is insufficient for identifying similar users.

**CONTENT-BASED FILTERING ALGORITHM:** The content-based approach uses additional information about users and/or items. This filtering method uses item features to recommend other items similar to what the user likes and also based on their previous actions or feedback that are given. The main idea of the content-based filtering is to build a model based on the features that are available.

**COLLABORATIVE FILTERING ALGORITHM:** Collaborative filtering method for recommender systems is a method that is solely based on the past interactions that have been recorded between users and items, in order to produce new recommendations. Collaborative Filtering tends to find what similar users would like and the recommendations to be provided and in order to classify the users into clusters of similar types and recommend each user according to the preference of its cluster.

**9.Novelty:**

Many recommendation techniques, including collaborative filtering, knowledge-based filtering, content-based filtering, and perhaps graph-based models, are combined in a hybrid recommendation system.

Here, we use Hybrid recommendation system, a combination of collaborative and content-based filtering for recommending books for a particular user. By considering each user's unique preferences, behavior, and context, this hybrid method enables improved customization. Collaborative Filtering: To generate recommendations, collaborative filtering examines the interests and actions of several individuals. It finds users who have your likes and suggests books that these users have either highly rated or enjoyed. It is possible to base collaborative filtering on either item-based or user-based methodologies. Content-Based Filtering: This method suggests books according to their inherent qualities, like genre, author, synopsis, and keywords. By taking into account the books the user has enjoyed or given high ratings, it creates a user profile and suggests books that are comparable to the user's tastes. Instead of taking user behaviour or preferences into account, content-based filtering concentrates on the characteristics of the items themselves. Hybrid Approach: To enhance recommendation quality, a hybrid recommendation system combines content-based and collaborative filtering techniques. By identifying users with similar tastes through collaborative filtering, the system can use their preferences to provide recommendations. For a more customised experience, it can also include content-based filtering, which takes into account the particular qualities of books that match the user's interests. Certain drawbacks of separate approaches can be overcome by the hybrid system by merging both content-based and collaborative filtering. When there is little or no user data available, collaborative filtering can help with the "cold start" issue and offer suggestions based on group intelligence. Content-based filtering can accommodate certain user preferences and provide a wider range of recommendations.

By utilising the complementing qualities of both strategies, the hybrid recommendation system enhances suggestion coverage, accuracy, and personalisation. It offers highly customised book recommendations by considering the individual tastes, actions, and circumstances of each user.

**10. Result and discussion:**

**Data collection:** Dataset has a major role in machine learning model when given as input and Based on the information in the dataset, the output is anticipated. Datasets used are collected from Kaggle which are in “.csv” format i.e. “book\_tags.csv”, “books.csv”, “ratings.csv”, “to\_read.csv”, “tags.csv”.

**Data Cleaning:** Upon examining the data, I discovered that there are many ratings for specific user and book combinations. So I removed users who have rated fewer than 3 books.

**Content based recommender**: Content-based filtering leverages item information, such as prior activities or explicit feedback, to suggest other things that are like the user's favourites.

**Cosine Similarity:** It is the dot product of 2 data points used to measure the similarity between two books.

A mathematical equation with black text

Description automatically generated

**Weighted average rating:** A computation known as a weighted average account for the different weights assigned to each integer in a data set. Each value in the data set is multiplied by a predefined weight before the final computation is conducted to produce the weighted average. A simple average, where each value in a data collection is given the same weight, may not be as accurate as a weighted average.

A math equation with numbers and symbols

Description automatically generated

where, v is the number of ratings for the book.

m is the minimum ratings required to be listed in the chart

R is the average rating of the book

C is the mean rating across the whole report

**Collaborative based recommender:** Collaborative Filtering is based on the idea that users similar to other user that can be used to predict how much the user will like a particular product or service those users have used/experienced.

**User-Based:** which calculates the similarity between users and other users.

**Item-based:** which calculates the similarity between the items that users’ rate or interact with and other items.

**Hybrid Recommender:** User will try to build a simple hybrid recommender that brings together techniques we have implemented in the content based and collaborative filter-based engines. This is how it will work:

**Input:** User ID and the Title of a Book

**Output:** Similar books sorted based on expected ratings by that user.

**11. Result and Conclusion:**

**Cosine similarity**

A screenshot of a computer program

Description automatically generated

**Weighted Average Rating**

A screenshot of a computer

Description automatically generated

**User-based**

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

**Item-based**

A white background with many squares

Description automatically generated with medium confidence

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

**Hybrid Recommendation:**

A screenshot of a computer

Description automatically generated

**Comparison graphs:**

A chart of comparison of methods

Description automatically generated

A graph of different colored squares

Description automatically generated

**CONCLUSION:** To sum up, hybrid algorithms in recommender systems present a viable way to overcome the drawbacks and improve the effectiveness of conventional recommendation techniques. Hybrid algorithms combine the best features of content-based and collaborative filtering to give users recommendations that are more varied, precise, and context-aware. Common issues in recommender systems, like the absence of serendipity, data sparsity, accuracy-diversity trade-off, and cold start problem, are addressed by the hybrid approach. Hybrid algorithms have the ability to produce recommendations that are more relevant and personalised by utilising contextual data, item content information, and user-item interactions. Hybrid algorithms also enable recommendation systems to be more adaptable and customised. In order to adjust to varied user preferences, item attributes, and application domains, they make it possible to incorporate alternate weighting methods, switching mechanisms, or feature combination strategies. Although hybrid algorithms seem to have a lot going for them, research is still being done to increase how successful they are and solve new problems like scalability, privacy, and fairness. Future innovations might include improvements in explainable AI, deep learning techniques, managing unbalanced data, and combining privacy-aware and fairness-aware methods. All things considered, hybrid algorithms offer a strong foundation for developing recommender systems that can get beyond the drawbacks of distinct recommendation techniques. Hybrid algorithms have the ability to provide more precise, varied, and context-aware recommendations by fusing collaborative filtering and content-based filtering. This could improve user experience and encourage greater interaction across a range of fields.

**12. References:**

1. Manisha Chandak, Sheetal Girase, Debajyoti Mukhopadhyay, Introducing Hybrid Technique for Optimization of Book Recommender System, Procedia Computer Science, Volume 45, 2015, Pages 23-31, ISSN 1877-0509, <https://doi.org/10.1016/j.procs.2015.03.075>.
2. Achraf Gazdar, Lotfi Hidri, A new similarity measure for collaborative filtering-based recommender systems, Knowledge-Based Systems, Volume 188,2020,105058, ISSN0950-7051,

https://doi.org/10.1016/j.knosys.2019.105058.

1. Kukkar, Gurpreet & Khandare, Rohan & Ali, Syed. (2023). BOOK RECOMMENDATION SYSTEM USING COLLABORATIVE FILTERING. IJARCCE. 12. 10.17148/IJARCCE.2023.12735.
2. Okon, Emmanuel & Eke, Bartholomew & Asagba, Prince. (2018). An Improved Online Book Recommender System using Collaborative Filtering Algorithm. 10.13140/RG.2.2.24240.46086.
3. Rajalakshmi, S.., Indumathi, G.., Elias, A.., Priya, G. S.., & Muthulakshmi. R, V. (2024). Personalized Online Book Recommendation System Using Hybrid Machine Learning Techniques. International Journal of Intelligent Systems and Applications in Engineering, 12(15s), 39–46. Retrieved from <https://ijisae.org/index.php/IJISAE/article/view/4712>.
4. Wayesa, F., Leranso, M., Asefa, G. *et al.* Pattern-based hybrid book recommendation system using semantic relationships. *Sci Rep* **13**, 3693 (2023). https://doi.org/10.1038/s41598-023-30987-0
5. Adomavicius, G., & Tuzhilin, A. (2005). Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. IEEE Transactions on Knowledge and Data Engineering, 17(6), 734–749. <https://doi.org/10.1109/TKDE.2005.99>.
6. Aggarwal, C. C. (2016a). An introduction to recommender systems. In C. C. Aggarwal (Ed.), Recommender systems (1st ed., pp. 1–28). Cham: Springer Retrieved from [https://link.springer.com/book/10.1007 %2F978-3-319-29659-3](https://link.springer.com/book/10.1007/978-3-319-29659-3).
7. Aggarwal, C. C. (2016b). Evaluation of recommender systems. In C. C. Aggarwal (Ed.), Recommender systems (1st ed.). Cham: Springer Retrieved from [https://link.springer.com/book/10.1007%2F978-3-31929659-3.](https://link.springer.com/book/10.1007/978-3-319-29659-3)
8. Amatriain, X., Jaimes, A., Oliver, N., & Pujol, J. (2010). Data mining methods for recommender systems. In F. Ricci, L. Rokach, B. Shapira, & P. Kantor (Eds.), Recommender systems handbook (pp. 39–71). Boston, MA: Springer. <https://doi.org/10.1007/978-0-387-85820-3_2>.
9. Batmaz, Z., Yurekli, A., Bilge, A., & Kaleli, C. (2018). A review on deep learning for recommender systems: Challenges and remedies. Artificial Intelligence Review, 52, 1–37. [https://doi.org/10.1007/s10462-0189654-y.](https://doi.org/10.1007/s10462-018-9654-y)
10. Chen, H., Li, Z., & Hu, W. (2015). An improved collaborative recommendation algorithm based on optimized user similarity. The Journal of Supercomputing, 72(7), 2565–2578. [https://doi.org/10.1007/s11227-0151518-5.](https://doi.org/10.1007/s11227-015-1518-5)