

## A novel approach for e-health recommender systems

Adeeb M. Alsaaidah<sup>1</sup>, Qusai Y. Shambour<sup>2</sup>, Mosleh M. Abualhaj<sup>1</sup>, Ahmad Adel Abu-Shareha<sup>3</sup>

<sup>1</sup>Department of Networks and Cybersecurity, Faculty of Information Technology, Al-Ahliyya Amman University, Amman, Jordan

<sup>2</sup>Department of Software Engineering, Faculty of Information Technology, Al-Ahliyya Amman University, Amman, Jordan

<sup>3</sup>Department of Data Science and Artificial Intelligence, Faculty of Information Technology, Al-Ahliyya Amman University, Amman, Jordan

### Article Info

#### Article history:

Received Oct 20, 2023

Revised Feb 13, 2024

Accepted Feb 22, 2024

#### Keywords:

Collaborative filtering

Content filtering

Data sparsity

E-health

Healthcare

Multi-criteria

Recommender systems

### ABSTRACT

The increasing use of the internet for health information brings challenges due to the complexity and abundance of data, leading to information overload. This highlights the necessity of implementing recommender systems (RSs) within the healthcare domain, with the aim of facilitating more effective and precise healthcare-related decisions for both healthcare providers and users. Health recommendation systems can suggest suitable healthcare items or services based on users' health conditions and needs, including medications, diagnoses, hospitals, doctors, and healthcare services. Despite their potential benefits, RSs encounter significant limitations, including data sparsity, which can lead to recommendations that are unreliable and misleading. Considering the increasing significance of health recommendation systems and the challenge of sparse data, we propose an effective approach to improve precision and coverage in recommending healthcare items or services. This aims to assist users and healthcare practitioners in making informed decisions tailored to their unique needs and health conditions. Empirical testing on two healthcare rating datasets, including sparse datasets, illustrate that our proposed approach outperforms baseline recommendation methods. It excels in improving both the precision and coverage of health-related recommendations, demonstrating effective handling of extremely sparse datasets.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



### Corresponding Author:

Qusai Y. Shambour

Department of Software Engineering, Faculty of Information Technology, Al-Ahliyya Amman University

Amman, Jordan

Email: q.shambour@ammanu.edu.jo

## 1. INTRODUCTION

Given the issues of information overload and the lack of personalized health services, there is a growing need to implement recommender systems (RSs) that can address these challenges and support both patients and medical professionals in making informed healthcare decisions. RSs aim to suggest items that are highly suitable for users based on their profiles or historical preferences. In the context of health, these items can include medications, medical treatments, doctors, exercise routines, diagnoses, diets, or other healthcare services. A health recommender system (HRS) is a type of recommender system that provides users with personalized health information tailored to their specific health conditions and needs. The goal of HRS is to enable users to engage in important health decision-making with minimal time and effort invested, ultimately facilitating a transition towards a more personalized paradigm in the healthcare industry [1]-[3].

Collaborative filtering (CF) is the foundation of the most popular recommendation systems. CF involves assessing items based on the opinions and preferences of other users. Despite their efficacy, CF approaches face the data sparsity problem, which occurs when there is insufficient data to make precise

recommendations. It describes a situation in which the number of user-item interactions or ratings is relatively low in comparison to the total number of users and items in the system [4]-[6].

Recent studies have investigated the integration of additional information to improve the accuracy of recommendations in CF-based RSs by addressing the issues of data sparsity. In this regard, content-based recommendation approaches have emerged as a promising solution. Content-based approaches consider the attributes and characteristics of items, such as genre, keywords, or metadata, to generate recommendations. By relying on item attributes, content-based filtering can offer a more effective approach to recommend items with limited ratings, as it can identify relevant items even when rating data is scarce. Therefore, it is essential to consider integrating item-based CF with content-based recommendation approaches to leverage their respective strengths, improve overall system performance, and provide a solution for the aforementioned challenges [7], [8].

Furthermore, most CF-based recommendation methods rely on a single rating criteria, which reflects the item's overall rating, to infer user preferences. However, these techniques overlook the detailed users' preferences for specific features or aspects of an item, even though various websites permit users to rate items based on more than one criteria [9]-[11]. For instance, in the healthcare domain, patients on websites like WebMD.com can rate medications according to a number of criteria such as ease of use, effectiveness, and satisfaction. Similarly, patients on RateMDs.com can rate doctors on criteria like, punctuality, staff, knowledge, and helpfulness. Therefore, there is a need to develop multi-criteria (MC) recommendation approaches that utilize the additional rating data to gain a more accurate understanding of user preferences and provide precise and effective recommendations [12].

Over the past decade, RSs have experienced a surge in popularity and have demonstrated their efficacy in numerous domains, such as e-commerce, e-learning, and e-tourism [13]. At present, health RSs are employed in various health domains, including but not limited to doctor recommendations, drug recommendations, physical activity promotion, diet recommendations, and other related healthcare applications [1]-[3]. Such systems possess the capacity to enhance the well-being of users and aid healthcare practitioners in rendering more accurate patient-centric decisions.

There has been a notable growth in the quantity of medical information that is accessible to the general public in recent years. Accordingly, patients may encounter certain challenges when attempting to identify appropriate doctors who possess the best expertise to effectively address their health concerns. Narducci *et al.* [14] presented a recommendation system focused on semantics that suggests hospitals and doctors to patients based on their profiles. The system initially calculates semantic similarities among patients, then creates a prioritized list of hospitals and doctors that best match the profile of the patient. The system is integrated within the HealthNet social network, with the primary objective of facilitating knowledge sharing, identifying comparable patients, and reviewing their experiences. Han *et al.* [15] have developed a hybrid recommender system that offers reliable recommendations of doctors. The system models patient trust in doctors using a comprehensive dataset of consultation histories. Furthermore, it calculates similarities between patients and doctors by utilizing their metadata. The objective of the proposed system is to improve the primary care health service of a prominent healthcare provider in Portugal. The primary goal is to mitigate the search burden experienced by patients. Mondal *et al.* [16] present a doctor recommendation system that enables patients to choose a doctor from a selection of available healthcare providers. The system incorporates three key criteria, namely, the patient's presenting symptoms, the geographical proximity of the doctor, and the patient's degree of trust in the doctor's competence. The system recommends a doctor for a patient by utilizing these three specific criteria. The system incorporates a multilayer graph to represent the patient-doctor relationship, which serves the dual purpose of facilitating efficient data retrieval and modeling the trust factor based on the underlying database.

Medication errors represent a significant category of medical errors that pose a grave risk to patients' well-being. The proliferation of drug information has resulted in medication errors, posing challenges in identifying pertinent drugs. Drug RSs have been devised in this particular context to aid users and healthcare professionals in the identification of the most appropriate medications for a particular disease. Based on implicit feedback and cross-recommendation methods, Chen *et al.* [17] proposed a drug recommendation system for epilepsy treatments. The proposed system aims to examine the medical histories of epileptics to find possible relations between the syndromes and the drugs. When compared to a baseline system using artificial neural network (ANN), the proposed system outperforms ANN in terms of recall rate by up to 30%. Overall, the proposed system outperforms ANN in terms of performance. To aid patients in making informed drug selections, Hossain *et al.* [18] developed a drug recommender system that employs sentiment analysis technologies on drug reviews. First, a sentimental measurement approach is used to generate drug ratings based on drug reviews. Second, the extent to which drug reviews are valuable to users, the patient's conditions, and the sentiment polarity of the dictionary are considered. Then, all of that information was combined in the recommendation system to generate a shortlist of drugs that the patient would do best on. Zhang *et al.* [19] investigated the medicine recommendation problem, which involves predicting a set of

medications based on a set of symptoms. The researchers developed MedRec, which incorporates a medical knowledge graph and a medicine attribute graph, to address the issue of sparse data. The medical knowledge graph mines the relationships between medicine, symptoms, diseases, and physical examinations, while the medicine attribute graph mines the relationships between medications.

In summary, health recommendation systems primarily rely on knowledge-based approaches, employing knowledge graphs and ontologies for information storage [14], [16], [19]. Nevertheless, these systems are constrained by factors such as the requirement for extensive domain knowledge, human labor, proficiency in knowledge representation, and continuous updates and the maintenance of knowledge bases [20]. Furthermore, model-based CF methods have been suggested as possible remedies for health recommendations [15], [17], [18]. However, it is important to note that these approaches requires the collection of extensive features, reduction of dimensionality, and the availability of extensive data, which could potentially have a negative impact on recommendation effectiveness [21]. Conversely, there has been limited focus on implementing memory-based CF approaches in health recommendation systems. Although memory-based CF methods have limitations like sparsity, they offer benefits such as simplicity, justification, and stability, even with the addition of new items and users. A notable concern lies in the lack of research on MC RSs within the realm of health recommendations.

Accordingly, in order to overcome the aforementioned challenges and assist patients in making well-informed healthcare decisions tailored to their specific needs and health conditions, this study proposes an intelligent recommendation approach. This approach can be utilized in e-health RSs, such as assisting patients in locating doctors who fulfill their specific treatment requirements and aiding patients in efficiently identifying the most appropriate medication for their medical conditions. To boost the effectiveness of recommendations and address the issue of sparse data, the proposed approach combines content-based filtering and MC-based CF. Experimental results using real-world healthcare MC rating datasets demonstrate the effectiveness of the proposed approach in terms of predictive accuracy and coverage compared to other baseline recommendation approaches. The succeeding sections of this study are formulated as follows: section 2 delivers an overview of previous research on various health RSs, section 3 outlines the structure and components of the proposed approach, section 4 delves into the detailed presentation of experimental evaluation. Finally, the study is concluded in section 4, which also includes recommendations for further research.

## 2. DESIGN OF THE PROPOSED APPROACH

The identification of relevant items for users based on their needs and preferences can be approached as a recommendation problem, as discussed by Adomavicius and Tuzhilin [22]. In this context, let  $U$  represent the set of all users (patients), and  $I$  represent the set of all possible items, such as doctors or drugs that can be recommended. To measure the suitability of an item  $i \in I$  to a user  $u \in U$ , we can define a utility function  $T: U \times I \rightarrow T$ , where  $T$  represents a nonnegative real number falling within a specific range. Then, the purpose of the proposed recommendation approach is to identify an item  $i$  for the user  $u$  to maximize the utility function  $T_i^u$ , which can be formulated as (1):

$$\forall u \in U, i = \arg \max_{i \in I} T_i^u \quad (1)$$

Next, we outline the primary steps of the proposed approach in this study, which consists of five major steps. From this point forward, we refer to patients as users and doctors or drugs as items. Initially, we assume that each user provides ratings for an initial set of items to express their preferences. Subsequently, the proposed approach is employed to predict additional items that a user might find appealing based on their expressed preferences and other items' features.

### 2.1. Compute item-based content similarity

Relying solely on ratings to calculate the similarity between items may have limitations due to a lack of an adequate number of ratings. This scarcity of data, known as data sparseness, poses a challenge in providing effective item recommendations solely based on user ratings. This is where content-based recommendations offer a solution. Content-based algorithms consider the attributes and characteristics of items, such as genre, keywords, or metadata, to generate recommendations. By relying on item attributes, content-based filtering can offer a more effective approach to recommending items with limited ratings, as it can identify relevant items even when rating data is scarce.

In this stage, it is assumed that all items have been categorized with predefined meta-information. The categorization allows for the identification of relationships or similarities between items. When two

items fall into the same category, they are considered related or similar to each other. Leveraging this categorical information, the item-based content similarity approach utilizes the categorical representations of a set of items that have been rated by an active user. This information is then used to find related items from a new set of items that the active user has not explored yet. These related items, based on their category similarity, have the potential to be recommended to the active user. By considering the categorical relationships between items, the system can recommend items that are related to the user's personal preferences and exhibit similarity to items they have already shown interest in.

In this context, the assumption is made that, for example, are classified according to specific specialty categories. Doctors who share the same specialty are considered similar according to the categorization of their respective areas of specialization. The item-based content similarity approach is utilized to compare the category representations of doctors who have received ratings from an active patient. This comparison is used to recommend new doctors that the patient has not visited before. For instance, referring to Table 1, we have a list of doctors with their corresponding specialty categories.

Table 1. Types of doctor's specialization

ID	Specialty category
D1	Psychiatrist
D2	Orthopedist and Rheumatologist
D3	Psychiatrist
D4	Cardiologist
D5	Rheumatologist
D6	Psychiatrist
D7	Nephrologist
D8	Orthopedist

In this case, each doctor is assigned to a specific specialty category. For example, doctor D1 is categorized as a Psychiatrist, while doctor D2 has multiple specialties, namely Orthopedist and Rheumatologist. By considering these specialty categories, we can identify similarities between doctors. Doctors belonging to the same specialty category are generally considered to possess similar attributes and characteristics due to their shared expertise in that specialty.

To compute the content-based similarity of items through their associated categories, the first step involves representing each item as a binary number vector. This vector indicates the presence or absence of a particular category for a given item, as illustrated in Table 2. As an illustration, considering Table 1, the binary vector representation of doctors according to their specialization would appear as follows.

Table 2. Binary vector representation of doctors based on their specialization

ID	Specialty category	Binary vector representation
D1	Psychiatrist	[1, 0, 0, 0, 0]
D2	Orthopedist and Rheumatologist	[0, 1, 1, 0, 0]
D3	Psychiatrist	[1, 0, 0, 0, 0]
D4	Cardiologist	[0, 0, 0, 1, 0]
D5	Rheumatologist	[0, 0, 1, 0, 0]
D6	Psychiatrist	[1, 0, 0, 0, 0]
D7	Nephrologist	[0, 0, 0, 0, 1]
D8	Orthopedist	[0, 1, 0, 0, 0]

In the representation shown in Table 2, each binary vector has a length equal to the total number of specialty categories. A value of 1 indicates that the corresponding category is associated with the item, while a value of 0 indicates the absence of that category. By representing items in this binary vector form, we can perform calculations to measure the item-based content-based similarity between items based on their category representation.

For this purpose, the Sorgenfrei metric [23] is used to measure the item-based content similarity between items based on their binary vector representations. The Sorgenfrei metric calculates the similarity between two binary vectors by taking into account the number of elements where the vectors match (both have a value of 1) and where they are different (one vector has a value of 1 while the other vector has a value of 0). The formula for the Sorgenfrei metric is as (2):

$$Sim_{i,j}^{Content} = \frac{A^2}{(A+B)*(A+C)} \quad (2)$$

Where  $A$  is the number of matching elements where both elements of vectors' (items)  $i$  and  $j$  have a value of 1;  $B$  is the number of non-matching elements where the element of vector  $i$  has a value of 0 and element of vector  $j$  has a value of 1; and  $C$  is the number of non-matching elements where the element of vector  $i$  has a value of 1 and element of vector  $j$  has a value of 0. The Sorgenfrei metric is used to calculate a similarity score between two binary vectors, ranging from 0 to 1. A score of 0 indicates no similarity, while a score of 1 indicates complete similarity.

## 2.2. Compute item-based MC CF similarity

In this particular step, we begin by examining a set of items that have been rated by an active user, aiming to identify items that reveal the maximum similarity to the user's preferences based on their MC ratings. To measure the partial implicit similarity between items  $i$  and  $j$  with respect to each of the rating criteria  $c$ , the mean square differences (MSD) method is employed, which evaluates the predictive accuracy of item  $i$  as a sole recommender for item  $j$  with respect to each of the rating criteria  $c$ . To calculate the predicted rating of an item with respect to the rating criteria  $c$  utilizing only a single neighbor item, we employ Resnick's prediction method [24] as (3):

$$P_{u,i}^c = \overline{r_i^c} + (r_{u,i}^c - \overline{r_j^c}) \quad (3)$$

Where  $r_{u,i}^c$  represents the rating given by user  $u$  to item  $i$  with respect to criteria  $c$ ;  $\overline{r_i^c}$  and  $\overline{r_j^c}$  represent the average ratings of items  $i$  and  $j$  on criteria  $c$ , respectively. Subsequently, the MSD method is employed to measure the partial implicit similarities between items  $i$  and  $j$  with respect to each of the rating criteria  $c$  by taking the average of the squared differences between the predicted and actual rating scores of shared users who have rated both items  $|U_i \cap U_j|$ . To ensure that the  $MSD_{i,j}^c$  values fall within the range of [0,1], both the predicted rating  $P_{u,i}^c$  and the actual rating  $r_{u,i}^c$  are normalized using the max-min normalization approach.

$$MSD_{i,j}^c = 1 - \left( \frac{\sum_{u=1}^{|U_i \cap U_j|} (P_{u,i}^c - r_{u,i}^c)^2}{|U_i \cap U_j|} \right) \quad (4)$$

Next, the worst-case similarity [25] is utilized as an aggregation technique for the partial similarities. This allows us to determine the overall similarity value between a given target item  $i$  and its neighboring item  $j$  as (5):

$$Sim_{i,j}^{MSD} = \min_{c=1,\dots,k} MSD_{i,j}^c \quad (5)$$

Where  $MSD_{i,j}^c$  is the value of the partial similarity between items  $i$  and  $j$  based on criterion  $c$ ,  $k$  represents the number of individual criteria considered in the aggregation process.

One approach to mitigate the limitations of the MSD measure and improve the accuracy of predictions is to consider asymmetric weights. This involves considering the compromise factor [26], represented by (6), which takes into account the percentage of common users when calculating the similarity among items. The compromise factor is based on the proportion of co-rated users among items, normalized by the number of users who have rated a target item.

The MSD measure often does not consider biases that can influence the relationships between items. For instance, certain items may tend to be rated by more users compared to others. These biases can affect the accuracy of the MSD measure. By incorporating the compromise factor, the influence of such biases can be taken into account, leading to more accurate similarity calculations. By combining the MSD metric with the compromise factor, a more comprehensive similarity measurement is achieved, providing a better representation of the underlying relationships between items. This integration, as shown in (7), aims to mitigate the issue of misrepresentation of similarity among items and ultimately improve the accuracy of predictions.

$$CFactor_{i,j} = 1 - \exp\left(-\frac{|U_i \cap U_j|}{|U_i|}\right) \quad (6)$$

$$Sim_{i,j}^{CF} = Sim_{i,j}^{MSD} \times CFactor_{i,j} \quad (7)$$

### 2.3. Compute global item similarity

In order to enhance the proposed approach's predictive capability for unobserved items, a global similarity score for each item has been introduced. This score plays a crucial role in overcoming the challenges posed by sparsity, resulting from a restricted amount of nearest neighbors (NN). The global similarity score of the target item is influenced by the count of similar items present in the item-item MC CF similarity matrix that was created in the preceding step, as well as the average difference in ratings between the item and the user's average rating, as (8):

$$GIS_i = \exp\left(-\frac{\sum_{u \in U_i} |T_i^u - \bar{r}_u|}{|U_i|}\right) \times \sqrt{\frac{|I_i|}{|I|}} \quad (8)$$

Where  $\bar{r}_u$  is the average rating of user  $u$ , and  $|U_i|$  represents the total number of users who have rated item  $i$ .  $|I_i|$  denotes the cardinality of the set of items that exhibit similarity associations with item  $i$ , while  $|I|$  represents the cardinality of the entire set of items in the dataset.  $T_i^u$  signifies the total utility of user  $u$  for item  $i$ , which is attained by employing the following aggregate function [27].

$$T_i^u = \sum_{c=1}^k w_{u,c} \times r_{u,c}$$

where  $\sum_{c=1}^k w_{u,c} = 1, w_{u,c} > 0$  (9)

Where  $r_{u,c}$  represents the rating of user  $u$  on item  $i$  in relation to criteria  $c$ , and  $w_{u,c}$  is a weighting factor that signifies the significance of criterion  $c$  as assigned by user  $u$ .

### 2.4. Compute item-based hybrid similarity

In this step, the proposed approach incorporates the switching hybridization strategy [8] to switch between the three proposed similarity measures: item-based content similarity, item-based MC CF similarity, and global item similarity, based on specific conditions. The selection criterion for choosing the appropriate similarity measure is its ability to generate a similarity value between the target item and other items. In cases where both item-based content similarity and item-based MC CF similarity can provide similarity values, the arithmetic-harmonic mean is used to combine these values. The harmonic mean function is used to guarantee a high total similarity value is achieved only when both item-based content similarity and item-based MC CF similarity yield high similarity values, as (10):

$$Sim_{i,j}^{final} = \begin{cases} GIS_j & ; \text{if } Sim_{i,j}^{CF} = 0 \text{ and } Sim_{i,j}^{Content} = 0 \\ Sim_{i,j}^{CF} & ; \text{if } Sim_{i,j}^{CF} \neq 0 \text{ and } Sim_{i,j}^{Content} = 0 \\ Sim_{i,j}^{Content} & ; \text{if } Sim_{i,j}^{CF} = 0 \text{ and } Sim_{i,j}^{Content} \neq 0 \\ \frac{2 \times Sim_{i,j}^{CF} \times Sim_{i,j}^{Content}}{Sim_{i,j}^{CF} + Sim_{i,j}^{Content}} & ; \text{if } Sim_{i,j}^{CF} \neq 0 \text{ and } Sim_{i,j}^{Content} \neq 0 \end{cases} \quad (10)$$

### 2.5. Compute predicted ratings

For computing predictions, the weighted sum of deviations metric is utilized to generate predicted ratings by considering the deviations from the mean ratings, as (11):

$$PredRating_{u,i} = \bar{r}_i + \frac{\sum_{j \in NN} Sim_{i,j}^{final} \times (T_j^u - \bar{r}_j)}{\sum_{j \in NN} Sim_{i,j}^{final}} \quad (11)$$

Where  $NN$  represents the *Top-N* set of nearest neighbors of items to the target item  $i$  based on the final similarity.  $T_j^u$  signifies the total utility of user  $u$  for item  $j$ .

## 3. PERFORMANCE EVALUATION

### 3.1. Datasets description

In order to assess the efficacy of the proposed approach, two MC rating datasets were employed: the RateMDs MC dataset and the WebMD MC rating dataset:

- The RateMDs MC dataset has been collected from the *ratemds.com* healthcare portal, which functions as a platform where patients can provide reviews of doctors, utilizing a rating scale that spans from 1 to 5. The ratings determined by four criteria: punctuality, staff, knowledge, and helpfulness. The dataset comprises 31,180 MC ratings, which have been contributed by 3,464 patients for a total of 3,118 doctors. The dataset contains 21 different specialties representing the doctors, such as cardiologist, gynecologists, rheumatologist, pediatricians, physiatrists, and family doctors, and more.
- The WebMD MC dataset has been collected from the *webmd.com* healthcare portal, functions as a platform where patients can provide reviews of medicines, utilizing a rating scale that spans from 1 to 5 across three criteria: ease of use, medication effectiveness, and satisfaction. The dataset includes 32,054 ratings for 845 medications, which have been contributed by 2,136 patients in total. The medicines in the dataset are represented by 915 main categories, which correspond to various medical conditions. These categories include but are not limited to nerve pain, joint pain, psoriasis of the scalp, backache, and many more.

In both datasets, we allocated 80% of the data to the training set, ensuring that it contained a substantial portion of the ratings for training. The remaining 20% of the data was reserved for the test set, which served as an independent dataset for evaluation purposes.

### 3.2. Evaluation metrics

In order to evaluate the performance of the proposed approach and baseline recommendation methods, we employed three evaluation metrics: mean absolute error (MAE), root-mean-square error (RMSE), and coverage [28]. MAE is a widely utilized measure of evaluation in RSs for assessing prediction accuracy. It computes the mean absolute deviation between the predicted ratings and the actual ratings given for items by users. A lower MAE value indicates better prediction accuracy, indicating that the recommendation method's predictions align more closely with the true ratings. RMSE is another popular evaluation metric in RSs. It calculates the square root of the average of the squared differences between the predicted ratings and the actual ratings. RMSE provides a measure of the overall prediction error and is commonly used alongside MAE. Similar to MAE, a lower RMSE value indicates better prediction accuracy. Coverage is also a widely utilized measure of evaluation in RSs, especially when dealing with sparsity. It assesses the ability of a recommendation method to produce predictions for a varied set of items, comprising those that are unrated or new. It calculates the proportion of unrated items that are yet recommended by the recommendation approach. Increased coverage indicates that the recommendation approach can recommend a broader variety of items in the collected data, even including new or unrated ones. This implies that the method can offer more personalized recommendations to users and is less affected by data sparsity [28].

### 3.3. Baselines

For the purpose of conducting a comparative analysis, the following item-based recommendation methods have been included as baselines. These methods represent different strategies for item-based recommendation approaches, incorporating single-criteria ratings, MC ratings, and semantic relationships among items.

- SC-ICF [29] is based on single-criteria ratings to make personalized recommendations. The cosine similarity is used to calculate the similarity between items in order to identify the NN.
- MC-ICF [25] is an expansion of the SC-ICF that incorporates MC ratings. MC-ICF aims to enhance prediction accuracy and provide more accurate and relevant recommendations.
- MC-SeCF [9] incorporates both MC ratings and semantic relations among items. MC-ISCF addresses challenges such as data sparsity and cold-start item problems by incorporating item semantics into the recommendation process.

### 3.4. Comparison results

In order to demonstrate that the proposed method is more effective than the baseline methods in terms of improving prediction accuracy and coverage, a number of experiments were carried out, as outlined.

#### 3.4.1. Performance comparison on the RateMDs MC dataset

Figures 1 and 2 depict the results of the experimental evaluation of the proposed approach, along with the baseline methods, on the RateMDs MC dataset. The focus of these experiments is to evaluate the prediction accuracy using the MAE and RMSE metrics across different NN values, specifically 5, 10, 15, 20, 25, 30, and 50. In Figure 1, it is evident that the proposed approach achieves the best performance (lowest MAE) across different NN values. When compared to the SC-ICF, MC-ICF, and MC-SeCF methods, the proposed approach showcases significant improvements of 90%, 90%, and 86% in terms of MAE,

respectively. Furthermore, Figure 2 demonstrates that the proposed approach also outperforms other baseline methods in terms of RMSE, displaying the minimum RMSE values with respect to various values of NN. In relation to RMSE, the proposed approach exhibits notable improvements of 80%, 80%, and 76% when compared to the SC-ICF, MC-ICF, and MC-SeCF methods, respectively. Overall, the results of the experimental evaluation show that the proposed approach outperforms the baseline methods in terms of both MAE and RMSE. This means that the proposed approach can predict the ratings more accurately and precisely than the baseline methods on the RateMDs MC dataset.

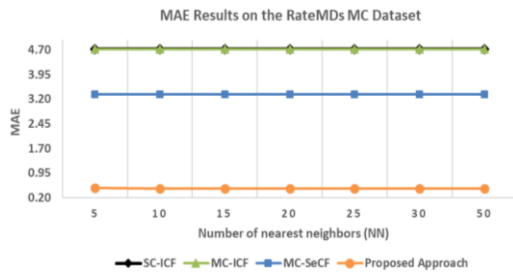


Figure 1. The MAE of different approaches with different NN on RateMDs dataset

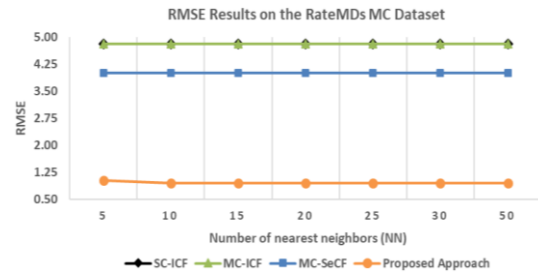


Figure 2. The RMSE of different approaches with different NN on RateMDs dataset

**3.4.2. Performance comparison on the WebMD MC dataset**

Figures 3 and 4 illustrate the outcomes of the experimental evaluation conducted on the WebMD MC dataset, comparing the proposed approach with the baseline methods. These experiments aim to evaluate the prediction accuracy using the MAE and RMSE metrics across various NN values: 5, 10, 15, 20, 25, 30, and 50. Figure 3 clearly demonstrates that the proposed approach achieves the best performance (lowest MAE) across different values of NN. Comparatively, when compared to the SC-ICF, MC-ICF, and MC-SeCF methods, the proposed approach demonstrates significant improvements of 63%, 56%, and 41% in terms of MAE, respectively. Additionally, Figure 4 presents that the proposed approach also surpasses other baseline methods in terms of RMSE, displaying the minimum RMSE values with respect to various values of NN. In relation to RMSE, the proposed approach exhibits notable enhancements of 58%, 52%, and 41% when compared to the SC-ICF, MC-ICF, and MC-SeCF methods, respectively. Overall, the results of the experimental evaluation indicate that the proposed approach surpasses the baseline methods in both MAE and RMSE metrics. This implies that the proposed approach demonstrates superior accuracy and precision in predicting ratings compared to the baseline methods when applied to the WebMD MC dataset.

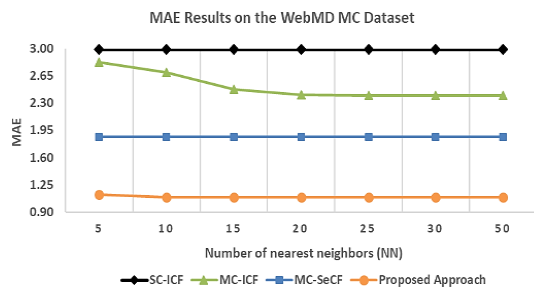


Figure 3. The MAE of different approaches with different NN on WebMD dataset

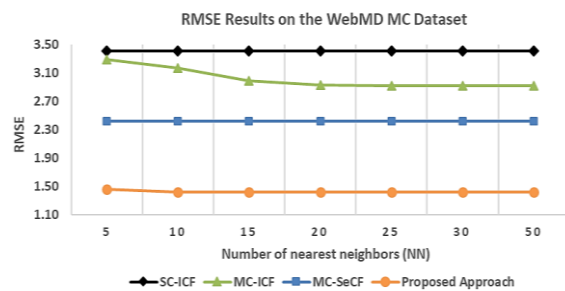


Figure 4. The RMSE of different approaches with different NN on WebMD dataset

**3.4.3. Performance comparison on a dataset of various sparsity levels**

Experiments were carried out to validate the practicality of the proposed approach in addressing the issue of data sparsity. These experiments involved the creation of six datasets with different levels of sparsity, varied from 99.8% to 98%. The primary goal of these experiments was to evaluate the capacity of the compared methods to cope with different levels of sparsity and assess their performance accordingly.



Figure 5 compares the MAE results of the proposed approach to those of the baseline methods. The performance of the proposed method is estimated to be 67%, 61%, and 31% superior to that of the baseline methods, respectively. This emphasizes the enhanced efficacy of the proposed approach in precisely predicting ratings amidst diverse levels of data sparsity. Figure 6 exhibits that the proposed approach improves not only MAE but also prediction coverage. The proposed approach exhibits notable increases in prediction coverage of 57%, 45%, and 13% relative to the baseline methods. This demonstrates that the proposed approach is also superior at predicting ratings not previously observed. The perceived improvements in MAE and prediction coverage provide compelling evidence of the robustness of the proposed approach, especially when handling highly sparse datasets.

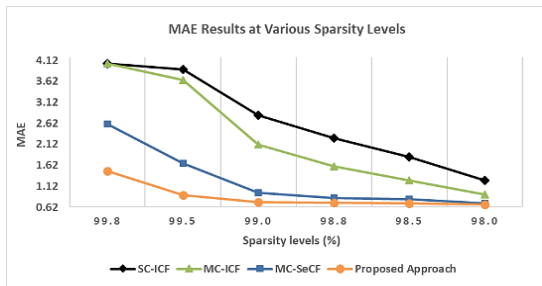


Figure 5. MAE results at various sparsity levels

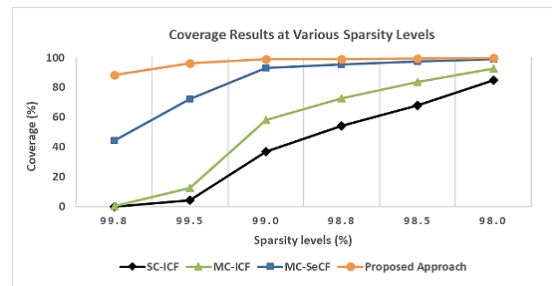


Figure 6. Prediction coverage results at various sparsity levels

### 3.5. Discussion and implications

The results from the experimental evaluation above provide compelling evidence supporting the effectiveness of our proposed approach in enhancing both prediction accuracy and coverage in health-related recommendations. Notably, it exhibits robust performance in handling extremely sparse datasets. However, the SC-ICF and MC-ICF methods depend only on ratings for identifying neighbors and generating predictions, resulting in lower accuracy and coverage, especially in datasets with very sparse ratings.

While the MC-SeCF approach, which closely resembles our proposed method, demonstrates superior performance compared to SC-ICF and MC-ICF in accuracy for rating prediction and coverage, particularly in situations with sparse datasets, it achieves this by integrating MC ratings and semantic associations among items. However, our proposed approach goes a step further by incorporating reputation scores along with content-based filtering and MC-based CF during the neighbor selection phase. This leads to better predictive rating accuracy and coverage, especially when working with highly sparse datasets. For instance, our proposed approach showed a 45% improvement in MAE and a 50% improvement in coverage compared to the MC-SeCF approach in a dataset demonstrating a sparsity score of 99.8%. Moreover, in the dataset that has a sparsity score of 99.5%, our approach shows an average improvement of 43% and 25% compared to the MC-ICF approach with respect to MAE and coverage, correspondingly.

The ramifications of implementing the proposed recommendation approach are multifaceted. Firstly, it can empower patients by facilitating their search for healthcare providers who specialize in addressing their specific requirements, thereby ensuring treatment that is more tailored and efficient. Furthermore, it can play a crucial role in assisting patients in promptly identifying the most suitable medication that corresponds to their specific medical conditions. In the future, the incorporation of this health recommendation approach has the capacity to revolutionize patient care through the provision of a reliable and personalized decision support system. This system has the potential to be extremely valuable in the future, as it provides users and healthcare practitioners with a tool that not only tackles the problem of having too much information but also improves the efficiency and effectiveness of healthcare decision-making.

## 4. CONCLUSION

The abundance and intricacy of health-related information available on various online healthcare platforms pose challenges for patients in making effective healthcare-related decisions. To address this issue, this study proposes an effective health recommendation approach that is promising in aiding healthcare practitioners and users alike to make well-informed decisions pertaining to their specific health conditions and needs. The utilization of this approach may facilitate the assistance of patients in locating doctors who

can satisfy their particular needs for efficacious treatment and in promptly determining the optimal medication that corresponds with their distinct medical conditions. The proposed approach integrates content-based filtering with MC-based CF to enhance the efficacy of recommendations and tackle the data sparsity challenge. The content-based filtering improves the recommendation process by utilizing item attributes to generate recommendations for items that have sparse ratings. The MC CF approach utilizes an innovative similarity metric that takes into account both structural and distance similarities, in addition to a global item similarity measure, in order to address the difficulties presented by data sparsity. The effectiveness of the proposed approach in relation to predictive accuracy and coverage, when compared to other baseline recommendation methods, was demonstrated through experimentation using real-world healthcare MC rating datasets. In the future, this work can be expanded by incorporating user review sentiments into the proposed approach's recommendation process in order to better comprehend user preferences and generate more accurate recommendations.




## REFERENCES

- [1] T. N. T. Tran, A. Felfernig, C. Trattner, and A. Holzinger, "Recommender systems in the healthcare domain: state-of-the-art and research issues," *Journal of Intelligent Information Systems*, vol. 57, no. 1, pp. 171-201, 2021, doi: 10.1007/s10844-020-00633-6.
- [2] Y. Cai, F. Yu, M. Kumar, R. Gladney, and J. Mostafa, "Health recommender systems development, usage, and evaluation from 2010 to 2022: A scoping review," *International Journal of Environmental Research and Public Health*, vol. 19, no. 22, pp. 1-15, 2022, doi: 10.3390/ijerph192215115.
- [3] Y. Sun, J. Zhou, M. Ji, L. Pei, and Z. Wang, "Development and Evaluation of Health Recommender Systems: Systematic Scoping Review and Evidence Mapping," *Journal of Medical Internet Research*, vol. 25, p. e38184, 2023, doi: 10.2196/38184.
- [4] C. C. Aggarwal, "Neighborhood-Based Collaborative Filtering," *Recommender Systems: The Textbook*, pp. 29-70, Cham: Springer International Publishing, 2016, doi: 10.1007/978-3-319-29659-3\_2.
- [5] H. Ko, S. Lee, Y. Park, and A. Choi, "A Survey of Recommendation Systems: Recommendation Models, Techniques, and Application Fields," *Electronics*, vol. 11, no. 1, pp. 1-48, 2022, doi: 10.3390/electronics11010141.
- [6] Q. Shambour, A. H. Hussein, M. Abualhaj, and Q. Kharma, "An effective hybrid content-based collaborative filtering approach for requirements engineering," *Computer Systems Science and Engineering*, vol. 40, no. 1, pp. 113-125, 2022, doi: 10.32604/csse.2022.017221.
- [7] E. Çano, and M. Morisio, "Hybrid recommender systems: A systematic literature review," *Intelligent Data Analysis*, vol. 21, no. 6, pp. 1487-1524, 2017, doi: 10.3233/IDA-163209.
- [8] R. Burke, "Hybrid Web Recommender Systems," *The Adaptive Web: Methods and Strategies of Web Personalization*, pp. 377-408, 2007, doi: 10.1007/978-3-540-72079-9\_12.
- [9] Q. Shambour, M. Hourani, and S. Fraihat, "An Item-based Multi-Criteria Collaborative Filtering Algorithm for Personalized Recommender Systems," *International Journal of Advanced Computer Science and Applications*, vol. 7, no. 8, pp. 274-279, 2016, doi: 10.14569/IJACSA.2016.070837.
- [10] Q. Shambour, and J. Lu, "Integrating multi-criteria collaborative filtering and trust filtering for personalized recommender systems," in *Proc. of the 2011 IEEE Symposium on Computational Intelligence in Multicriteria Decision-Making*, Paris, France, 2011, pp. 44-51, doi: 10.1109/SMDCM.2011.5949274.
- [11] Q. Y. Shambour, M. M. Abu-Alhaj, and M. M. Al-Tahravi, "A hybrid collaborative filtering recommendation algorithm for requirements elicitation," *International Journal of Computer Applications in Technology*, vol. 63, no. 1-2, pp. 135-146, 2020, doi: 10.1504/IJCAT.2020.107908.
- [12] Q. Y. Shambour, A. A. Abu-Shareha, and M. M. Abualhaj, "A hotel recommender system based on multi-criteria collaborative filtering," *Information Technology and Control*, vol. 51, no. 2, pp. 390-402, 2022, doi: 10.5755/j01.itc.51.2.30701.
- [13] J. Lu, D. Wu, M. Mao, W. Wang, and G. Zhang, "Recommender system application developments: A survey," *Decision Support Systems*, vol. 74, pp. 12-32, 2015, doi: 10.1016/j.dss.2015.03.008.
- [14] F. Narducci, C. Musto, M. Polignano, M. d. Gemmis, P. Lops, and G. Semeraro, "A Recommender System for Connecting Patients to the Right Doctors in the HealthNet Social Network," in *Proc. of the 24th International Conference on World Wide Web*, Florence, Italy, 2015, pp. 81-82, doi: 10.1145/2740908.2742748.
- [15] Q. Han, M. Ji, I. M. d. R. d. Troya, M. Gaur, and L. Zejnilovic, "A Hybrid Recommender System for Patient-Doctor Matchmaking in Primary Care," in *Proc. of the 2018 IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA)*, Turin, Italy, 2018, pp. 481-490, doi: 10.1109/dsaa.2018.00062.
- [16] S. Mondal, A. Basu, and N. Mukherjee, "Building a trust-based doctor recommendation system on top of multilayer graph database," *Journal of Biomedical Informatics*, vol. 110, p. 103549, 2020, doi: 10.1016/j.jbi.2020.103549.
- [17] C. Chen, L. Zhang, X. Fan, Y. Wang, C. Xu, and R. Liu, "A Epilepsy Drug Recommendation System by Implicit Feedback and Crossing Recommendation," in *Proc. of the IEEE SmartWorld, Ubiquitous Intelligence & Computing*, Guangzhou, China, 2018, pp. 1134-1139, doi: 10.1109/SmartWorld.2018.00197.
- [18] M. D. Hossain, M. S. Azam, M. J. Ali, and H. Sabit, "Drugs Rating Generation and Recommendation from Sentiment Analysis of Drug Reviews using Machine Learning," in *Proc. of the 2020 Emerging Technology in Computing, Communication and Electronics*, 2020, pp. 1-6, doi: 10.1109/ETCCE51779.2020.9350868.
- [19] Y. Zhang, X. Wu, Q. Fang, S. Qian, and C. Xu, "Knowledge-enhanced attributed multi-task learning for medicine recommendation," *ACM Transactions on Information Systems*, vol. 41, no. 1, pp. 1-24, 2023, doi: 10.1145/3527662.
- [20] I. Tiddi and S. Schlobach, "Knowledge graphs as tools for explainable machine learning: A survey," *Artificial Intelligence*, vol. 302, 2022, doi: 10.1016/j.artint.2021.103627.
- [21] H. Papadakis, A. Papagrigoriou, C. Panagiotakis, E. Kosmas, and P. Fragopoulou, "Collaborative filtering recommender systems taxonomy," *Knowledge and Information Systems*, vol. 64, no. 1, pp. 35-74, 2022, doi: 10.1007/s10115-021-01628-7.
- [22] G. Adomavicius and A. Tuzhilin, "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*, vol. 17, no. 6, pp. 734-749, 2005, doi: 10.1109/TKDE.2005.99.
- [23] S.-S. Choi, S.-H. Cha, and C. C. Tappert, "A survey of binary similarity and distance measures," *Journal of systemics, cybernetics and informatics*, vol. 8, no. 1, pp. 43-48, 2010.




- [24] P. Resnick, N. Iacovou, M. Suchak, P. Bergstrom, and J. Riedl, "GroupLens: an open architecture for collaborative filtering of netnews," in *Proc. of the 1994 ACM Conference on Computer Supported Cooperative Work*, USA, 1994, pp. 175-186, doi: 10.1145/192844.192905.
- [25] G. Adomavicius and Y. O. Kwon, "New recommendation techniques for multicriteria rating systems," *IEEE Intelligent Systems*, vol. 22, no. 3, pp. 48-55, 2007, doi: 10.1109/MIS.2007.58.
- [26] P. Pirasteh, D. Hwang, and J. E. Jung, "Weighted Similarity Schemes for High Scalability in User-Based Collaborative Filtering," *Mobile Networks and Applications*, vol. 20, no. 4, pp. 497-507, 2015, doi: 10.1007/s11036-014-0544-5.
- [27] J. S. Dyer, "MAUT—Multiattribute Utility Theory," *Multiple Criteria Decision Analysis: State of the Art Surveys*, New York, NY: Springer New York, pp. 265-292, 2005, doi: 10.1007/0-387-23081-5\_7.
- [28] C. C. Aggarwal, "Evaluating Recommender Systems," *Recommender Systems: The Textbook*, C. C. Aggarwal, ed., Cham: Springer International Publishing, pp. 225-254, 2016, doi: 10.1007/978-3-319-29659-3\_7.
- [29] B. Sarwar, G. Karypis, J. Konstan, and J. Riedl, "Item-based collaborative filtering recommendation algorithms," in *Proceedings of the 10th international conference on World Wide Web*, Hong Kong, China, Apr. 2001, pp. 285-295, doi: 10.1145/371920.372071.

## BIOGRAPHIES OF AUTHORS






**Adeeb M. Alsaaidah**    received the Bachelor's degree in Computer Engineering from the Faculty of Engineering, ALBalqa Applied University, the Master's degree in Networking and Computer Security from NYIT University, and the Ph.D. degree in Computer Network from USIM, Malaysia. He is currently an Assistant Professor in Department of Network and Cybersecurity at Al-Ahliyya Amman University (AAU). His research interests include network performance, multimedia networks, network quality of service (QoS), the IoT, network modeling and simulation, network security, and cloud security. He can be contacted at email: a.alsaaidah@ammanu.edu.jo.






**Qusai Y. Shambour**    holds a B.Sc. degree in Computer Science from Yarmouk University, Jordan, in 2001, an M.Sc. degree in Computer Networks from the University of Western Sydney, Australia, in 2003, and a Ph.D. degree in software engineering from the University of Technology Sydney, Australia, in 2012. Currently, he is a Professor at the Department of Software Engineering, Al-Ahliyya Amman University, Jordan. His research interests include information filtering, recommender systems, VoIP, machine learning, and data science. He can be contacted at email: q.shambour@ammanu.edu.jo.



**Mosleh M. Abualhaj**    is a senior lecturer at Al-Ahliyya Amman University. He received his first degree in Computer Science from Philadelphia University, Jordan, in 2004, his Master degree in Computer Information System from the Arab Academy for Banking and Financial Sciences, Jordan, in 2007, and his Ph.D. in Multimedia Networks Protocols from Universiti Sains Malaysia in 2011. His research area of interest includes VoIP, multimedia networking, and congestion control. He can be contacted at email: m.abualhaj@ammanu.edu.jo.



**Ahmad Adel Abu-Shareha**    received his first degree in Computer Science from Al Al-Bayt University, Jordan, in 2004, Master degree from Universiti Sains Malaysia (USM), Malaysia, in 2006, and his Ph.D. degree from USM, Malaysia, in 2012. His research focuses on data mining, artificial intelligent, and multimedia security. He investigated many machine learning algorithms and employed artificial intelligence in a variety of fields, such as network, medical information processing, knowledge construction, and extraction. He can be contacted at email: a.abushareha@ammanu.edu.jo.