Journal of Applied Data Sciences
Vol. 5, No. 3, September 2024, pp. 1134-1146

Mitigating Healthcare Information Overload: a Trust-aware Multi-Criteria Collaborative Filtering Model

Qusai Y. Shambour[1,*], Mosleh M. Abualhaj[2], Ahmad Abu-Shareha[3], Abdelrahman H. Hussein[4], Qasem M. Kharma[5]

[1,5]*Department of Software Engineering, Faculty of Information Technology, Al-Ahliyya Amman University, Jordan*

[2,4]*Department of Networks and Information Security, Faculty of Information Technology, Al-Ahliyya Amman University, Jordan*

[3]*Department of Data Science and Artificial Intelligence, Faculty of Information Technology, Al-Ahliyya Amman Universit, Jordan*

(Received: May 30, 2024; Revised: June 24, 2024; Accepted: July 22, 2024; Available online: August 5, 2024)

Abstract

The rapid growth of online health information resources has made it difficult for users, as well as providers of healthcare, to cope with large volumes of information that are becoming increasingly complex. Hence, there is an urgent demand for developing new advanced recommendation techniques in the healthcare domain to enhance decision-making processes. However, most current health recommendation systems, which recommend personalized healthcare services and items such as diagnoses, medications, and doctors based on users' health conditions and needs, are hindered by the data sparsity issue that compromises the reliability of their recommendations. In this paper, we intend to address this issue by proposing a Trust-aware Multi-Criteria Collaborative Filtering model for recommendation services in the healthcare domain. This model leverages multi-criteria ratings and integrates user-item trust relationships to improve the precision and coverage of recommendations, thus facilitating more informed healthcare choices that align closely with their individual needs. Our empirical analysis on two healthcare multi-criteria rating datasets, including those with sparse data, shows the proposed model's superior performance over existing baseline methods. On the RateMDs dataset, our model improved the average MAE by 24% and RMSE by 19% compared to baseline methods. For the WebMD dataset, it enhanced the average MAE by 6% and RMSE by 2%. In sparse data scenarios, the model boosted the average MAE by 18% and Coverage by 6% compared to baseline approaches.

Keywords: Recommender Systems, Collaborative Filtering, Healthcare, Multi-Criteria, Trust Relationships, Data Sparsity

1. Introduction

The healthcare industry is facing a critical challenge: information overload. The sheer volume of data and resources available to patients and medical professionals has created a pressing need for personalized healthcare services that can effectively support decision-making. Health recommender systems (HRS) have emerged as a promising solution, designed to suggest personalized health-related items or services, such as medications, treatments, healthcare providers, and lifestyle recommendations tailored to individual users' needs and conditions. By leveraging user profiles and historical preferences, HRS aim to streamline decision-making processes for both patients and healthcare professionals, thus facilitating a more personalized healthcare experience [1], [2].

One of the key challenges in developing effective HRS is the scarcity of high-quality user interaction data. This scarcity, known as data sparsity, arises when users interact infrequently with specific health-related items, making it difficult to accurately predict their preferences. This issue is particularly pronounced in the healthcare domain, where user interactions with health-related items may be infrequent or highly individualized [2], [3], [4], [5], [6], [7]. For instance, a medication recommender system may struggle to provide accurate recommendations for rare diseases because few patients have rated or reviewed these medications. Similarly, new medications or users may lack sufficient rating data, posing a challenge for personalized recommendations.

To overcome data sparsity, researchers have turned to innovative approaches that incorporate additional information beyond user-item ratings. One such approach is trust-aware recommender systems, which leverage trust relationships

***Corresponding author: Qusai Y. Shambour (q.shambour@ammanu.edu.jo)**
DOI: https://doi.org/10.47738/jads.v5i3.297

between users to improve recommendation accuracy. In the healthcare domain, trust can play a crucial role, as patients often rely on the experiences and recommendations of others whom they trust. By integrating trust metrics into the recommendation process, systems can better capture the reliability and relevance of user-generated content, thereby enhancing the quality of the recommendations [8].

Furthermore, most existing recommendation approaches employ a single-criterion approach to explore the correlation between historical feedback and model predictions. However, the single-criterion methods ignore the users' multi-criteria (MC) behavioral characteristics in item evaluation and selection decisions [9], [10], [11]. Therefore, there is a growing effort to develop multi-criteria recommender systems capable of generating recommendations that resonate with the complex and varied preferences of individual users [12].

Recent advancements in artificial intelligence and machine learning have revolutionized the development of HRS, enabling the integration of complex data sources and sophisticated algorithms to improve recommendation accuracy. For instance, deep learning models with attention mechanisms have been employed to capture critical information about patients' conditions from text data, such as consultation records and textual reviews [13], [14]. Opinion-mining techniques have been applied to analyze drug review sentiments and determine medication effectiveness [15], [16]. Moreover, matrix factorization and adversarial knowledge graphs have been leveraged to optimize recommendations for accuracy and safety [17], [18]. Despite these advancements, many current HRS rely on model-based recommendation methods that require extensive feature collection, dimensionality reduction, and large datasets, potentially compromising recommendation effectiveness [19].

This study introduces a novel trust-aware multi-criteria collaborative filtering (TAMCCF) model that addresses the dual challenges of data sparsity and single-criterion limitations in order to offer accurate, personalized, and diverse healthcare recommendations. This can empower patients to make informed healthcare decisions that align with their individual needs and health status. The proposed model integrates multi-criteria ratings with trust relationships to address the dual challenges of data sparsity and the need for high-quality, personalized recommendations. By capitalizing on the additional data provided by multi-criteria ratings and the trust relationships among users, the TAMCCF model aims to deliver more accurate, relevant, and reliable recommendations to patients. Evaluations using real-world healthcare recommendation datasets demonstrate that our proposed model achieves superior performance in terms of both prediction accuracy and coverage compared to existing benchmark recommendation methods.

2. Related Works

Recent studies have proposed innovative recommendation methods as potential solutions to the overabundance of online medical information by leveraging machine and deep learning, data mining, and hybrid filtering techniques. For instance, in the realm of doctor recommendation systems, several studies have explored deep learning models that incorporate attention mechanisms to capture critical information about patients' conditions from text data such as consultation records and textual reviews [13], [14]. For example, Nie et al. [13] proposed a hierarchical attention network (HAN) to construct doctor-patient models from consultation records, using attention to enhance doctor-patient interaction. Their recommendation scheme based on patient ratings achieved a hit rate of 79.7%. Similarly, Kulshrestha et al. [14] developed a hierarchical attention bidirectional long short-term memory network to predict online doctor ratings from textual reviews, outperforming benchmark models. Other studies have focused on hybrid filtering approaches that combine collaborative filtering (CF) and content-based filtering to match patients with suitable doctors. Mani and Thilagamani [20] integrated demographic, collaborative, and content-based filtering along with a fuzzy analytic hierarchy process to rank doctors based on patients' preferred criteria. Sofia et al. [21] implemented a hybrid filtering method for telemedicine, using collaborative filtering with rating parameters and content-based filtering with doctor profiles, achieving 91% accuracy.

In the context of drug recommendations, researchers have investigated opinion-mining techniques to analyze drug review sentiments and determine medication effectiveness [15], [16]. Keikhosrokiani et al. [15] proposed a system that applies opinion mining to drug reviews and uses a hybrid filtering method to overcome the limitations of content-based and collaborative filtering, assisting healthcare professionals in drug decision-making. Begum and Sree [16] utilized a recurrent neural network for sentiment analysis on drug reviews to predict appropriate medications. Other studies have focused on leveraging matrix factorization and adversarial knowledge graphs of drug-drug interactions to optimize recommendations for accuracy and safety. Symeonidis et al. [17], [18] extended matrix co-factorization and singular value decomposition algorithms to incorporate information from patient's electronic health records and drug-drug interaction knowledge graphs, reducing the toxicity scores of recommended drug combinations while maintaining acceptable efficacy.

While prior studies have made valuable contributions to developing health recommendation systems, several critical research gaps remain. Existing efforts have mainly focused on individual aspects, such as doctor or drug recommendations, rather than integrating them into unified models within a single framework. This integration could greatly improve the efficiency and effectiveness of healthcare decision-making processes. Additionally, there has been limited consideration of trust factors, such as the reputation of doctors and drugs, as well as the patient-patient relationship, all of which can significantly influence the acceptance of recommendations. Moreover, many current health recommendation systems rely on model-based CF that require extensive feature collection, dimensionality reduction, and large datasets, potentially compromising recommendation effectiveness [19]. Finally, there is a lack of research on developing multi-criteria recommender systems tailored specifically for health recommendations.

3. Design of the TAMCCF Model

This study presents an effective TAMCCF model designed specifically for health recommendation systems, comprising three primary components. Firstly, the user-based trust-aware multi-criteria collaborative filtering component incorporates trust relationships among users and multi-criteria ratings assigned by users to generate recommendations. Secondly, the item-based trust-aware multi-criteria collaborative filtering component integrates trust relationships among items and multi-criteria ratings associated with items to generate recommendations. Finally, the hybrid prediction component merges the predicted ratings derived from the preceding two components, leveraging their respective strengths to deliver personalized health recommendations.

3.1. The User-based Trust-Aware Multi-Criteria CF Component

This component utilizes the direct, propagated, and global trust scores to produce user-based trust-aware predictions. Initially, trust values are computed based on user ratings, thereby establishing the initial trust network with direct connections between users. Subsequently, leveraging this network facilitates the propagation of indirect trust among users who lack direct connections. Finally, each user's global trust score is derived from their average rating deviation from item means and their connectivity within the propagated trust network.

Step 1: Calculate direct trust scores between users

This study conceptualizes a user's "trustworthiness" as their reliability in delivering accurate recommendations to others, drawing upon previous research indicating a strong correlation between user similarity and trust within online communities. Accordingly, evaluating the trustworthiness of a pair of users entails assessing the accuracy of the ratings predicted by one user for the other, considering their respective past ratings [10]. To facilitate this process, the following prediction formula is initially employed to compute the predicted ratings for each user pair, as elucidated below:

$$P_{m,d} = \bar{r}_m + (U^n(d) - \bar{r}_n) \tag{1}$$

Here, $\bar{r}_m$ and $\bar{r}_n$ represent the mean rating of users m and n respectively. $U^n(d)$ signifies the overall utility of user n on item d, as defined below:

$$U^n(d) = \sum_{a=1}^{k} w_a^n(d) \times r_a^n(d), \quad \text{where} \sum_{a=1}^{k} w_a^n(d) = 1 \tag{2}$$

Where $w_a^n(d)$ denotes the weight assigned by user n to criterion a for item d, reflecting its importance. $r_a^n(d)$ represents the rating provided by user n on criterion a for item d.

Following the prediction, two complementary metrics are employed: (1) Mean Squared Differences (MSD) [22], which compares prediction errors on co-rated items, and (2) Bhattacharyya Coefficient (BC) [23], which leverages the entire distribution of ratings provided by users. To account for the confidence in the similarity assessment, an asymmetry trust factor is introduced, which is based on the proportion of co-rated items between users. The direct trust score is then obtained by combining the MSD and BC metrics using a sigmoid function, weighted by an asymmetry factor.

$$\mathrm{Sim}_{m,n}^{MSD} = 1 - \left(\frac{\sum_{d=1}^{x}(P_{m,d} - U^{m}(d))^2}{x}\right) \tag{3}$$

where $P_{m,d}$ and $U^m(d)$ denote the predicted rating and total utility of item d with respect to user m, respectively. x is the total number of co-rated items between users m and n. To ensure MSD values range from 0 to 1, we normalize the predicted and overall utility rating values using Max-Min normalization.

While the MSD metric effectively captures similarities between users based on co-rated items, it becomes less reliable in sparse datasets where the overlap of rated items is limited or non-existent. To overcome this limitation, we employ the BC as a complementary similarity measure that utilizes the entire distribution of recorded ratings from both users to compute their similarity, as follows:

$$\mathrm{Sim}_{m,n}^{BH} = \sum_{d \in I_m} \sum_{f \in I_n} BC(d,f) \times loc(r_{m,d}, r_{n,f}) \tag{4}$$

Here, I_m and I_n are the sets of items rated by users m and n, respectively.

BC and loc represent two distinct similarity metrics. However, they leverage different information sources: BC relies on global information, while loc focuses on local information. Accordingly, BC is used as the initial step to calculate partial similarities between items d and f, as rated by users m and n for each rating criterion a, as follows:

$$BC_{d,f}^{a} = \sum_{v=1}^{z} \sqrt{\left(\frac{\#v}{\#d}\right)\left(\frac{\#v}{\#f}\right)} \tag{2}$$

Where z represents the total number of possible rating values (e.g., 1 to 5), #v represents the number of users rating an item with value v, and #d and #f represent the total number of users who rated items d and f, respectively. Then, we employ a worst-case aggregation function to compute the overall similarity score between given item pair, as follows:

$$BC(d,f) = \min_{a=1,\ldots,k} BC_{d,f}^{a} \tag{3}$$

$loc(r_{m,d}, r_{n,d})$ represents the local similarity among two ratings based on their deviation from the median rating value, as shown below:

$$loc(r_{m,d}, r_{n,d}) = \frac{(r_{m,d}, r_{med})(r_{n,d}, r_{med})}{\sqrt{\sum_{s \in I_m}(r_{m,s}, r_{med})^2}\sqrt{\sum_{s \in I_n}(r_{n,s}, r_{med})^2}} \tag{7}$$

Here, r_{med} represents the value of the median rating, which in our case is 3.

Previous research [9], [24], [25] highlights the importance of considering the number of commonly rated items between users when measuring their level of trust. We account for this by introducing an asymmetry trust factor (TF) [24] that weights the final trust score based on the proportion of co-rated items, as given below:

$$TF_{m,n} = \frac{1}{1 + \exp\left(-\frac{|I_m \cap I_n|}{|I_m|}\right)} \tag{4}$$

where $|I_m \cap I_n|$ is the number of co-rated items between users m and n, and $|I_m|$ is the total number of items rated by user m. To sum up, the overall direct trust between users m and n is obtained by combining the similarity metrics and the asymmetry trust factor using a sigmoid function, as follows:

$$\mathrm{Trust}_{m,n}^{\mathrm{Direct}} = \frac{1}{(1+\exp(-(\mathrm{Sim}_{m,n}^{\mathrm{MSD}} + \mathrm{Sim}_{m,n}^{\mathrm{BH}})))} \times \mathrm{TF}_{m,n} \tag{5}$$

Step 2: Propagate trust scores between not-connected users

Despite the initial direct trust scores computed between users, the resulting trust network may suffer from sparsity issues. This is because users in recommender systems typically provide ratings for only a limited subset of items. To overcome this sparsity challenge and maximize the utility of the trust network, our approach adopts the concept of trust propagation, which is commonly observed in social networks. Trust propagation allows trust relationships to be transmitted through intermediary users, thereby establishing new indirect connections within the network. This process expands the trust network, enabling the modeling of more intricate trust relationships that go beyond direct interactions between users. Accordingly, we propose an aggregation function that incorporates confidence weights when measuring trust propagated between users. This function quantifies the level of trust from a user *m* to a user *u*, mediated through a common neighbor *n*, as follows:

$$\mathrm{Trust}_{m,u}^{\mathrm{Prop}} = \frac{\sum_{n \in \mathrm{intermediary(m\ and\ u)}} (\mathrm{Trust}_{m,n}^{\mathrm{Direct}} \times \mathrm{TF}_{m,n}) + (\mathrm{Trust}_{n,u}^{\mathrm{Direct}} \times \mathrm{TF}_{n,u})}{\sum_{n \in \mathrm{intermediary(m\ and\ u)}} \mathrm{TF}_{m,n} + \mathrm{TF}_{n,u}} \tag{6}$$

Step 3: Calculate global trust scores for users

Traditional recommender systems often struggle with data sparsity, where limited ratings make it difficult to generate accurate recommendations. Our model tackles this challenge by incorporating user reputation, quantified by a global trust score, as a supplementary factor [26]. When a user has few or no ratings, our model can still generate personalized recommendations by leveraging the reputation of similar users. This is achieved by incorporating the global trust score into the model to provide relevant recommendations even in the absence of sufficient rating data.

This step, which leverages user reputation through a global trust score calculation, sets our model apart from existing methods by providing a more comprehensive and robust solution. The global trust score for each user score plays a key role in improving recommendation accuracy, especially for active users with limited interaction history. The score is derived from two factors [26]: the average deviation between the user's ratings and the item's average rating, which captures the alignment of the user's ratings with the broader consensus and indicates the trustworthiness of their evaluations, and the number of users who trust the given user, reflecting the extent to which the user is trusted by their peers within the propagated users' trust network

A user with a smaller average rating deviation and higher connectivity receives a higher global trust score. The global trust score for a user *m* is calculated as follows:

$$\mathrm{GTS}_m = \exp\left(-\frac{\sum_{d \in I_m} \left|r_{m,d} - \bar{r}_d\right|}{|I_m|}\right) \times \sqrt{\frac{|U_m|}{|U|}} \tag{7}$$

where $r_{m,d}$ represents average ratings across all criteria given by user *m* given to item *d*, $\bar{r}_d$ denotes mean ratings across all criteria for of item *i* given by all users, and $|U_m|$is the number of users connected to user *m* within the users' trust network.

Step 4: Calculate the user-based trust-aware predicted ratings

The final step involves generating predicted ratings for an active user and a target item by leveraging the direct, propagated, and global trust scores. A mean-centering approach [27] is employed, where the predicted rating is computed based on the weighted average of the target item's mean rating and the deviations of similar users' ratings from their respective means, weighted by the corresponding trust scores, as shown below:

$$P_{m,d}^{User} = \begin{cases} \bar{r}_m + \dfrac{\sum_{n\in NN(m)} Trust_{m,n} \times (r_{n,d} - \bar{r}_n)}{\sum_{n\in NN(m)} Trust_{m,n}}; & \text{if } Trust_{m,n} \neq 0 \\ \bar{r}_m + \dfrac{\sum_{n\in NN(m)} GTS_n \times (r_{n,d} - \bar{r}_n)}{\sum_{n\in NN(m)} GTS_n} \quad ; & \text{if } Trust_{m,n} = 0 \end{cases} \tag{8}$$

Here, NN(m) represents the set of nearest neighbors in relation to user m based on the users' trust network.

3.2. The Item-based Trust-Aware Multi-Criteria Collaborative Filtering Component

This component utilizes the trust relationships between items in the items' trust network, along with each item's global trust score, to generate item-based trust-aware recommendations. The component achieves this in three key steps:

Step 1: Calculate trust scores between items

Building on the rationale used in the previous component, this approach employs two complementary metrics: (1) MSD, which compares prediction errors of co-rated users, and (2) BC, which leverages the entire distribution of ratings for both items. To account for the confidence in similarity assessment, an asymmetry trust factor (TF) is used again, based on the proportion of co-rated users. The trust score is then calculated by combining the MSD and BC metrics using a sigmoid function, weighted by an asymmetry factor.

$$\mathrm{Sim}_{d,f}^{MSD} = 1 - \left(\frac{\sum_{u=1}^{t} (P_{u,d} - U^u(d))^2}{t} \right) \tag{9}$$

where $P_{u,d}$ and $U^u(d)$ denote the predicted rating and total utility of item *d* with respect to user *u*, respectively. *t* is the total number of users who have commonly rated items *d* and *f*. MSD values are normalized to a range of 0 to 1 using max-min normalization to ensure comparability.

To address the limitation of MSD, we employ the BC as a complementary similarity measure that utilizes the entire distribution of recorded ratings for both items to compute their similarity, as shown by (5) and (6).

Similar to the previous component, the asymmetry TF emphasizes the importance of considering the number of users who have commonly rated both items when measuring their level of trust is utilized as follows:

$$\mathrm{TF}_{d,f} = \frac{1}{1 + \exp\left(- \frac{|U_d \cap U_f|}{|U_d|} \right)} \tag{10}$$

where $|U_d \cap U_f|$ is the number of users who co-rated items *d* and *f*, and $|U_d|$ is the total number of users rated item *d*. The overall trust between items d and f is then obtained by combining the similarity metrics using a sigmoid function, weighted by the asymmetry trust factor, as follows:

$$\mathrm{Trust}_{d,f} = \frac{1}{(1 + \exp(-(\mathrm{Sim}_{d,f}^{MSD} + \mathrm{BC}(d,f))))} \times \mathrm{TF}_{d,f} \tag{11}$$

Step 2: Calculate global trust scores for items

To enhance the accuracy of predictions in a recommender system, especially for sparsely rated items, we introduce a global trust score for each item. The global trust score for an item is determined by two factors: the count of connected items and the average rating deviation. The count of connected items is based on the item-item trust network, with a higher count indicating its overall level of engagement. The average rating deviation measures the difference between the target item's rating and the average rating given by users. A higher connectivity and smaller deviation contribute to a higher global trust score for an item, as defined below:

ISSN 2723-6471

$$\mathrm{GTS_d} = \exp\left(-\frac{\sum_{m\in U_d}\left|r_{m,d}-\bar{r}_m\right|}{|U_d|}\right)\times\sqrt{\frac{|I_d|}{|I|}} \tag{16}$$

Where $\bar{r}_m$ represents the average rating given by user m across all items he has rated, and $|U_d|$ represents the total number of users who have rated item d. $|I_d|$ represents the number of items that are connected to item d within the items' trust network, while $|I|$ represents the total number of items in the entire dataset.

Step 3: Calculate the item-based trust-aware predicted ratings

This step involves predicting the rating an active user would give to a target item. This prediction leverages the direct and global trust scores of items. To account for potential biases in items' ratings, we employ a mean-centering approach [27]. This approach is explained in the following formula:

$$P^{Item}_{m,d} = \begin{cases} \bar{r}_d + \dfrac{\sum_{f\in NN(d)} Trust_{d,f}\times(r_{m,d}-\bar{r}_f)}{\sum_{f\in NN(d)} Trust_{d,f}}; & \text{if } Trust_{d,f}\neq 0 \\ \bar{r}_d + \dfrac{\sum_{f\in NN(d)} GTS_f\times(r_{m,d}-\bar{r}_f)}{\sum_{f\in NN(d)} GTS_f}; & \text{if } Trust_{d,f}= 0 \end{cases} \tag{12}$$

Here, NN(d) represents the set of nearest neighbors in relation to item d based on the items' trust network.

3.3 The Hybrid Prediction Component

Inspired by the effectiveness of combining various recommendation techniques [28], this component utilizes a switch hybridization scheme. This scheme dynamically selects the most appropriate recommendation approach based on the current context, with the ultimate goal of enhancing recommendation accuracy and coverage. The core principle behind the selection process is the ability to predict ratings for unseen items. If both candidate approaches can predict ratings for unseen items, the Root Mean Square metric is employed to combine their outputs. This metric offers a valuable advantage by quantifying the level of agreement or disagreement between the two predicted ratings, ultimately informing a more robust recommendation selection.

$$P^{Final}_{m,d} = \begin{cases} 0 & ;\ \text{if } P^{User}_{m,d}=0 \text{ and } P^{Item}_{m,d}=0 \\ P^{User}_{m,d} & ;\ \text{if } P^{User}_{m,d}\neq 0 \text{ and } P^{Item}_{m,d}=0 \\ P^{Item}_{m,d} & ;\ \text{if } P^{User}_{m,d}=0 \text{ and } P^{Item}_{m,d}\neq 0 \\ \sqrt{\dfrac{(P^{User}_{m,d})^2+(P^{Item}_{m,d})^2}{2}} & ;\ \text{if } P^{User}_{m,d}\neq 0 \text{ and } P^{Item}_{m,d}\neq 0 \end{cases} \tag{13}$$

4. Experimental Setup and Evaluation

4.1. Experimental Setup

For the evaluation of our proposed model, we employed two multi-criteria (MC) rating datasets from the healthcare domain: the RateMDs and the WebMD rating datasets. The RateMDs MC dataset was sourced from the ratemds.com healthcare platform, where patients can review doctors using a rating scale from 1 to 5 across four criteria: punctuality, staff, knowledge, and helpfulness. This dataset consists of 31,180 MC ratings contributed by 3,464 patients for 3,118 doctors. The WebMD MC dataset, on the other hand, was gathered from the webmd.com healthcare platform, where patients can review and rate medications on a scale of 1 to 5 based on three criteria: medication effectiveness, ease of use, and satisfaction. This dataset contains 32,054 ratings provided by 2,136 patients on a collection of 845 medications. Notably, our datasets are clean and do not require any preprocessing, as they do not contain any missing values or inconsistencies.

To provide a comprehensive assessment of our model's performance, we employed three widely adopted metrics in recommender systems: mean absolute error (MAE), root mean square error (RMSE), and Coverage. MAE is a prevalent measure for measuring prediction accuracy, computed by taking the absolute difference between the predicted rating and the actual rating for each user-item pair and then averaging these differences across all the pairs. RMSE is calculated as the square root of the average of the squared differences between predicted and actual ratings. This makes RMSE more sensitive to large errors, which can be beneficial in recommender systems where large prediction errors may be more detrimental to the user experience. The lower the MAE and RMSE, the better the accuracy of the recommender system. Coverage is another crucial metric, especially in the context of data sparsity. It evaluates a recommendation method's capacity to generate predictions for a wide array of items, including those that are unrated or new. Specifically, it measures the proportion of unrated items that the recommendation method can effectively recommend. Higher coverage implies that the recommendation model can suggest a wider range of items within the collected data, even encompassing new or unrated ones, thereby enabling more personalized recommendations and mitigating the impact of data sparsity [29].

In our evaluation, we compared our model's performance against two benchmark approaches: the MC trust-based collaborative filtering (MC-TCF) approach [30] and the MC user-item trust-enhanced collaborative filtering (MCUITeCF) approach [31]. The MC-TCF approach aims to enhance predictive accuracy and address data sparsity and cold-start user issues by leveraging multi-criteria ratings and inferred trust relationships among users. The MCUITeCF approach integrates multi-criteria ratings and user-item trust relationships to improve recommendation quality while tackling data sparsity and cold-start problems.

4.2. Experimental Results

This section provides a comprehensive comparative analysis of the results obtained through a series of experiments. These experiments were designed to evaluate the performance of the proposed model in comparison to benchmark methods in terms of both prediction accuracy and coverage.

4.2.1. Performance Evaluation on the RateMDs Dataset

Figure 1 displays the MAE performance of the TAMCCF model on the RateMDs dataset. The figure shows nearest neighbor sizes (5-70) on the x-axis and MAE values on the y-axis. TAMCCF consistently outperforms the MC-TCF and MCUITeCF benchmarks across all neighbor sizes, with the most significant improvements at smaller sizes - crucial for sparse data scenarios. TAMCCF achieves an average MAE of 0.44, compared to 0.48 for MCUITeCF and 0.75 for MC-TCF, representing improvements of 7.14% and 41.07% respectively. These results demonstrate TAMCCF's superior recommendation accuracy in terms of MAE, particularly in challenging data conditions.

Figure 2 presents the RMSE results on the RateMDs dataset. The x-axis represents the size of the nearest neighbors from 5 to 70, while the y-axis shows the RMSE values. Similar to MAE, TAMCCF outperforms the benchmarks, achieving an RMSE of 0.89 for a neighbor size of 5, compared to 1.03 for MCUITeCF and 1.30 for MC-TCF. The percentage improvements, on average, are 6.89% over MCUITeCF and 31.14% over MC-TCF. These results further confirm TAMCCF's superior predictive accuracy, particularly in scenarios with limited neighbor data.

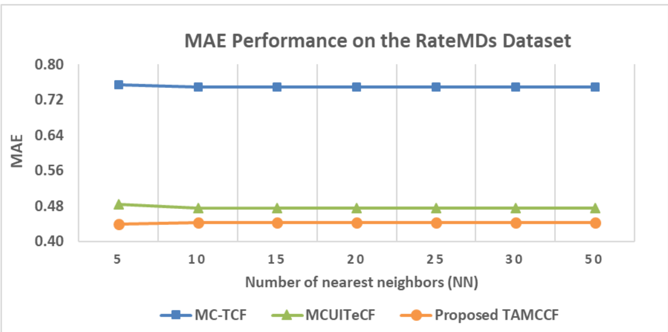

Figure 1. The MAE performance on the RateMDs dataset with different sizes of nearest neighbors.

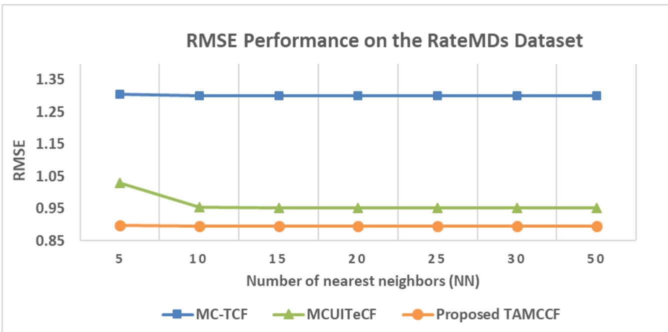

Figure 2. The RMSE performance on the RateMDs dataset with different sizes of nearest neighbors.

ISSN 2723-6471

A detailed analysis of the MAE and RMSE results on the RateMDs dataset reveals that the TAMCCF model's superior performance can be attributed to its incorporation of multi-criteria ratings and trust relationships. The reduction in MAE and RMSE across varying neighbor sizes further confirms the reliability of TAMCCF in providing precise recommendations even with smaller neighbor sizes. TAMCCF's use of trust relationships among users and between items uncovers latent preferences and makes more accurate predictions, even for users or items with limited interaction history. Additionally, the incorporation of multi-criteria ratings provides deep insights into user preferences, resulting in more accurate and personalized recommendations. Consequently, TAMCCF proves to be an effective tool for personalized doctor recommendations, with the potential to significantly enhance user satisfaction and decision-making in healthcare contexts, highlighting its practical importance.

4.2.2. Performance Evaluation on the WebMD Dataset

Figure 3 presents the MAE performance of TAMCCF on the WebMD dataset, comparing it to MC-TCF and MCUITeCF benchmarks. The x-axis shows nearest neighbor sizes (5-70), while the y-axis displays MAE values. The results demonstrate significant improvements in recommendation accuracy over the benchmark methods as indicated by lower MAE values. This highlights the effectiveness of the TAMCCF model in providing more accurate recommendations across all neighbor sizes, with the most significant improvements at smaller sizes. On average, the TAMCCF model achieved an average MAE of 1.08, compared to 1.12 for MCUITeCF and 1.16 for MC-TCF, reflecting improvements of 4.04% and 7.58%, respectively.

Figure 4 illustrates the RMSE performance of TAMCCF on the WebMD dataset, comparing it with MC-TCF and MCUITeCF benchmarks. The figure plots nearest neighbor sizes (5-70) on the x-axis against RMSE values on the y-axis. The RMSE for TAMCCF at a neighbor size of 5 is 1.45, compared to 1.47 for MCUITeCF and 1.51 for MC-TCF. On average, the percentage improvements are 0.86% over MCUITeCF and 3.35% over MC-TCF. TAMCCF maintains lower RMSE scores across all neighbor sizes, demonstrating its consistent superiority.

A thorough analysis of the MAE and RMSE results on the WebMD dataset underscores the TAMCCF model's robustness in harnessing the power of trust relationships and multi-criteria ratings to boost predictive accuracy. The consistent reduction in MAE and RMSE values across varying neighbor sizes confirms the model's reliability in delivering precise recommendations, even in scenarios with sparse user-item interaction data. By considering trust relationships among users and between items, the model can uncover latent preferences and make more accurate predictions, even for users or items with limited interaction history, thereby enhancing recommendation reliability. Furthermore, the incorporation of multi-criteria ratings provides granular insights into user preferences, by capturing the diverse aspects of user preferences, resulting in more accurate and personalized recommendations. The TAMCCF model's ability to provide effective personalized medication recommendations has significant implications for enhancing user satisfaction and decision-making in healthcare contexts, with broad implications for improving healthcare outcomes, highlighting its practical importance and potential to drive meaningful impact.

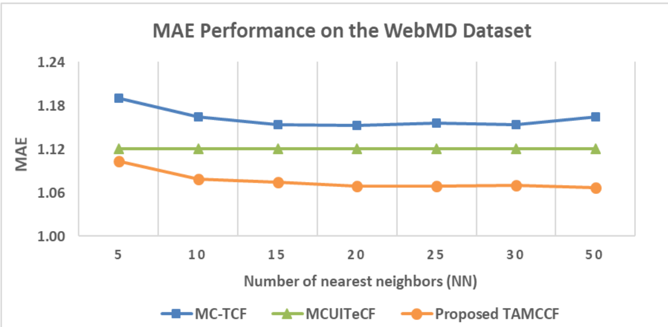

Figure 3. The MAE performance on the WebMD dataset with different sizes of nearest neighbors.

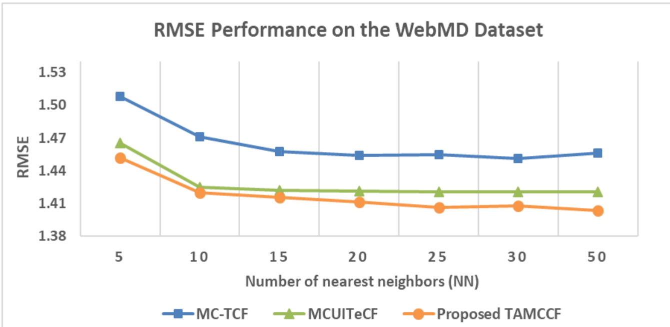

Figure 4. The RMSE performance on the WebMD dataset with different sizes of nearest neighbors.

4.2.3. Performance Evaluation on a Dataset of Different Levels of Sparsity

To assess the viability of our proposed approach in tackling data sparsity, we designed and executed a series of experiments. Six datasets with varying levels of sparsity, ranging from 99.8% to 98%, were created for this purpose. We aimed to thoroughly evaluate how various methods, including our own, performed when confronted with these diverse sparsity levels. This experimental setup allowed us to measure the robustness and effectiveness of each approach under different data conditions, providing insights into their relative strengths and limitations in handling different sparsity scenarios.

Figure 5 and figure 6 illustrate the evaluation of our proposed TAMCCF model on a dataset with varying levels of sparsity. The TAMCCF model demonstrates its superior performance in terms of MAE and Coverage compared to the benchmark methods MC-TCF and MCUITeCF. As shown by figure 5, the TAMCCF model consistently achieved lower MAE across all sparsity levels, with the most significant improvement observed at 99.80% sparsity, where it recorded an MAE of 1.32, compared to 2.73 for MC-TCF and 1.45 for MCUITeCF. On average, TAMCCF reduced MAE by 32.58% relative to MC-TCF and by 4.28% relative to MCUITeCF. In terms of coverage, as depicted in Figure 6, the TAMCCF also outperformed the benchmarks, particularly at higher sparsity levels, achieving 89.33% Coverage at 99.80% sparsity, versus 40.55% for MC-TCF and 83.54% for MCUITeCF. Overall, the TAMCCF model improved average Coverage by 10.34% over MC-TCF and by 1.27% over MCUITeCF, underscoring its robustness and effectiveness in delivering accurate and comprehensive recommendations in sparse data scenarios.

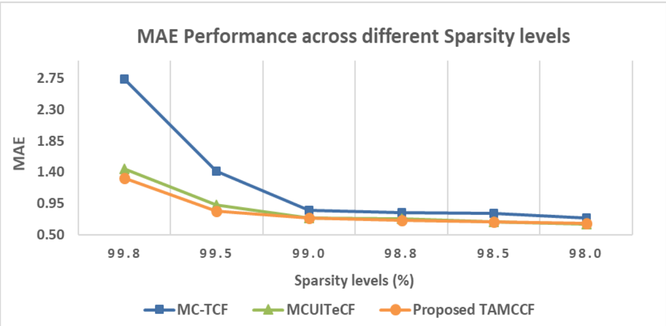

Figure 5. The MAE performance at different levels of sparsity.

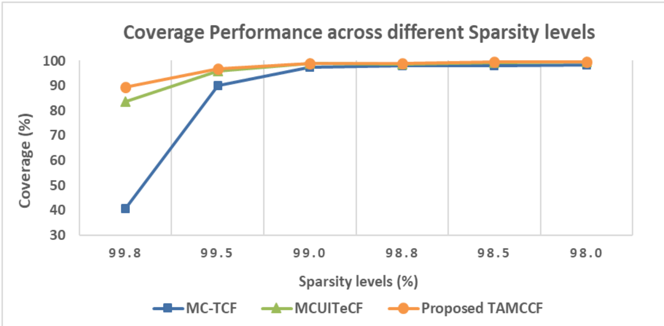

Figure 6. The Prediction coverage performance at different levels of sparsity.

A comprehensive analysis of MAE and Coverage results across datasets with varying sparsity levels demonstrates the TAMCCF model's exceptional robustness. The model effectively leverages trust relationships, global user/item trust scores, and multi-criteria ratings to significantly enhance both predictive accuracy and coverage. TAMCCF consistently reduces MAE and improves Coverage across all sparsity levels, with the most remarkable performance observed at 99.80% sparsity level. This underscores the model's reliability in delivering precise recommendations even in extremely sparse user-item interaction scenarios, a common challenge in real-world applications. By incorporating trust relationships among users and between items, along with global user/item trust scores, TAMCCF expands its effective neighborhood. This enables accurate predictions even for users or items with minimal interaction history, addressing the sparsity problem prevalent in many recommendation systems. Moreover, the integration of multi-criteria ratings provides a granular understanding of user preferences, leading to more accurate and personalized recommendations that reflect the complexity of real-world decision-making processes.

5. Conclusion

The transformative role of deep learning techniques in revolutionizing lower limb prosthetics for improved activity recognition was explored. In this study, the employed approach involved the utilization of an optimized deep learning technique model to enhance the recognition of lower limb activities. The suggested methodology encompassed three key steps. The images that were gathered underwent pre-processing through the application of an enhanced wavelet denoising technique as well as Empirical mode decomposition. Subsequent to the pre-processing stage, features were extracted from the processed data by means of an enhanced sliding window approach along with a time domain feature.

The extracted features were subsequently employed in the process of feature classification. The classification of the features was accomplished utilizing an Optimized LSTM model. The optimization of the LSTM model was achieved through the utilization of the black window Optimization algorithm. By calculating metrics like accuracy, precision, recall, F-score, specificity, sensitivity, MCC, NPV, FPR, FNR, the Optimized LSTM achieved higher percentages of accuracy (98.65%), precision (98.71%), sensitivity (98.70%), specificity (98.8%), and F-score (98.48%). Sustainable development principles offer a promising framework for advancing prosthetic leg innovation and improving the lives of amputees globally. Implementing advanced prosthetic technologies, such as those integrating microprocessors, neural interfaces, and advanced materials, comes with several potential challenges. By addressing the challenges through concerted efforts from researchers, healthcare providers, policymakers, and manufacturers, the benefits of advanced prosthetic technologies can be more widely realized, improving the quality of life for individuals with limb loss. The exploring multi-modal sensor fusion, adapting models through transfer learning, integrating advanced sensors, implementing edge computing, optimizing for diverse user populations, and addressing ethical considerations for user-centric lower limb prosthetics activity recognition using deep learning.

6. Declarations

6.1. Author Contributions

Conceptualization: Q.Y.S., M.M.A., A.A.-S., A.H.H., and Q.M.K.; Methodology: Q.Y.S.; Software: M.M.A. and A.A.-S.; Validation: Q.Y.S. and M.M.A.; Formal Analysis: Q.Y.S. and M.M.A.; Investigation: Q.Y.S.; Resources: A.H.H., and Q.M.K.; Data Curation: A.A.-S.; Writing Original Draft Preparation: Q.Y.S. and M.M.A.; Writing Review and Editing: M.M.A., Q.Y.S., A.A.-S., A.H.H., and Q.M.K.; Visualization: A.H.H., and Q.M.K.; All authors have read and agreed to the published version of the manuscript.

6.2. Data Availability Statement

The data presented in this study are available on request from the corresponding author.

6.3. Funding

The authors received financial support for the publication of this article from Al-Ahliyya Amman University.

6.4. Institutional Review Board Statement

Not applicable.

6.5. Informed Consent Statement

Not applicable.

6.6. Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] 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.

[2] 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, no. Jan, pp. 1-13, 2023, doi: 10.2196/38184.

[3] J. Lu, Q. Shambour, and G. Zhang, “Recommendation technique-based government-to-business personalized e-services,” *in Proc. 2009 Annual Meeting of the North American Fuzzy Information Processing Society, Cincinnati, OH, USA*, vol. 2009, no. July, pp. 1-6, doi: 10.1109/NAFIPS.2009.5156456.

[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, "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] M. Kolhar, F. Al-Turjman, A. Alameen, and M. M. Abualhaj, "A Three Layered Decentralized IoT Biometric Architecture for City Lockdown During COVID-19 Outbreak," *IEEE Access,* vol. 8, no. Sep., pp. 163608-163617, 2020, doi: 10.1109/ACCESS.2020.3021983.

[8] Q. Shambour, N. Turab, and O. Adwan, "An Effective e-Commerce Recommender System Based on Trust and Semantic Information," *Cybernetics and Information Technologies,* vol. 21, no. 1, pp. 103-118, 2021, doi: 10.2478/cait-2021-0008.

[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,* vol. 2011, no. July, pp. 44-51, doi: 10.1109/SMDCM.2011.5949274.

[11] Q. Y. Shambour, M. M. Abu-Alhaj, and M. M. Al-Tahrawi, "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] C. R. Nie Hui, "Online Doctor Recommendation System with Attention Mechanism," *Data Analysis and Knowledge Discovery,* vol. 7, no. 8, pp. 138-148, 2023-08-25, 2023, doi: 10.11925/infotech.2096-3467.2022.0761.

[14] A. Kulshrestha, V. Krishnaswamy, and M. Sharma, "A deep learning model for online doctor rating prediction," *Journal of Forecasting,* vol. 42, no. 5, pp. 1245-1260, 2023, doi: 10.1002/for.2953.

[15] P. Keikhosrokiani, K. Balasubramaniam, and M. Isomursu, "Drug Recommendation System for Healthcare Professionals' Decision-Making Using Opinion Mining and Machine Learning," Digital Health and Wireless Solutions, Digital Health and Wireless Solutions M. Särestöniemi, P. Keikhosrokiani, D. Singh et al., eds., pp. 222-241, Cham: Springer Nature Switzerland, 2024, doi: 10.1007/978-3-031-59091-7_15.

[16] S. G. Begum, and P. K. Sree, "Drug Recommendations Using a Reviews and Sentiment Analysis" by a Recurrent Neural Network," *Indonesian Journal of Multidisciplinary Science,* vol. 2, no. 9, pp. 3085-3094, 2023, doi: 10.55324/ijoms.v2i9.530.

[17] P. Symeonidis, L. Bellinazzi, C. Berbague, and M. Zanker, "Safe and Effective Recommendation of Drug Combinations based on Matrix Co-Factorization," *in Proc. 2023 IEEE 36th International Symposium on Computer-Based Medical Systems (CBMS), L'Aquila, Italy,* vol. 2023, no. July, pp. 634-639, doi: 10.1109/CBMS58004.2023.00292.

[18] P. Symeonidis, G. Manitaras, and M. Zanker, "Accurate and Safe Drug Recommendations based on Singular Value Decomposition," *in Proc. 2023 IEEE 36th International Symposium on Computer-Based Medical Systems (CBMS), L'Aquila, Italy,* vol. 2023, no. July, pp. 163-168, doi: 10.1109/CBMS58004.2023.00210.

[19] 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.

[20] V. Mani, and S. Thilagamani, "Hybrid Filtering-based Physician Recommender Systems using Fuzzy Analytic Hierarchy Process and User Ratings," *International Journal of Computers Communications and Control,* vol. 18, no. 6, pp. 1-17, 2023, doi: 10.15837/ijccc.2023.6.5086.

[21] Y. Sofia, F. Renaldi, and I. Santikarama, "Recommendation System for Matching Patients and Doctors in Telemedicine Based on Hybrid Filtering," *in Proc. Proceedings of the 3rd Asia Pacific International Conference on Industrial Engineering and Operations Management, Johor Bahru, Malaysia,* vol. 2022, no. Sep., pp. 420-431, doi: 10.46254/AP03.20220068.

[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] B. K. Patra, R. Launonen, V. Ollikainen, and S. Nandi, "A new similarity measure using Bhattacharyya coefficient for collaborative filtering in sparse data," *Knowledge-Based Systems*, vol. 82, no. 2015, pp. 163-177, 2015, doi: 10.1016/j.knosys.2015.03.001

[24] Y. Wang, J. Deng, J. Gao, and P. Zhang, "A hybrid user similarity model for collaborative filtering," *Information Sciences*, vol. 418-419, no. 12, pp. 102-118, 2017, doi: 10.1016/j.ins.2017.08.008.

[25] J. Feng, X. Feng, N. Zhang, and J. Peng, "An improved collaborative filtering method based on similarity," *PLOS ONE*, vol. 13, no. 9, pp. 1-18, 2018, doi: 10.1371/journal.pone.0204003.

[26] H. Song, Q. Pei, Y. Xiao, Z. Li, and Y. Wang, "A Novel Recommendation Model Based on Trust Relations and Item Ratings in Social Networks," *in Proc. 2017 International Conference on Networking and Network Applications (NaNA), Kathmandu, Nepal,* vol. 2017, no. Jan., pp. 17-23, doi: 10.1109/NaNA.2017.17.

[27] 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, Chapel Hill, North Carolina, United States,* vol. 1994, no. Oct., pp. 175 - 186, doi: 10.1145/192844.192905.

[28] R. Burke, "Hybrid Web Recommender Systems," The Adaptive Web: Methods and Strategies of Web Personalization, P. Brusilovsky, A. Kobsa and W. Nejdl, eds., pp. 377-408, Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, doi: 10.1007/978-3-540-72079-9_12.

[29] C. C. Aggarwal, "Evaluating Recommender Systems," Recommender Systems: The Textbook, C. C. Aggarwal, ed., pp. 225-254, Cham: Springer International Publishing, 2016, doi: 10.1007/978-3-319-29659-3_7.

[30] Q. Shambour, "A user-based multi-criteria recommendation approach for personalized recommendations," *International Journal of Computer Science and Information Security*, vol. 14, no. 12, pp. 657-663, 2016.

[31] A. Hussein, and Q. Shambour, "A Trust-enhanced Recommender System for Patient-Doctor Matchmaking in Online Healthcare Communities," *International Journal of Intelligent Engineering and Systems*, vol. 16, no. 6, pp. 684-694, 2023, doi: 10.22266/ijies2023.1231.57.