An Neural Collaborative Filtering (NCF) based Recommender System for Personalized Rehabilitation Exercises

Yoon-Seop Chang
Intelligent Human Twin Research Center
Electronics and Telecommunications Research Institute
Daejeon, Republic of Korea
ychang76@etri.re.kr

Boosun Jeon
Intelligent Human Twin Research Center
Electronics and Telecommunications Research Institute
Daejeon, Republic of Korea
bsjeon@etri.re.kr

NohSam Park
Intelligent Human Twin Research Center
Electronics and Telecommunications Research Institute
Daejeon, Republic of Korea
siru23@etri.re.kr

Mikyong Han
Intelligent Human Twin Research Center
Electronics and Telecommunications Research Institute
Daejeon, Republic of Korea
mkhan@etri.re.kr

Jae-Chul Kim
Intelligent Human Twin Research Center
Electronics and Telecommunications Research Institute
Daejeon, Republic of Korea
kimjc@etri.re.kr

***Abstract*—Rehabilitation exercises should be prescribed and performed by carefully selecting the type of exercise and intensity for each patient. If rehabilitation exercises are prescribed incorrectly, they may be ineffective or cause dangerous situations such as secondary injuries. In this study, we developed a personalized rehabilitation exercise recommender system based on data from a clinical trial of rehabilitation exercise intervention conducted at Chungnam National University Sejong Hospital. We generated a training dataset from the clinical trial data of rehabilitation exercise intervention for patients with shoulder adhesive capsulitis, rotator cuff injury, and back pain, and developed a rehabilitation exercise recommender system based on NCF's NeuMF model to verify its performance. The results show that the developed rehabilitation exercise recommender model outperforms the methods based on GMF and MLP models only. The personalized rehabilitation exercise recommender system in this study is expected to help rehabilitation patients recover their functions, improve their health, and return to society quickly and safely.**

Keywords—rehabilitation exercise, recommender system, collaborative filtering, neural network, personalization

I. INTRODUCTION

Rehabilitation exercises should be prescribed and performed carefully, taking into account the individual patient's health status, diseases, injured body parts, injury severity, etc. In particular, the type and intensity of exercises should be carefully adjusted for each patient. If exercises are wrongly prescribed to patients, there will be little effect or it may bring adverse effects such as secondary injuries. Therefore, rehabilitation exercises are usually prescribed by professionals such as physical therapists. An enormous amount of health data makes it difficult for staff and patients to make decisions, and it increases the time and cost spent on decision making [1][2].

Healthcare recommender systems help people manage and promote their health by providing personalized recommendation results. These systems provide useful and accurate information for disease severity prediction, disease diagnosis and treatment, health management and promotion, lifestyle improvement, and so on. Machine learning-based recommender systems can be applied to the prescription of rehabilitation exercises.

In this study, a recommender system based on Neural Collaborative Filtering (NCF) was developed to recommend personalized rehabilitation exercises for patients with musculoskeletal disorders. The recommender model of the system was trained and tested using real data from the hospital's clinical trial. As a result, the system has shown a hit ratio (HR) performance of more than 70%. The ensemble model Neural Matrix Factorization (NeuMF) has shown better performance than the case of Generalized Matrix Factorization (GMF) or Multi-Layer Perceptron (MLP) model only.

II. DATASETS

A. Rehabilitation Clinical Trial Data

The raw data used to generate the model training dataset were collected from a clinical trial in which rehabilitation exercises were performed on patients with musculoskeletal disorders. The clinical trial was conducted by the Department of Physical and Rehabilitation Medicine of Chungnam National University Sejong Hospital. The clinical trial, in detail, was conducted for adults aged 19 to 65 years with low back pain or shoulder disorder including adhesive capsulitis or rotator cuff injury.

Patients with shoulder disorder underwent rehabilitation exercises 4 times a week for 4 weeks in the hospital and once a week for 8 weeks in the gym. Aerobic, muscular, and flexibility exercises for the shoulders, elbows, hands, and ankles of the upper extremity were prescribed and performed for these patients. Patients with low back pain did rehabilitation exercises 4 times a week in the hospital for 4 weeks and once a week in the gym for 4 weeks. Aerobic, muscle and flexibility, balance, and gait exercises for the neck and trunk were prescribed and performed for these patients.

TABLE I. A SUMMARY OF REHABILITATION CLINICAL DATA

	Shoulder disease	Low back pain	Total
Date	2022.04.18~2022.11.25	2022.05.02~2022.11.25	2022.04.18~2022.11.25
Patients	13	33	46
Exercise Programs	111	132	243
Exercises (pool)	125	103	204
Exercises (execution)	880	1,063	1,943

TABLE II. AN EXAMPLE OF DATASET FOR MODEL TRAINING

Disease code	Patient ID	Stage	Num	Exercise ID	Date
1	10101	1	1	1662603213	2022-04-18
1	10101	1	2	1662603200	2022-04-18
1	10101	2	1	1662603120	2022-04-18
1	10101	2	2	1662603109	2022-04-18
...					
3	22601	2	5	1662603022	2022-08-24
3	22601	3	1	1662602886	2022-08-24

Rehabilitation exercises were prescribed to the patients in the form of exercise programs consisting of several individual exercises. Exercise programs consist of three stages - warm-up, main stage and cool-down. Individual exercises are added to one of these three stages and set up with attributes such as the number of repetitions, number of sets, pause time between sets, and load of the exercise, etc.

Each exercise includes information such as its ID, name, description, exercise part, exercise part details, exercise type, functional movement type, posture, use of equipment, level of difficulty, etc. Each exercise program includes information such as its ID, name, description, disease code, exercise part details, exercise purpose, target age groups, exercise places, use of tools and machines, and sets of individual exercises divided into warm-up, main and cool-down stages.

Clinical trial data collected from April 18, 2022 through November 25, 2022 were used to generate the dataset for model training. Table 1 provides a summary of this clinical trial data. Data collection began on April 18, 2022 for patients with shoulder disorders and on May 2, 2022 for patients with low back pain. The number of patients enrolled in the clinical trial for each condition is 13 and 33, respectively. The number of exercise programs is 111 and 132 for each condition. 125 and 103 individual exercises were provided as exercise pools for exercise program prescription for each condition, and a total of 204 exercises, excluding duplicates, were used for the rehabilitation clinical trial. Exercises from these pools were prescribed to patients 880 times for shoulder disorders and 1,063 times for low back pain. A total of 1,943 prescription histories were used to generate the dataset for model training.

B. Dataset Generation for Model Training

The above clinical rehabilitation trial data were stored in the AI-based rehabilitation coaching service system of this study as divided into DB tables for patients' health status data, exercise and exercise program attribute data, each patient's prescription history, and prescribed exercise program performance results, respectively. Datasets for model training were generated from these DB tables and updated periodically.

The recommender system and its dataset were designed to recommend individual exercises, not to recommend exercise programs in this study. This was because the overlap between the exercise programs prescribed to each patient was insignificant according to the prescription history, while the overlap or similarity between the users' interactions with items affects the performance of collaborative filtering algorithms. In fact, almost all different exercise programs were prescribed to patients each time. In the case of individual exercises, on the other hand, 204 exercises form a kind of pool, and these exercises were prescribed to patients 1,943 times, as shown in Table 1. There was overlap and similarity between the user interactions, which made it possible to recommend personalized exercises based on collaborative filtering.

The dataset for model training was generated from the whole data of prescribed and performed individual exercises as mentioned above. The dataset was generated with a focus on user-exercise interactions, i.e., the prescription and execution of exercises by users. Table 2 shows an example of the generated dataset for model training. It includes disease code, patient ID, exercise stage, sequential number in each stage, exercise ID, execution date.

Patient IDs were processed through an additional conversion procedure during dataset generation. The reason for this is that users may have a different health status at any given time. This may affect the prescription of exercises, and different appropriate exercises should be prescribed to patients considering their different health status. Therefore, in this study, each health status for one user was treated as the concepts of different users, and each health status even for one user, i.e., for each prescription, was assigned new user IDs. For example, patients with shoulder disorders are assigned original user IDs such as 'RS01', 'RS02', '...' and patients with low back pain are assigned 'RB01', 'RB02', '...'. These IDs have been converted to '101', '102', '...' and '201', '202', '...' by replacing 'RB' and 'RS' with '1' and '2' respectively. The final new user IDs are then converted to '10101', '10102', '...' by appending the sequential number '01', '02', '...' for each prescription. Table 2 contains these new IDs as patient IDs. This conversion may also have the effect of increasing the size of the user dimension in the user-item interaction matrix.

Table 3 shows the summary of patient-exercise interaction data, i.e., user-item interaction data, in the generated dataset for each disease. The data for shoulder disorders contained 870 interactions for 109 patients and 125 exercises, and the data for low back pain contained 1,058 interactions for 130 patients and 103 exercises. There was a total of 1,928 interactions for 239 patients and 204 exercises.

TABLE III. SUMMARY OF USER-ITEM INTERACTION DATA

	Patients	Exercises	Interactions
Shoulder disease	109	125	870
Low back pain	130	103	1,058
Total	239	204	1,928

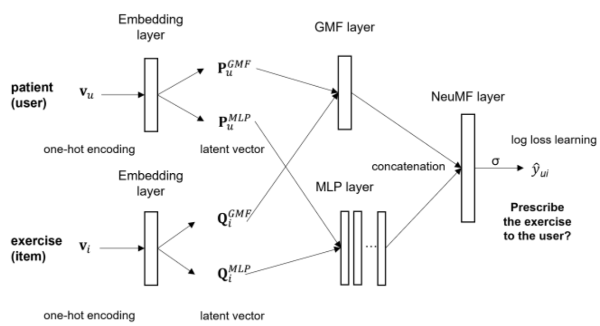

Fig 1. Overview of the rehabilitation exercises recommender model [3]

III. SYSTEM DESIGN

A. Overview of the Recommender Model

The goal of this study was to develop a system that learns the history of rehabilitation exercises prescribed to all patients and recommends the top K new exercises for a given patient that the patient has not yet performed but is appropriate to perform. Patients, exercises, and prescriptions correspond to users, items, and user-item interactions in collaborative filtering, respectively. The different health statuses of a given patient at each prescription time are treated as the concepts of several different users. In this way, appropriate exercises are recommended to patients according to their changing health status.

Fig. 1 shows the overview and flows of the rehabilitation exercises recommender model based on the NCF framework[3][4]. One-hot encoded vectors v_u, v_i each for user and item, i.e., patient and exercise, respectively, are input to the embedding layers, and latent vectors P_u^{GMF}, P_u^{MLP} and Q_i^{GMF}, Q_i^{MLP} each for user and item, respectively, are randomly initialized. Then, embedding layers are trained and latent vectors are learned for all users and items during the training of the whole model. Then, the resulting latent vectors P_u^{GMF} and Q_i^{GMF} are input to the GMF layer, and P_u^{MLP}, Q_i^{MLP} are input to the MLP layer. Finally, the output vectors of the GMF and MLP layers are input into the NeuMF layer, and these vectors are concatenated, and the prediction score $\hat{y}_{ui}$ is calculated by applying the sigmoid (σ) function to the concatenated vector. The ground truth y_{ui} is an implicit interaction, and it represents whether each exercise has been previously prescribed to the given user. The recommender model is trained until the log loss between y_{ui} *and* $\hat{y}_{ui}$ is minimized. Each recommended exercise is decided to be prescribed or not to be prescribed to each user according to the value of $\hat{y}_{ui}$.

B. Rehabilitation Exercise Recommender System

The rehabilitation exercise recommender system of this study was developed based on NCF in order to improve and solve the problems of the matrix factorization (MF) model, which has been so often adopted in collaborative filtering recommender systems. The nonlinear model using multi-layer perceptron is first presented, and then the ensemble model, which adopts matrix factorization along with it, is presented to explain the system.

The input layer of the recommender system consists with one-hot encoded vectors $\mathbf{v}_u^U$ and $\mathbf{v}_i^I$ for each patient and exercise ID. These sparse vectors are converted into denser vectors by the following embedding layers. $\mathbf{v}_u^U$ and $\mathbf{v}_i^I$, in detail, are converted into $\mathbf{P}^T\mathbf{v}_u^U$ and $\mathbf{Q}^T\mathbf{v}_i^I$ by operations with latent factor matrices $\mathbf{P} \in \mathbb{R}^{M\times K}$ and $\mathbf{Q} \in \mathbb{R}^{N\times K}$ in the embedding layers each for patient and exercise, respectively. When these embedded latent vectors are input to the MLP layer, and the prediction score $\hat{y}_{ui}$ is calculated as in (1) [4]. $\emptyset_{out}$ and $\emptyset_x$ in (1) are mapping functions each for the output layer and x^{th} hidden layer of the MLP, respectively. The value of this prediction score is evaluated to determine whether or not to prescribe the exercise to the given user.

$$\begin{aligned}\hat{y}_{ui} &= f(\mathbf{P}^T\mathbf{v}_u^U, \mathbf{Q}^T\mathbf{v}_i^I) \\ &= \emptyset_{out}(\emptyset_L(\ldots\emptyset_2(\emptyset_1(\mathbf{P}^T\mathbf{v}_u^U, \mathbf{Q}^T\mathbf{v}_i^I))\ldots))\end{aligned} \tag{1}$$

The detailed structure of the above MLP layers shows their nonlinear characteristics as in (2). Rectified Linear Unit (RELU) and Sigmoid functions are applied as the activation function a_x each for the x^{th} hidden layer and the output layer, respectively [4]. These nonlinear characteristics solve the limit derived from the linear characteristics of the MF model, and can more effectively learn the complicated user-item interaction functions.

$$\begin{gathered}\mathbf{z}_1 = \emptyset_1(\mathbf{p}_u, \mathbf{q}_i) = \begin{bmatrix}\mathbf{p}_u \\ \mathbf{q}_i\end{bmatrix}, \\ \emptyset_2(\mathbf{z}_1) = a_2(\mathbf{W}_2^T\mathbf{z}_1 + \mathbf{b}_2), \\ \ldots \\ \emptyset_L(\mathbf{z}_{L-1}) = a_L(\mathbf{W}_L^T\mathbf{z}_{L-1} + \mathbf{b}_L), \\ \hat{y}_{ui} = \sigma(\mathbf{h}^T\emptyset_L(\mathbf{Z}_{L-1}))\end{gathered} \tag{2}$$

The inner product equation of the MF model in (3) can be represented as (4) by transforming the above equation in (2). ⊙ is an element-wise product, and (4) becomes identical to the inner product of the MF model if the activation function a_{out} is assumed to be an identity function and all elements of the weighting vector h are assumed to be 1. The NCF model can thus reproduce the MF model, and this MF model is called the GMF model and is used together with the MLP model to implement an ensemble model NeuMF [4].

$$\hat{y}_{ui} = f(u, i|\mathbf{p}_{u,}\mathbf{q}_i) = \mathbf{p}_u^T\mathbf{q}_i = \sum_{k=1}^{K} p_{uk}\, q_{ik} \tag{3}$$

$$\begin{gathered}\emptyset_1(\mathbf{p}_u, \mathbf{q}_i) = \mathbf{p}_u \odot \mathbf{q}_i, \\ \hat{y}_{ui} = a_{out}(\mathbf{h}^T(\mathbf{p}_u \odot \mathbf{q}_i))\end{gathered} \tag{4}$$

Performance improvement can be expected by varying the embedding vectors input to the GMF and MLP models each other. Equation (5) shows this procedure and the integration of two models together. $\mathbf{p}_u^G$ and $\mathbf{q}_i^G$ are embedding vectors for input to the GMF layer, and $\mathbf{p}_u^M$ and $\mathbf{q}_i^M$ are embedding vectors for input to the MLP layer. $\emptyset^{GMF}$ and $\emptyset^{MLP}$ are output vectors from the GMF and MLP layers, and these vectors are input to the NeuMF layer to compute the prediction score $\hat{y}_{ui}$. The rehabilitation exercise recommender system of this study was developed based on such an ensemble model.

$$\emptyset^{GMF} = \mathbf{p}_u^G \odot \mathbf{q}_i^G,$$
$$\emptyset^{MLP} = a_L(\mathbf{W}_L^T(a_{L-1}(\dots a_2(\mathbf{W}_2^T \begin{bmatrix} \mathbf{p}_u^M \\ \mathbf{q}_i^M \end{bmatrix} + \mathbf{b}_2)\dots)) + \mathbf{b}_L), \quad (5)$$
$$\hat{y}_{ui} = \sigma(\mathbf{h}^T \begin{bmatrix} \emptyset^{GMF} \\ \emptyset^{MLP} \end{bmatrix})$$

Hyperparameter initialization is very important during the training of the NeuMF model. The GMF and MLP models were previously trained and optimized, and the resulting parameters are used as initial parameters for training the NeuMF model. Parameters from two previous models are set to be equally weighted.

IV. Experiments and Results

A. Experimental Environment

Experiments of the developed recommender system were conducted on the clinical trial data of rehabilitation exercises prescribed to patients with musculoskeletal disorders. Experiments were conducted separately for two types of diseases: shoulder disorders and low back pain. There were 870 interaction data in the 109 x 125 patient-exercise interaction matrix for the case of shoulder disorders, and 1,058 interaction data in the 130 x 103 interaction matrix for the case of low back pain.

The rehabilitation exercises prescribed to each user were separated into the training and test datasets, and stratified k-fold cross-validation was performed on the experiments. More specifically, the training and test datasets were separated to 75% : 25%, and stratified 4-fold cross-validation was performed. The NeuMF ensemble model, which integrates the GMF and MLP, was trained, and the resulting performance was compared with that of the GMF and MLP models.

B. Hyperparameter Tuning

Hyperparameter tuning and model optimization were performed during the training of the recommender model. Hyperparameter tuning was performed for the number of latent factors, negative samples, and hidden layers of the model. Hit Ratio (HR) and Normalized Discounted Cumulative Gain (NDCG) were observed as performance indices during the optimization.

The performance of the model was observed while varying the number of latent factors such as [4, 8, 16, 32, 64]. The performance improvement of the model decreased after the number of latent factors 16 for the case of shoulder disorders as shown in Fig. 2. The performance improvement of the model was insignificant after the number of latent factors 16 for the case of low back pain as shown in Fig. 3. After that, the number of latent factors of the recommender model was set to 16 for both cases of two diseases.

Hyperparameter tuning and model optimization were also performed for the number of negative samples [2, 4, 6, 8, 10, 12, 14] and for the number of hidden layers [3, 4, 5, 6]. As a result, the number of negative samples and the number of hidden layers were set to 8 and 5, respectively.

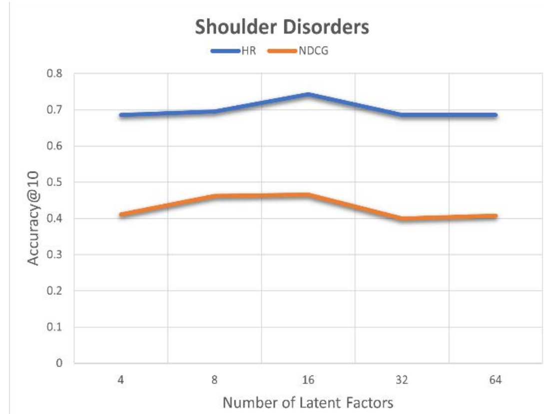

Fig 2. Performance of the model according to the variation of the number of latent factors (for shoulder disorders)

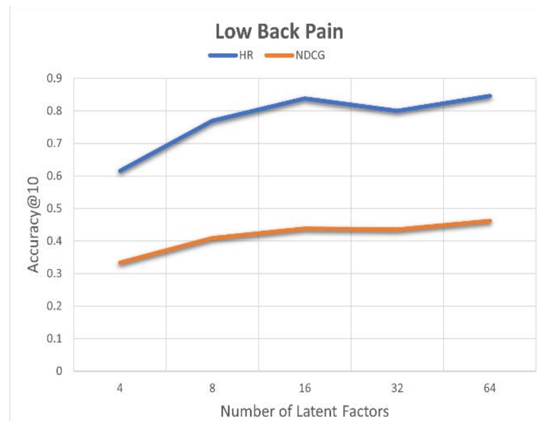

Fig 3. Performance of the model according to the variation of the number of latent factors (for low back pain)

C. Training the recommender model

Fig. 4 shows the result of the training of the rehabilitation exercise recommender model for patients with shoulder disorders. The performance of the model, i.e., HR and NDCG at k = 10, was converged after about 21 epochs. The maximum value of HR was 0.736, and the value of NDCG was 0.465 at that time.

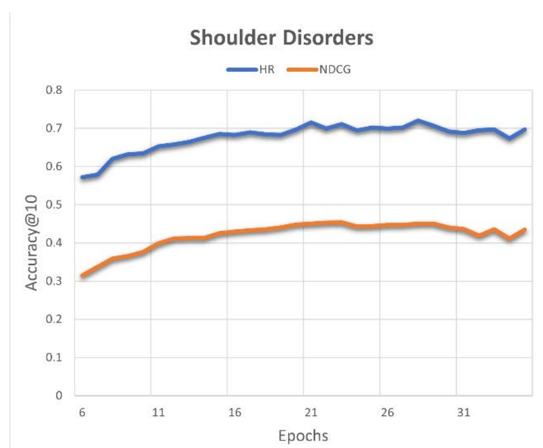

Fig 4. The training of the recommender model for shoulder disorders

TABLE IV. PERFORMANCE OF RECOMMENDER MODELS FOR SHOULDER DISORDERS

	K	HR	NDCG
NeuMF	5	**0.602**	**0.420**
	7	**0.657**	0.424
	10	**0.736**	**0.465**
GMF	5	0.423	0.268
	7	0.474	0.281
	10	0.574	0.315
MLP	5	0.593	0.413
	7	0.636	**0.430**
	10	0.701	0.453

TABLE V. PERFORMANCE OF RECOMMENDER MODELS FOR LOW BACK PAIN

	K	HR	NDCG
NeuMF	5	**0.705**	**0.378**
	7	**0.759**	**0.409**
	10	**0.829**	0.419
GMF	5	0.489	0.236
	7	0.534	0.252
	10	0.578	0.270
MLP	5	0.651	0.363
	7	0.722	0.386
	10	0.790	**0.427**

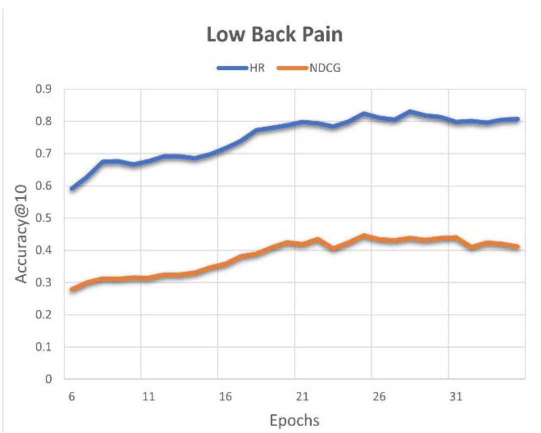

Fig 5. The training of the recommender model for low back pain

Fig. 5 shows the result of the recommender model for low back pain patients. The performance of the model, i.e., HR and NDCG at k = 10, was converged after about 25 epochs. The maximum value of HR was 0.829, and the value of NDCG was 0.419.

D. Comparison of Model Performance

The ensemble model NeuMF showed a better performance than the other two models, i.e. GMF and MLP model only, in the case of shoulder disorders as shown in Table 4. The HR values of the NewMF model were higher than the other models in all cases of k = 5, 7 and 10. The NDCG value at k = 7 showed the highest value in the MLP model, but the difference with the NeuMF model value was not significant. The highest values of HR and NDCG of the NeuMF model were 0.736 and 0.465 at k=10.

The NeuMF model also showed better performance than the other two models in the case of low back pain, as shown in Table 5. The HR values of the NeuMF model were higher than the other models in all cases of k = 5, 7 and 10. The NDCG value at k = 10 showed the highest value in the MLP model, but the difference with the NeuMF model value was also not insignificant. The highest values of HR and NDCG of the NeuMF model were 0.829 and 0.419 at k=10.

With the expansion of the training data, we believe that the performance of the developed recommender model can be gradually improved. Therefore, this study is currently continuing with clinical trials and inviting other hospitals to participate in more large-scale demonstrations. In the future, we will continue to expand the clinical data and plan additional studies and experiments.

V. CONCLUSIONS

In this study, we developed a personalized rehabilitation exercise recommender system based on rehabilitation exercise intervention clinical trial data. We generated a training dataset for the development of a collaborative filtering-based recommender system from the data of rehabilitation exercise intervention clinical trials conducted at Chungnam National University Sejong Hospital for patients with shoulder adhesive capsulitis, rotator cuff injury, and low back pain. Using the generated training dataset, we developed a rehabilitation exercise recommender algorithm based on NCF's NeuMF model and verified its performance. We compared the developed rehabilitation exercise recommender algorithm with methods based on GMF and MLP models and confirmed that it has better performance. It is expected that future studies such as solving the problem of data sparsity problem, using explicit interaction data, and using attribute information of patients and exercises to solve the new user problem, along with the continuous collection of additional clinical data, will help rehabilitation patients recover their functions and return to society.

Acknowledgment

This research was supported by Culture, Sports and Tourism R&D Program through the Korea Creative Content Agency grant funded by the Ministry of Culture, Sports and Tourism in 2023 (Project Name: Development of Intelligent Personalized Rehabilitation Service Technology, Project Number: SR202104001, Contribution Rate: 100%)

References

[1] B. Stark, C. Knahl, M. Aydin, and K. Elish, "A literature review on medicine recommender systems," Int. J. Advanced Comput. Sci. and Appl., vol. 10, no. 8, pp. 6-13, 2019.

[2] M. Wiesner, and D. Pfeifer, "Health recommender systems: Concepts, requirements, technical basics and challenges," Int. J. Environ. Res. and Public Health, vol. 11, pp. 2580-2607, 2014.

[3] Y.S. Chang, B.S. Jeon, N.S. Park, M.K. Han, and J.C. Kim, "Development of a Rehabilitation Exercise Recommender System Based on Neural Collaborative Filtering (NCF)," J. Korean Institute of Communication and Information Sciences, vol. 47, no. 11, pp. 1843-1858, 2022.

[4] X. He, L. Liao, H. Zhang, L. Nie, X. Hu, and T.S. Chua, "Neural Collaborative Filtering," Proc. 26th International Conference on World Wide Web, pp.173-182, 2017.