Physical Fitness Recommender Framework for Thyroid Patients using Restricted Boltzmann Machines

Vaishali S Vairale\*  Samiksha Shukla

\*CHRIST (Deemed to be University), India

* Corresponding author’s Email: vairale.sheshrao@res.christuniversity.in

Abstract: These days, people can easily acquire the information from online sources. Individuals are generally using recommendation services before buying products considering the availability of online. Recommendation systems propose the relevant services or products to users. But sometimes people face issues while retrieving health related information from the recommender systems. A focus on keeping people healthy is one way to address the serious societal concern of healthcare domain. A health-based physical recommender system suggests workout plans for users using their activity level and health condition. A personalized approach is the most effective solution for the fitness-based recommender framework based on user’s desired characteristics. This article presents a personalized fitness recommender system for thyroid patients. The proposed fitness recommender model integrates the user’s data like personal and health profile, preferences, calorie intake, and activity level. The proposed hybrid model is built using Restricted Boltzmann Machines (RBM) integrating content based and matrix factorization techniques. The results of experiments prove that the proposed hybrid model outperforms content based, pure RBM and matrix factorization recommendation techniques. The current proposal achieves the personalization approach by incorporating user’s thyroid health condition and exercise preferences in recommendation process. The recommended result of hybrid RBM method is revised based on user's new preferences.

Keywords: Fitness recommender system, Collaborative filtering techniques, Content based methods, Probabilistic matrix factorization, Restricted boltzmann machines, Thyroid disorder, Exercise plans.

1 Introduction

Over the past few years, the lifestyle of people is changing significantly due to the civilization and industrialization. People are suffering from various diseases like diabetes, cardiovascular, cancer, thyroid, obesity and other disorders due to the changes in food intake and lifestyle as explained in [1]. It has been observed nowadays that people are sitting at a one place for more time and consume unhealthy foods instead of balanced meal. Consumption of unhealthy food and less or no exercise in a day lead towards the serious societal challenge like obesity. Practicing regular exercises can reduce the impact of diseases like diabetes, thyroid, coronary heart disorders and cure some diseases like obesity as investigated in [2]. Though people are aware about their health condition, they neglect to exercise regularly. Some individual pay attention to keep themselves fit but they cannot find the suitable list of exercises as per their choices. In this context hiring a fitness professional or joining a fitness club may improve the fitness level of an individual. It has been noticed that people delay exercise considering the cost and the time. If this facility is available in personalised form on their figure tip, it will be more effective. Regular practice of exercises increases the fitness levels of individuals and reduces the negative effects of chronic diseases as discussed in [3]. Medical experts prescribe the exercise therapies based on disease type to improve the patients' health.

One of the solutions to address this serious societal issue is to provide easy and simple way of accessing workout programs at user's own space. Computer-aided systems can design information filtering frameworks that can empower the users to manage their personal health issues. Recommender
system is one of the information filtering approach which suggests services to users of their interests. Health improvement recommender systems provide suggestions for nutrition, disease prevention therapy, and recipes. There are health improvement recommender systems which suggest regular workout regime to enhance the health condition as designed in [4, 5].

Existing studies provided the recommendations either in generalized way or only considering user's exercise preferences or health status alone. The current proposal considers user's thyroid health status, exercise preferences and calories intake together. The proposed system extracts information from patient profile such as physiological data, thyroid disease profile, preference for exercise type, activity level, and calorie intake. The proposed hybrid model designs a fitness recommender framework using hybrid RBM model to generate exercise programs for thyroid patients. The hybrid model designs a workout plans by calculating preference score using matrix factorization method and similarity score using content-based method. The proposal hybrid fitness recommender model works as follows: i) A hybrid model is proposed for designing a fitness recommender system that generates exercise plans based on user's thyroid disease level and exercise preferences, ii) The proposed framework is implemented by combining preference score computed using Probabilistic Matrix Factorization (PMF) method and similarity score obtained from Content-based method, iii) The proposed framework incorporates user's physiological status, preferences on exercise style, thyroid disease level in designing exercise plans, and, iv) The proposed study is evaluated and validated using exercise data extracted from various online assessable websites.

The article is presented as follows. The analysis of reviewed studies is reported in Section 2. The proposed hybrid fitness recommender framework is provided in Section 3. Experimental setup is analysed in Section 4 using online extracted dataset along with explanation of obtained results. Section 5 presents the conclusion and further works.

1.1 Impact of regular exercises on thyroid disorders

Exercise influences the movement of various glands of a human body and secretion of glands' hormones. Thyroid gland is one of them which positively get affected by practicing regular exercises. A butterfly shaped like thyroid gland is an endocrine gland placed underneath the larynx. Thyroid gland releases two distinct thyroid hormones namely Tetraiodothyronine (T4) and Triiodothyronine (T3) which are the product of amino corrosive iodine bound. These hormones control the body's functionalities like growth, metabolism, growth, gene expression and tissue differentiation [6]. Thyroid hormones control the body's energy expenditure, basal metabolism, glucose, and lipid mechanism. Thyroid disorder patients have been reported with obesity due to fluctuations in secretion of thyroid hormones. Body mass index (BMI) is associated with low serum levels of T4 and high serum hormonal levels of TSH [6].

Various observational studies showed that practicing regular exercise can positively influence the thyroid function. Strength, flexibility and core stability are achieved by performing Pilates exercise, Tai-chi and Yoga as these exercises control the posture, body movement, and breathing [7, 8]. Above-mentioned exercises provide positive effects on thyroid hormones and body metabolism. The study [9] examined the effect Pilates workout on thyroid hormones like TSH, T3 and T4 in sedentary women for eight weeks. Authors studied the effect of practicing physical exercises with moderate intensity on thyroid patients. The study found that Serum TSH was significantly reduced in patients following regular exercises [10]. A pilot study was conducted to observe the diet and exercise effects on the hormonal production of endocrine gland. The study reported that the type of exercise, its intensity and duration affected the production of thyroid hormones [11].

2 Related work

Digital health recommender system is recently an emerging area for many researchers by integrating healthcare services with information tools and technology [12]. These recommender systems attract a greater number of users by providing easy and convenient way of selecting health related information. Relatively very less research work is carried out on health-related recommender systems, particularly in the fitness domain. This section narrates a brief review on some recommender strategies applied in the fitness domain.

LifeStyle platform is developed in Ref. [13] to enhance the communication among the health experts and patients. The system provides interface for patients to track and manage their health conditions. The medical professional tracks the online information remotely. The data is located at the cloud-based platform and use for integration using a web interface. Patients can manage their information like sleep, blood pressure, weight, and mood. The
study suggested the exercise programs based on health status only and should need to integrate user's preferences.

Authors designed CoCare, a mobile application-based fitness recommendation framework to support health improvement and for preventing diseases [14]. This system recommended physical activity videos using users' profiles. The system applied decision tree and case-based reasoning approaches. The system classified and tagged the videos based on predictions using textual description from user's profiles. CoCare system calculated similarity among video categories and user profiles. The system extracted a small set of videos and presented a generic recommendation solution. The system would require a proper profiling and categorizing process to recommend suitable activity videos.

A prototype platform 'Fitness that Fits', was designed to recommend physical workout videos based on labelled video data obtained from Youtube-8M dataset [15]. The model integrated collaborative filtering and basic content-based and mechanisms using user preferences recent viewing behaviours. An iterative video replacement process was introduced using neighbourhood-collaborative filtering approach to promote diversified recommendations for users. The system recommended distinct types of fitness programs to enhance the user's health condition. The system should adopt user's preferences and calories intake to provide personalized approach.

Authors implemented HealthRecSys as a content-based recommendation model [16]. The system provided personalized links to extract online health videos with educational content. Metadata of YouTube videos was extracted to produce Medline Plus1 links. The system needs to recommend the health videos based on user's acceptance and perspectives like health status and age.

The system was developed PRO-Fit, a personalized fitness recommender framework to motivate users in physical activities by incorporating a social factor [17]. PRO-fit collected data from tracking devices and classified user's activity type. The framework incorporated a social element using collaborative filtering similarity measure for activity data and user profile. PRO-Fit generated personalized workout schedules based on users' availability and fitness goals. The study suggested the activities based on social impact on user's behaviour. The system needs to consider user's health status in recommendation process.

The study proposed an endurance sports-based fitness recommender system using machine learning strategies [18]. The recommender system focused on providing fitness suggestions for athletes to perform endurance sports like mountain biking, marathon running and triathlons. Author investigated an exercise solution by classifying and predicting relevant endurance training programs for athletes. The model should recommend the exercises considering user's age, health status and preferences.

Authors presented a collaborative filtering-based health care recommendation model, COHESY [19]. This model found the dependency between user's health status and performed physical activities. Authors implemented a recommender algorithm to discover a set of activities which would affect the user's health parameters positively using neighbour similarity method. The system identified the effectiveness of the exercises based on user's health parameter and recommended these activities to other similar users. But the influence of specific exercise may change based on individual's health condition.

A selfBack recommender system was implemented using case-based reasoning and genetic algorithm [20]. Case-based recommender algorithm created exercise lists for patients with back pain, while genetic algorithm used to boost the personalized approach to recommend exercise plans. Experiments were conducted based on real patients' data with system generated and expert's recommended solutions. This recommendation solution should integrate user's physiological features and intensities of exercises based on age and gender.

The study developed an expert advisory system by applying fuzzy inference engine to design workout programs [21]. The system gathered user's personal information, exercise preferences, and available time. The fuzzy inference model generated a customized exercise plan and provided a schedule according to user's availability. The recommendation process should consider user's health condition to generate effective set of workout programs.

Authors designed HyperRecSysPA, a physical fitness recommender model for hypertensive patients [22]. The Physical Activity Result Indexes were used to analyse the importance of physical activity. The model was composed of 32 parameters and classified into three groups to build the user profiles for generating recommendations. A specific set of questionnaires was verified by physicians and obtained conclusive results in the assessments performed. The model should incorporate exercise preferences and hence personalized approach is missing in the recommendation process.

The analysis of reviewed studies is resulted in finding few research gaps as follows: i) No specific exercise recommender model was found for Thyroid patients, ii) There were studies proposing generalized
fitness recommendations which may not be appropriate for every person, and, iii) The proposed recommendation system integrates features such as user's health condition, exercise preferences and calories intake together. The current model integrates the results of content based and PMF model as inputs for the RBM model to develop exercise plans specifically for thyroid patients. The main contribution of this recommendation model is to extend the traditional recommender models like collaborative filtering and content-based by introducing deep learning RBM model. The proposed framework obtains the useful data from users to provide meaningful and diverse exercise recommendations. The hybrid RBM model achieves its goal by generating personalized, effective, and accurate recommendations for thyroid patients.

3 Proposed fitness recommender framework

The proposed hybrid recommendation framework is designed using Content-based, Probabilistic Matrix Factorization and Restricted Boltzmann Machines approaches. A content-based model extracts the data from user personal and health profiles to identify the list of exercises. These workout plans are suggested based on user's thyroid disease level, food intake, exercise preferences and level of activity. A collaborative filtering based RBM model identifies the diverse set of exercises based on user's preferences. The intensity of exercise is quantified using MET as Metabolic Equivalent of Task. MET is measured as the total amount of the oxygen an individual consumes per unit of human body's weight for one minute in resting position [23]. For this proposal, one MET is used to represent the number of kilocalories spent per unit body weight for an hour of physical activity. For example, if physical exercise has 5 as a MET value then it shows that the intensity of this exercise is 5 times more than the resting position. Exercise recommendations should consider the calorie intake of the user [24]. Consumed calories intake is estimated using Eq. (3).

Daily calorie requirement is computed as follows:

\[ \text{Daily Calorie intake} = \text{BMR} \times \text{Fitness Factor} \tag{3} \]

where, BMR is computed for male and female as per Eq. (1) and Eq. (2) respectively.

A health-based fitness recommender framework is designed for thyroid patients based on their fitness level and thyroid disease type as illustrated in Fig. 1. The proposed architecture is classified into five parts: data collection, transformation, analysis, recommendation, and data visualization. In data collection phase, user profiles and exercise data sources are extracted from various online fitness portals. This phase contains information such as user's personal data, thyroid disease type, activity level, preferences, exercise type, exercise intensity and duration. Data transformation process includes attribute selection and normalization techniques. Data analysis process identifies the correlation

\[ \text{BMR} = 13.397W + 4.799H - 5.677A + 88.36 \tag{1} \]

\[ \text{BMR} = 9.247W + 3.098H - 4.330A + 447.593 \tag{2} \]
Table 2. Attributes used in proposed framework

<table>
<thead>
<tr>
<th>Attributes</th>
<th>Type</th>
<th>Values/Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender (G)</td>
<td>Categorical</td>
<td>Male, Female</td>
</tr>
<tr>
<td>Age (A)</td>
<td>Numerical</td>
<td>18 &lt;= age &gt;= 75</td>
</tr>
<tr>
<td>Height (H)</td>
<td>Numerical</td>
<td>Height in centimetres</td>
</tr>
<tr>
<td>Weight (W)</td>
<td>Numerical</td>
<td>Weight in kilograms</td>
</tr>
<tr>
<td>BMI</td>
<td>Numerical</td>
<td>BMI (kg/m²)</td>
</tr>
<tr>
<td>Activity-level (AL)</td>
<td>Categorical</td>
<td>Sedentary, Low, Moderate, Intense, Very Intense</td>
</tr>
<tr>
<td>Exercise Type</td>
<td>Categorical</td>
<td>Yoga, Pilate, High Intensity, Walking, Jogging, Swimming, Cycling</td>
</tr>
<tr>
<td>Daily Calorie Intake (DCI)</td>
<td>Numerical</td>
<td>BMR × Fitness factor</td>
</tr>
<tr>
<td>Thyroid Disease-type (TDT)</td>
<td>Categorical</td>
<td>Hypothyroid (hypo), Hyperthyroid (hyper)</td>
</tr>
</tbody>
</table>

Table 3. Rating dataset

<table>
<thead>
<tr>
<th>Uid1</th>
<th>Ex1</th>
<th>Ex2</th>
<th>Ex3</th>
<th>Ex4</th>
<th>Ex5</th>
<th>Ex6</th>
<th>Ex7</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>5</td>
<td>1</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>1</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

among users and disease-related preventable exercises. This process finds the optimal features to detect the exercise types related to user’s thyroid disease type. During recommendation process physical exercises are generated such as yoga, swimming, aerobics, Pilates, running and walking. The intensity and duration of recommended workouts provide the quick recovery for thyroid patients using their fitness level and preferences. Data visualization part shows how the recommended exercise plans should be presented.

3.1 Dataset description

This research work extracted the data for 94 users from various online fitness blogs and portals like thefitindian.com and healthfitnessindia.in to evaluate the proposed model. 7 major types of exercises with their intensities and duration suitable for thyroid health improvement were collected from online portals. Few important attributes of user are identified and analysed with the help of fitness experts to design the current proposal.

Table 2 describes the features of the generated dataset used in current proposal and Table 3 shows the sample dataset. Each user’s calorie intake is estimated using Eq. (3). BMI of each user is computed using height-weight ratio. The proposed recommendation system constructs a user-exercise rating matrix. Table 3 shows a rating matrix where each row-column entry presents exercise preference scale ranging from highly preferred as 5 to not preferred as 1.

This rating matrix is expressed as hidden factors to reduce errors for predicting unknown data. The proposed model estimates the missing values in the rating matrix by applying Stochastic Gradient Descent for the weight vectors.

3.2 Proposed method

The process flow of the fitness recommendation process is presented in Fig. 2. The framework develops a hybrid deep learning-based approach by integrating Content based method and Probabilistic Matrix Factorization as a collaborative filtering model. To achieve this goal, three methods are introduced: (i) Estimating user-exercise similarity score, (ii) Computing user-exercise preference score (iii) Deep learning-based hybrid model. The following steps are involved to integrate user personal and health profiles in order to generate personalized recommendations:

- Extract user’s information such as personal and medical profile, activity level, and preferences.
- Estimate the fitness status of user by calculating daily calorie requirements, Body Mass Index (BMI), Basic Metabolic Rate (BMR), and Metabolic Equivalent of Task (MET) values.
- Determine the intensity level and duration of selected exercises based on user's fitness status and current level of activity.
3.2.1. Estimating user-exercise based similarity score

User’s physiological data and thyroid disease type are provided as inputs to identify the health-related exercises. The user specified exercise preferences are considered as topics using content-based method. For instance, user \( u \), has preferred exercises like aerobics, yoga, and high intensity workouts. These preferences are fed to the model as topics. Let \( k \) be the topics available in the proposed model. The user profile \( u \) is constructed using \( k \)-dimensional binary vector,

\[
U_i = (u_{i1}, u_{i2}, u_{i3}, \ldots, u_{ik})
\]

where, \( u_{ij} = 1 \) in case of the \( j^{\text{th}} \) topic preferred by user \( u \); otherwise \( u_{ij} = 0 \).

The similarity score for user \( u \) is estimated using similarity measure as presents in Eq. (4). It computes the similarity based on exercise preferences and the thyroid disease type related exercises. The final exercise list has been ordered using the relevance among user preferences and health-related exercises. Similarity score is calculated using associative coefficients in Vector Space content-based model. This score is computed using normalized inner product of exercise vectors. Similarity measure like Cosine coefficient computes the angle among exercise and user vectors and correlation values are found between [-1,1].

\[
S(U, E) = \frac{\sum_{i,j} U_{ij} \times E_{ij}}{\sqrt{\sum_{i,j} U_{ij}^2 \times \sum_{i,j} E_{ij}^2}}
\]

Similarity score is calculated using cosine similarity among two vectors, user \( U_i \) and exercise \( E_i \) vectors as represented in Eq. (4).

3.2.2. Computing user-exercise based preference score

Matrix factorization methods are the most effective collaborative filtering techniques. These methods use latent hidden factors and build the interactions among users and products. Matrix factorization method generates two or more than two matrices from the original matrix. Probabilistic matrix factorization method finds the latent features using probability of selection of product by user or not. In current proposal, PMF predicts the probability that user selects a latent random attribute and a random exercise belongs to that attribute [26]. For example, user prefers a random latent attribute such as Hath Yoga and selects a random Yoga as exercise type which belongs to Hath Yoga feature. PMF computes the probability of exercise preference of user for the selected type of exercise.

\[
p(u|e) = \sum_f p(u|f) \times p(f|e)
\]

where, \( p(u|e) \) : Probability of user \( u \) preferred the exercise \( e \); \( p(u|f) \) : User selects a random latent attribute; \( p(f|e) \) : Exercise belongs to that selected latent feature; \( f \) : Number of latent attributes.

3.2.3. Deep leaning-based hybrid model

Lots of research investigations are available to develop recommender systems using content-based, neighbourhood-based and matrix factorization-based collaborative filtering approaches. The research work on recommender framework using deep learning needs attention. The reason of implementation of RBM in this hybrid model is its unique procedure of learning. RBM learn the conditional joint probability distribution values from its input. This neural network consists of two layers only as hidden (H) and visible (V) input layers. Unlike typical neural networks, RBM have no connectivity between the same layer input units’ neurons as well as among hidden unit’s neurons. The joint distribution \( p \) for \( V \) and \( H \) can be computed as:

\[
p(V, H) = \frac{1}{z} e^{-E(V,H)}
\]

where, \( z \) and energy function \( E(V, H) \) are described as below:

\[
z = \sum_{V,H} e^{-E(V,H)}
\]
**3.3 Evaluation methods**

The recommendation system is generally evaluated using some performance metrics to validate a model’s accuracy. The quality of the recommendation model can be validated using statistical accuracy metrics like Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) [28]. Model is considered as more accurate with lesser values of MAE and RMSE as defined in Eq. (14) and Eq. (15), respectively. These error metrics are computed as follows:

\[
MAE = \frac{1}{N} \sum_{i=1}^{N} |a_i - p_i| 
\]  

\[
RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (a_i - p_i)^2} 
\]

where, \(a_i\) and \(p_i\) are actual and predicted exercise preferences respectively; \(N\) is total number of instances in dataset.

**4 Experiments and results**

The proposed framework for fitness recommender system is built using content based, PMF and RBM methods. These models are implemented using Python programming language with Surprise recommendation package and TensorFlow framework. The similarity among the user and exercises is calculated based on thyroid disease type using content-based method. PMF estimates the exercise preference score based on the latent hidden features which are indirectly related to users and exercises. Stochastic Gradient Descent function is applied to PMF to decrease the error in recommendation process.

**4.1 Training phase**

The proposed model creates a training data with rating matrix provided by each user to all exercise types. These rating preferences are normalized in interval of 0 and 1. Value 0 is assigned to all the
exercises that user has not preferred and 1 to preferred exercises. Each rating is divided by 5 to normalize the data before feeding to RBM model. The proposed Hybrid RBM model has extra two visible units along with input and hidden units. The dataset has 94 data instances, so the visible unit have 94 input nodes. The choice of nodes for hidden layer is random and mostly considers as power of 2. The framework uses 16 units of hidden layer. Each node of input and hidden layers has a bias and weight along the interconnection. So, the model is trained with sigmoid activation function by using weight matrix of dimensions 94×16.

4.2 Testing phase

The trained model calculates the probability of new user's data using hidden binary state vectors against its visible input units’ vectors p(v) using Eq. (13). The proposed system minimizes the objective log-likelihood function defined in Eq. (7) by updating biases and the weight vectors w. The system uses stochastic gradient descent to optimized weight using negative and positive gradients using Eqs. (10)-(12). These gradients reflect the density distribution of probabilities computed by the proposed model. Negative gradient depends on model and positive one is based on input data. The model uses CD to approximate the negative gradient and Gibbs sampling for model distribution. In each epoch, for each batch RBM model computes CD, updates the biases and weights and calculates error.

Fig. 4 presents the error rate computed by the model as the difference among the input data and its reconstruction. The proposed model predicts the ratings for similar exercises and extracts the exercise types, their intensities, and durations data. The model then generates the top 3 exercise plans based on sorted predicted scores. The system can adopt the user's new preferences as new ratings and reconstructs the model.

The experiments are conducted to compare the performance of content based, PMF, and RBM using MAE and RMSE as shown in Table 4. The experimental results are illustrated in Fig. 5, which depicts the performance of recommender models used in this study. The proposed hybrid model using deep learning-based approach outperforms the traditional Content-based, Matrix Factorization-based collaborative filtering and RBM models.

5 Conclusion

Presently, most of the people prefer to adopt a healthy lifestyle by using recommendation platforms along with the medical counselling. Health recommender systems generate relevant and effective recommendations to users using advance e-health platforms. Due to the availability of generic health content, users may not be get benefited in certain circumstances. This article presents a hybrid fitness recommender system based on user’s preferences and thyroid disorder type. Every user has distinct characteristics such as individual’s preferences, activity level, and health condition. The features like preferred exercise intensity, duration, and health improvement exercises may not be available in dataset for all the users. Such hidden attributes are more effective to find similarity among users and exercises. A deep learning approach is required to identify these hidden features from user and exercise
profiles. The proposed hybrid model incorporates exercise preference and similarity values as inputs to Restricted Boltzmann Machines. The predicted list of exercises is generated and recommended Top-3 exercises list to user. The proposed system uses the feedback method to adjust the recommended exercises as per user’s requirement. The proposed fitness recommender framework can be extended by suggesting healthy food items to investigate the joint effect of balanced food intake and fitness recommendations. Recommendation of workout videos is another future direction for this proposed framework.

Conflicts of Interest
The authors declare no conflict of interests. This article is author's original work and not has been published elsewhere. The authors warrant the validity of data and the obtained results.

Author Contributions
The substantial contributions of the authors is reported in this article as follows: (i) Conceptualization- V S Vairale (ii) Methodology- V S Vairale, (iii) Software- V S Vairale, (iv) Validation- V S Vairale and S Shukla (v) Data curation- V S Vairale (vi) Writing- original draft preparation- V S Vairale (vii) Visualization- V S Vairale and S Shukla.

References


