Recommender systems in the healthcare domain: state-of-the-art and research issues

Thi Ngoc Trang Tran[1] · Alexander Felfernig[1] · Christoph Trattner[2] · Andreas Holzinger[3]

Abstract

Nowadays, a vast amount of clinical data scattered across different sites on the Internet hinders users from finding helpful information for their well-being improvement. Besides, the overload of medical information (e.g., on drugs, medical tests, and treatment suggestions) have brought many difficulties to medical professionals in making patient-oriented decisions. These issues raise the need to apply recommender systems in the healthcare domain to help both, end-users and medical professionals, make more efficient and accurate health-related decisions. In this article, we provide a systematic overview of existing research on healthcare recommender systems. Different from existing related overview papers, our article provides insights into recommendation scenarios and recommendation approaches. Examples thereof are food recommendation, drug recommendation, health status prediction, healthcare service recommendation, and healthcare professional recommendation. Additionally, we develop working examples to give a deep understanding of recommendation algorithms. Finally, we discuss challenges concerning the development of healthcare recommender systems in the future.

Keywords Health recommender systems · Food recommendation · Drug recommendation · Health status prediction · Healthcare service recommendation · Healthcare professionals recommendation

1 Introduction

In the past decades, a considerable amount of clinical data representing patients' health status (e.g., medical reports, laboratory results, and disease treatment plans) have been collected. This has remarkably increased digital information available for *patient-oriented decision making*. Such digital information is often scattered across different sites, which hinders users from finding useful information for their well-being improvement. Besides,

✉ Thi Ngoc Trang Tran
   ttrang@ist.tugraz.at

Extended author information available on the last page of the article.

more drugs, tests, and treatment recommendations are available for medical staff daily, which triggers difficulties in deciding appropriate remedies for patients (Stark et al. 2019; Wiesner and Pfeifer 2014). In this context, recommender systems for medical use should be implemented to bridge these gaps and support both, patients and medical professionals, to make better healthcare-related decisions. Recommender systems have been integrated into online retailers, streaming services, and social networks to facilitate users' item selection process (Felfernig and Gula 2006; Tran et al. 2018). Recently, these systems have been widely applied to the healthcare domain (so-called *Health Recommender Systems - HRS*) to better support medical suggestions. Different from the precursors in the same domain (e.g., medical expert systems), HRS offer a better personalization that increases the details of provided recommendations and improves users' understanding of their medical condition. These systems also provide patients with a better experience, improve their health condition, and motivate them to follow a healthier lifestyle. Moreover, they also assist healthcare professionals with disease predictions/treatments (Holzinger et al. 2016; Pincay et al. 2019; Sahoo et al. 2019; Schäfer et al. 2017; Wiesner and Pfeifer 2014). HRS should analyze patients' health status and recommend personalized diets, exercise routines, medications, disease diagnoses, or other healthcare services. HRS's great concern is to send the necessary information to patients at the right time while ensuring the accuracy, trustworthiness, and privacy of patient information (Sahoo et al. 2019). Moreover, these systems are expected to minimize the cost of the healthcare-related decision making process (in terms of *time* and *effort*) (Valdez et al. 2016).

Although several studies have been carried out on HRS, they target a specific disease or recommendation context. This raises a need for a comprehensive overview that provides a "*full landscape*" of recommendation scenarios supported by HRS (Pincay et al. 2019). In the current literature, only a few endeavors summarize current approaches to designing and implementing HRS. For instance, Sezgin and Özkan (Sezgin and Özkan 2013) and Wiesner and Pfeifer (Wiesner and Pfeifer 2014) discussed some recommendation scenarios (e.g., *drug recommendations, medical information suggestions*, and *disease predictions*) and various methods to evaluate the effectiveness of HRS. Calero Valdez et al. (Valdez et al. 2016) provided a literature review, in which a framework to develop HRS was proposed. Stark et al. (Stark et al. 2019) analyzed 13 existing studies and categorized them according to criteria, such as *disease, data storage, interface, data collection, data preparation, and recommendation techniques*. Finally, Pincay et al. (Pincay et al. 2019) presented an overview of the methods and techniques used to design and implement HRS. Compared to the mentioned related work, our article presents a *broader picture* of recommendation scenarios supported by HRS with a different set of considered studies. The discussed scenarios focus on two types of users: *end-users* (healthy users and patients) and *healthcare professionals* (e.g., doctors, nurses, clinicians, or physicians). For end-users, HRS provide *nutritional information, medications, treatment plans, diagnoses/disease predictions, physical activities* or *other healthcare services* (e.g., finding good doctors or proper medical services for patients) (Wiesner and Pfeifer 2014). For healthcare professionals, HRS use medical resources to assist them in creating more precise suggestions for patients. For each recommendation scenario, we summarize recommendation algorithms and develop corresponding working examples. Besides analyzing existing studies, we discuss research challenges as well as potential directions for future HRS. The remainder of the article is structured as follows. In Section 2, we present the used methodology for our literature analysis. In Section 3, we discuss basic approaches that are frequently used in recommender systems and necessary adaptations to apply these in the healthcare domain. In Section 4, we present

recommendation scenarios supported by HRS and used recommendation techniques. In Section 5, we summarize evaluation methods employed in the mentioned approaches. Finally, we discuss open issues for future work in Section 6 and conclude the article in Section 7.

2 Research methodology

This study was performed based on a *bibliographic review* method, which provides a systematical analysis of domain-specific knowledge. We first collected a set of studies concerning HRS using the following keywords: *"health recommender systems", "medicine recommender systems", "recommender systems in the wellness domain"*, and *"e-Health systems"*. To have a deeper look at recommendation scenarios in the healthcare domain, we searched for references using additional keywords: *"food recommendation", "nutrition recommendation", "drug recommendation", "heath status prediction", "healthcare service recommendation", "physical activity recommendation",* and *"doctor recommendation"*. Besides, to ensure the quality of the references, we selected the papers using some criteria proposed by Stark et al. (Stark et al. 2019): *(i)* published from 2000 onward; *(ii)* well referenced with more than 15 sources, *(iii)* providing logical and reasonable findings of the domain, and *(iv)* presenting a detailed discussion on recommendation techniques. We searched for references in well-known digital libraries, such as Google Scholars[1], Springer[2], ACM[3], ResearchGate[4], Science Direct[5], and PubMed[6]. In this context, we checked the title, keywords, abstract, conclusion, tables, and figures of the collected papers. Finally, we filtered out *98 papers* that meet the mentioned criteria and have a strong relationship with our work. From these, we selected and analyzed 37 studies, which provide detailed discussions on recommendation approaches in the healthcare domain. These studies are summarized in Section 4: *eight* papers related to food recommendation, *18* papers on drug recommendation, *three* papers related to health status prediction, *four* papers on physical activity recommendation, and *four* papers on healthcare professional recommendation. Additionally, we analyzed 32 papers to find out open issues for HRS and potential directions for future work. The remaining papers are cited in other sections of this article. Most papers cited in our work were published in prestigious conferences, such as ACM Conference on Recommender Systems, ACM Conference on User Modelling, Adaptation and Personalization, IEEE International Conference on e-Health Networking, and International Conference on Software Engineering and Knowledge Engineering. For journal articles, we selected the ones published in *journals on computer science* (e.g., Journal of Computer Applications, Journal of Expert System Applications, Journal of Computing Sciences in Colleges, Journal of Data Mining and Knowledge Discovery, and Journal of Engineering and Technology) and *on medicine* (e.g., Journal of Science Translational Medicine, International Journal of Basic Science in Medicine, and Journal of Biomedical Semantics, and Journal of Biomedical Informatics).

---

[1] https://scholar.google.at/

[2] https://link.springer.com/

[3] https://dl.acm.org/

[4] https://www.researchgate.net/

[5] https://www.sciencedirect.com/

[6] https://pubmed.ncbi.nlm.nih.gov/

3 Basic techniques in recommender systems

Collaborative filtering (Aberg 2006; Bankhele et al. 2017; Berkovsky and Freyne 2010; Davis et al. 2009; Dharia et al. 2016; Han et al. 2018; Narducci et al. 2015; Nasiri et al. 2016; Stark et al. 2017; Zhang et al. 2016), content-based (Aberg 2006; Dharia et al. 2016; Han et al. 2018), knowledge-based (Ali et al. 2018; Doulaverakis et al. 2012; Mahmoud and Elbeh 2016), and hybrid approaches (Aberg 2006; Dharia et al. 2016; Han et al. 2018) are the basic recommendation techniques that can be used in HRS. Besides, other algorithms are also applied to generate recommendations in the healthcare domain, such as ant colony algorithm (Rehman et al. 2017), classification (Hussein et al. 2012; Shimada et al. 2005), clustering (Rokicki et al. 2015), decision tree (Bresso et al. 2013), logistic regression (Huang et al. 2011), natural language processing (Gujar et al. 2018), inductive logic programming (Bresso et al. 2013), ontologies (Chen et al. 2011; Chen et al. 2012; Donciu et al. 2011; Doulaverakis et al. 2012; Faiz et al. 2014; Mahmoud and Elbeh 2016), sparse canonical correlation (Yamanishi et al. 2012), support vector machines (Huang et al. 2011), semantic technologies (Donciu et al. 2011; Faiz et al. 2014; Medvedeva et al. 2007), multi-criteria decision making (Chen et al. 2011; Chen et al. 2012), graph-based recommendations (Stark et al. 2017), context-aware recommendation (Ali et al. 2018), and matrix factorization (Zhang et al. 2015). In this section, we present basic recommendation techniques applied in the healthcare domain. Other techniques will be discussed in Section 4.

3.1 Aspects of recommender systems

There are three main aspects that need to be considered in recommender systems: *usage context, users*, and *items* (Sánchez-Bocanegra et al. 2015). *Usage context* describes the environment where all elements (e.g., items, users, and their relationship) interact with each other. *Users* are the end-users of recommender systems, and *items* are the elements that users are looking for. In the healthcare domain, additional aspects concerning the mentioned elements should be considered to generate more precise recommendations.

Usage context The usage context in HRS consists of *contextual factors* and *multi-factorial goal settings* that can influence on how items are recommended or presented. *Contextual factors* indicate *dynamic attributes* which might affect a specific activity (e.g., *time* to take medicine - the optimal time to take fat-soluble vitamins is with dinner) and *dynamic factors* from users (such as emotional states). The inclusion of such contextual information in the sequence of recent contexts of a user can help to better understand contexts that led to the user's current behavior and preferences. For *multi-factorial goal settings*, different domain-specific criteria should be considered when evaluating an item. In e-commerce domains, people might naively think that the "*most preferred items*" are more likely to be recommended to users. However, this idea needs to be reconsidered in the healthcare domain since items that are the best for this user might not be good for others (Valdez et al. 2016). For instance, although diuretics and blood pressure-lowering medicines are good for patients suffering from hypertension, these drugs can be dangerous for diabetes or gout patients. Besides, even patients have the same diseases, remedies for this patient could not always be suggested to others since they might have different health conditions.

Users HRS are able to support two types of users: *end-users* and *healthcare professionals*. *End-users* could be healthy users or patients. For each end-user, the system has to save a user profile describing his/her health condition. For instance, the profile of a cardiovascular

patient includes the following information: *name, birthday, weight, height, cardiovascular type*, and *blood pressure measurement*. This information helps HRS identify appropriate medications for the user. *Healthcare professionals* can be *doctors, nurses, physicians, clinicians,* or *pharmacists*. Besides, *medical researchers* and *policy makers* can also benefit from HRS (Valdez et al. 2016).

Items HRS can offer recommendations concerning different categories, such as *diets* to optimize nutrition, *physical activities/sports* that match the user's requirements and needs, *recommended diagnoses* of patients to doctors or nurses, *treatments/medications* for a specific disease, and *medical information/sources* that motivate(s) users to follow a healthy lifestyle and improve their well-being (Valdez et al. 2016).

3.2 Basic recommendation techniques

The information of the mentioned elements can be the input of algorithms that generate personalized recommendations to patients.

Collaborative Filtering (CF) CF recommends items to a user based on the following idea: "If users shared the same interests in the past, then they would have similar tastes" (Jannach 2011). In the context of HRS, this approach can be interpreted as follows: "If patients share similar disease profiles/health conditions, then they would have similar treatments/healthcare services".

Content-based Filtering (CB) This approach looks for items similar to those that the user liked in the past and match the user profile (Lops et al. 2011; Ricci et al. 2010; Sánchez-Bocanegra et al. 2015). In HRS, this approach suggests healthcare services that fit the patient's health condition/disease situation and are similar to those assigned to him/her in the past.

Knowledge-based Recommendation (KB) This approach is applied to domains where the quantity of available item ratings is quite limited (e.g., *cars*, *apartments*, and *financial services*) or when the user wants to define his/her requirements for items explicitly (e.g., "the food should not contain cheese since I am allergic to milk products"). This approach creates recommendations based on *knowledge about the items*, *explicit user preferences*, and *a set of constraints* describing the dependencies between users' preferences and items' properties (Felfernig and Burke 2008).

Hybrid Recommendation (HyR) The idea of this approach is to combine the aforementioned recommendation techniques to make use of the advantages of one approach and fix the disadvantages of another approach (Ricci et al. 2010). For instance, CF usually faces a *cold-start* problem triggered when a new item is added to the system and has no user ratings, whereas CB can tackle this issue since the prediction for new items is generally based on available descriptions of these items.

4 Recommendation scenarios in the healthcare domain

HRS offer users various types of recommendations that help to improve their well-being. These systems also assist healthcare professionals in making more precise patient-oriented

decisions. In the following subsections, we provide a detailed discussion on recommendation scenarios and corresponding recommendation approaches (see also Table 1).

4.1 Food recommendation

Due to the extensive growth of food variety and busy lifestyles, people have been facing the issue of making healthy food decisions to reduce the risk of chronic diseases (Ge et al. 2015; Robertson et al. 2004). In this context, *food recommender systems* can motivate users to change their eating behaviors or suggest healthier food choices (Tran et al. 2018; Trattner and Elsweiler 2017; Yang et al. 2017). In the following, we summarize scenarios where food recommender systems support users in optimizing their nutrition intake. Studies on *food recommendation* were presented in our earlier survey (Tran et al. 2018). However, different from (Tran et al. 2018) where these studies were grouped based on recommendation techniques, in this article, they are grouped according to dietary needs. Besides, we include additional studies on *"food-substitutes suggestions"* to increase the coverage of the article. For the studies already presented in our earlier survey, we shortly mention the general idea of recommendation algorithms, and for further details, we refer to (Tran et al. 2018).

Recommend proper diets Many people are suffering from health problems concerning inappropriate eating habits. Thus, one of the main functions of food recommender systems is to understand the eating behavior and recommend proper diets to users. There exist in the current literature some systems that fulfill this function. For instance, Aberg et al. (Aberg 2006) developed a *menu-planning tool* to deal with the malnutrition of the elderly. Rehman et al. (Rehman et al. 2017) highlighted the appropriateness of selected diets by proposing a *cloud-based food recommender system* so-called DIET-RIGHT. This system uses an *ant colony algorithm* to generate an optimal food list and to suggest proper food for users according to their pathological reports.

Prevent/Cure Food-based Illness Lacking the nutritional understanding of users leads to poor selections of ingredients and causes *food-related diseases*. To prevent these issues, food recommender systems have been developed to provide nutrition recommendations that consider both the preferences and health conditions of patients. For instance, Rokicki et al. (Rokicki et al. 2015) suggested menus that best match the patients' tastes and dietary restrictions. Ueta et al. (Ueta et al. 2011) proposed a *goal-oriented recipe recommendation* to provide a user with a nutrient list for treating his/her health problem.

Suggest Food Substitutes Another approach of food recommender systems is to identify a substitute relationship between food pairs as the first step towards *"similar but healthier"* food recommendations (Achananuparp and Weber 2016). In this approach, foods are assumed to be similar dietarily if they are consumed in similar contexts. For instance, "a chicken sandwich can be a substitute for a turkey sandwich if they are both consumed with french fries and salad" (Achananuparp and Weber 2016). This approach analyzes the real-world self-reported food consumption of users created by the MyFitnessPal[7]. The consumed food items and corresponding contexts are represented in a *food-context matrix*, in which the rows represent food items, and the columns represent a context. The similarity between two food items is measured using the *Cosine similarity* between the corresponding row vectors in the matrix. The higher the similarity score, the higher the probability of food pairs to

---

[7]https://en.wikipedia.org/wiki/MyFitnessPal

Table 1 A summary of key studies discussed in the article

Rec. scenarios	Papers	Rec. techniques	Functionalities
Food recommendation	(Aberg 2006)	CF, CB	resolves the malfunction of the elderly
	(Achananuparp and Weber 2016)	context-based rec.	proposes healthy food for users
	(Berkovsky and Freyne 2010; Elahi et al. 2014)	CF, group rec.	generates food recommendations to a family
	(Elsweiler et al. 2017)	ingredient network	generates similar but healthier recipe replacements
	(Rehman et al. 2017)	ant colony algorithm	generates optimal food lists for users
	(Rokicki et al. 2015)	clustering	recommends healthy menus to users
	(Ueta et al. 2011)	goal-oriented recipe rec.	provides suitable nutrient to treat a user's health problems
Drug recommendation	(Bankhele et al. 2017)	CF	suggests proper drugs to diabetes patients
	(Bresso et al. 2013)	decision trees, inductive logic programming	specifies side-effect profiles of drugs
	(Cami et al. 2011; Galeano and Paccanaro 2018)	Pharmacosafety Networks	provides accurate drug side-effects predictions
	(Chen et al. 2011; Chen et al. 2012)	ontologies, multi-criteria decision making	proposes anti-diabetic drug recommendations
	(Doulaverakis et al. 2012; 2014)	ontologies, rule-based rec.	recommends proper drugs
	(Fliri et al. 2006; Fukuzaki et al. 2009)	structure-activity, structure - property relationships	predicts drug side effects
	(Huang et al. 2011)	support vector machine, logistic regression	predicts drug side effects

Table 1 (continued)

Rec. scenarios	Papers	Rec. techniques	Functionalities
	(Mahmoud and Elbeh 2016)	ontologies, rule-based decision making	suggests medicines with dose prescriptions
	(Medvedeva et al. 2007)	case-similarity retrieval system	assists doctors in optimizing treatments for their patients
	(Rodríguez et al. 2009)	semantic web	provides patients with drugs to heal a pathology
	(Shimada et al. 2005)	risk-level classification	assists doctors in selecting first-line drugs
	(Stark et al. 2017)	CF, graph-based rec.	provides precise drugs to migraine-disease patients
	(Yamanishi et al. 2012)	sparse canonical correlation	predicts potential side-effect profiles of drug candidate molecules
	(Zhang et al. 2016)	CB	predicts missing side effects of approved drugs

Table 1 (continued)

Rec. scenarios	Papers	Rec. techniques	Functionalities
Health status prediction	(Hussein et al. 2012)	random forest classification	predicts disease risks of patients
	(Davis et al. 2009; Nasiri et al. 2016)	CF	predicts risk factors of chronic-disease patients
Physical activity recommendation	(Ali et al. 2018)	context-aware rec., knowledge-based rec.	provides patients with physical activities
	(Dharia et al. 2016)	CF, CB	suggests personalized workout-session recommendations
	(Donciu et al. 2011; Faiz et al. 2014)	ontologies, semantics technologies	recommends proper exercises to users
Healthcare professionals recommendation	(Gujar et al. 2018)	coreNLP techniques	generates doctor recommendations to patients
	(Han et al. 2018)	CF, CB	recommends family-doctor to patients
	(Narducci et al. 2015)	CF	suggests doctors and hospitals according to the patient's profile
	(Zhang et al. 2016)	hybrid matrix factorization	predicts the rating of doctors for patients

be suitable substitutes for each other. Elsweiler et al. (Elsweiler et al. 2017) investigated the feasibility of replacing recipes consumed by a user with similar and healthier alternatives. To find appropriate recipe substitutions, the authors applied an *ingredient-network* approach (Teng et al. 2012) to establish recipe pairs based on their pairwise similarities. Thereafter, they looked at the distribution of health features across pairs to identify healthier replacements. Finally, the rating distributions within pairs were considered to find out replacements with higher ratings than the original recipes' rating.

Food Recommendation for Groups In many real-world scenarios, food consumption is a good example of a group activity, for instance, deciding on the menu for a birthday party with friends or daily meals with family members (Elahi et al. 2014; Felfernig et al. 2018). In these scenarios, recommendations should be tailored in such a way that assures the maximum satisfaction of each member and the group as a whole (O'Connor et al. 2001). Berkovsky and Freyne (Berkovsky and Freyne 2010) examined the applicability of CF strategies to discover the best strategy for group recommendations. The authors discussed two group-based recommendation strategies: *aggregated models strategy* and *aggregated prediction strategy*[8]. These strategies recommend a list of recipes to the whole family by considering the task of recommending *top-k* recipes. Similar work was conducted by Elahi et al. (Elahi et al. 2014), where a novel interactive environment for groups was developed in planning meals through a conversational process based on critiquing (Chen and Pu 2012). For an example of food recommendation for groups, we refer to our earlier survey (Tran et al. 2018).

4.2 Drug recommendation

4.2.1 Drug recommendation for curing diseases

Medication errors are one of the most serious medical errors that could threaten patients' life (Charkhat Gorgich et al. 2015). More than 42% of these errors are caused by doctors who have limited experiences/knowledge about drugs and diseases (Bao and Jiang 2016). Another reason lies in the increasing number of available drug information, which has brought obstacles concerning the discovery of *relevant drugs* and *drug-disease interactions* (Doulaverakis et al. 2012). In this context, drug recommender systems have been developed to assist end-users and healthcare professionals in identifying accurate medications for a specific disease.

Diabetes disease Diabetes is one of the most popular diseases caused by busy lifestyles with a lack of physical activities and unhealthy eating habits (Bankhele et al. 2017; Mahmoud and Elbeh 2016). Plenty of drug recommender systems have been developed to help end-users effectively control diabetes and avoid future complications. These systems also assist medical professionals in giving precise medicine recommendations to patients. Chen et al. (Chen et al. 2011) created *anti-diabetic drug recommendations* based on *patient ontology knowledge* and *multi-criteria decision making*. Mahmoud et al. (Mahmoud and Elbeh 2016) utilized ontologies to represent knowledge about patients' profiles and anti-diabetes drugs. This system additionally combines ontologies with *rule-based decision making* to provide restrictions on target treatment goals and medicines with dose prescription. The defined rules select drugs for each patient based on his/her profile. An example rule of selecting drugs can be described as follows: "If a patient is under 60 years old, suffering a

---

[8]For details of these strategies, we refer to (Felfernig et al. 2018; Tran et al. 2018).

liver problem, and used Sulfonzlureas (Glipizide), then starting dose should be 2.5mg daily" (Mahmoud and Elbeh 2016). Medvedeva et al. (Medvedeva et al. 2007) developed a *web-based case-similarity retrieval system* to enable doctors to share their knowledge with the community and to optimize disease treatments for their patients. In this system, patient histories are utilized by doctors to select suitable treatment plans for patients. Bankhele et al. (Bankhele et al. 2017) proposed a recommendation approach based on the *CF technique* to suggest proper medications to diabetes patients. A patient has to register in the system and then enters a predefined set of attributes, such as *age, insulin, glucose, BMI, BP,* and *triceps thickness*, which are then analyzed to create personalized recommendations. A *user-based CF* is applied to find patients whose attributes best match the active patient's attributes. This matching is done using Formula (1), where P is the attribute set of patients a and b; $r_{a,p}$ is the value of patient a for attribute p with $\overline{r_a}$ as the mean over set P of attributes p; $r_{b,p}$ is the value of patient b for the attribute p with $\overline{r_b}$ as the mean over set P of attributes p.

$$sim(a, b) = \frac{\sum_{p \in P}(r_a - \overline{r_a})(r_b - \overline{r_b})}{\sqrt{\sum_{p \in P}((r_a - \overline{r_a})^2)}\sqrt{\sum_{p \in P}((r_b - \overline{r_b})^2)}} \tag{1}$$

Example 1 For demonstration purposes, we introduce an example describing the drug recommendation process using the approach presented in (Bankhele et al. 2017). Assume, *Tom* is an active patient who has entered to the system some attributes of his health status (see Table 2). The data of patients who share similar attributes with *Tom* ($patient_1...patient_4$) is summarized in Table 2. Based on Formula (1), the similarity scores of patients with regard to current user attributes are calculated as follows:

$sim(Tom, patient_1) = \textbf{0.86} \checkmark$; $sim(Tom, patient_2) = 0.51$

$sim(Tom, patient_3) = 0.49$; $sim(Tom, patient_4) = 0.04$

The calculations show that $patient_1$ is the most similar to *Tom*. Thus, the drugs prescribed for this patient can be recommended to *Tom*.

Migraine Disease Stark et al. (Stark et al. 2017) proposed a drug recommender system assisting doctors in writing more appropriate and accurate prescriptions to migraine-disease patients. This system uses a graph database to store patients' information. The database is organized as *nodes* and *edges*. Nodes represent patients' information, diseases, allergies, and drugs, whereas edges represent the relationships between nodes. Using a CF approach, drug recommendations are created as follows:

- Filter out patients who are similar to the active patient in terms of gender (male/female), aura (yes/no), and the type of migraine (acute/chronic).
- Calculate the similarity level between each neighbor and the active patient according to the following features: age, allergies, disease history, preexisting conditions, current drug prescription, and blood pressure. Each feature is weighted depending to its importance. For instance, *age* and *disease history* are more important than other features. Therefore, these features have a higher weight compared to others: $w_{age} = w_{diseaseHistory} = 3$ and $w_{allergies} = w_{preexistingConditions} = w_{bloodPressure} = 1$.
- Sum up all features' scores. Only drugs consumed by the patients who are at least 80% similar to the current patient will be included in the recommendation.

Infectious Diseases Shimada et al. (Shimada et al. 2005) developed a recommender system that helps doctors select proper *first-line drugs* for patients suffering from infectious diseases. Before giving suggestions, doctors have to know the ability of patients to protect

Table 2 Example 1 - The input attributes of *Tom* (active patient) and other patients in the database

	age	insulin	glucose	BMI	BP	triceps thickness
Tom (active patient)	25	25.4	86.0	27.9	67	23
patient₁	27	25.9	166	25.8	72	19
patient₂	48	33	118	46	82	41
patient₃	27	28	89	31	70	20
patient₄	53	43.2	195	30.7	69.7	45

themselves from risk factors. For this, a *risk-level classification* method that utilizes clinical information of patients is applied. This method assigns each risk factor to a *score* representing its impact degree on the patient. Besides, a knowledge base consisting of risk factors and their impact degrees are also constructed. The system returns risk levels that are helpful for precisely predicting the patient's health condition and then recommending to him/her appropriate drugs.

Other Diseases Besides drug recommendations for specific diseases, plenty of recommender systems have been developed for undefined diseases. For instance, GALENOWL (Doulaverakis et al. 2012) allows doctors to search for drug information and recommends suitable drugs to patients based on their disease, allergies, and drug interactions in the past. This system uses *ontologies* and *ICD-codes* to store rules about drugs and their interactions. These rules are the system input to generate the most fitting drugs for patients. Based on the GALENOWL system, a semantic framework so-called PANACEA (Doulaverakis et al. 2014) was developed to assist physicians in prescribing drugs according to drugs' active substance indications and contraindications. PANACEA generates drug recommendations based on standardized medical terminologies and rules describing drug-drug and drug-disease interactions. This system outperforms GALENOWL while guaranteeing the same recommendation quality. Similar to PANACEA, SEMMED (Rodríguez et al. 2009) was developed based on semantic web technologies. This system provides patients with correct drugs and treatment recommendations that are proper to heal a concrete pathology. Besides, it helps healthcare professionals avoid mistakes in the drug interaction process and discard factors causing risks to patients, such as drug allergies or contraindications.

4.2.2 Predict drug side effects

Drug side effects or *adverse drug reactions (ADR)* are one of the leading causes of morbidity and mortality in health care (Galeano and Paccanaro 2018). As reported by the American Institute of Medicine, unexpected drug side effects cause 100,000 deaths annually in the USA (Gurwitz et al. 2003). Thus, medical researchers have taken heed of developing systems for drug discovery (Zhang et al. 2016). One of the first ideas of drug side-effect predictions is to utilize structure-activity or quantitative structure-property relationships. For instance, Fliri et al. (Fliri et al. 2006) translated adverse effect data derived from 1,045 prescription drug labels into effect spectra, and then showed their utility for diagnosing induced effects of drugs. Fukuzaki et al. (Fukuzaki et al. 2009) designed a model to list drug side effects by searching for cooperative pathways shared among gene expression profiles. The general idea of this work is: "A drug is produced to affect a specific gene. However, if the drug inadvertently activates other genes, then it might cause side effects". In this approach, each *pathway* is represented as a *graph* with vertexes and edges. Each *vertex* represents a *gene* that indicates an *item-set* showing a set of drugs or conditions activating the gene. Each *edge* indicates a gene interaction. Based on this graph, sub-pathways showing side effects can be found based on the item-sets (i.e., activation conditions) shared between them.

Recently, some methods based on machine-learning have been employed to predict potential side effects of drugs. *"In silico"* is the most common method which creates side-effect predictions based on the *structure chemistry* and *biological features* of drugs, such as target proteins, protein-protein interactions, or gene ontology annotations (Zhang et al. 2015). Bresso et al. (Bresso et al. 2013) used this method to characterize side-effect profiles shared by several drugs. Huang et al. (Huang et al. 2011) utilized the drug targets, protein-protein interactions, and gene ontology annotations, and after that applied the *support vector*

machine and *logistic regression* techniques to create predictions. Yamanishi et al. (Yamanishi et al. 2012) combined *drug structures* (from chemical profiles) and *target proteins* (from biological profiles) and then adopted the *Sparse Canonical Correlation* to predict potential side-effect profiles of drug candidate molecules.

The prediction methods mentioned above face some limitations concerning the availability of chemistry structures, considerable required computational power, and a high amount of false positives (Deshpande and Butte 2011). Besides, they are usually done in clinical trials, where many side effects could not be detected until drugs are approved. This raises a critical need to predict *potential* or *missing side effects* for drugs (Zhang et al. 2016). A few drug recommender systems were developed to address this need. One example thereof was proposed by Zhang et al. (Zhang et al. 2016), in which the potential side-effect prediction is formed as a recommendation task. An *integrated neighborhood-based method* is applied to make predictions. This method is an extension of the *classic neighborhood-based recommendation*, which utilizes known side effects of similar drugs. We will present the detail of this recommendation method using the following example.

Example 2 Given a target drug d, a list of four *approved* drugs $\{d_1, d_2, d_3, d_4\}$, and corresponding side effects as shown in Table 3, we predict the probability of s_5 and s_6 to be the side effects of drug d. The prediction is formulated as a recommendation problem, in which drugs, side effects, and drug-side effects associations are combined. The prediction process is conducted in the following steps:

- *Step 1*: Calculate drug-drug similarity based on side effect profiles. Given two drugs d_i and d_k whose side effect profiles are S_i and S_k, the *Jaccard* similarity is used to calculate their similarity $sim(i, k)$ (see Formula (2)).

$$sim(i, k) = \frac{|S_i \cap S_k|}{|S_i \cup S_k|} \tag{2}$$

$sim(d, d_1) = \frac{2}{4} = \mathbf{0.5}\checkmark; sim(d, d_2) = \frac{1}{4} = 0.25$
$sim(d, d_3) = \frac{3}{4} = \mathbf{0.75}\checkmark; sim(d, d_4) = \frac{1}{4} = 0.25$

- *Step 2*: A set of neighbor drugs of the target drug d are determined by filtering similarity scores with a pre-defined threshold θ. In this example, we assume $\theta = 0.5$, which means only drugs d_1 and d_3 are selected to be the neighbors of d.

Table 3 Example 2 - An example of drugs and corresponding side effects. Values '0' or '1' represents the absence or presence of side effects s_j for drug d_i

Drugs	Side effects					
	s_1	s_2	s_3	s_4	s_5	s_6
d (target drug)	1	1	0	0	?	?
d_1	0	1	0	1	0	1
d_2	1	0	1	1	0	1
d_3	1	1	0	1	1	1
d_4	0	1	1	1	0	1

- *Step 3*: Calculate the probability of drug d_i inducing side effect s_j - $prob(d_i, s_j)$ by aggregating the known side effect s_j of its neighbors (see Formula (3)).

$$prob(d, sj) = \frac{\sum_{k=1, k \neq i, S_{k,i} \geq \theta}^{n} M_{k,j} \times sim(i, k)}{\sum_{k=1, k \neq i, S_{k,i} \geq \theta}^{n} sim(i, k)} \qquad (3)$$

$$prob(d, s_5) = \frac{0 \times 0.5 + 1 \times 0.75}{0.5 + 0.75} = 0.6; \ prob(d, s_6) = \frac{1 \times 0.5 + 1 \times 0.75}{0.5 + 0.75} = 1\checkmark$$

The probability $prob(d, s_6) = 1 > prob(d, s_5)$, meaning that s_6 is chosen as the potential side effect of the target drug d.

4.3 Health status prediction

In the past decades, predicting risks concerning specific diseases has become an intensive research topic (Davis et al. 2009), where the number of related studies for chronic diseases has been increasing significantly. The reason lies in the rapid growth of these diseases worldwide (Hussein et al. 2012). Long-term diseases prevent patients from physical activities and trigger burdens concerning time and money of the treatment process (Nasiri et al. 2016). To help patients avoid these diseases, HRS have been developed to detect disease symptoms as early as possible. Moreover, they can assist healthcare professionals in making proper treatment plans for patients. Davis et al. (Davis et al. 2009) and Nasiri et al. (Nasiri et al. 2016) proposed recommender systems to predict *risk factors* (e.g., *potential complications* or *further diseases*) that a target patient with a chronic disease would face in the future. These systems applied CF, which is based on the intuitive assumption "patients who share similar diseases and health status might face the same risk factors". Predictions of disease risks can be generated based on a set of similar patients' information. The traditional CF technique is modified to make it suitable for the healthcare domain. The reason for this modification lies in the rating values of items. The patients' ratings are non-ordinal values; they are binary (*1/0* - the patient is *facing/not facing* a risk factor j). For this approach, given an active patient a, a set of patients I, and a set of risk factors J, the risk factor prediction is generated in the following steps:

- *Step 1*: Calculate the similarity between patient a and each patient $i \in I$ using Formula (4), where $v_{a,j}$ is the vote of patient a for risk factor j:

$$w(a, i) = \sum_{j \in J} \frac{v_{a,j}}{\sqrt{\sum_{k \in J_a} v_{a,k}^2}} \frac{v_{i,j}}{\sqrt{\sum_{k \in J_i} v_{i,k}^2}} \qquad (4)$$

- *Step 2*: Find *the most similar patients* to patient a based on the similarity scores. The most similar patient has the greatest similarity score.
- *Step 3*: Calculate the prediction score of a risk factor j which have not been faced by patient a using Formulae (5) - (7), where $p(a, j)$ is the prediction score for the patient a on risk factor j, $\overline{v_j}$ is the average vote of all patients who have faced risk factor j, $w(a, i)$ is the similarity between patients a and i (see Formula (4)), and $|I_j|$ is the number of patients who have faced risk factor j. The normalized constant k ensures the prediction score within the range of possible votes.

$$p(a, j) = \overline{v_j} + k_a(1 - \overline{v_j}) \sum_{i \in I_j} w(a, i) \qquad (5)$$

$$\overline{v_j} = \frac{|I_j|}{|I|} \tag{6}$$

$$k_a = \frac{1}{\sum_{i \in I} w(a, i)} \tag{7}$$

Example 3 For a better understanding, we exemplify the mentioned approach using a specific example. Assume *Maria* is an active patient suffering from the diabetes disease. She is now facing some risk factors, such as *nerve damage*, *eye damage*, *slow healing*, *skin issues*. $p_1...p_4$ are the patients who share similar profiles to *Maria* (see Table 4). To predict which risk factors that *Maria* might face in the future, the prediction scores for the not-been-faced risk factors are calculated: *p(Maria, kidneyDamage) = 0.89, p(Maria, hearingImpairment) = **0.90**, p(Maria, heartDisease) = 0.69*. This shows Maria might face *"hearing impairment"* in the near future.

Besides recommendation techniques, machine learning approaches have been employed to generate disease predictions. For instance, Lafta et al. (Lafta et al. 2015) proposed an innovative *time series prediction algorithm* to support the decision making process of heart-disease patients. Particularly, the algorithm helps to decide whether a medical measurement, such as a heart-rate test, needs to be taken today based on the patient's measurement readings for the past *k* days. Hussein et al. (Hussein et al. 2012) presented a *Chronic Disease Diagnosis (CDD)* recommender system using the Random Forest - RF classification model (Özcift 2011).[9] This system requires *three* types of input information to build up predictions for undiagnosed patients: (1) *training data* consisting of medical records of previous diabetic patients; (2) *demographic data* showing the patient's profile, such as *name, age,* and *education level*; and (3) *the medical data of active patient* referring to two types of tests: *home-tests* (e.g., blood sugar level, blood pressure, and weight) and *lab-tests* from the laboratory. The system's output information includes a *prediction* representing the patient's disease risks and a *recommendation* showing disease risk status acknowledgment that the patient is looking for.

4.4 Physical activity recommendation

Besides recommendations of disease treatment plans, suggestions on physical activities have become another focus of HRS. Physical-activity recommendations help to decrease the probability of becoming frail of patients and prevent them from further health complications (Valdez et al. 2016). Moreover, they also encourage users to follow daily activities that meet their calorie-burn goals. RUNNER (Donciu et al. 2011) and SHADE (Faiz et al. 2014) provide users with food and exercise recommendations to keep them stay healthily. These studies generate a recommendation based on the fact that "what and when you eat during and after exercise can be just as important" (Donciu et al. 2011). These recommendations are tailored based on users' health status, goals, and preferences, which are usually collected from different sources, such as foods, physical activities, elderly/diabetes/runner domains, user-health state, and user preferences. Therefore, ontologies and semantics technologies (Orgun and Vu 2006) are utilized to address the heterogeneous issues of user data. The recommendation process can be done as follows: First, an initial set of exercises for

---

[9]Other decision tree classification models such as J48 (Kaur and Raghava 2003), Decision Stump (Noh et al. 2003), REP tree (Park et al. 2006) have been applied to predict the risk factors of chronic diseases.

Table 4 Example 3 - A summary of risk factors that *Maria* (active patient) and similar patients are suffering from. The similarity between *Maria* and each patient is shown in the last column

similar patients	nerve damage	eye damage	slow healing	skin issues	kidney damage	hearing impairment	heart disease	$w(Maria, p_i)$
Maria	1	1	1	1	?	?	?	
p_1	1	1	1	1	1	0	0	0.82
p_2	1	1	1	1	0	1	0	0.89
p_3	1	1	1	1	1	1	1	0.82
p_4	1	1	1	1	1	1	1	0.76

the user is selected based on his/her physical health status and exercise goal. Thereafter, the usage history and prior feedback regarding *difficulty* and *enjoyment* levels are used to adapt the selected exercises before sending it to the user.

Dharia et al. (Dharia et al. 2016) proposed a system to suggest personalized workout session recommendations based on the contextual data of users, such as past activities, preferences, and physical state. The recommendation process is performed as follows. The user first enters his/her contextual data. Thereafter, the system collects all the contacts and calendars events from the user's device and employs a *hybrid approach* to recommend fitness sessions to the user. This approach combines CB and CF recommendations, in which the CB considers the user's preferences, and the CF considers the preferences of similar users. The system also offers available slots in the user's calendar so that he/she can re-schedule sessions anytime.

Imran Ali et al. (Ali et al. 2018) developed a hybrid framework that provides physical activity and diet recommendations using context-aware recommendation (Verbert et al. 2012) and knowledge-based recommendation (Burke 2000). The proposed framework consists of a *multi-stage recommender system* which supports the following modules:

- *Module 1 (Data acquisition and processing)*, which stores the demographic information and preferred activities of users collected from sensory devices.
- *Module 2 (Context generation)*, which saves the current activity, location, weather conditions, and emotional state of the user.
- *Module 3 (Expert knowledge repository)*, which represents rules as IF-THEN form, which are then adopted to create recommendations. For instance, "IF a patient is pregnant and facing the gestational diabetes mellitus, THEN she should do a 20-30 minute moderate-intensity exercise on almost every day of the week" (Colberg et al. 2016).
- *Module 4 (Multi-stage recommender)*, which utilizes the user information collected from Modules 1 and 2 to create a comprehensive recommendation to the user. The recommendation process is done in two stages. In *Stage 1*, the system calculates the user's calorie-burn, in-take targets, and a generic set of physical activity recommendations. Additionally, a *case-based reasoning* mechanism is used to infer the most relevant rules from the knowledge-base. In *Stage 2*, the recommendations generated in *Stage 1* are refined in a personalized manner. A contextual matrix is created to recommend suitable activities to the user at a given time. This matrix is calculated based on the user's surveyed results to filter out proper physical activities in different contexts. For instance, "since the user is now staying at home, *stretching* seems to be more appropriate for him than running".
- *Module 5 (Explanations of suggested activities)*, which are sent together with recommendations to describe as to why a specific physical activity has been recommended to the user. For instance, "you should run at least one hour daily to improve your current health condition and meet one of your calorie-burn targets". Additional explanations based on the context can also be provided, e.g., "it is quite cold today, hence consider to bring a sports jacket with you before going out".

4.5 Healthcare professional recommendations

In recent years, there has been a significant increase in the amount of available medical information, which results in some difficulties for patients when searching for suitable doctors. What concerns patients greatly is *how to find medical professionals with the best expertise for resolving their health issues* (Hoens et al. 2010; Narducci et al. 2015). Most

existing healthcare providers do not provide patients with full infrastructure or service design implementations that assist them in fulfilling this task. This gap raises an open topic on patient-doctor matchmaking, in which patients can find the right doctors to build a trust relationship (Han et al. 2018). Han et al. (Han et al. 2018) proposed a hybrid recommender system, in which family-doctor recommendations are made based on the level of available information about users. The authors discussed *three* use cases of generating recommendations:

- *Use case 1* (New patient): The patient has recently joined the network, and only basic demographic information is available. The CB recommendation is used to create recommendations based on similar demographic profiles.
- *Use case 2* (Existing patient with no interactions with primary care doctors): The patient has already visited specialists or hospitals, but has not visited family doctors yet. The activities of other patients in previous visits are utilized to narrow down the doctor list. Besides, a complementary data set describing hospital inpatient procedures and certain types of diseases of patients are used to create the patient profiles and then generate recommendations using the CB recommendation approach.
- *Use case 3* (Existing patient with prior interactions with primary care doctors): The CF recommendation approach is applied to look for doctors visited by similar patients (i.e., patients who have visited the same doctors earlier).

Zhang et al. (Zhang et al. 2016) proposed an *iDoctor* system to provide users with personalized doctor recommendations. This system explores the emotions and preferences of users about doctors through their ratings and reviews. Three modules are integrated into the system: *sentiment analysis, topic modeling,* and *hybrid matrix factorization*. The *sentiment analysis* module calculates the emotional offset from user reviews. The *topic modeling* module extracts user preferences and doctor features (e.g., specialty, fee range, and prescribing habits) from user reviews. The extracted information is used in the *hybrid matrix factorization* module to predict the rating for doctors.

Gujar et al. (Gujar et al. 2018) proposed a recommendation approach based on the *"word of mouth"* recommendation (e.g., asking friends or relatives), which is often used in reality to find doctors. The authors developed a recommender system to identify the location, contact, and other necessary information of medical specialists. They used the CoreNLP technique to generate doctor recommendations based on the review of previous users. The recommended doctors are filtered out based on some criteria, such as *fewer fees, more experienced, nearest location*, and *feedback reviews of doctors*. Different from the studies discussed earlier, this system allows patients to give their feedback on recommended doctors, which are then used to improve the quality of future recommendations. Based on a similar recommendation mechanism as mentioned in (Gujar et al. 2018), Narducci et al. (Narducci et al. 2015) presented a social network so-called HEALTHNET, where a recommendation component is integrated to suggest doctors and hospitals which best fit a specific patient profile. In HEALTHNET, a patient enters his/her health data, such as *conditions*, *treatments* (drugs, surgeries, or side effects), *health indicators* (blood pressure, body weight, laboratory analysis, etc.), *consulted doctors,* and *hospitalizations*. Based on the input data, the system search for similar patients stored in the database. The similarity between the active patient p and another patient p' in the database is estimated using Formula (8), where:

- k and n are the numbers of conditions of patients p and p' respectively.
- z and r are the numbers of treatments of patients p and p' respectively.
- p_c and p'_c are the conditions of patients p and p' respectively.

- p_t and p'_t are the treatments of patients p and p' respectively.
- $s_c(p_{c_i}, p'_{c_j})$ is the similarity score between the condition c_i of patient p and the condition c_j of patient p' (see Formula (9)). If these two conditions are the same, then this score is the logarithm of the ratio between the number of conditions in the database (#C) and the number of patients affected by that condition (P_{c_i}). Otherwise, the s_c is computed as the number of edges in the shortest path s_p, which connects the two conditions in the disease hierarchy[10]. The idea of this rule is to figure out whether two patients are affected by similar disease conditions. For instance, *dilated cardiomyopathy* and *coronary artery* conditions of two patients can be considered the same since they both refer to heart-muscle failures. In this context, the experiences of consulted doctors/hospitals of this patient could be useful for another (Narducci et al. 2015).

- $s_t(p_{t_i}, p'_{t_j})$ is the similarity score between the treatment t_i of patient p and the treatment t_j of patient p' (see Formula (10)).
- α refers to the contribution of conditions and treatments to patients' similarity.
- β indicates the weight of the community (patients) and the ministry indicator.

$$s(p, p') = \alpha \frac{\sum_{i=1}^{k} \sum_{j=1}^{n} s_c(p_{c_i}, p'_{c_j})}{k + n} + (1 - \alpha) \frac{\sum_{i=1}^{z} \sum_{j=1}^{r} s_t(p_{t_i}, p'_{t_j})}{z + r} \qquad (8)$$

$$s_c(p_{c_i}, p'_{c_j}) = \begin{cases} log \frac{\#C}{\#p_{c_i}}, & if c_i = c_j \\ \frac{1}{sp(c_i, c_j)}, & otherwise \end{cases} \qquad (9)$$

$$s_t(p_{t_i}, p'_{t_j}) = \begin{cases} 1, & if t_i = t_j \\ 0, & otherwise \end{cases} \qquad (10)$$

Given an active patient p_i, the relevant doctors and hospitals for this patient can be estimated using Formulae (11) and (12).

$$scoreDoctor(d_z, p_i) = \sum_{j=1}^{p_i} s(p_i, p_j).r_j(d_z) \qquad (11)$$

$$scoreHospital(h_m, p_i) = \beta \left(\sum_{j=1}^{p} s(p_i, p_j) * r_j(h_m) \right) + (1 - \beta).q_i(h_m) \qquad (12)$$

Example 4 For the demonstration purposes, we present a simple example to show how relevant doctors and hospitals can be suggested to the patient using the mentioned approach. Assume an active patient p_i is having heart disease and suffering from condition c so-called *dilated cardiomyopathy*. This patient was applied the treatment t - *nitrates*. Now, he needs doctor and hospital recommendations that can effectively resolve his health problems. These recommendations can be generated based on the relevant information from other patients in the system. Assume two heart-disease patients (p_1 and p_2) have visited doctors X and Y from hospitals A and B respectively. The information of these patients is summarized in Table 5.

Patient p_1 has a higher similar disease condition to patient p_i compared to patient p_2 (p_i and p_1 have trouble with heart muscle failures). Therefore, we assume the distance in the disease hierarchy tree between p_i and p_1 is 2 and between p_i and p_2 is 3. Based on Formula (9), the condition similarities would be $s_p(c_i, c_1) = 1/2$ and $s_p(c_i, c_2) = 1/3$. We assume disease conditions and treatments have the same impacts on the patient similarity scores

---

[10] http://apps.who.int/classifications/icd10/browse/2010/en

Table 5 Example 4 - The information of the patients who are suffering heart disease stored in the system

			consulted doctor		visited hospital	
patient	condition	treatment	name	rating	name	rating
p_1	coronary artery	statins	X	4.1	A	4.2
p_2	mitral regurgitation	surgery	Y	4.9	B	4.3

(i.e., $\alpha = 0.5$), and the community and ministry have the same weights (i.e., $\beta = 0.5$). The ratings of the ministry for hospitals A and B are $q(A) = 4.0$ and $q(B) = 4.5$ respectively. The necessary calculations are presented below, which show that doctor X and hospital A are recommended to patient p_i.

$s_c(p_i, p_1) = 1/2$; $s_c(p_i, p_2) = 1/3$; $s_t(p_i, p_1) = 0$; $s_t(p_i, p_2) = 0$
$s(p_i, p_1) = (0.5 * 0.5)/2 = 0.125$; $s(p_i, p_2) = (0.5 * 0.33)/2 = 0.083$
$score Doctor(X, p_i) = 0.125 * 4.1 = \mathbf{0.513}\checkmark$
$score Doctor(Y, p_i) = 0.083 * 4.5 = 0.375$
$score Hospital(A, p_i) = 0.5 * 0.125 * 4.2 * 0.5 * 4 = \mathbf{0.525}\checkmark$
$score Hospital(B, p_i) = 0.5 * 0.083 * 4.8 * 0.5 * 4.5 = 0.45$

5 Evaluating health recommender systems

The most common evaluation method applied in the aforementioned recommendation approaches is *offline evaluation* (Trattner et al. 2018), estimating the prediction quality of a recommendation approach using existing data sets. With this method, *accuracy metrics* are used to compare recommendations determined by a recommender system with a predefined set of real-world user opinions (also known as *ground truth* (Shani and Gunawardana 2011)). For instance, Achananuparp et al. (Achananuparp and Weber 2016) constructed a real-world food consumption from MyFitnessPal's public food diary entries[11] and obtained group truth judgments of food substitutes from a crowdsourcing service. The authors used classification metrics *"precision", "mean average precision",* and *"normalized discounted cumulative gain"* to measure the method accuracy. Similar evaluation methods were applied in (Galeano and Paccanaro 2018; Hussein et al. 2012; Ueta et al. 2011; Yamanishi et al. 2012; Zhang et al. 2016), where metrics *"precision"* and *"recall"* were used to evaluate and compare the prediction performance of the recommendation algorithms. Besides classification metrics, *error metrics* (Trattner et al. 2018) were also employed to measure the error made by a recommender system when predicting an item rating (Galeano and Paccanaro 2018; Hussein et al. 2012; Narducci et al. 2015; Zhang et al. 2016). For instance, Narducci et al. (Narducci et al. 2015) carried out a preliminary evaluation, where the *"Mean Absolute Error"* was computed to compare their semantic approach based on the disease hierarchy to a simple string matching baseline. Another offline evaluation approach was *cross validation* (Dubitzky 2009) that allows to evaluate the performance of recommendation algorithms (Bresso et al. 2013; Elsweiler et al. 2017; Han et al. 2018; Zhang et al. 2016). Han et al. (Han et al. 2018) determined hyper-parameters for their model by performing a *temporal cross-validation*, which chronologically splits the data into train and test sets over the years. To consider item relevance and item position in a recommendation list, Achananuparp et

---

[11]https://www.myfitnesspal.com

al. (Achananuparp and Weber 2016) computed the "*discounted cumulative gain (DCG)*" metric based on the idea that items appearing lower in a recommendation list should be personalized by downgrading relevance values logarithmically (Trattner et al. 2018).

Compared to offline evaluation, much lesser number of studies employed *online evaluation* to test recommendation algorithms' accuracy in HRS (Berkovsky and Freyne 2010; Donciu et al. 2011; Ueta et al. 2011). The idea of online evaluation is to use A/B testing or laboratory studies to evaluate an algorithm, a user interface, or the whole system (Trattner et al. 2018). For instance, using a dataset of explicit recipe ratings, Berkovsky et al. (Berkovsky and Freyne 2010) conducted a user study to observe families' interaction with an experimental eHealth portal. This study aimed to uncover a recommendation strategy that could be most suitable for implementing an aggregation strategy in a group recommender system. Another approach was *"direct test"*, which was employed in some studies on drug recommendations (Doulaverakis et al. 2012; Mahmoud and Elbeh 2016). These tests were conducted with medical experts (e.g., doctors, clinicians, physicians, or nurses), where they were asked for feedback on the preciseness of recommendation outcomes. Mahmoud et al. (Mahmoud and Elbeh 2016) carried out a study in which experts evaluated recommendation results of the developed recommender system using a specific number of data sets. After collecting experts' feedback, the precision was calculated. This value indicates the exactness or the quality of recommendation results. A *true positive rate* shows that the expert agreed with the recommendation result. In contrast, a *false positive rate* indicates that the expert disagreed with the recommendation result (Mahmoud and Elbeh 2016).

In summary, most of the existing studies discussed in this article use evaluation methods that have been previously developed in traditional recommendation domains and mainly focus on evaluating the accuracy of recommendation algorithms. However, the quality of HRS should be further evaluated according to other features beyond the accuracy, such as *trust, causability, robustness, privacy, ethics, user satisfaction, uncertainty, effectiveness*, and *in-situ evaluation*. Up to now, how to employ evaluation methods considering the mentioned evaluation perspectives has remained an open issue. For further discussion on the mentioned perspectives, see Section 6.4.

6 Open issues for future work

Although the current literature has shown many benefits of HRS to improve their health conditions, there still exist some gaps regarding developing and evaluating HRS that need to be bridged. In the following, we discuss some research challenges that HRS face and corresponding solutions to tackle them.

6.1 Constructing user profiles

In HRS, besides *user preferences* that are typically used in recommender systems, further user information should be collected to obtain relevant, diverse, and precise recommendations. Thereof are demographic information, current health condition, diseases/allergies, treatments/surgeries/diagnoses experienced in the past, physical activities, nutrition needs, eating habits, feelings, and experiences. Although many sources exist to accumulate this information, recording such information could not avoid faults (Mika 2011). Hence, it is critical to require a standard concerning data formats, the authenticity of data sources, and automated update intervals (Schäfer et al. 2017) to ensure the quality of obtained information. Besides, the user profile parameters could be conflict with each other (e.g., user

preferences vs. health conditions). To guarantee optimal suggestions in terms of *balancing* between user satisfaction and healthiness, parameters in the user profile need to be deemed wholly and appropriately. In some cases, the parameters regarding health conditions should be taken precedence over those concerning user preferences. For instance, a user diagnosed with a high risk of having diabetes should not be recommended any food containing trans-fats, even this recommendation goes against his preference.

6.2 Early disease detection

Many reports show that patients suffering from chronic diseases or cancers are often not well perceived about their disease or treatment options until the disease situations fall in late stages (Schäfer et al. 2017). The late detection of such diseases causes a low probability of completely curing the disease, and in some cases, this could threaten patients' lives (e.g., the late stage of cancers). In such a context, besides assisting patients in finding suitable treatment methods (Davis et al. 2009; Nasiri et al. 2016), HRS should offer users a *health education* functionality that helps to improve the perception of users concerning the diseases. Stettinger et al. (Stettinger et al. 2020) developed an e-learing application so-called KNOWLEDGECHECKR[12], which provides intuitive learning contents and suggests learning units in a personalized fashion. With this application, helpful information concerning disease descriptions and related symptoms can be transferred to patients. Moreover, HRS should analyze the *underlying health condition* of patients and predict early diseases that users might face in the near future. Also, necessary diagnoses and information of corresponding healthcare professionals should be delivered to patients. For chronic and life-threatening diseases, such early disease detection can minimize disease complications and the treatment process' burdens.

6.3 Persuasive recommendations

One focus of HRS is to track users' daily activities and motivate them to adjust their routines or habits positively. However, it seems to be a big challenge to change habits that have become so deeply entrenched over the years (Tran et al. 2018). Therefore, researchers have recently paid attention to develop persuasive systems, in which various strategies and persuasion principles are discovered to encourage users to adopt and maintain beneficial behaviors and attitudes. For instance, Thomas et al. (Thomas et al. 2017) investigated argument-based approaches in which motivating arguments are created to change the eating habit of users healthily. These studies indicate that it is necessary to produce persuasive arguments based on user attributes such as age, gender, or personality. Although these studies show positive effects on the behavior changing of users, it does not guarantee full acceptance of changes. The argumentation-based approaches have been proved to be sufficiently effective for patients in the late stages of the disease, whereas they show a lower effect for the patients in the early stages of the disease (Nguyen and Masthoff 2008). This raises an open issue of developing arguments that are *strong, relevant,* and *convincing* enough to bring actual changes for those in the early phases of health risks. Besides, while many efforts have been conducted to estimate the arguments' perceived persuasiveness, measuring the *actual persuasiveness* of arguments is still an open issue. In fact, what people perceive to be persuasive is not necessarily what will persuade them to act. In the

---

[12]https://www.knowledgecheckr.com

healthcare domain, this can be interpreted that users might be unwilling to change their behavior, even though they are aware of the risks triggered by unhealthy habits (Nguyen and Masthoff 2008). For instance, although some people may perceive the harmful effects of smoking, they are not ready to give it up. On the other hand, changing users' behavior or attitude is a long-term process with plenty of steps. In this context, the question is *"how to generate persuasive arguments that motivate users as much as possible"*. In the healthy food domain, the answer to this question could be to develop food recommender systems, where theories from health psychology are integrated to stimulate users to comply with healthy eating behaviors (Schäfer and Willemsen 2019). One approach is to apply a simple change at a specific time until the user behavior becomes habitual. Another approach is to compare nutrients consumed by the user to the ones acquired from reliable sources (e.g., USDA, DACH[13]) (Snooks 2009).

6.4 Further aspects for evaluating health recommender systems

Typically, the evaluation of recommender systems emphasizes the accuracy metrics (Powers 2011) (see also Section 5). However, in the healthcare domain, recommender systems' quality needs to be measured based on aspects beyond the accuracy objectives (Valdez et al. 2016).

Trust is one of the most important criteria that should be considered when evaluating recommender systems (O'Donovan and Smyth 2005). This is even more critical for HRS to convince patients to follow health-related recommendations. This aspect can be enhanced by providing explanations for recommendations (Tran et al. 2019). Similar to other domains, explanations in the healthcare domain should show how a suggestion has been created for the user (Elahi et al. 2014), e.g., "According to the tests you did last week, we have detected that the level of uric acid in your blood is still really high. Therefore, r_i seems to be the most appropriate recipe for you since it has no ingredients containing purines". Besides, effective visualizations should be included in HRS to further explain recommended items (Valdez et al. 2016). For instance, in food recommender systems, a table showing the description of the nutrition value of food items should be provided to the users to emphasize the healthiness of a recommended recipe (Tran et al. 2018).

Causability helps users understand why specific recommendations have been made. This criterion is useful in many domains and especially crucial in the medical domain to enhance trust in the results and enable domain experts to retrace, understand, and explain why a particular recommendation was given. This does not necessarily mean that everything must be explained automatically, but that a domain expert has a chance to understand it on demand. To measure the understandability of recommendations, the concept of causability can be helpful. In the same way that usability encompasses measurements for the quality of use, causability encompasses measurements for the quality of explanations (Holzinger et al. 2019).

Robustness is related to the *trustworthiness* of a recommender system. In HRS, sometimes, end-users could not be differentiated from potential attackers, which causes a degradation of trust in the objectivity and accuracy of the system (Valdez et al. 2016). To ensure secure HRS for users, future studies should model potential attacks and investigate the impacts of such attacks on recommendation algorithms (Mobasher et al. 2007).

---

[13]http://www.sge-ssn.ch.

Privacy is referred to as the ability of HRS to preserve patients' preferences and medical information. The leak of such information raises the doubts of patients and consequently decreases the willingness to share their sensitive medical data with HRS (Valdez et al. 2016). The most common approach to address the privacy concern is *data encryption* that provides data confidentiality while utilizing the user data to generate precise recommendations (Hoens et al. 2010). However, this method requires highly overhead computation- and communication-wise, which significantly decreases the performance of HRS (Verhaert et al. 2018). Although there exist some studies to improve the data encryption approach, some of them still face the issue concerning the low efficiency of the system (Hoens et al. 2010; Verhaert et al. 2018). Up to now, developing HRS ensuring trade-offs between the privacy and the high performance of recommendation algorithms is still an open issue.

Ethics has been raised in recommender systems to help users pick up morally appropriate items during the post-recommendation process (Tang and Winoto 2016). In HRS, ethics should be considered more strictly to prevent recommendations from directions that could harm the healthiness of patients (Valdez et al. 2016). The principle of *"first do no harm"* should be kept in mind when developing HRS to minimize potential risks and maximize benefits for users. The healthiness of patients is the most crucial criterion when creating recommendations, even this might be against patients' preferences (Tang and Winoto 2016).

User satisfaction with recommendations can be different depending on user diversity. Some studies have been performed recently to have a deep look at modeling user satisfaction for the purpose of predicting satisfaction models (Chen et al. 2019; Nguyen et al. 2017). In HRS, it would make sense to investigate the relationship between health-related recommendations and users' satisfaction from different user groups, e.g., patients, doctors, nurses, physicians, and medical researchers (Valdez et al. 2016). The differences in expertise, overview knowledge, and recommendation tasks of these users could influence their satisfaction with recommended items.

Uncertainty in HRS links to potential risks, such as imprecise predictions since user preferences are not always captured well, or the inability to find a perfect pattern because of incomplete data. The risks could result in a reduced quality of the patient's life. Therefore, when developing HRS, system designers should find ways to visualize uncertainty in a set of recommendations, allowing users to evaluate the option adequately before making a final decision (Valdez et al. 2016).

Effectiveness is referred to as the ability of HRS to help patients meet their desired changes in health. To measure this aspect, we need to consider which health parameters to be assessed or which medical tests to be employed to ensure medical effectiveness. For instance, in HRS that support the lose-weight targets of users, the effectiveness should be assessed based on both short-term and long-term recommendations. The reason is that, in some cases, short-term recommendations could burden or conflict with long-term ones. For instance, crash diets could help a patient lose weight quickly since it cuts the calories too low and makes drastic changes regarding food types to be consumed. However, this reduces the metabolism of the patient's body and consequently burdens the long-term weight loss (Valdez et al. 2016).

In-Situ Evaluation indicates real-life non-laboratory settings that have to be evaluated to prove the worthiness of HRS. This evaluation paradigm should be able to precisely evaluate the ability of HRS to improve the quality of care (concerning *accuracy, relevance,* and *early diagnosis*) and reduce the cost of care. Besides, it should be capable of evaluating the robustness to false information and the ability to consider potential health risks based on various dimensions (e.g., age, culture, ethnicity, etc.). Moreover, long-term behavioral

effects must also be investigated in-situ evaluation to address the complexity of health and health behaviors (Schäfer et al. 2017).

6.5 Bundle recommendation

In the healthy food domain, users might require recommendations of a complete meal with the combination of many recipes or a food schedule for more than one day (e.g., foods for next week). This issue is known as *bundle recommendation*, which is a new research branch of recommender systems. The idea here is to recommend a sequence of items instead of separated ones. Recommending a complete meal is quite complicated since the system has to consider not only the preferences of users but also other aspects, such as the *meal variety, weather and season*, *the healthiness of recipes, health problems,* or *nutrition needs*. Thus, approaches to generate bundle recommendations in the healthy food domain have remained an open issue.

6.6 Group decision making

For some scenarios (i.e., diet recommendation), recommended items could involve groups of users rather than individual users (e.g., recommend a menu for a Christmas party). The current literature shows a limited number of studies on food recommender systems for groups. Therefore, it is still an open topic that needs to be analyzed in future research. Group recommender systems usually attach the requirements/preferences of different users into group recommendation. This is the crucial idea discussed in many related studies (Berkovsky and Freyne 2010; Felfernig et al. 2018; O'Connor et al. 2001). Recommending a joint meal for a group of users is a complicated task since different goals and dietary constraints of group members should be taken into account. While we have a solution for merging the constraints (Atas et al. 2019), a solution for merging goals is still an open issue. Besides, recommendations generated for groups should assure fairness among group members, which means negotiation and argumentation mechanisms have to be developed to support group members in expressing acceptable trade-offs (Felfernig et al. 2014). For instance, in a meal plan for a week, users' preferences ignored in previous meals should have a stronger influence on the upcoming meals. On the other hand, although different aggregation approaches have been applied to generate group recommendations, they do not ensure that recommended items reflect a high agreement level among group members (Castro et al. 2015). In this context, a consensus making process is needed to bring individual preferences closer to each other before delivering group recommendations. Further issues need to be considered to accelerate such a process. One promising solution is to enrich user interfaces that allow group members to share their preferences (Nguyen and Ricci 2017). Besides, psychological aspects (e.g., personality and emotions) beyond group members' preferences are also crucial to be taken into account in group decision making. This draws an open topic regarding the influence of group members' personality and emotions on group recommendation strategies (Quijano-Sanchez et al. 2013).

7 Conclusion

Health recommender systems have emerged as tools to support patients and healthcare professionals to make better health-related decisions. In this article, we have given insights into recommendation scenarios offered by these systems, such as food recommendation, drug

recommendation, health status prediction, physical activity recommendation, and health-care professional recommendation. For each recommendation scenario, various algorithms have been employed, which are based on recommendation techniques (e.g., CF, CB, KB, HyR, and context-based recommendations) or machine learning techniques (e.g., classification, clustering, decision tree, natural language processing, logic programming, ontologies, and semantic technologies). Although the proposed HRS bring many benefits in terms of health-related improvements, there still exist a number of challenges that need to be tackled for the better development of these systems in the future.

References

Aberg, J. (2006). Dealing with malnutrition: a meal planning system for elderly. In *AAAI Spring symposium: argumentation for consumers of healthcare, (pp. 1–7). AAAI*.

Achananuparp, P., & Weber, I. (2016). Extracting food substitutes from food diary via distributional similarity. In *Proceedings of the. 10th ACM Conference on Recommender Systems, (pp. 1–5)*.

Ali, S.I., Amin, M.B., Kim, S., Lee, S. (2018). A hybrid framework for a comprehensive physical activity and diet recommendation system. In Mokhtari, M., Abdulrazak, B., Aloulou, H. (Eds.) *Smart homes and health telematics, designing a better future: urban assisted living (ICOST 2018), (pp 101–109). Springer International Publishing, Cham*.

Atas, M., Tran, T.N.T., Felfernig, A., Polat-Erdeniz, S., Samer, R., Stettinger, M. (2019). Towards similarity-aware constraint-based recommendation. In *Advances and trends in artificial intelligence, lecture notes in computer science, (pp. 287–299). Springer*.

Bankhele, S., Mhaske, A., Bhat, S. (2017). V., s.: a diabetic healthcare recommendation system. *International Journal of Computer Applications*, *167*, 14–18.

Bao, Y., & Jiang, X. (2016). An intelligent medicine recommender system framework. In *2016 IEEE 11Th conference on industrial electronics and applications (ICIEA), (pp. 1383–1388)*.

Berkovsky, S., & Freyne, J. (2010). Group-based recipe recommendations: Analysis of data aggregation strategies. In *Proceedings of the Fourth ACM Conference on Recommender Systems, RecSys '10, (pp 111–118). ACM, New York, NY, USA*.

Bresso, E., Grisoni, R., Marchetti, G., Karaboga, A., Souchet, M., Devignes, M.D., Smail, M. (2013). Integrative relational machine-learning approach for understanding drug side-effect profiles. *BMC bioinformatics*, *14*, 207.

Burke, R. (2000). Knowledge-based recommender systems. In *Encyclopedia of library and information systems, vol. 69, (pp. 180–200). Marcel dekker*.

Cami, A., Arnold, A., Manzi, S., Reis, B. (2011). Predicting adverse drug events using pharmacological network models. *Science translational medicine*, *3*, 114–127.

Castro, J., Quesada, F.J., Palomares, I., Martínez-López, L. (2015). A consensus-driven group recommender system. *International Journal of Intelligent Systems*, *30*(8), 887–906.

Charkhat Gorgich, E.A., Barfroshan, S., Ghoreishi, G., Yaghoobi, M. (2015). Investigating the causes of medication errors and strategies to prevention of them from nurses and nursing student viewpoint. *Global Journal of Health Science*, *8*, 220.

Chen, L., & Pu, P. (2012). Critiquing-based recommenders: survey and emerging trends. *User Modeling and User-Adapted Interaction*, *22*(1-2), 125–150.

Chen, L., Yang, Y., Wang, N., Yang, K., Yuan, Q. (2019). How serendipity improves user satisfaction with recommendations? a large-scale user evaluation. In *The world wide web conference, WWW'19, (pp 240–250). Association for Computing Machinery, New York, NY, USA.*

Chen, R.C., Chiu, J.Y., Bau, C.T. (2011). The recommendation of medicines based on multiple criteria decision making and domain ontology — an example of anti-diabetic medicines. In *The Proceedings of the International Conference on Machine Learning and Cybernetics, vol. 1, (pp. 27–32).*

Chen, R.C., Huang, Y.H., Bau, C.T., Chen, S.M. (2012). A recommendation system based on domain ontology and SWRL for anti-diabetic drugs selection. *Expert Systems with Applications, 39*(4), 3995–4006.

Colberg, S., Sigal, R., Yardley, J., Riddell, M., Dunstan, D., Dempsey, P., Horton, E., Castorino, K., Tate, D. (2016). Physical activity/exercise and diabetes: a position statement of the american diabetes association. *Diabetes Care, 39,* 2065–2079.

Davis, D.A., Chawla, N.V., Christakis, N.A., Barabási, A.L. (2009). Time to care: a collaborative engine for practical disease prediction. *Data Mining and Knowledge Discovery, 20,* 388–415.

Deshpande, T., & Butte, A. (2011). Exploiting drug-disease relationships for computational drug repositioning. *Briefings in bioinformatics, 12,* 303–11.

Dharia, S., Jain, V., Patel, J., Vora, J., Chawla, S., Eirinaki, M. (2016). Pro-fit: a personalized fitness assistant framework. In *SEKE.*

Donciu, M., Ionita, M., Dascalu, M. (2011). Trausan-matu, S.: The runner - recommender system of workout and nutrition for runners. *2011 13th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC),* (pp. 230–238).

Doulaverakis, C., Nikolaidis, G., Kleontas, A., Kompatsiaris, I. (2012). Galenowl: Ontology-based drug recommendations discovery. *Journal of biomedical semantics, 3,* 14.

Doulaverakis, C., Nikolaidis, G., Kleontas, A., Kompatsiaris, I. (2014). Panacea, a semantic-enabled drug recommendations discovery framework. *Journal of biomedical semantics, 5,* 13.

Dubitzky, W. (2009). *Fundamentals of data mining in genomics and proteomics.* Berlin: Springer.

Elahi, M., Ge, M., Ricci, F., Massimo, D., Berkovsky, S. (2014). Interactive food recommendation for groups. In *RECSYS, vol. 1247.*

Elsweiler, D., Trattner, C., Harvey, M. (2017). Exploiting food choice biases for healthier recipe recommendation. In *Proceedings of the 40th Conference on Research and Development in Information Retrieval, SIGIR '17, (pp 575–584). Association for Computing Machinery, NY, USA.*

Faiz, I., Mukhtar, H., Khan, S. (2014). An integrated approach of diet and exercise recommendations for diabetes patients. In *IEEE 16th International Conference on e-Health Networking, Applications and Services (Healthcom), pp 537–542.*

Felfernig, A., Atas, M., Helic, D., Tran, T.N.T., Stettinger, M., Samer, R. (2018). *Group Recommender systems: An Introduction, chap. Algorithms for Group Recommendation,* (pp. 27–58). Berlin: Springer.

Felfernig, A., & Burke, R. (2008). Constraint-based recommender systems: Technologies and research issues. In *Proceedings of the 10th International Conference on Electronic Commerce, ICEC'08, (pp. 3:1–3:10). ACM, New York, NY, USA.*

Felfernig, A., & Gula, B. (2006). An empirical study on consumer behavior in the interaction with knowledge-based recommender applications. In *The 8th IEEE international conference on e-commerce technology and the 3rd IEEE international conference on enterprise computing, e-commerce, and e-services (CEC/EEE'06), (pp. 37–37).*

Felfernig, A., Stettinger, M., Ninaus, G., Jeran, M., Reiterer, S., Falkner, A.A., Leitner, G., Tiihonen, J. (2014). Towards open configuration. In *Proceedings of the 16th International Configuration Workshop, Novi Sad, Serbia, September 25-26, 2014., (pp. 89–94).*

Fliri, A., Loging, W., Thadeio, P. (2006). Analysis of drug-induced effect patterns to link structure and side effects of medicines. *Nature chemical biology, 1,* 389–97.

Fukuzaki, M., Seki, M., Kashima, H., Sese, J. (2009). Side effect prediction using cooperative pathways. In *IEEE International conference on bioinformatics and biomedicine, (pp. 142–147).*

Galeano, D., & Paccanaro, A. (2018). A recommender system approach for predicting drug side effects. In *IJCNN 2018: International joint conference on neural networks, (pp. 1–7). IEEE xplore.*

Ge, M., Ricci, F., Massimo, D. (2015). Health-aware food recommender system. In *Proceedings of the 9th ACM Conference on Recommender Systems, RecSys '15, (pp. 333–334). Association for Computing Machinery, New York, NY, USA.*

Gujar, D., Biyani, R., Bramhane, T., Bhosale, S., Vaidya, T.P. (2018). Disease prediction and doctor recommendation system. *International Research Journal of Engineering and Technology (IRJET), 5,* 3207–3209.

Gurwitz, J., Field, T., Harrold, L., Rothschild, J., DeBellis, K., Seger, A., Cadoret, C., Fish, L., Garber, L., Kelleher, M., Bates, D. (2003). Incidence and preventability of adverse drug events among older persons in the ambulatory setting. *JAMA: Journal of the American Medical Association, 289,* 1107–16.

Han, Q., Ji, M., Martínez de Rituerto de Troya, I., Gaur, M., Zejnilovic, L. (2018). A hybrid recommender system for patient-doctor matchmaking in primary care. In *The 5th IEEE international conference on data science and advanced analytics (DSAA), (pp. 1–10)*.

Hoens, T.R., Blanton, M., Chawla, N.V. (2010). Reliable medical recommendation systems with patient privacy. In *Proceedings of the 1st ACM International Health Informatics Symposium, IHI '10, (pp. 173–182). Association for Computing Machinery, New York, NY, USA*.

Holzinger, A., Carrington, A., Müller, H. (2019). Measuring the quality of explanations: The system causability scale (scs) - comparing human and machine explanations. *KI - Künstliche Intelligenz (German Journal of Artificial intelligence) Special Issue on Interactive Machine Learning 24*(2).

Holzinger, A., Valdez, A.C., Ziefle, M. (2016). Towards interactive recommender systems with the doctor-in-the-loop. In Weyers, B., & Dittmar, A. (Eds.) *Mensch und computer 2016 – workshopband. Gesellschaft für Informatik e.V., Aachen*.

Huang, L.C., Wu, X., Chen, J. (2011). Predicting adverse side effects of drugs. *BMC genomics 12 Suppl*, 5, S11.

Hussein, A.S., Omar, W.M., Li, X., Ati, M. (2012). Accurate and reliable recommender system for chronic disease diagnosis. In *The first international conference on global health challenges - global health 2012, (pp. 113–118)*.

Jannach, D. (2011). Recommender systems : an introduction Cambridge University Press.

Kaur, H., & Raghava, G.P.S. (2003). A neural-network based method for prediction of gamma-turns in proteins from multiple sequence alignment. *Protein Society*, 12, 923–929.

Lafta, R., Zhang, J., Tao, X., Li, Y., Tseng, V.S. (2015). An intelligent recommender system based on short-term risk prediction for heart disease patients. In *2015 IEEE/WIC/ACM International conference on web intelligence and intelligent agent technology (WI-IAT), vol. 3, (pp. 102–105)*.

Lops, P., Gemmis, M.D., Semeraro, G. (2011). *Content-based Recommender systems: State of the Art and Trends*, (pp. 73–105). Boston: Springer.

Mahmoud, N., & Elbeh, H. (2016). IRS-T2D: Individualize recommendation system for type2 diabetes medication based on ontology and SWRL. In *Proceedings of the 10th International Conference on Informatics and Systems, INFOS '16, (pp. 203–209). Association for Computing Machinery, NY, USA*.

Medvedeva, O., Knox, T., Paul, J. (2007). Diatrack: web-based application for assisted decision-making in treatment of diabetes. *Journal of Computing Sciences in Colleges*, 23, 154–161.

Mika, S. (2011). Challenges for nutrition recommender systems. In *Workshop Proceedings on Context Aware Intelligent Assistance, (pp. 25–33). CEUR-WS.org*.

Mobasher, B., Burke, R., Bhaumik, R., Williams, C. (2007). Toward trustworthy recommender systems: An analysis of attack models and algorithm robustness. *ACM Transactions on Internet Technology*, 7(4), 23–es.

Narducci, F., Musto, C., Polignano, M., de Gemmis, M., Lops, P., Semeraro, G. (2015). A recommender system for connecting patients to the right doctors in the healthnet social network. In *WWW'2015 Companion, (pp. 81–82)*.

Nasiri, M., Minaei, B., Kiani, A. (2016). Dynamic recommendation: Disease prediction and prevention using recommender system. International Journal of Basic Science in Medicine, (pp. 13–17).

Nguyen, H., & Masthoff, J. (2008). Designing persuasive dialogue systems: Using argumentation with care. In *H. Oinas-kukkonen, p.f.v. hasle, M. Harjumaa, K. Segerståhl, P. Øhrstrøm (eds.) Persuasive Technology, Third International Conference, PERSUASIVE 2008, Oulu, Finland, June 4-6, 2008. Proceedings, Lecture Notes in Computer Science, vol. 5033, (pp. 201–212). Springer*.

Nguyen, T., Harper, F., Terveen, L., Konstan, J. (2017). User personality and user satisfaction with recommender systems. *Information Systems Frontiers*, 20, 1–17.

Nguyen, T.N., & Ricci, F. (2017). A chat-based group recommender system for tourism. *Information Technology & Tourism*, 18(2), 5–28.

Noh, H.J., Kwak, M.J., Han, I.G. (2003). Handling incomplete data problem in collaborative filtering system. *Journal of Intelligent Information Systems*, 9, 51–63.

O'Connor, M., Cosley, D., Konstan, J.A., Riedl, J. (2001). Polylens: a recommender system for groups of users. In *Proceedings of the Seventh Conference on European Conference on Computer Supported Cooperative Work, ECSCW'01, (pp. 199–218). Kluwer Academic Publishers, Norwell, USA*.

O'Donovan, J., & Smyth, B. (2005). Trust in recommender systems. In *International Conference on Intelligent User Interfaces, Proceedings IUI, (pp. 167–174)*.

Orgun, B., & Vu, J. (2006). Hl7 ontology and mobile agents for interoperability in heterogeneous medical information systems. *Computers in Biology and Medicine*, 36(7), 817–836. Special Issue on Medical Ontologies.

Özcift, A. (2011). Random forests ensemble classifier trained with data resampling strategy to improve cardiac arrhythmia diagnosis. *Computers in biology and medicine*, 41, 265–71.

Park, J., Tyan, H., Kuo, C.C.J. (2006). Internet traffic classification for scalable qos provision. In *2006 IEEE International conference on multimedia and expo, (pp. 1221–1224)*.

Pincay, J., Terán, L., Portmann, E. (2019). Health recommender systems: a state-of-the-art review. In *2019 Sixth international conference on edemocracy egovernment (ICEDEG), (pp. 47–55)*.

Powers, D.M.W. (2011). Evaluation: from precision, recall and f-measure to roc., informedness, markedness & correlation. *Journal of Machine Learning Technologies*, *2*(1), 37–63.

Quijano-Sanchez, L., Recio-Garcia, J.A., Diaz-Agudo, B., Jimenez-Diaz, G (2013). Social factors in group recommender systems. *ACM Transactions on Intelligent Systems and Technology*, *4*(1), 8:1–8:30.

Rehman, F., Khalid, O., Haq, N., Khan, A.U.R., Bilal, K., Madani, S. (2017). Diet-right: a smart food recommendation system. *KSII on Internet and Information Systems*, *11*, 2910–2925.

Ricci, F., Rokach, L., Shapira, B., Kantor, P.B. (2010). *Recommender systems handbook*, 1st edn. Berlin: Springer.

Robertson, A., Tirado, M., Lobstein, T., Jermini, M., Knai, C., Jensen, J., Ferro-Luzzi, A., James, W. (2004). Food and health in europe: a new basis for action. WHO regional publications. European series (pp. i–xvi), 1.

Rodríguez, A., Jiménez, E., Fernández, J., Eccius, M., Miguel Gómez, J., Alor-Hernández, G., Posada-Gómez, R., Laufer, C. (2009). Semmed: Applying semantic web to medical recommendation systems. *2009 First International Conference on Intensive Applications and Services*, (pp. 47–52).

Rokicki, M., Herder, E., Demidova, E. (2015). Vol What's on my plate: Towards recommending recipe variations for diabetes patients. In Cristea, A.I., Masthoff, J., Said, A., Tintarev, N. (Eds.) *UMAP Workshops, CEUR Workshop Proceedings, vol 1388. CEUR-WS.org*.

Sahoo, A.K., Pradhan, C., Barik, R.K., Dubey, H. (2019). Deepreco: Deep learning based health recommender system using collaborative filtering. *Computation*, *7*, 25.

Sánchez-Bocanegra, C.L., Sánchez-Laguna, F., Sevillano, J.L. (2015). Introduction on health recommender systems. *Methods in molecular biology (Clifton N.J.)*, *1246*, 131–46.

Schäfer, H., Hors-Fraile, S., Karumur, R.P., Calero Valdez, A., Said, A., Torkamaan, H., Ulmer, T., Trattner, C. (2017). Towards health (aware) recommender systems. In *Proceedings of the 2017 International Conference on Digital Health, DH '17, (pp. 157–161). ACM, New York, NY, USA*.

Schäfer, H., & Willemsen, M. (2019). Rasch-based tailored goals for nutrition assistance systems. In *Proceedings of the 24th International Conference on Intelligent User Interfaces, IUI '19, (pp. 18–29). Association for Computing Machinery, New York, NY, USA*.

Sezgin, E., & Özkan, S. (2013). A systematic literature review on health recommender systems. In *2013 E-health and bioengineering conference, EHB 2013, (pp. 1–4)*.

Shani, G., & Gunawardana, A. (2011). Evaluating recommendation systems. In Ricci, F., Rokach, L., Shapira, B., Kantor, P.B. (Eds.) *Recommender systems handbook, (pp. 257–297). Springer*.

Shimada, K., Takada, H., Mitsuyama, S., Ban, H., Matsuo, H., Otake, H., Kunishima, H., Kanemitsu, K., Kaku, M. (2005). Drug-recommendation system for patients with infectious diseases. *Annual Symposium proceedings. AMIA Symposium*, (pp. 11–12).

Snooks, M. (2009). Health psychology: Biological, Psychological, and Sociocultural Perspectives, chapter 5: Applications of Health Psychology to Eating behaviors: Improving health through nutritional changes Jones & Bartlett Learning.

Stark, B., Knahl, C., Aydin, M., Elish, K. (2019). A literature review on medicine recommender systems. *International journal of advanced computer science and applications 10*(8).

Stark, B., Knahl, C., Aydin, M., Samarah, M., Elish, K.O. (2017). Betterchoice: a migraine drug recommendation system based on neo4j. In *2017 2Nd IEEE international conference on computational intelligence and applications (ICCIA), (pp. 382–386)*.

Stettinger, M., Tran, T., Pribik, I., Leitner, G., Felfernig, A., Samer, R., Atas, M., Wundara, M. (2020). KNOWLEDGECHECKR: Intelligent techniques for counteracting forgetting. In *G.d. giacomo, A. Catalá, B. Dilkina, M. Milano, S. Barro, A. Bugarín, J. Lang (eds.) ECAI 2020 - 24th European Conference on Artificial Intelligence, 29 August-8 September 2020, Santiago de Compostela, Spain, August 29 - September 8, 2020 - Including 10th Conference on Prestigious Applications of Artificial Intelligence (PAIS 2020), Frontiers in Artificial Intelligence and Applications, vol. 325, (pp. 3034–3039). IOS Press*.

Tang, T.Y., & Winoto, P. (2016). I should not recommend it to you even if you will like it: the ethics of recommender systems. *New Review of Hypermedia and Multimedia*, *22*(1-2), 111–138.

Teng, C.Y., Lin, Y.R., Adamic, L.A. (2012). Recipe recommendation using ingredient networks. In *Proceedings of the 4th Annual ACM Web Science Conference, WebSci '12, (pp. 298–307). Association for Computing Machinery, New York, NY, USA*.

Thomas, R.J., Masthoff, M., Oren, N., de Vries, P.W., Oinas-Kukkonen, H., Siemons, L., Jong, N.B., van Gemert-Pijnen, L. (2017). Adapting healthy eating messages to personality. In *Persuasive Technology: Development and Implementation of Personalized Technologies to Change Attitudes and Behaviors*

- 12th International Conference, PERSUASIVE 2017, Amsterdam, The Netherlands, April 4-6, 2017, Proceedings, Lecture Notes in Computer Science, vol. 10171, (pp. 119–132). Springer.

Tran, T.N.T., Atas, M., Felfernig, A., Le, V.M., Samer, R., Stettinger, M. (2019). Towards social choice-based explanations in group recommender systems. In *Proceedings of the 27th ACM Conference on User Modeling, Adaptation and Personalization, UMAP '19, (pp. 13–21). Association for Computing Machinery, New York, NY, USA.*

Tran, T.N.T., Atas, M., Felfernig, A., Stettinger, M. (2018). An overview of recommender systems in the healthy food domain. *Journal of Intelligent Information Systems*, *50*(3), 501–526.

Trattner, C., & Elsweiler, D. (2017). Food recommender systems: Important contributions, challenges and future research directions. CoRR abs/1711.02760.

Trattner, C., Said, A., Boratto, L., Felfernig, A. (2018). *Evaluating Group Recommender Systems*, (pp. 59–71). Manhattan: Springer International Publishing Cham.

Ueta, T., Iwakami, M., Ito, T. (2011). A recipe recommendation system based on automatic nutrition information extraction. In *Proceedings of the 5th international conference on knowledge science, engineering and management, KSEM'11, (pp. 79–90). Springer: Berlin.*

Valdez, A.C., Ziefle, M., Verbert, K., Felfernig, A., Holzinger, A. (2016). Recommender systems for health informatics: State-of-the-art and future perspectives. *Lecture Notes in Computer Science*, *9605*.

Verbert, K., Manouselis, N., Ochoa, X., Wolpers, M., Drachsler, H., Bosnic, I., Duval, E. (2012). Context-aware recommender systems for learning: a survey and future challenges. *IEEE Transactions on Learning Technologies*, *5*(4), 318–335.

Verhaert, D., Nateghizad, M., Zekeriya, E. (2018). An efficient privacy-preserving recommender system for e-healthcare system. In *International conference on security and cryptography, (pp. 1–12).*

Wiesner, M., & Pfeifer, D. (2014). Health recommender systems: Concepts, requirements, technical basics and challenges. *International Journal of Environmental Research and Public Health*, *11*, 2580–2607.

Yamanishi, Y., Pauwels, E., Kotera, M. (2012). Drug side-effect prediction based on the integration of chemical and biological spaces. *Journal of chemical information and modeling, 52.*

Yang, L., Hsieh, C., Yang, H., Pollak, J., Dell, N., Belongie, S., Cole, C., Estrin, D. (2017). Yum-me: A personalized nutrient-based meal recommender system. *ACM Transactions on Information Systems 36*(1).

Zhang, Q., Zhang, G., Lu, J., Wu, D. (2015). A framework of hybrid recommender system for personalized clinical prescription. In *2015 10Th international conference on intelligent systems and knowledge engineering (ISKE), (pp. 189–195).*

Zhang, W., Zou, H., Luo, L., Liu, Q., Wu, W., Xiao, W. (2016). Predicting potential side effects of drugs by recommender methods and ensemble learning. *Neurocomputing*, *173*(P3), 979–987.

Zhang, Y., Chen, M., Huang, D., Wu, D., Li, Y. (2016). idoctor: Personalized and professionalized medical recommendations based on hybrid matrix factorization. *Future Generation Computer Systems*, *66*, 30–35.

Affiliations

Thi Ngoc Trang Tran[1] ⬤ · Alexander Felfernig[1] · Christoph Trattner[2] · Andreas Holzinger[3]

Alexander Felfernig
alexander.felfernig@ist.tugraz.at

Christoph Trattner
christoph.trattner@uib.no

Andreas Holzinger
andreas.holzinger@medunigraz.at

[1]   Institute of Software Technology, Graz University of Technology, Graz, Austria

[2]   Information Science and Media Studies Department, University of Bergen, Bergen, Norway

[3]   Institute for Medical Informatics, Statistics and Documentation, Medical University of Graz, Graz, Austria