Survey on Recommender Systems for Biomedical Items in Life and Health Sciences

MATILDE PATO*, FIT-ISEL, Departamento de Engenharia Electrónica e Telecomunicações e de Computadores, Portugal and LASIGE, Departamento de Informática, Faculdade de Ciências, Universidade de Lisboa, Portugal
MÁRCIA BARROS, LASIGE, Departamento de Informática, Faculdade de Ciências, Universidade de Lisboa, Portugal
FRANCISCO M. COUTO, LASIGE, Departamento de Informática, Faculdade de Ciências, Universidade de Lisboa, Portugal

The generation of biomedical data is of such a magnitude that its retrieval and analysis have posed several challenges. A survey of recommender system (RS) approaches in biomedical fields is provided in this analysis, along with a discussion of existing challenges related to large-scale biomedical information retrieval systems. We collect original studies, identify entities, models, and how knowledge graphs (KG) can improve results. As a result, most of the papers used model-based collaborative filtering algorithms, most of the available datasets did not follow the standard format < user, item, rating >, and regarding qualitative evaluations of RSs use mainly classification metrics. Finally, we have assembled and coded a unique dataset of 60 papers — Sur-RS4BioT, available for download at DOI:10.34740/kaggle/ds/2346894

CCS Concepts: • **Information systems** → **Collaborative filtering**; *Similarity measures*; **Personalization**; **Information extraction**; • **Applied computing** → **Health informatics**; **Genetics**; **Bioinformatics**.

Additional Key Words and Phrases: Recommender Systems, Biomedical Informatics, Information Retrieval, Knowledge-graphs, Reproducibility datasets

1 INTRODUCTION

The growth of the Internet has transformed medicine, enabling healthcare to be delivered more efficiently, improving patient outcomes, and increasing access to medical information and services. As technology advances, we can expect to see even more innovative uses of the Internet in medicine, with significant implications for healthcare delivery and patient outcomes.

The data collected by smart medical devices and clinical databases can be used to improve their recommendations. The large amounts of data generated by these devices can be analyzed using Artificial Intelligence (AI) and Machine Learning (ML) algorithms to identify patterns and make more accurate predictions about a patient's health status and treatment outcomes. Multi-patient data analysis will enable healthcare organizations to identify trends and patterns that might otherwise go undetected. It may help to identify new therapies, improve the precision of diagnoses and optimize individual patient care plans. In addition to improving patient outcomes, data from smart medical devices and clinical databases can improve the design of future medical devices and

Authors' addresses: Matilde Pato, matilde.pato@isel.pt, FIT-ISEL, Departamento de Engenharia Electrónica e Telecomunicações e de Computadores, Rua Conselheiro Emídio Navarro, 1, Lisboa, Portugal, 1959-007 and LASIGE, Departamento de Informática, Faculdade de Ciências, Universidade de Lisboa, Lisboa, 1749–016, Portugal; Márcia Barros, LASIGE, Departamento de Informática, Faculdade de Ciências, Universidade de Lisboa, Lisboa, 1749–016, Portugal, marciabarros@edu.ulisboa.pt; Francisco M. Couto, LASIGE, Departamento de Informática, Faculdade de Ciências, Universidade de Lisboa, Lisboa, 1749–016, Portugal.

ACM 0360-0300/2024/1-ART
https://doi.org/10.1145/3639047

treatments. By identifying patterns and trends in patient data, manufacturers can design more effective devices and treatments tailored to individual patient needs.

In addition, due to technological advances, patients' data (treatment history and genetic data) and novel drugs, researchers have focused their attention on the application of recommender system (RS) in life and health sciences. Similar to other applications, RS could help doctors and researchers make a better-informed decision from understanding users' tastes and past experiences.

To make it clearer, let us begin by defining biomedical items as all entities that belong to the biomedical fields and are characterized by attributes that can be modeled, and an entity of research, where biomedical field is defined as an area that explores the effects of drugs and medical techniques on biological systems. The purpose of this survey is to collect information on the state of RS for biomedical items over the last decade, in order to answer the following research questions (RQs):

RQ1: What are the real experiences with RS for biomedical items? Which kind of users or items are used?

RQ2: Which recommender techniques are being used across different biomedical items?

RQ3: What role does the knowledge graph-based play as side information in RS?

RQ4: What is the best way to evaluate the RS?

In order to address these questions comprehensively, we take a multidisciplinary approach to the existing RS solutions for biomedical items and compare them according to some criteria. While other survey papers focus on the existing solutions for healthcare providers, our survey tries to understand which RS approach is most frequent, and if a knowledge graph (KG) is explored, the type of evaluation, the source of datasets, the availability of the datasets, and whether they are public or not. Faced to the specificity of these items (biomedical), researchers do not have available datasets to assess RS because there are no datasets to guide their choices [6]. We hope to highlight the importance of RS in this topic with this survey. Together, we will demonstrate existing solutions and give researchers a glimpse into the research of RS.

The remainder of this manuscript is structured as follows. In section 2 we give an overview of RS in general: methods, algorithms, KG-based RS, evaluation methods and data. Section 3 describes the methodology used to select the most relevant papers on the RS field in biomedical items. Finally, in section 4 we present the results, discuss some of the limitations associated and draw conclusions in section 5. The conclusions section also indicates our perspective about the near future and challenges addressed to be with KG-based recommendation.

2 BACKGROUND

2.1 Recommender systems: concepts

RSs are information filtering systems that suggest items to users based on their prior knowledge and mathematical-statistical methods. Based on some information about each user's preferences, the system lists recommendation rankings and proposes items related to each user. Items could be anything, for instance: books, movies, or even biomedical items like genes, diseases, drug response, and health information. Table 1 briefly overviews the description, strengths, and limitations of the primary RS approaches: (1) **collaborative filtering** (CF), (2) **content-based** (CB), and (3) **hybrid filtering** [1, 78].

In short, the CB is primarily based on utilizing the side/content information of users and items to predict ratings and make recommendation; by contrast, the CF recommendation does not use the content information about users and items, and it considers only ratings/preferences information across users and items. Accordingly to a commonly accepted taxonomy, both CF and CB can be grouped into two classes: (a) memory-based and (b) model-based [1, 11].

A memory-based method can make recommendations over an entire rating matrix (R) or content matrix (C), if necessary. Using these matrices, on the other hand, model-based approaches estimate user preferences and

Table 1. Main approaches of recommender system techniques

Technique	Description	Strengths	Limitations
Collaborative filtering (CF)	CF analyzes the behavior of multiple users to make recommendations to an individual user. Depending on the model, it can be either memory-based, based on similar user preferences, or model-based, based on ML algorithms.	CF is able to recommend popular items among similar users even if the user hasn't expressed a specific interest in those items. CF works well with ample user data, clear user behavior patterns and preferences.	CF suffer from "cold start" problem (new user/item). Data sparsity can also limit CF. In addition, CF may suffer from "popularity bias", whereby popular items are recommended more often, and others are ignored.
Content-based filtering (CB)	Recommendations are based on the characteristics of the items and the user's preferences. The algorithm determines the items' major characteristics, and suggests comparable items to users based on their prior choices.	Ability to make personalized recommendations based on users' explicit preferences and interests. CB can be effective when there are many items and clear patterns in their characteristics.	Overspecialization, limited content analysis, lack of serendipity and diversity are some of the limitations of the CB.
Hybrid approaches	Combine CF and CB methods, in an attempt of minimizing their challenges, and improve the recommendations. Implementing these models can be done in various ways, including merging the results of two different models or adding characteristics to another model. Seven types of hybrid recommendation were introduced in Burke' study [15]	Hybrid models can be effective when both user and item are available, and when complex patterns of user behavior and preferences cannot be easily identified.	Hybrid models can be complex to implement and require significant data to train. In addition, hybrid models can suffer from CF and CB limitations, such as data sparsity, the "cold start" and the overspecialization problem.

then make recommendations accordingly on them. Memory-based CF methods use historical ratings to compute the similarity between users or items. Methods can be classified as (i) used-based — predicts items a user might like by looking at ratings given to that item by users with similar tastes to the target user (e.g., "users who are similar to you also liked ..."); and (ii) item-based — looks for similar items depending on the items users have already liked or positively interacted with (e.g., "users who liked this item also liked ..."). Memory-based typically uses similarity metrics to calculate the distance between two users or two items based on their ratios. Using ML algorithms, a model generates predictions about how users might rate items that have not been rated. The different approaches are based on (i) using dimensionality reduction techniques (e.g., singular value decomposition — SVD, principal component analyses — PCA, matrix factorization); (ii) clustering (e.g., K-means, K-nearest neighbor — KNN, density-based spatial clustering of applications with noise, etc.); (iii) neural networks (e.g., long short-term memory — LSTM, generative adversarial networks — GAN, convolutional neural network — CNN); (iv) graphs (e.g., graph convolutional network — GCN, graph neural networks — GNN) ; and (v) probabilistic methods (e.g., naive-bayes — NB).

Finally, hybrid filtering uses a combination of CF with CB to make use of the benefits of both techniques [3, 78]. When creating the hybrid RS, we may use (a) monolithic, (b) ensemble, or (c) mixed designs. The monolithic

does not clearly distinguish between CB and CF modules. For example, monolithic can use feature augmentation, where the features from various sources are aggregated, and meta-level, where one RS uses as input the model created by another RS. The ensemble design consists of combining the results of two different recommendation algorithms. Weighted methods combine the scores of different recommender algorithms into a final score by weighing the scores.

Several survey papers on RS have been published in the last few years. For example, some medical recommendation engines are described by Stark et al. [86] in their survey, and future research directions are presented. Pincay et al. [75] examined 249 papers published between 2006 and 2018, which provides insights about trends and methods regarding the design and development of health recommender system (HRS). Recently, De Croon et al. [23] discuss the various subdomains of HRS that are used, also the different RS algorithms, the different manners they are evaluated, and how they present recommendations to the user. To the best of our knowledge, our survey is the first to address but also on HRS issues to something more specific, such as recommending biomedical items in Life and Health Sciences.

2.2 Knowledge Graphs-based Recommender Systems

Over the past years, there has been considerable research conducted on KGs, especially in the Semantic Web community as can be read in the preface of the 13th International Semantic Web Conference Proceedings (2014): *"Linked Data is pervasive: from enabling government transparency to helping integrate data in life sciences and enterprises, to publishing data about museums and integrating bibliographic data. Significantly, major companies, such as Google, Yahoo, Microsoft, and Facebook, have created their own "knowledge graphs" that power semantic searches and enable smarter processing and delivery of data: The use of these knowledge graphs is now the norm rather than the exception."* [63].

A KG is a structure that identifies and disambiguates entities in text, enriches search results with semantic summaries, and provides links to related entities in exploratory search, all to improve the search engine's functionality and improve user experience [34]. The information is gathered and displayed from multiple sources. Using KG in various areas has led researchers to develop KG-based recommendation methods.

The KG describes the objective world's concepts, entities and their relationships in the form of graphs. Item attributes can be mapped into the KG to determine the relationships between them [4]. Further, the KG can be used to store user information, including information about users and items, and even user preferences, which makes relations between members of the KG possible. Recently, KGs have been proposed for recommendation in addressing two of the classic problems of RS: (a) the limited content analysis problem, which is caused by the lack of content-based features that describe the items, and (b) the overspecialization problem, which is caused by the triviality of the recommendations, which are frequently too similar to the items the user already likes [34]. In addition, KG can also help address some limitations of traditional recommender approaches, such as the "cold start" problem, with insufficient data to recommend new users or items accurately. KG help recommender engines leverage knowledge about items and users and make more informed recommendations without historical data.

In the life and health sciences, the biomedical knowledge graph (BMKG) connects biomedical entities (e.g., genes, proteins, drugs, diseases, biological pathways, etc) through defined relationships. BMKGs are important tools to solve computational problems associated with biomedical knowledge. There have been numerous applications of BMKGs in multiple tasks, including identifying disease mechanisms [43], extracting disease biomarkers [89], and predicting the efficacy of a drug over a placebo [42], or a drug discovery [81], all of which could lead to further refinement in precision medicine and clinical decision support. For instance, Cong et al. [19] propose a method for generating a BMKG based on Semantic MEDLINE Database and Linked Open Data. In addition, Gong et al. [31] propose a novel framework, called safe medicine recommendation (SMR), that aims to provide

safe medicines for patients with multiple diseases. It combines the capabilities of electronic medical records and medical KGs to build a high-quality graph and then embedded the related relationships between patients and medicines.

There are several publicly available BMKG, such as the Unified Medical Language System (UMLS), the Medical Subject Headings (MeSH) ontology, the Human Phenotype Ontology (HPO), and the DrugBank. Kyoto Encyclopedia of Genes and Genomes (KEGG) [48] can also be considered a BMKG database since it represents the relationships between these entities as nodes and edges in a graph. These BMKGs typically combine manual curation with automated techniques like natural language processing (NLP) and ML to extract information from biomedical literature and databases.

The relationships between different types in a KG can be used to enhance recommender accuracy and diversify recommended things. Using KGs improves the accountability of RS. The current methods for developing KG-based RS may be divided into three categories: (1) **embedding-based**, (2) **connection-based techniques**, and (3) **unified methods**.

Embedding-based approaches employ KG methods to pre-process the KG embedding, which may be either an item graph or a user-item graph, to produce the embedding of entities and relations, which is then used in the RS. However, this technique ignores the graph's informative connection patterns, and only a few studies can give reasons for the recommended outcomes. Accordingly to Wang et al. [93], the KG embedding algorithms can be divided into two classes: (a) the translation distance models such as TransE [13], TransH [95], TransR [57] and TransD [46], etc; and, (b) the semantic matching models such as RESCAL [68], DistMult [98] and HolE [67]. On the one hand, the translation distance models are used to calculate the probability of a fact as the distance between the two entities from distance-based scoring functions. On the other hand, the semantic matching models measure the likelihood of a point by matching the latent semantics of entities and relations in their vector space representations with the similarity-based scoring functions.

Connection-based approaches use the graph's connection patterns to guide the suggestion. The user-item KG is used in most studies to explore the relationships between entities in the graph. KG connection-based can be approached in two ways [34]: (a) the meta-structure based method, such as user-user, item-item or user-item similarities; and, (b) the path-embedding based method. The meta-structure-based method can restrict user and item representations or forecast user preferences based on similar users or items in the interaction history. In the path-embedding approach, the connection pattern between a user and an item is combined into latent vectors, allowing the mutual influence of the target user and the candidate item to be considered. Most models can also identify and mine connection patterns without specifying meta-structures because they count and select the most meaningful pathways. Therefore, expressive link patterns are likely to be captured.

Unified techniques combine the semantic representation of entities and relations and connectivity information to fully use the KG information for improved recommendations. The embedding propagation concept underpins the suitable technique. With the help of the KG connective structure, these methods enhance the entity representation. To fully leverage information from both sides, a new research trend is to combine the embedding-based technique with the path-based method [34].

2.3 Qualitative evaluations metrics of recommendation systems

For a RS to be effective, it must be evaluated according to certain criteria. The evaluation of RS algorithm has been based on information retrieval [84]. Depending on the available resources and the goal of the RS, there are various ways of evaluating its performance. Two methods co-exist: (1) **offline** and (2) **online** evaluation. Offline systems are evaluated using a pre-collected dataset and are used to measure the accuracy of RSs [33]. The datasets allow us to simulate users' behavior, predicting preferences based on historical data, either implicitly or explicitly, and evaluating RS algorithm performance. During the offline evaluation, the dataset is divided into

training and test sets. There are several advantages to modeling and testing algorithms, including their speed and simplicity. However, some bias is inevitable since the results are not directly correlated to newer users.

Offline evaluation is divided into three groups: (a) the accuracy of predicted ratings, i.e. difference between the prediction and the real rating, for instance by measuring the mean absolute error — MAE, mean squared error — MSE and root mean square error — RMSE; (b) the accuracy of recommended items based on classification metrics are hit-ratio — HR, precision — PPV, recall/sensitivity (summarized as recall in the following) — RE, receiver operating characteristic — ROC, area under the receiver operating characteristic curve — AUROC, F-measure — F1; and, (c) the accuracy based on ranking metrics, i.e. correlation between the prediction and the real classification, such as by looking at mean reciprocal rank — MRR, half-life utility — HLU, Pearson correlation coefficient — PCC, Spearman correlation, Matthews correlation coefficient — MCC, and normalized discounted cumulative gain — nDCG.

An online evaluation, the A/B-testing[1] (or multivariate testing), is different from an offline evaluation in that it measures the observed satisfaction of the user [33]. The assumptions about what a user will interact with may differ slightly from the actual interactions within a different context (when experimenting with discovering new interests or with a limited number of items). The primary issue is defining the user's satisfaction since the results depend on clicking on an item (the click-through rate or CTR). Even though offline evaluation is easy to conduct, repeatable, fast, and allows for arbitrary models to be incorporated, it has been suggested that it is impossible to mirror well the true utility of RS as seen in online experiments. Alternatively, A/B testing on live systems is quite time-consuming since the time required scales linearly with the number of approaches evaluated since users see harmful recommendations.

Another way to include the method evaluation is to have the user's feedback. Feedback is information that a recommender can collect from its users. One of the most common ways a recommender can collect this type of feedback is through explicit feedback, users' input regarding their interest in an item. An example would be for users to enter their ratings on a numeric scale based on how much they liked or disliked the content. A feature like this can be challenging to implement due to the cognitive load in generating accurate ratings. Implicit feedback, on the other hand, can avoid the restrictions associated with rating systems since the information can be gathered after observing the users' behavior. Hybrid feedback combines both types of feedback to generate a recommendation. The ratings could allow RS to provide better and more accurate recommendations.

2.4 Biomedical database

One of the main advantages of using RS is its ability to provide accurate and interpretable recommendations. This feature can be easily incorporated into various applications. Next, we provide an overview of the most common datasets used in life and biomedical science. DrugBank [96] database consists of information about drugs, their molecular target, and their pharmacological effects. The database contains information on thousands of drugs, including prescription, over-the-counter, and experimental drugs. KEGG [48] is a large database that provides a wide variety of biological data, including genes, proteins, biological processes, and human diseases. KEGG is widely used in bioinformatics, drug discovery [64, 91], and systems biology [47, 49]. ChEMBL [30] was initially not created as a "drug-target" database, but instead as a collection of bioactive chemicals. PubChem [52] is a public database of chemical compounds and their biological activities. The database contains over 100 million chemical compounds, including structures, properties, and biological activities. The Genomics of Drug Sensitivity in Cancer (GDSC) is the largest public repository for information on molecular markers of drug response and drug sensitivity in cancer cells [101]. There are more than 75,000 experiments on drug sensitivity in GDSC, which describe the response to 518 anticancer drugs across almost 1000 cancer cell lines. In addition to identifying new

[1]Also known as split testing, control/treatment testing, bucket testing, randomized experiments, and online field experiments.

marker-driven cancer dependencies, the Cancer Cell Line Encyclopedia (CCLE) provides an unbiased framework for studying genetic variants, target candidates, small molecules, and biologic therapeutics [5].

In the context of a RS, for instance, PubChem can be a valuable resource for predicting chemical compounds' biological activity and identifying potential drug candidates. By using information from PubChem in conjunction with other data sources, such as clinical databases or electronic health records, RSs can make more accurate predictions about drug effectiveness for specific diseases or conditions. They can suggest personalized treatment plans for individual patients. Additionally, PubChem can identify novel chemical entities not yet tested in clinical trials, leading to new drugs and treatments for various diseases. DrugBank is useful for RS because it provides detailed information on molecular drug targets and pharmacological properties. This information can be used to develop more accurate drug recommendations based on a user's medical history, health status, and other factors. Based on a drug's molecular targets and known pharmacology, a prescription drug recommendation system can identify drugs with a high likelihood. GDSC has been used in some RS to predict drug sensitivity for new cell lines based on genomic features. These models can expect how sensitive cell lines will be to different drugs. The GDSC database can be handy in RSs because it contains many drug sensitivity data for various cancer cell lines and accompanying genomic data. This allows the development of more accurate models for predicting drug sensitivity. These are brief examples of applications in RSs that we will develop later.

3 METHODOLOGY

This systematic review was built upon the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines [73] and is limited to the papers identified using the RS in biomedical items. First, we define a set of search terms (ST) that we consider relevant and generic in these studies. Once the ST has been described, the search algorithm is constructed using the logical operators AND and OR to combine them:

- The [title/abstract] field MUST contains RS approach
  - (*recommender* OR *recommendation*) AND (*system* OR *engine*)
  - *collaborative* AND (*filtering* OR *approach*)
  - *content-based* AND (*filtering* OR *approach*)
- The [title/abstract] fields MUST contains biomedical items, for instance
  - *drugs* OR *medication*,
  - *chemical compounds*,
  - *disease*,
  - *genes/proteins* (summarized as *gene* in the following),
  - *health* OR *patients*,
  - *health information*

Although we filter by "Title" and "Abstract", some systems (i.e., journal databases) also include the "Keywords" of the manuscripts. Likewise, with some other recommendation domains, we should initially get the recommended item. By the information from users and items, the division is made into fields: (1) **Biology**, (2) **Chemistry** (3) **Genetic**, and (4) **Health**, as shown in Fig. 1. Later, we split each domain by the pair < item, user >type. For instance, biology includes cell lines and drugs. Otherwise, the items or users for chemistry should be proteins, chemical compounds, reagents, drug response, and drug targets. Genetic includes diseases, genes, and microRNA. Finally, health is related to health information, professionals, patients, treatments, or advice activities.

Afterward, an electronic journal database search is conducted to provide a comprehensive list of scientific papers on RS: ACM Digital Library, IEEEXplore, PubMed, ScienceDirect, and Springer Link.

The search was conducted initially on June 2021 and updated on November 2021. Only peer-reviewed journal papers and conference proceedings with complete text are included in the search from early 2015 to 2021. We have also included well-documented and well-written English studies that clearly stated their findings and arguments

Fig. 1. Biology, Chemistry, Genetic and Health are the user-defined fields from the items and users found in the surveyed papers.

with a minimum of 10 sources. Technical reports, surveys, and master's and Ph.D. dissertations are excluded. Upon completing the database search, duplicate papers are removed. During the analysis, we focus on obtaining information about the purpose and methodology of each study by noting the most critical aspects in the research papers' "Method" and "Discussion" sections.

The search structure in the above databases is listed in Table 2 and presents the numbers of papers from the initial phase (n = 1883 in total). The ST "drugs OR medication", "disease" and "health information" yielded an average of 27, 34 and 34 papers, respectively, out of the 4 databases for potential review when combined with RS, following the CF. The ST "genes/proteins" (summarized as *gene* in the following) yielded an average of 8 papers out of the 5 databases for potential review when combined with CF following the RS. The highest value is found with the ST of "health OR patients", which was 85 papers combined with RS following the CF. According to our analysis of algorithmic approaches, and as shown in Table 2, CB approaches are used in a small number of reviewed papers. At this stage, there was no concern for the search for a hybrid approach, leaving this analysis for later.

4 RESULTS AND DISCUSSION

A total of 1878 papers are initially included as shown in Table 2, leaving only 60 research papers remaining from 58 journals after the PRISMA guidelines. These guidelines include: (i) identification of records, removal of duplicates; (ii) screening by title and abstract, not complete proceedings; and (iii) continuing with the eligibility, for instance, reports with no full text available, technical reports, surveys that are well-documented, and many others. A summary of the studies included in this review is provided in the appendix (Table 6) as a ready-reference summary of the existing research.

First, we identify a broader range of research studies that provide insights into the current state-of-the-art. We also aim to discover recent experiences with RS for biomedical items (RQ1) through a comprehensive analysis,

Table 2. Total number of papers found for each search term in the five databases: ACM Digital, IEEEXplore, PubMed, ScienceDirect, and Springer.

Search terms	ACM Digital	IEEEXplore	PubMed	ScienceDirect	Springer
ST#I: { Recommender OR Recommendation } AND { Systems OR Engine }					
ST#1: Combined ST#I AND drugs OR medication OR chemical compounds	7	17	5	83	26
ST#2: Combined ST#I AND disease	26	55	13	36	38
ST#3: Combined ST#I AND genes	21	2	3	-	15
ST#4: Combined ST#I AND health OR patients	111	92	28	90	100
ST#5: Combined ST#I AND health information	3	53	2	22	89
ST#II: Collaborative AND { Filtering OR Approach }					
ST#1: Combined ST#II AND drugs OR medication OR chemical compounds	1	3	18	9	15
ST#2: Combined ST#II AND disease	3	5	58	9	19
ST#3: Combined ST#II AND genes	7	5	6	20	19
ST#4: Combined ST#II AND health OR patients	2	20	168	22	38
ST#5: Combined ST#II AND health information	243	8	2	14	32
ST#III: Content-based AND { Filtering OR Approach }					
ST#1: Combined ST#II AND drugs OR medication OR chemical compounds	-	-	-	1	15
ST#2: Combined ST#II AND disease	-	-	-	3	42
ST#3: Combined ST#II AND genes	-	-	-	-	15
ST#4: Combined ST#II AND health OR patients	-	-	1	3	71
ST#5: Combined ST#III AND health information	-	-	-	-	49
Total:	**424**	**260**	**304**	**312**	**583**

Legend: ST means search term

Fig. 2. Temporal distribution for the referenced papers related to each field: biology (n = 7), chemistry (n = 16), genetic (n = 8) and health (n = 29).

identify which RSs can be used successfully in the domain (RQ2), how KG-based RS is an efficient way to leverage and connect a user's and an item's knowledge (RQ3) and assess how RS are being evaluated (RQ4). Finally, we have assembled and coded a unique dataset of 60 papers — Sur-RS4BioT, free available for download in DOI:10.34740/kaggle/ds/2346894. Details are described below.

4.1 (RQ1) Real experiences with RS

After collecting all the research papers, we split them by fields as shown in Fig. 1 based on users and items. Fig. 2 shows the annual distribution of such papers between 2015 to 2021 and by each field. The data shows two peaks, the first in 2018 (16) and the second in 2021 (13) with similar curves across all areas. Furthermore, health is the one that collects the most data, followed by chemistry.

Biology. Personalized therapies, or "precision medicine", as described by Sulphavilai et al. [87] provide the most appropriate regimen for each patient, while their responses may differ. For example, several authors [14, 24, 54, 59, 87, 92, 105] use a drug response prediction algorithm to the anti-cancer effects of different drugs based on the cell line similarity and the drug's chemical structure. Zhang et al. [105] estimate the baseline similarity score for the various cell line and drug pairs based on the available responses. They use the known correlation

coefficient to find the most similar neighbors. Sulphavilai et al. [87] have proposed a matrix factorization based RS (CaDRReS), which considers essential genes for drug-response prediction. Liu et al. [59] adopt a neighbor-based CF with global effect removal (NCFGER), removing the global effect and shrinking the similarity score for each cell line and each drug pair. They use the K-similarity score to predict the unknown ones after removing the global impact. In contrast, Wang et al. [92] introduce a new model using dual-layer strengthened collaborative topic regression (DS-CTR) that combines the knowledge of pharmacogenomics data and cell line similarity network. Emdadi and Eslahchi [24] present a novel method for cancer drug sensitivity named drug sensitivity prediction using logistic matrix factorization (DSPLMF) based RS. The motivation of DSPLMF is to find the features of cell lines that are sensitive to drugs since similar cell lines can also improve the prediction of drug response and gene expression profile. Like Sulphavilai et al. [87], the PCC of predicted drug responses and pathway activity scores infer drug-pathway associations. Koras et al. [54] propose a deep neural network RS-based approach (DEERS) to the problem of kinase inhibitor sensitivity. Using autoencoders and neural network-based prediction, DEERS combines dimensionality reduction and hidden representations of the cell line and drug features. An utterly alternative solution is proposed by Brandão et al. [14], in which experiments show that wavelet-transformed DNA microarray images produce better results, not only in terms of evaluation metrics but also in terms of execution time, by improving the search for cancer-cell lines with similar profiles to the new cell line. Lastly, to obtain the cancer cell lines profiles, and a drug-response matrix are acquired from the GDSC database [14, 54], and using the drug structure information from PubChem database in case of Wang et al. [92], or even the CCLE [5] datasets.

Chemistry. Scientists in the pharmaceutical industry have been focusing on developing novel drugs (or, therapeutics) by utilizing expertise on existing drugs [39, 107]. Drug discovery begins with the identification of drug-target interactions (e.g. genes), which can be reliably done by *in vitro* experiments, and one of the biggest risks is the possibility of unexpected or unintended interactions between drugs and off-target proteins [25, 102, 106]. *In silico* techniques are becoming more popular as a means of reducing temporal and monetary costs [74]. For *in silico* prediction of drug–target interactions (DTIs, also called compound–protein interactions), ML approaches are a solution. Knowledge about drugs, targets (i.e. protein), and already confirmed DTIs makes up features of ML methods (for instance, feature-based, matrix factorization, deep learning and network-based approaches) [26, 55, 83, 106, 107], which then form the basis for training a predictive model, which can determine interactions between new drugs and/or targets. Recent methods that use matrix factorization algorithms outperform other ML methods in terms of efficiency [29, 39, 74, 80]. DTI prediction is best done using a combination of chemical and genomics information using RS approaches [1]. To predict DTI, the most popular methods include drug–drug and target–target similarity measurements through similarity or distance functions. Nearest neighbor algorithms define "nearness" in various ways based on distance functions [25, 39, 80]. There are a wide variety of feature-based methods that perform DTI prediction. These include SVM, tree-based and other kernel-based methods [83]. Also, deep learning methods show good performance [102]. Barros et al. [6] designed a method called LIBRETTI to create implicit feedback datasets of scientific entities such as clusters of stars and chemical compounds. Later, the authors propose a framework to recommend chemical compounds based on ontologies, a new method for calculating similarities between large numbers of entities (over 16000 chemicals) [8]. The items/entities are from 4 distinct ontologies: chemical compounds from Chemical Entities of Biological Interest Ontology (CHEBI) [41]; functions of genes from Gene Ontology (GO) [20]; phenotype abnormalities from Human Phenotype Ontology (HPO) [53]; and, diseases from Disease Ontology (DO) [82]. Many drug-related databases have been set up to support the aforementioned methods. These databases provide forms of drug-related data and are important resources for in *silico* DTI predictions, for example, DrugBank (6) [96], KEGG (1) [48], ChEMBL (2) [30] and PubChem (2)[52] (see Table 3).

Genetic. Protein subcellular localization (SCL) has a role in identifying potential drug targets (i.e., protein) and genome annotating because proteins have distinct functions in individual cells. In this field, we have identified several novel methods that we will describe in detail below. Mehrabad et al. [62] use a personal RS protein multiple location prediction based on RS (PMLPR) to predict a list of locations for each protein, and it successfully solves the significant location prediction issue. To overcome the cold start problem, PMLPR uses protein interaction scores. This approach creates a bipartite network of users and items; in this example, the network is built using data from SWISS-PROT and the cellular component in GO. Kim et al. [51] propose gene selection using the expression heterogeneity (GSEH) method, which combines the concept of gene expression heterogeneity with the analysis of biological processes related to diseases. Gene expression heterogeneity refers to samples from the same class that might have different amounts of gene expression. This concept could be used to identify disease-associated genes. GSEH is divided into two steps: (1) creates a new matrix with a CF pattern, then selects the target genes based on their expected scores; and, (2) compares the data obtained in the first step with the original data to compute each gene's prioritization score; it then picks genes based on their scores. Zeng et al. [104] introduce a deep collaborative filtering model that combines Bayesian stacked denoising autoencoders (SDAE) and matrix completion. This technique provides a scalable platform for incorporating numerous gene and disease characteristics. The presented quantitative findings outperform existing state-of-the-art baselines by utilizing deep architectures. Weighted imputed neighborhood-regularized tri-factorization (WINTF) is a tool for predicting transcription factor-gene associations that apply one-class CF techniques [56]. The tool allows users to specify different low ranks for items and users separately. With a collection of known associations, it can also be applied to more tissue-specific tasks to predict new TF-gene associations. To study protein-domain interaction networks (PDIs), a further collaborative filtering model-based method (CFMM) has been proposed recently by Zhu et al. [109]. The authors propose a calculative method for inferring potential essential proteins to achieve this goal. This method is based on an improved PageRank algorithm, which integrated the original PDI network's topological features with the proteins' biological characteristics. RNAcommender tool [21] assists researchers in identifying potential interacting candidates for most RNA-binding proteins (RBPs) with uncharacterized binding preferences. In recent follow-up work of RNAcommender, the ProbeRating method [100] is designed to predict binding profiles for unknown or poorly characterized RBPs based on the binding profiles of their homologous RBPs that are currently known.

Health. Increasingly, health information systems are playing an important role in healthcare services [4, 28, 44, 45, 69, 69, 71, 79]. Physical activities are frequently customized based on individual preferences [28]. In addition to physical activities, Rohani et al. [79] develop a smartphone-based system for "behavioral activation" (MUBS); a personalized patient model is created by storing activity features along with the patients' ratings after an activity has been completed. In another example, Chen et al. [16] motivate users to stop smoking by providing them with tailored messages called computer-tailored health communication (CTHC), such as "In 5 to 15 years of living smoke-free, your risk of stroke goes down to a nonsmokers risk. Congratulations on a job well done!". Additionally, other studies focus on Personalized trustworthy health care information per se [12, 71], or personalized access to general health information [18, 60]. Many others focused on specific health conditions [2, 4, 44, 45, 65, 66, 76, 88, 94, 103]. For example, Torrent-Fontbona and López [88] build a knowledge-based (KB) RS to assist diabetes patients in numerous cases, Mustaqeem et al. [66] propose an improvised algorithm for recommending medical advice to cardiac patients, Ormel et al. [71] and Iatraki et al. [44] apply personal health record system constructed for cancer patients. Personalized cancer care involves relating genomics markers to treatment outcomes based on genomics information; Zhang et al. [105] build a clinical patient drug RS, and the authors suggest having only one drug model for each training sample instead of having multiple models for different drugs.

Table 3. Top@10 of the datasets by each user-defined field.

	Biology	Chemistry	Genetic	Health
CCLE	4			
ChEMBL		2		
CheRM-20		2		
Clinical Data		1		**11**
dbDEMC			2	
DrugBank		**6**		1
GDSC	**7**			
Personal Data				**7**
PubChem	1	2		1
SIDER		2		1

Legend: Cancer Cell Line Encyclopedia (CCLE), Genomics of Drug Sensitivity in Cancer (GDSC), Side effect resource (SIDER)

Summary. Personalized therapies have emerged to tailor medical treatments based on a patient's unique characteristics, including genetics, medical history, and lifestyle. Some examples illustrated above, include CaDRReS [107], DSPLMF [24] and DEERS [54]. Based on a patient's genomic data, a CB may recommend customized medicine. In contrast, a CF might recommend a therapy based on the treatment history of similar patients. A hybrid RS could use a patient's genomic data and medical history to recommend personalized treatment. In drug discovery, recommendation engines can recommend new molecules or targets for drug development. DrugBank and PubChem are two examples of databases. In particular, an CB may suggest new compounds with similar biological properties to known drugs by using compositional descriptors as a prior knowledge [83]. Based on preclinical and clinical trial data, RSs can predict drug efficacy and safety. For instance, a CF could use clinical trial data to predict which patient groups may be more likely to benefit from a new drug. Clinical data can benefit RSs by providing necessary information about a patient's medical history, current health status, and previous treatments (e.g., [28, 50]). As illustrated in Table 3, this is the one with the highest number of papers analyzed (11). For example, a recommendation system incorporating clinical data might use this information to identify the most effective treatments for a particular disease or condition based on a patient's medical history and current symptoms. The system could also consider a patient's age, gender, and other relevant demographic information when making treatment recommendations. However, some challenges are associated with using clinical data in RSs, which we describe later. Every available tool referenced in this section is also included in Table 4. Out of the 60 articles analyzed, only 14 of them provide direct access to the source code on GitHub. The lack of comprehensive and reliable documentation undermines the reproducibility of recommender studies and hinders validation and extension. A vital aspect of this challenge is the inadequate documentation of RS tools, which includes algorithms, frameworks and software used in system development. Researchers struggle to understand these tools' functionality, parameters and implementation details, making replication and comparison difficult. The insufficiency of clear documentation also makes it challenging to reproduce experiments and assess the impact of tools on system performance, for example, using them in a shared and fair evaluation using the same objective and dataset. Poor documentation has consequences beyond replication challenges, such as real-world adoption.

Table 4. Overview of all available software. The first column shows the paper's author(s) and publication year. The second column lists the field/area we defined above, and the next two list the name of the tool and URL link

Author(s) and Year	Field	Tool' name	url
Kim et al. (2016) [51]	Genetic	GSEH	http://embio.yonsei.ac.kr/files/hjkim/gseh.zip
Peska et al. (2017) [74]	Chemistry	BRDTI	http://www.ksi.mff.cuni.cz/~peska/BRDTI
Yang et al. (2017) [99]	Health	Yum-Me	https://github.com/ylongqi/fooddist
Zhang et al. (2018) [105]	Biology	HIWCF	https://github.com/laureniezhang/HIWCF
Yasuo et al. (2018) [102]	Chemistry	CoDe-DTI	https://github.com/sekijima-lab/CoDe-DTI
Zeng et al. (2019) [104]	Genetic	DCF	https://github.com/xzenglab/DCF
Embadi and Eslahchi (2020) [24]	Biology	DSPLMF	https://github.com/emdadi/DSPLMF
Barros et al. (2020) [7]	Chemistry	CheRM	https://github.com/lasigeBioTM/CheRM
Lim and Xie (2020) [56]	Genetic	WINTF	https://github.com/XieResearchGroup/WINTF
Barros et al. (2021) [8]	Chemistry	ChemRecSys	https://github.com/lasigeBioTM/ChemRecSys
Koras et al. (2021) [54]	Biology	DEERS	https://github.com/kkoras/rec-system-for-drug-response
Sadeghi et al. (2021) [80]	Chemistry	NMF-DR	https://github.com/sshaghayeghs/NMF-DR
Barros et al. (2021) [9]	Health		https://github.com/lasigeBioTM/blah7
Zhu et al (2021) [109]	Genetic		http://dip.doe-mbi.ucla.edu

After defining the biomedical items, another relevant question is *who are the users for the domain?*. In this scenario, we split into two categories: health and others. Regarding health, a RS should be designed to be used by an end-user who can be either a patient or a healthy personal as shown in Fig. 3. Aside from doctors, other health professionals, such as nurses and pharmacists, could also benefit from the system. For others, there is a greater dissipation between drugs, genes, cell lines, or diseases.

The availability of datasets is another topic of the domain that is usually neglected. Despite the advantages of having public data, such as DrugBank or PubChem, this resource is rare for developing health recommendations in particular (more than 30%) as shown in Fig. 4. Some of the issues originate from health data being inherently privacy sensitive. One challenge is the need for patient privacy and data security protection [27]. Clinical data contains sensitive information about patients. It must be handled carefully to avoid data breaches or unauthorized access. Another challenge is the quality and completeness of the data. Health data can be complex and challenging to interpret, as well as they are stored in multiple formats and systems, making integration and analysis difficult. Additionally, this data may be incomplete or inaccurate, leading to incorrect or ineffective treatment recommendations. In the case of chemistry, the most significant data sources are DrugBank, PubChem, and SIDER, as shown in Table 3. All of them are public, but these are real recommendation datasets. In contrast, the dataset proposed by Barros et al. [7–9] follows the standard format < user, item, rating >, where the items are scientific entities, the users are authors from research papers, where these items are mentioned, and the ratings are the number of articles an author wrote about an entity. All datasets are available.

Recommender systems help reduce information overload by extracting user preferences or interests from relevant datasets. The most commonly used datasets[2] for RS are Netflix [10], Pinterest [58], MoviLens [40], Amazon Product Data, MIND by Microsoft, Yelp Dataset and many others [11], all of them available in the Kaggle platform[3] and follows the standard format. Scientific databases have emerged as one of the milestones in the modern scientific enterprise. Three databases can be mentioned in the biomedical area, all of them being

[2]Andreas Chandra, "Common Datasets Benchmark for Recommendation System" in Medium (April 2021)
[3]https://www.kaggle.com/datasets?search=recommender+system

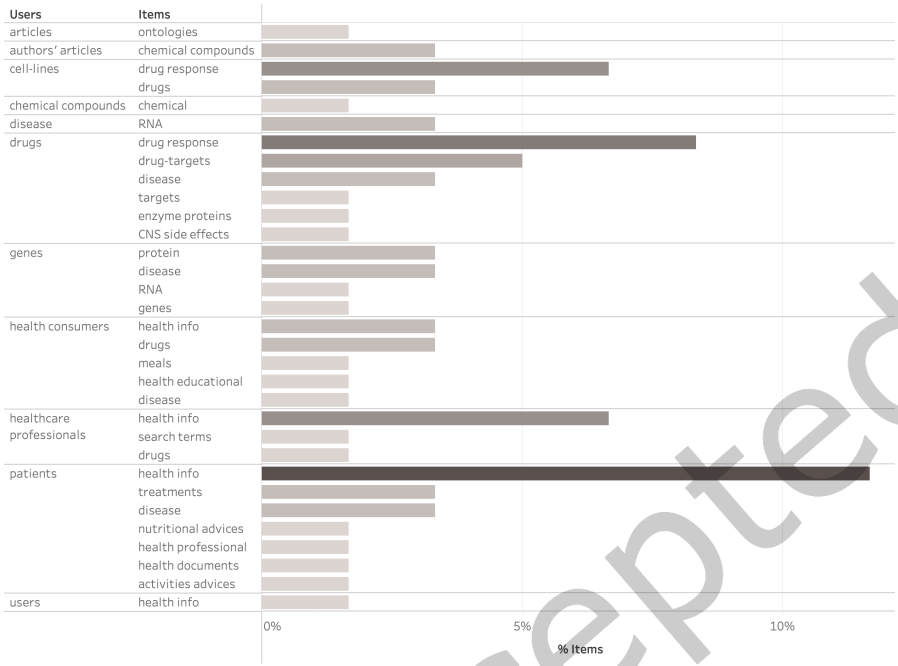

Fig. 3. Pair < users, items >interactions in the recommender system techniques used in this survey.

open-source and providing bioinformatics and cheminformatics resource: (1) DrugBank online [97] database containing information on drugs and drug targets (protein); (2) GDSC database [101] with data on the sensitivity of genomically characterized cancer cell lines to selected compounds; and (3) CCLE with about 1000 cancer cell lines. Although, much research in the health field is based on proprietary and non-public datasets.

4.2 (RQ2) Recommender system techniques

The recommender techniques are usually classified into three main categories and briefly described in the previous section. As shown in Fig. 5, the CF is the most popular approach in the studies of this survey, 41 in total, followed by CB with 10, 7 for hybrid filtering and 2 for others.

Collaborative filtering. CF models produce recommendations using a collaborative process that utilizes multiple users' ratings. A common matrix factorization is found in the vast majority of research findings [14, 24, 32, 51, 87, 103]. If, on the one hand, Han et al. [37] and Zhang et al. [105] use the weighted matrix factorization, on the other hand, Ha et al. [35] consider the probabilistic matrix factorization, and later, the same authors added the miRNA functional similarity scores to avoid cold start problem from MF (IMIPMF) [36], finally Embadi and Eslahchi [24] apply the logistic matrix factorization. Yue et al. [103] propose a modified CF based on user-based and the item-based. On the other hand, Hao et al. [39] and Liu et al. [59] design approaches derived from neighbor-based to infer potential drug candidates for targets of interest. Ezzat et al. [25] present an ensemble model-based with weighted KNN and graph regularized matrix factorization (GRMF). Otherwise, Galeano and Paccanaro [29] suggest that latent factor models can be useful for detecting unknown adverse drug events early

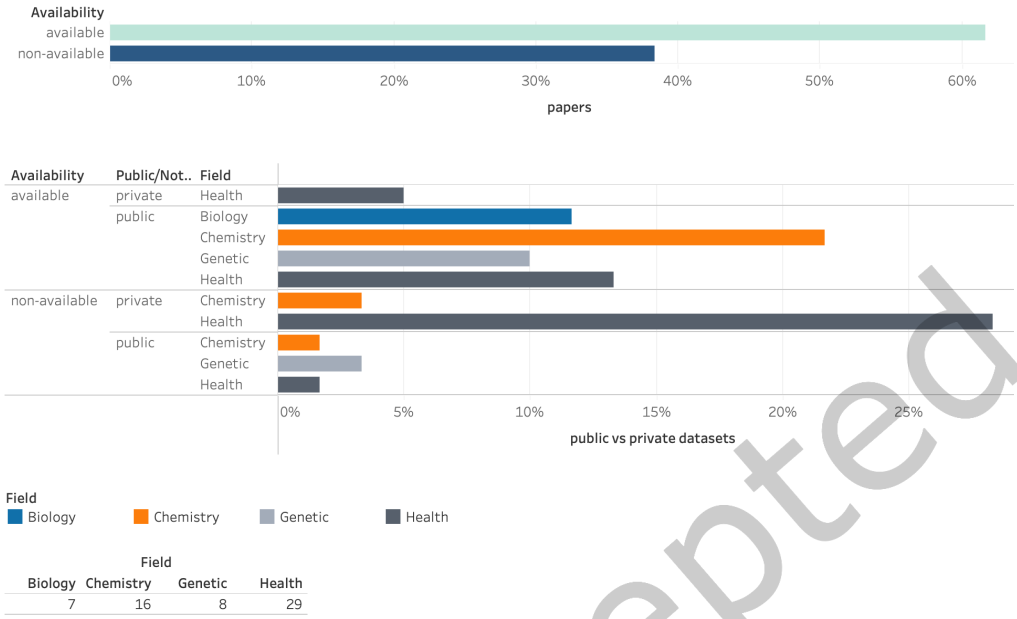

Fig. 4. Availability of datasets: a global view above and a distribution by fields below. Available and public data in health are rare (less than 15%), chemistry is the field that provides more available datasets even though they do not follow the standard format.

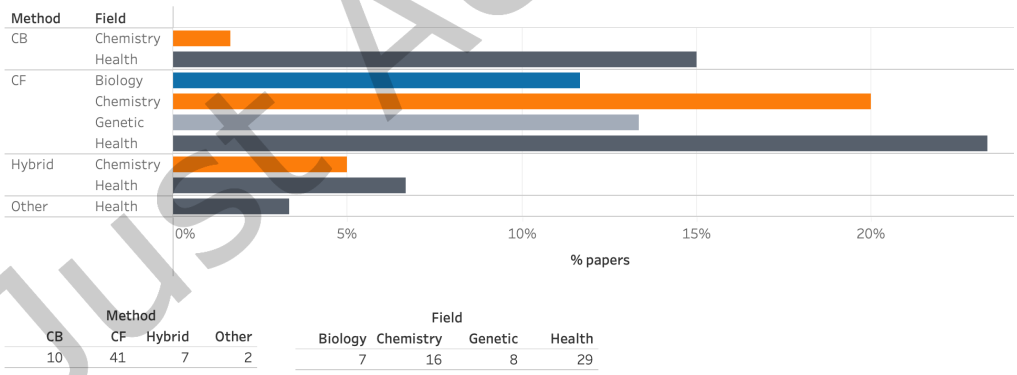

Fig. 5. Overview of the different recommender techniques by fields (above) and total below.

and accurately. Based on structured electronic health record data from a tertiary academic hospital, Chen et al. [17] train an order recommendation system (item-based) analogous to Netflix or Amazon's. Barros et al. [9], with LIBRETTI methodology, found the relations between entities and recommended entities of interest for a particular researcher. They selected the alternating least squared (ALS) as RS method.

A total of 19 studies reported using a model-based CF like clustering, SVM, NB, elastic net, random forest, and others [18, 21, 26, 38, 45, 55, 61, 62, 65, 66, 72, 74, 80, 92, 105–109]. Ren et al. [77] use two methods: HCFMH and cpHCFMH. HCFMH recommends search terms based on the most recent search terms and encounters, while cpHCFMH suggests words based on a patient's past encounters. One-class collaborative filtering (OCCF) problem is present too [56, 90].

Deep learning is a growing field with applications spanning several use cases. This survey identifies 6 papers. Collaborative deep learning uses a combination of CF and the probabilistic matrix factorization (PMF) with denoising autoencoder (DAE) [102], or stacked denoising autoencoders [104], or deep feed-forward neural network, DeepSurv [50], and DEERS [54], to replace the dot product for modeling the user-item interactions in the latent space, and capture the complex user-item interactions in the hidden space, word embeddings of NLP. Yang et al. [100] develop a two-stage framework: the first stage involved encoding the protein and nucleic acid sequences into distributed feature vectors, and the second stage involved recommending binding preferences for new proteins; a multilayer neural net [70]. Oszoy et al. [72] use the Pareto dominance and CF approaches to predict future venue preferences (i.e., check-in locations) of target users.

Content-based filtering. CB is most commonly used when a lot of attribute information is available. Like in CF, it works with data that the user provides, either explicitly (rating) or implicitly (clicking on a link). As a result, CB is especially well adapted to making suggestions in text-heavy and unstructured domains. As expected, it is in the health field since (i) information about a patient's health is collected in the form of an electronic medical record (EMR) [60, 76, 88], or (ii) electronic health record (EHR) [12, 28, 44, 71, 79, 99] in medical centers, hospitals, and pharmacies authorized to do so. Seko et al. [83] tailor a descriptor-based RS to estimate the relevance of chemical compositions where crystals can be formed based on existing inorganic crystal structure database. The model-based algorithms were logistic regression, gradient boosting, and random forest.

Hybrid. Several RS combine CF and CB methods, which helps to avoid certain limitations among these. Nouh et al. [69] propose a smart RS of hybrid learning (SRHL) for personal well-being services regarding health food service. SRHL includes: resolving the cold start problem for new users by transitioning between CB and CF; detecting user context inside dynamic filtering; and integrating profile learner approaches to reflect user input. Sosnina et al. [85] apply RS approaches in the antiviral drug discovery context with the CF algorithm implemented in Surprise package and sparse-group inductive matrix completion (SGIMC) implementation of CB. Surprise Python package operates only the interaction matrix elements: KNN, clustering algorithms, and matrix factorization. The RS proposed by Chen et al. [16] use a hybrid ML algorithm, which combines CF and CB ranking to select messages that are most suitable for individual smokers. Ammar et al. [4] describe a personal health library, namely mHealth app, that provides hybrid RS by incorporating KG and linked data. For recommending chemical compounds, Barros et al. [8] developed a hybrid semantic recommendation model suitable for implicit feedback datasets and focused on retrieving ranked lists based on the relevance of the items. In this model, the authors incorporate CF for implicit feedback (ALS and Bayesian personalized ranking) and a new CB based on semantic similarities among chemical compounds in ChEBI ontology.

Others. In this survey, we found two more studies that used other RS techniques, such as the KB. Wang et al. [94] propose the KB RS for helping people with chronic diseases manage their health by recommending educational materials. Through an ontology, it could link patient characteristics to the content of the materials. Another one is proposed by Agapito et al. [2] for adaptive nutrition content delivery to patients with diet-related chronic diseases and healthy subjects (DIETOS).

Summary. Generally, the RS techniques choice depends on the biomedical application and the available data. Different techniques can be combined to improve the accuracy and effectiveness of recommendations. Among the papers reviewed, recommendation based on CF (more precisely, model-based) prevailed, regardless of the

field and the data type (health data or data from public databases, e.g., GDSC and CCLE). Because of its specificity, it was expected that the CB model would be used to recommend therapeutics in more significant numbers. For example, a CB system can recommend a treatment based on its mechanism of action or reported side effects. This approach is advantageous when patient data is limited, or there are well-defined criteria for treatment. Another application could be drug discovery to predict drug efficacy and safety based on a candidate's chemical and biological properties. In addition, the use of KB to model therapy-patient-disease relationships will be increasingly used. KB can help identify patterns and relationships that may not be obvious from patient data alone, leading to more accurate and effective recommendations.

4.3 (RQ3) Knowledge-graph based recommendation algorithms

Knowledge-graph recommendation leverages the connections among the entities of the user, the items, and their interaction to determine the best recommendations. The algorithms use explicit or implicit connections to find items that may be interesting or valuable to the users. The relationships give extra essential information to the KG-based recommender, allowing it to use inference between nodes to uncover novel connections. Three approaches were described above: (1) embedding-based, (2) connection-based, and (3) unified method. As presented in Fig. 6, 13 out of 60 studies use KG to improve the results in the RS. A curiosity is that the field of genetics resorts to using this methodology, a value we do not find in other areas. It is also worth noting that the hybrid technique has a higher propensity to the KG regardless of field.

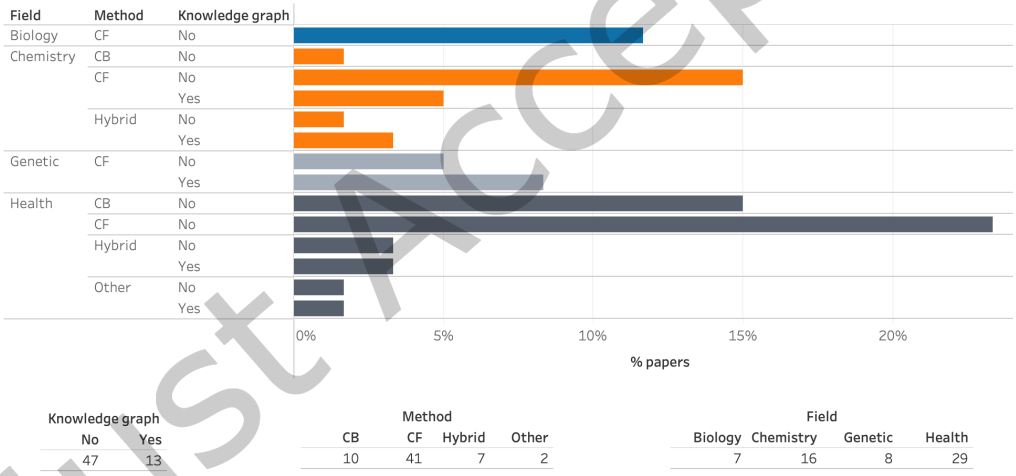

Fig. 6. Knowledge-graph in numbers: overview of papers distribution by field and recommendation system methods.

Embedding-based approaches can be divided in three phases: (i) representing entities and relations, (ii) constructing a scoring mechanism, and (iii) learning entity and relation representations. Wang et al. [93] categorize such embedding techniques into two groups: translational distance, and semantic matching [4, 108] models.

For instance, Zheng et al. [108] first designed a pretraining method based on neural CF to get the initial embeddings for patients and drugs. The drug interaction graph will be initialized using the medical records and domain knowledge. The proposed drug package recommendation aims to build a Personalized scoring function for each patient. Ammar et al. [4] created a Resource Description Framework (RDF) representation of a personal KG that maintains a digital health state of each patient from a historical perspective.

Connection-based approaches use the user-item graph to find path-level similarities between items by pre-defining meta-paths or automatically mining connective patterns. Users can also get an explanation for the outcome using the path-based method. As mentioned above, two methods are described: the meta-structured based [25, 36, 90, 104, 108], and the path-embedding based [21, 35, 94].

For instance, Ezzat et al. [25] propose a method for tackling the drug-target interaction with the GRMF to prevent overfitting. The authors derived a p-nearest neighbor graph from each drug and target similarity matrices. Wang et al. [90] integrate heterogeneous datasets from genomics (ZINC, ChEMBL, and DrugBank databases) into a multi-layered network model. In this model, each node is either a chemical entity (drugs and other chemicals), a biological entity (genes or proteins that it encodes), or a phenotype entity (disease and side effects). Nodes in the same entity class are linked by similarity (e.g., chemical-chemical similarity) or interactions (e.g., protein-protein interactions). Nodes that belong to different entity classes reside in different network layers and are linked by known associations (e.g., drug-target interactions, disease-gene associations). Chemical-chemical, gene-gene, and disease-disease similarity scores are inputs of the proposed tREMAP CF algorithm. Zeng et el. [104] adopt the Katz measure, a graph-based method to measure how similar two nodes are by computing based on how many paths of different lengths exist between them. Ha et al. [36] use the miRNA network as supplementary data to improve prediction accuracy. The miRNA network can be defined as a graph in which a node represents each miRNA, and an edge means each similarity weight. Corrado et al. [21] propose a CF that can also be interpreted as a feed-forward neural network with a Kroneker layer (second-order units). Briefly, the matrix factorization would map users and items to a latent feature space where a significant correlation (dot product) between latent vectors predicts an interacting user-item pair. Wang et al. [94] use the KB RS with a combination of ontology and several NLP techniques to recommend Chinese educational materials to chronic disease patients.

Summary. The papers discuss embedding-based approaches for representing entities and relations, constructing scoring mechanisms, and learning entity/relation representations. The techniques are categorized into translational distance and semantic matching models. On the one hand, based on pre-training methods and a drug interaction graph, these techniques can be used in personalized drug package recommendations for patients. On the other hand, they use a RDF representation to create personal KG, historically maintaining the digital health status of each patient. Embedding-based approaches have shown promise in personalized drug RS based on patient-specific information. A connection-based explores the user-item graph to find similarities between items by either mining connection patterns or predefining meta-paths. Two methods are used: meta-structured based and path-embedding based. For instance, drug-target interactions are addressed using meta-structured methods based on p-nearest neighbor graphs. Multilayer network models are used with path-embedding based methods for integrating heterogeneous genomic data.

4.4 (RQ4) Qualitative evaluation methods of recommendation systems

A RS's purpose is to predict how likely users would appreciate unknown items based on what the system already knows about them. The most common evaluation method is, as illustrated in the Fig. 7, the offline evaluation using existing datasets to estimate the *accuracy* measures of a RS.

Predictive accuracy represents the degree of similarity between the recommender's estimated and actual user ratings. This sort of measure is frequently used to evaluate non-binary ratings. For instance, Katzman et al. [50] consider the DeepSurv model with the concordance-index (C-index) and the MSE to quantify the difference between the model's predicted log-risk function and the true log-risk values. Otherwise, the experimental results of the SRHL model [69] are evaluated by using three absolute error measures: MSE, MAPE, and MAE.

Classification metrics aim to determine a recommendation algorithm's decision-making success. The performance of a RS may also be represented graphically using ROC, and the AUC indicates how well the model can

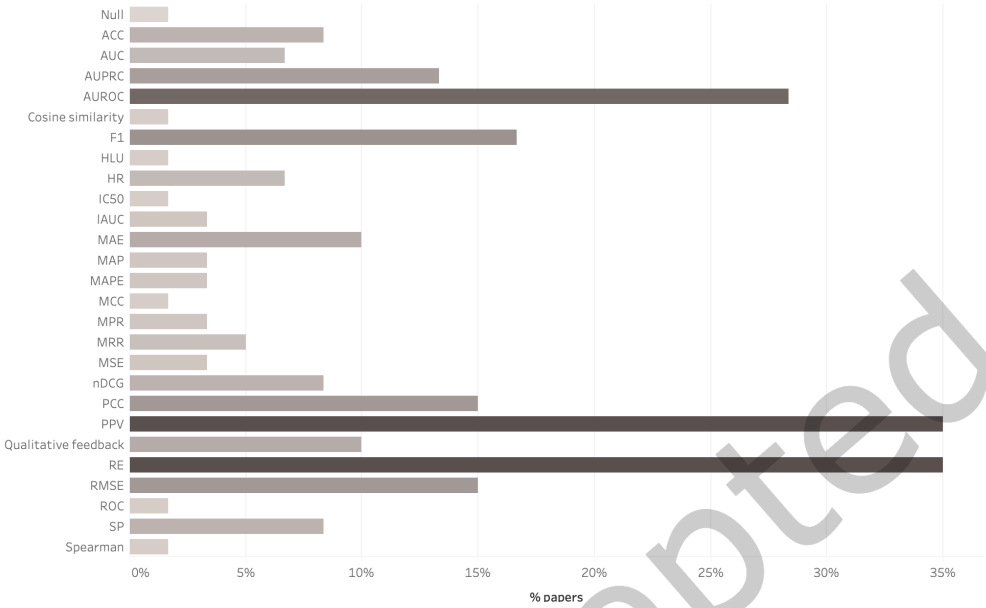

Fig. 7. Evaluation measures distribution in all surveyed papers, in case they have been disclosed.

Legend: Accuracy (ACC), Area Under Curve (AUC), Precision–Recall Curve (AUPRC), Area Under the Receiver Operating Characteristic curve (AUROC), Averaged Root Mean Square Error (RMSE), F-measure (F1), Half-Life Utility (HLU), Half-maximal inhibitory concentration (IC50), Hit Ratio (HR), limited Area Under the Curve (lAUC), Mean Absolute Error (MAE), Mean Average Precision (MAP), Mean Absolute Percentage Error (MAPE), Matthews Correlation Coefficient (MCC), Mean Percentile Ranking (MPR), Mean Reciprocal Rank (MRR), Mean Square Error (MSE), normalized Discounted Cumulative Gain (nDCG), Pearson Correlation Coefficient (PCC), Precision (PPV), Recall (RE), Specificity (SP)

distinguish between classes. As seen in Fig. 7, more than 25% use the above metrics. For instance, Sadeghi et al. [80] propose a RS-based method for drug repurposing to predict novel drug indications by integrating drug and diseases related data sources. The AUC performance is evaluated and compared with other methods using 5- and 10-fold cross-validation. The following performances are added to the previous works like (i) ROC curves to compare an ensemble extended neighborhood-based recommendation model [26], the CoDe-DTI method [102], the DS-CTR model [92], the IMIPMF method [35], and the efficiency of antiviral activity class prediction with Hybrid techniques [85] with other advanced models; and, (ii) ROC curves and AUPRC for highly imbalanced datasets such as for predicting side effects of marketed drugs [29], the improved prediction of miRNA-disease associations (IMDN) framework [36], the CFMM method [109]. Ezzat et al. [25] adopt the GRMF model using 10-fold cross validation in simulated "new drug" and "new target" cases. Hao and Blair [38] study a user-based CF on medical data with a categorical outcome in four publicly available datasets. Same evaluation metrics: recall and specificity, are applied for the DIETOS framework [2], a food RS for healthy people and individuals affected by diet-related chronic diseases. Pustozerov et al. [76] develop a RS infrastructures that incorporate personalized blood glucose prediction algorithms for diabetes patients. The model performance is estimated using standardized metrics (RMSE, MAE, and MAPE).

Recall and precision are the traditional evaluation metrics and the most widely used recommendation quality measures [11], as shown in Fig. 7, with approximately 33% for both. Precision is a measure of recommended items that are relevant. Otherwise, recall measures relevant items found in the recommendations items. Both help construct an "unbiased" test dataset and then score the resulting test dataset using a model [22]. With the above metrics, Macedo et al. [60] propose a software framework in the biomedical domain and recommend related scientific information to alert health professionals to promote preventive healthcare. However, qualitative analysis is carried out in addition to quantitative indicators. Zeng et al. [104] evaluate the performance of the deep CF model with five real-world datasets (see Appendix Table 6) in biology and compared with the others algorithms such as graph-based method, bagging SVM classifier, and many others. The following performances are added to the previously listed, like (i) F1-measure [55, 61, 62, 107, 108]; (ii) AUROC [17, 18, 21]; and, (iii) the aforementioned and AUPRC [72, 106], were used to weight the evaluation recommendation results. For instance, the performance of the RNAcommender [21] is computed in a leave-one-protein-out experiments. The model-based CF technique is something all mentioned papers have in common. Regarding to the DSPLMF model [24] on two datasets (GDSC and CCLE) the metrics are ACC, RE, PPV, SP, F1, MCC and AUC.

Ranking accuracy, also known as rank correlation measurement, measures the ability of a recommender to estimate the correct order of items based on the user's preferences. The PCC is one of the most popular means to evaluate how much two users are related in a CF approach. An example of this is the HyperRecSysPA model [28]. The HCFMH and cpHCFMH models [77] are evaluated with the HR@k to compare other recommendation methods. Performance and robustness of CaDRReS [87] using Spearman correlation, nDCG across 10 runs of 5-fold cross-validation and HR (number of sensitive drugs identified).

Besides classification metrics, error metrics were also employed to measure the error made by a RS when predicting an item rating [66, 70]. For instance, Mustaqeem et al. [66] present a hybrid model that gives cardiac patients illness predictions and treatment advice with a clinical dataset collected and labeled in consultation with medical experts. The prediction results are evaluated using three metrics, i.e., ACC, Kappa statistics, and RMSE. Ochoa et al. [70] implement RS that analyses the frequency of medical events in the EHR and delivers and the quality of the model was found with PPV, RE, ACC and RMSE.

On the other hand, both classification and ranking metrics are used to evaluate the relevance of the recommended item [7, 8, 12, 37, 39, 56, 74]. Regarding Barros et al. [7, 8], the recommending ranked of chemical compounds are evaluated with six metrics, i.e., PPV, RE, F1, MRR, nDCG and lAUC. Lim and Xie [56] identify target genes of transcription factors and the performances of the two methods: WINTF and REMAP, with four different metrics, for instance, AUROC, MAP, HLU and MPR.

The system for data-driven therapy decision support developed by Gräßer et al. [32] considers three different evaluation metrics: first, the individual RS produces a prediction of how the patient will respond to specific therapies with RMSE; second, the top-ranked therapies based on the affinity predictions are usually presented to the user after selection from the consultation with the precision; and, third, for the similarity computation, the Gower coefficient, cosine, Pearson and Spearman correlations are applied. The Gower's coefficient has the advantage of allowing for missing values and permits the introduction of a user-defined weighting scheme. To overcome the probability distribution with zero mean and constant variance assumptions, [32] applied the Spearman correlation. The online framework Yum-me [99] is evaluated both offline and online. Regarding offline, classification and error metrics are applied.

Online. RS emerged to model user preferences in various online applications to tackle the information overload problem. In general, RS have been developed to solve the problem of information increase and enhance the user experience on various online applications. Examples in health fields are PepperRec [44], PHIR [44], MUBS [79], HERS [71], CTHC [16] and personal health library–enabled mHealth [4] with personal and clinical data

Table 5. Top@10 of the evaluation metrics by each user-defined field.

	Biology	Chemistry	Genetic	Health
ACC	1	2	1	1
AUPRC		**5**	2	1
AUROC	2	7	5	3
F1	1	**6**	1	2
MAE				**6**
PCC	3		1	**5**
PPV	1	**5**	3	**12**
Qualitative feedback				**6**
RE	1	**6**	2	**12**
ARMSE	3			**6**

Legend: Accuracy (ACC), Area Under the Precision–Recall Curve (AUPRC), Area Under the Receiver Operating Characteristic curve (AUROC), Averaged Root Mean Square Error (ARMSE), F-measure (F1), Mean Absolute Error (MAE), Pearson Correlation Coefficient (PCC), Precision (PPV), Recall (RE)

Summary. There are several papers discussing attempts to improve the accuracy of RS results, for example RMSE, MAE, etc. Additionally, it is common to try to improve recommendations with PPV, RE, AUROC for instance. Recall and PPV (35%) and AUROC (28%) were the most commonly used offline evaluation metrics as shown in Fig. 7. Other popular offline evaluation metrics are accuracy-related measurements, such as F1, 16%, PCC, 16%, RMSE, 14%, MAE, 12%, and SP 12%. Measurements of the other metrics are inconsistent. Table 5 shows the top@10 common metrics by fields. Classification metrics are predominant in all fields, and as we expected online feedback is exclusive for Health.

5 CONCLUSIONS

In this survey paper, we examine RS for biomedical items and summarize the previous efforts on this topic. For this purpose, several papers published between 2015 and November 2021 from five scientific databases are retrieved for this purpose. After examining and selecting publications, 60 papers are categorized using a RS technique. The results show that in the last decade, the digital information (e.g., laboratory results, treatment plans, and medical reports) available for patient-oriented decision-making has increased dramatically. Because this information is scattered everywhere and in text form, one solution was to centralize it into personal health record systems, which can be managed like a classical information retrieval (IR) problem. A RS provides its users with medical information intended to be highly relevant for healthcare development. Most health data sets are not publicly available (≠ 30%) because they are sensitive and derived from private clinical data. In contrast, 60% of the datasets are available, but most lack key characteristics to enable good reproducibility and extensibility. These values result from some studies in chemistry with open databases such as CCLE and GDSC. The most significant data source is presented, with a prevalence of the GDSC, CCLE, and DrugBank databases. In general, the RS has made remarkable progress in recent years, developing various RS tools and datasets. However, amid this progress, the poor availability and quality of documentation for RS tools and datasets pose a significant challenge to the reproducibility of RS research. Replicability is essential for corroborating and expand the impact of research findings, but the lack of comprehensive and reliable documentation hinders this process. Poor documentation of RS tools makes it difficult for researchers to understand their functionality and replicate experiments. Similarly, poor documentation of datasets makes it difficult to assess their quality and reproduce experiments using the same data. Enforcing documentation standards, encouraging detailed information in research papers, and collaborating on best practices can improve the reproducibility and impact of RS research. As mentioned above, of the 60 articles analyzed, only 14 provide access to the source code on GitHub. An international shared-evaluation

would also be a boost to the mitigate this problem, like the Message Understanding Conferences (MUC) and Text REtrieval Conference (TREC) challenges were to IR. Following good examples such as BioCreative and BioASQ for biomedical text mining. An outstanding feature in this study is that most didn't follow the standard format <user, item, rating>, commonly used in RS. Another relevant point that deserves mention is that the model-based CF is the most used regardless of the field. This fact is primarily due to applying the ML algorithms (totaling 19 papers). Regarding the performance measurement of the recommendation techniques, the metrics remain offline, and the accuracy of recommended items is based on classification metrics (precision, recall, and AUROC). The research examines different approaches to utilizing KG as supplementary data to improve recommendation results and provide interpretable information in the recommendation process based on real-life experiences. New methods are emerging, proving that KG-based recommendation is a viable solution. Despite the numerous studies conducted in recent years, this is still an emerging field of research. More comprehensive studies are needed. We hope this survey paper can help readers better understand work in this area. Hopefully, new work can emerge in this area, minimizing the "cold-start" problem that is currently highly prevalent.

5.1 Future directions and challenges of the Knowledge-graph based recommendation

The future and potential of BMKG are vast and exciting. KGs provide a powerful way to organize, integrate, and analyze biomedical data meaningfully, given the rapid growth of biomedical data and the increasing need for personalized medicine and precision healthcare. They are a powerful tool for understanding complex relationships between entities in biomedicine. Some applications are identifying new drug targets, predict drug-drug interactions, and develop personalized patient treatment plans by representing these relationships as nodes and edges in a graph. In addition to their research and clinical applications, BMKG also have the potential to transform healthcare delivery by improving the interoperability of disparate electronic health record systems and enabling more accurate and efficient diagnosis and treatment. One of the challenges is the need for accurate and comprehensive data. While a wealth of biomedical data is available, much of it is still siloed in different databases and formats, and there are significant challenges in integrating and harmonizing this data. However, with the increasing adoption of standard data formats and the development of new data integration and analysis technologies, the potential for BMKG is enormous. Several challenges are associated with using BMKG. Some of the major challenges include: (1) **data integration**, (2) **KG quality**, (3) **scalability**, (4) **domain expertise**, (5) **interpretability**, and (6) **evaluation**. In summary, BMKG contain heterogeneous data from multiple sources. Integrating this data into a single KG can be challenging due to data formats, quality, and completeness differences. Incomplete or inaccurate information can lead to poor recommendations, and a growing KG can lead to slower recommendation times and reduced usability. Building and maintaining a KG requires significant biomedical and data science expertise, which can be a barrier for organizations with limited resources or expertise. As with traditional RS, the interpretability of the recommendation results is critical, and healthcare professionals need to understand how the recommendations are generated. Evaluating the performance of KG-based RS can be challenging. Developing appropriate evaluation metrics and benchmark datasets is critical to ensuring the quality and reliability of recommendations.

6 AUTHOR CONTRIBUTIONS STATEMENT

M.B. started by creating the dataset from a set of related papers. Afterwards, M.P. adds updated documents and categorizes them by areas, metrics, and methods. M.B and M.P. designed the survey structure. M.P. wrote the manuscript. All authors participated in the design and validation of the study.

7 ACKNOWLEDGMENTS

The authors thank the anonymous reviewers for their valuable suggestions. This work was supported by FCT through funding of Deep Semantic Tagger (DeST) project (ref. PTDC/CCI-BIO/28685/2017) and LASIGE Research Unit (ref. UIDB/00408/2020 and ref. UIDP/00408/2020).

8 DATA AVAILABILITY

All data relevant to the study are included in the paper or uploaded as supplementary information.

9 COMPETING INTERESTS

The authors declare they have no conflicts of interest.

Table 6. Overview of all surveyed papers which apply recommender systems and their approaches for biomedical items. The first column shows the paper's author(s) and publication year. The second column lists the field/area we defined above, and the following two list each paper's users and items. Column five shows the data sources in case they have been disclosed. The sixth column lists the paradigm(s) (i.e., recommendation system strategy) employed by the algorithm(s) in the paper. Column seven lists the presence/use of knowledge-graph for each paper. The eighth column shows the evaluation metrics used in the paper. The last two columns show whether the datasets are available for replication and public. The abbreviations used in the table can be found at the end of this.

Author(s) and Year	Field	User	Item	Data source	Str.	KG	Evaluation metric	Avail.	Public
Zhang et al. (2015) [107]	Chemistry	health consumers	drugs	Online drug store Walgreens	CF	No	PPV, RE, F1	n.a.	public
Zhang et al. (2016) [106]	Chemistry	drugs	drug response	SIDER, PubChem, DrugBank	CF	No	PPV, RE, F1, SP, ACC, AUPRC, AUROC	avail.	public
Macedo et al. (2016) [60]	Health	healthcare professionals	health information	PubMed	CB	No	PPV, RE	avail.	private
Hao and Blair (2016) [38]	Health	patients	health information	NHANES, SUPPORT, Chronic Kidney, Dermatology	CF	No	RE, SP	avail.	public
Corrado et al. (2016) [21]	Genetic	genes	RNA	AURA 2	CF	Yes	PPV, AUROC	n.a.	public
Chen at al. (2016)[18]	Health	healthcare professionals	health information	Clinical Data	CF	No	PPV, RE, AUROC	n.a.	private
Kim at al. (2016) [51]	Genetic	genes	disease	Singh, GSE15484, TCGA_PRAD	CF	No	AUC, PCC	avail.	public
Ezzat et al. (2016) [25]	Chemistry	drugs	protein	Enzyme, Ion channel, GPCR, Nuclear receptor	CF	Yes	AUPRC	avail.	public
Mustaqeem et al. (2017) [65]	Health	patients	health information	Clinical Data	CF	No	RE, SP, ARMSE	avail.	private
Bocanegra et al. (2017) [12]	Health	health consumers	health educational	Medical videos	CB	No	PPV, nDCG	n.a.	public
Gräßer et al. (2017) [32]	Health	patients	health information	Clinical Data	CF	No	PPV, ARMSE, PCC, Cosine similarity, SC	n.a.	private
Chen et al. (2017) [17]	Health	healthcare professionals	health information	Clinical Data	CF	No	PPV, RE, AUROC	n.a.	private
Fan et al. (2017) [26]	Chemistry	drugs	CNS side effects	PubChem, DrugBank, Clinical Data, KEGG	CF	No	AUROC, AUPRC	avail.	public
Medina-Moreira et al. (2017) [61]	Health	patients	health information	Clinical Data	CF	No	PPV, RE, F1	n.a.	private
Peska et al. (2017) [74]	Chemistry	drugs	drug-targets	Enzyme, Ion channel, GPCR, Nuclear receptor	CF	No	AUROC, nDCG	avail.	public

Continued on next page

Table 6 – *Continued from previous page*

Author(s) and Year	Field	User	Item	Data source	Str.	KG	Evaluation metric	Avail.	Public
Yang et al. (2017) [99]	Health	health consumers	meals	Personal Data	CB	No	PCC, ARMSE, MAE	avail.	public
Ozsoy et al. (2018) [72]	Health	drugs	disease	DrugBank, PubChem, UniProt, SIDER	CF	No	PPV, RE, F1, AUPRC, AUROC	avail.	public
Zhang et al. (2018) [105]	Biology	cell-lines	drug response	GDSC, CCLE	CF	No	PCC, ARMSE	avail.	public
Wang et al. (2018) [90]	Chemistry	drugs	disease	ZINC, ChEMBL, DrugBank	CF	Yes	RE	avail.	public
Seko et al. (2018) [83]	Chemistry	chemical compounds	chemical	ICSD	CB	No	ACC	avail.	public
Suphavilai et al. (2018) [87]	Biology	cell-lines	drug response	GDSC, CCLE	CF	No	nDCG, Spearman, HR	avail.	public
Pustozerov et al. (2018) [76]	Health	patients	disease	Personal Data	CB	No	AUC, ARMSE, MAE, MAPE	n.a.	private
Iatraki et al. (2018) [44]	Health	patients	health documents	Patient Preferences	CB	No	Qualitative feedback	n.a.	private
Agapito et al. (2018) [2]	Health	patients	nutritional advices	Personal Data	Other	No	RE, SP	avail.	private
Liu et al. (2018) [59]	Biology	cell-lines	drug response	GDSC, CCLE	CF	No	PCC, ARMSE	avail.	public
Katzman et al. (2018) [50]	Health	patients	treatments	WHAS, SUPPORT, METABRIC	CF	No	MSE	avail.	public
Yasuo et al. (2018) [102]	Chemistry	drugs	drug-targets	DrugBank	CF	No	AUROC	avail.	public
Hao et al. (2018) [39]	Chemistry	drugs	drug-targets	KEGG BRITE, BRENDA, SuperTarget, DrugBank	CF	No	AUROC, AUPRC, MPR	avail.	public
Galeano and Paccanaro (2018) [29]	Chemistry	drugs	drug response	SIDER	CF	No	AUROC, AUPRC	avail.	public
Han et al. (2018) [37]	Health	patients	health professional	Clinical Data	CF	No	PPV, HR	n.a.	private
Mehrabad et al. (2018) [62]	Genetic	genes	protein	RAT, FLY, HUMAN, Du et al., DBMLoc and Höglund	CF	No	PPV, RE, F1, ACC	avail.	public
Wang et al. (2018) [92]	Biology	drugs	drug response	GDSC, PubChem	CF	No	AUROC, ROC		public
Torrent-Fontbona and López (2019) [88]	Health	health consumers	drugs	Clinical Data	CB	No	Qualitative feedback	avail.	public
Nouh et al. (2019) [69]	Health	health consumers	health information	Personal Services	H	No	MSE, MAPE, MAE	n.a.	private
Jabeen et al. (2019) [45]	Health	health consumers	disease	Clinical Data	CF	No	PCC, RE, MAE	n.a.	private
Lan et al. (2019) [55]	Chemistry	drugs	enzyme proteins	LMMD	CF	No	F1	n.a.	
Zeng et al. (2019) [104]	Genetic	genes	disease	Microarray, HumanNet, Gene-phenotype associations, STRING	CF	Yes	PPV, RE	avail.	public
Sosnina et al. (2020) [85]	Chemistry	drugs	drug response	CHEMBL	H	No	AUROC	avail.	public
Ferretto et al. (2020) [28]	Health	patients	treatments	Personal Services	CB	No	PCC	n.a.	private
Mustaqeem et al. (2020) [66]	Health	patients	disease	Clinical Data	CF	No	PPV, RE, MAE	n.a.	private
Embadi and Eslahchi (2020) [24]	Biology	cell-lines	drug response	GDSC, CCLE	CF	No	PPV, RE, F1, SP, ACC, MCC	avail.	public

Continued on next page

Table 6 – *Continued from previous page*

Author(s) and Year	Field	User	Item	Data source	Str.	KG	Evaluation metric	Avail.	Public
Wang et al. (2020) [94]	Health	patients	health information	Clinical Data	Other	Yes	PPV, MAP	n.a.	private
Rohani et al. (2020) [79]	Health	patients	activities advices	Personal Data	CB	No	Qualitative feedback	n.a.	private
Ha et al. (2020) [35]	Genetic	disease	RNA	dbDEMC, HMDD, miR2Disease	CF	Yes	AUROC	avail.	public
Ha et al. (2020) [36]	Genetic	disease	RNA	dbDEMC, HMDD, miR2Disease	CF	Yes	AUROC, AUPRC	avail.	public
Barros et al. (2020) [7]	Chemistry	authors' articles	chemical compounds	CheRM-20	H	Yes	PPV, RE, F1, MRR, nDCG, lAUC	avail.	public
Lim and Xie (2020) [56]	Genetic	genes	genes	ChEA	CF	No	AUROC, MAP, HLU, MPR	n.a.	public
Ren et al. (2020) [77]	Health	healthcare professionals	search terms	Clinical Data	CF	No	HR	n.a.	private
Barros et al. (2021) [8]	Chemistry	authors' articles	chemical compounds	CheRM-20	H	Yes	PPV, RE, F1, MRR, nDCG, lAUC	avail.	public
Koras et al. (2021) [54]	Biology	cell-lines	drugs	GDSC	CF	No	PCC, AUROC, IC50, ARMSE	avail.	public
Brandão et al. (2021) [14]	Biology	cell-lines	drugs	GDSC	CF	No	HR	avail.	public
Zheng et al. (2021)[108]	Chemistry	healthcare professionals	drugs	EMR, DrugBank, YaoZhi	CF	Yes	PPV, RE, F1	n.a.	private
Yue et al. (2021) [103]	Health	patients	health information	SARA	CF	No	PCC, ARMSE, MAE	avail.	public
Ormel et al. (2021) [71]	Health	patients	health information	Personal Data	CB	No	Qualitative feedback	n.a.	private
Sadeghi et al. (2021) [80]	Chemistry	drugs	drug response	PREDICT, DrugNet, CDataSet, TL-HBGI	CF	No	AUC	avail.	public
Barros et al. (2021) [9]	Health	articles	ontologies	PubMed	H	Yes	PPV, RE, MRR	avail.	public
Ochoa et al. (2021) [70]	Health	healthcare professionals	health information	Pakistan	CF	No	PPV, RE, ACC, ARMSE	avail.	public
Chen et al. (2021) [16]	Health	users	health information	Personal Data	H	No	Qualitative feedback	n.a.	private
Ammar et al. (2021) [4]	Health	health consumers	health information	Personal Data	H	Yes	Qualitative feedback	n.a.	private
Zhu et al. (2021) [109]	Genetic	genes	protein	DIP, Krogan, Gavin, Pfam	CF	No	AUROC, AUPRC, AUC	avail.	public

Recommendation system strategy (Str.), content-based (CB), collaborative filtering (CF), Hybrid (H), Available (avail.), Non-available (n.a.)

Metrics: Accuracy (ACC), Area Under Curve (AUC), Precision–Recall Curve (AUPRC), Area Under the Receiver Operating Characteristic curve (AUROC), Root Mean Square Error (RMSE), F-measure (F1), Half-Life Utility (HLU), Half-maximal inhibitory concentration (IC50), Hit Ratio (HR), limited Area Under the Curve (lAUC), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), Mean Average Precision (MAP), Matthews Correlation Coefficient (MCC), Mean Percentile Ranking (MPR), Mean Reciprocal Rank (MRR), Mean Square Error (MSE), normalized Discounted Cumulative Gain (nDCG), Pearson Correlation Coefficient (PCC), Precision (PPV), Recall (RE), Specificity (SP).

Database: Database of Interacting Proteins (DIP), Atlas of UTR Regulatory Activity (AURA 2), Cancer Cell Line Encyclopedia (CCLE), ChIP Enrichment Analysis (ChEA), Electronic Medical Records (EMR), G protein-coupled receptors (GPCR), Gene Expression Dataset from prostate cancer patients (GSE15484), Genomics of Drug Sensitivity in Cancer (GDSC), Inorganic Crystal Structure Database (ICSD), Laboratory of Molecular Modeling and Design repository (LMMD), Kyoto Encyclopedia of Genes and Genomes (KEGG), National Health and Nutrition Examination Survey (NHANES), RNA Recognition Motif (RRM), Side effect resource (SIDER), Study to Understand Prognoses Preferences Outcomes and Risks of Treatment (SUPPORT), The Cancer Genome Atlas: Prostate Adenocarcinoma (TCGA_PRAD) and Worcester Heart Attack Study (WHAS).

REFERENCES

[1] Gediminas Adomavicius and Alexander Tuzhilin. 2005. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE transactions on knowledge and data engineering* 17, 6 (2005), 734–749. https://doi.org/10.1109/TKDE.2005.99

[2] Giuseppe Agapito, Mariadelina Simeoni, Barbara Calabrese, Ilaria Caré, Theodora Lamprinoudi, Pietro H Guzzi, Arturo Pujia, Giorgio Fuiano, and Mario Cannataro. 2018. DIETOS: A dietary recommender system for chronic diseases monitoring and management. *Computer methods and programs in biomedicine* 153 (2018), 93–104. https://doi.org/10.1016/j.cmpb.2017.10.014

[3] Charu C Aggarwal. 2016. Ensemble-Based and Hybrid Recommender Systems. In *Recommender systems*. Springer, 199–224. https://doi.org/10.1007/978-3-319-29659-3_6

[4] Nariman Ammar, James E Bailey, Robert L Davis, and Arash Shaban-Nejad. 2021. Using a Personal Health Library–Enabled mHealth Recommender System for Self-Management of Diabetes Among Underserved Populations: Use Case for Knowledge Graphs and Linked Data. *JMIR Formative Research* 5, 3 (2021), e24738. https://doi.org/10.2196/24738

[5] Jordi Barretina, Giordano Caponigro, Nicolas Stransky, Kavitha Venkatesan, Adam A Margolin, Sungjoon Kim, Christopher J Wilson, Joseph Lehár, Gregory V Kryukov, Dmitriy Sonkin, et al. 2012. The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* 483, 7391 (2012), 603–607. https://doi.org/10.1038/nature11003

[6] Márcia Barros, André Moitinho, and Francisco M Couto. 2019. Using research literature to generate datasets of implicit feedback for recommending scientific items. *IEEE Access* 7 (2019), 176668–176680. https://doi.org/10.1109/ACCESS.2019.2958002

[7] Márcia Barros, André Moitinho, and Francisco M Couto. 2020. Hybrid semantic recommender system for chemical compounds. *Advances in Information Retrieval* 12036 (2020), 94. https://doi.org/10.1007/978-3-030-45442-5_12

[8] Márcia Barros, André Moitinho, and Francisco M Couto. 2021. Hybrid semantic recommender system for chemical compounds in large-scale datasets. *Journal of cheminformatics* 13, 1 (2021), 1–18.

[9] Márcia Barros, Pedro Ruas, Diana Sousa, Ali Haider Bangash, and Francisco M Couto. 2021. COVID-19 recommender system based on an annotated multilingual corpus. *Genomics & Informatics* 19, 3 (2021). https://doi.org/10.5808%2Fgi.21008

[10] J. Bennett and S. Lanning. 2007. The Netflix Prize. In *Proceedings of the KDD Cup Workshop 2007*. ACM, New York, 3–6.

[11] Jesús Bobadilla, Fernando Ortega, Antonio Hernando, and Abraham Gutiérrez. 2013. Recommender systems survey. *Knowledge-based systems* 46 (2013), 109–132. https://doi.org/10.1016/j.knosys.2013.03.012

[12] Carlos Luis Sanchez Bocanegra, Jose Luis Sevillano Ramos, Carlos Rizo, Anton Civit, and Luis Fernandez-Luque. 2017. HealthRecSys: A semantic content-based recommender system to complement health videos. *BMC medical informatics and decision making* 17, 1 (2017), 1–10. https://doi.org/10.1186/s12911-017-0431-7

[13] Antoine Bordes, Nicolas Usunier, Alberto Garcia-Duran, Jason Weston, and Oksana Yakhnenko. 2013. Translating Embeddings for Modeling Multi-relational Data. *Advances in neural information processing systems* 26 (2013).

[14] Liliana Brandão, Fernando Paulo Belfo, and Alexandre Silva. 2021. Wavelet-based cancer drug recommender system. *Procedia Computer Science* 181 (2021), 487–494. https://doi.org/10.1016/j.procs.2021.01.194

[15] Robin Burke. 2002. Hybrid recommender systems: Survey and experiments. *User modeling and user-adapted interaction* 12, 4 (2002), 331–370. https://doi.org/10.1023/A:1021240730564

[16] Jinying Chen, Thomas K Houston, Jamie M Faro, Catherine S Nagawa, Elizabeth A Orvek, Amanda C Blok, Jeroan J Allison, Sharina D Person, Bridget M Smith, and Rajani S Sadasivam. 2021. Evaluating the use of a recommender system for selecting optimal messages for smoking cessation: patterns and effects of user-system engagement. *BMC public health* 21, 1 (2021), 1–13. https://doi.org/10.1186/s12889-021-11803-8

[17] Jonathan H Chen, Muthuraman Alagappan, Mary K Goldstein, Steven M Asch, and Russ B Altman. 2017. Decaying relevance of clinical data towards future decisions in data-driven inpatient clinical order sets. *International journal of medical informatics* 102 (2017), 71–79. https://doi.org/10.1016/j.ijmedinf.2017.03.006

[18] Jonathan H Chen, Tanya Podchiyska, and Russ B Altman. 2016. OrderRex: clinical order decision support and outcome predictions by data-mining electronic medical records. *Journal of the American Medical Informatics Association* 23, 2 (2016), 339–348. https://doi.org/10.1093/jamia/ocv091

[19] Qing Cong, Zhiyong Feng, Fang Li, Li Zhang, Guozheng Rao, and Cui Tao. 2018. Constructing Biomedical Knowledge Graph Based on SemMedDB and Linked Open Data. In *2018 IEEE international conference on bioinformatics and biomedicine (BIBM)*. IEEE, 1628–1631. https://doi.org/10.1109/BIBM.2018.8621568

[20] Gene Ontology Consortium. 2019. The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic acids research* 47, D1 (2019), D330–D338. https://doi.org/10.1093/nar/gky1055

[21] Gianluca Corrado, Toma Tebaldi, Fabrizio Costa, Paolo Frasconi, and Andrea Passerini. 2016. RNAcommender: genome-wide recommendation of RNA–protein interactions. *Bioinformatics* 32, 23 (2016), 3627–3634. https://doi.org/10.1093/bioinformatics/btw517

[22] Paolo Cremonesi, Yehuda Koren, and Roberto Turrin. 2010. Performance of recommender algorithms on top-n recommendation tasks. In *Proceedings of the fourth ACM conference on Recommender systems*. 39–46. https://doi.org/10.1145/1864708.1864721

[23] Robin De Croon, Leen Van Houdt, Nyi Nyi Htun, Gregor Štiglic, Vero Vanden Abeele, Katrien Verbert, et al. 2021. Health Recommender Systems: Systematic Review. *Journal of Medical Internet Research* 23, 6 (2021), e18035. https://doi.org/10.2196/18035

[24] Akram Emdadi and Changiz Eslahchi. 2020. DSPLMF: A Method for Cancer Drug Sensitivity Prediction Using a Novel Regularization Approach in Logistic Matrix Factorization. *Frontiers in genetics* 11 (2020), 75. https://doi.org/10.3389/fgene.2020.00075

[25] Ali Ezzat, Peilin Zhao, Min Wu, Xiao-Li Li, and Chee-Keong Kwoh. 2016. Drug-Target Interaction Prediction with Graph Regularized Matrix Factorization. *IEEE/ACM transactions on computational biology and bioinformatics* 14, 3 (2016), 646–656. https://doi.org/10.1109/TCBB.2016.2530062

[26] Jun Fan, Jing Yang, and Zhenran Jiang. 2018. Prediction of Central Nervous System Side Effects Through Drug Permeability to Blood–Brain Barrier and Recommendation Algorithm. *Journal of Computational Biology* 25, 4 (2018), 435–443. https://doi.org/10.1089/cmb.2017.0149

[27] Maria Fernandes, Jérémie Decouchant, and Francisco M. Couto. 2023. Chapter Two - Security, privacy, and trust management in DNA computing. In *Perspective of DNA Computing in Computer Science*, Suyel Namasudra (Ed.). Advances in Computers, Vol. 129. Elsevier, 39–81. https://doi.org/10.1016/bs.adcom.2022.08.009

[28] Luciano Rodrigo Ferretto, Ericles Andrei Bellei, Daiana Biduski, Luiz Carlos Pereira Bin, Mirella Moura Moro, Cristiano Roberto Cervi, and Ana Carolina Bertoletti De Marchi. 2020. A Physical Activity Recommender System for Patients With Arterial Hypertension. *IEEE Access* 8 (2020), 61656–61664. https://doi.org/10.1109/ACCESS.2020.2983564

[29] Diego Galeano and Alberto Paccanaro. 2018. A Recommender System Approach for Predicting Drug Side Effects. In *2018 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 1–8. https://doi.org/10.1109/IJCNN.2018.8489025

[30] Anna Gaulton, Anne Hersey, Michał Nowotka, A Patricia Bento, Jon Chambers, David Mendez, Prudence Mutowo, Francis Atkinson, Louisa J Bellis, Elena Cibrián-Uhalte, et al. 2017. The ChEMBL database in 2017. *Nucleic acids research* 45, D1 (2017), D945–D954. https://doi.org/10.1093/nar/gkw1074

[31] Fan Gong, Meng Wang, Haofen Wang, Sen Wang, and Mengyue Liu. 2021. SMR: Medical Knowledge Graph Embedding for Safe Medicine Recommendation. *Big Data Research* 23 (2021), 100174. https://doi.org/10.1016/j.bdr.2020.100174

[32] Felix Gräßer, Stefanie Beckert, Denise Küster, Jochen Schmitt, Susanne Abraham, Hagen Malberg, and Sebastian Zaunseder. 2017. Therapy Decision Support Based on Recommender System Methods. *Journal of healthcare engineering* 2017 (2017). https://doi.org/10.1155/2017/8659460

[33] Asela Gunawardana and Guy Shani. 2015. Evaluating Recommender Systems. In *Recommender systems handbook*. Springer, 265–308. https://doi.org/10.1007/978-1-0716-2197-4_15

[34] Qingyu Guo, Fuzhen Zhuang, Chuan Qin, Hengshu Zhu, Xing Xie, Hui Xiong, and Qing He. 2020. A Survey on Knowledge Graph-Based Recommender Systems. *IEEE Transactions on Knowledge and Data Engineering* (2020). https://doi.org/10.1109/TKDE.2020.3028705

[35] Jihwan Ha, Chihyun Park, Chanyoung Park, and Sanghyun Park. 2020. IMIPMF: Inferring miRNA-disease interactions using probabilistic matrix factorization. *Journal of Biomedical Informatics* 102 (2020), 103358. https://doi.org/10.1016/j.jbi.2019.103358

[36] Jihwan Ha, Chihyun Park, Chanyoung Park, and Sanghyun Park. 2020. Improved Prediction of miRNA-Disease Associations Based on Matrix Completion with Network Regularization. *Cells* 9, 4 (2020), 881. https://doi.org/10.3390/cells9040881

[37] Qiwei Han, Inigo Martinez de Rituerto de Troya, Mengxin Ji, Manas Gaur, and Leid Zejnilovic. 2018. A Collaborative Filtering Recommender System in Primary Care: Towards a Trusting Patient-Doctor Relationship. In *2018 IEEE International Conference on Healthcare Informatics (ICHI)*. IEEE, 377–379. https://doi.org/10.1109/ICHI.2018.00062

[38] Fang Hao and Rachael Hageman Blair. 2016. A comparative study: classification vs. user-based collaborative filtering for clinical prediction. *BMC medical research methodology* 16, 1 (2016), 1–14. https://doi.org/10.1186/s12874-016-0261-9

[39] Ming Hao, Stephen H Bryant, and Yanli Wang. 2018. A new chemoinformatics approach with improved strategies for effective predictions of potential drugs. *Journal of cheminformatics* 10, 1 (2018), 1–9. https://doi.org/10.1186/s13321-018-0303-x

[40] F Maxwell Harper and Joseph A Konstan. 2015. The movielens datasets: History and context. *ACM transactions on interactive intelligent systems (TIIS)* 5, 4 (2015), 1–19. https://doi.org/10.1145/2827872

[41] Janna Hastings, Gareth Owen, Adriano Dekker, Marcus Ennis, Namrata Kale, Venkatesh Muthukrishnan, Steve Turner, Neil Swainston, Pedro Mendes, and Christoph Steinbeck. 2016. ChEBI in 2016: Improved services and an expanding collection of metabolites. *Nucleic acids research* 44, D1 (2016), D1214–D1219. https://doi.org/10.1093/nar/gkv1031

[42] Rachel A Hodos, Brian A Kidd, Shameer Khader, Ben P Readhead, and Joel T Dudley. 2016. Computational Approaches to Drug Repurposing and Pharmacology. *Wiley interdisciplinary reviews. Systems biology and medicine* 8, 3 (2016), 186. https://doi.org/10.1002%2Fwsbm.1337

[43] Martin Hofmann-Apitius, Gordon Ball, Stephan Gebel, Shweta Bagewadi, Bernard De Bono, Reinhard Schneider, Matt Page, Alpha Tom Kodamullil, Erfan Younesi, Christian Ebeling, et al. 2015. Bioinformatics Mining and Modeling Methods for the Identification of Disease Mechanisms in Neurodegenerative Disorders. *International journal of molecular sciences* 16, 12 (2015), 29179–29206. https://doi.org/10.3390/ijms161226148

[44] Galatia Iatraki, Haridimos Kondylakis, Lefteris Koumakis, Maria Chatzimina, Eleni Kazantzaki, Kostas Marias, and Manolis Tsiknakis. 2018. Personal Health Information Recommender: implementing a tool for the empowerment of cancer patients. *ecancermedicalscience*

12 (2018). https://doi.org/10.3332%2Fecancer.2018.851

[45] Fouzia Jabeen, Muazzam Maqsood, Mustansar Ali Ghazanfar, Farhan Aadil, Salabat Khan, Muhammad Fahad Khan, and Irfan Mehmood. 2019. An IoT based efficient hybrid recommender system for cardiovascular disease. *Peer-to-Peer Networking and Applications* 12, 5 (2019), 1263–1276. https://doi.org/10.1007/s12083-019-00733-3

[46] Guoliang Ji, Shizhu He, Liheng Xu, Kang Liu, and Jun Zhao. 2015. Knowledge Graph Embedding via Dynamic Mapping Matrix. In *Proceedings of the 53rd annual meeting of the association for computational linguistics and the 7th international joint conference on natural language processing (volume 1: Long papers)*. 687–696.

[47] Minoru Kanehisa, Miho Furumichi, Yoko Sato, Masayuki Kawashima, and Mari Ishiguro-Watanabe. 2023. KEGG for taxonomy-based analysis of pathways and genomes. *Nucleic Acids Research* 51, D1 (2023), D587–D592. https://doi.org/10.1093/nar/gkac963

[48] Minoru Kanehisa, Miho Furumichi, Mao Tanabe, Yoko Sato, and Kanae Morishima. 2017. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic acids research* 45, D1 (2017), D353–D361. https://doi.org/10.1093/nar/gkw1092

[49] Minoru Kanehisa, Yoko Sato, and Masayuki Kawashima. 2022. KEGG mapping tools for uncovering hidden features in biological data. *Protein Science* 31, 1 (2022), 47–53. https://doi.org/10.1002/pro.4172

[50] Jared L Katzman, Uri Shaham, Alexander Cloninger, Jonathan Bates, Tingting Jiang, and Yuval Kluger. 2018. DeepSurv: personalized treatment recommender system using a Cox proportional hazards deep neural network. *BMC medical research methodology* 18, 1 (2018), 1–12. https://doi.org/10.1186/s12874-018-0482-1

[51] Hyunjin Kim, Sang-Min Choi, and Sanghyun Park. 2016. GSEH: A Novel Approach to Select Prostate Cancer-Associated Genes Using Gene Expression Heterogeneity. *IEEE/ACM transactions on computational biology and bioinformatics* 15, 1 (2016), 129–146. https://doi.org/10.1109/TCBB.2016.2618927

[52] Sunghwan Kim, Paul A Thiessen, Evan E Bolton, Jie Chen, Gang Fu, Asta Gindulyte, Lianyi Han, Jane He, Siqian He, Benjamin A Shoemaker, et al. 2016. PubChem Substance and Compound databases. *Nucleic acids research* 44, D1 (2016), D1202–D1213. https://doi.org/10.1093/nar/gkv951

[53] Sebastian Köhler, Michael Gargano, Nicolas Matentzoglu, Leigh C Carmody, David Lewis-Smith, Nicole A Vasilevsky, Daniel Danis, Ganna Balagura, Gareth Baynam, Amy M Brower, et al. 2021. The Human Phenotype Ontology in 2021. *Nucleic acids research* 49, D1 (2021), D1207–D1217. https://doi.org/10.1093/nar/gkaa1043

[54] Krzysztof Koras, Ewa Kizling, Dilafruz Juraeva, Eike Staub, and Ewa Szczurek. 2021. Interpretable deep recommender system model for prediction of kinase inhibitor efficacy across cancer cell lines. *bioRxiv* (2021). https://doi.org/10.1038/s41598-021-94564-z

[55] Chao Lan, Sai Nivedita Chandrasekaran, and Jun Huan. 2019. On the Unreported-Profile-is-Negative Assumption for Predictive Cheminformatics. *IEEE/ACM transactions on computational biology and bioinformatics* 17, 4 (2019), 1352–1363. https://doi.org/10.1109/TCBB.2019.2913855

[56] Hansaim Lim and Lei Xie. 2020. A New Weighted Imputed Neighborhood-Regularized Tri-Factorization One-Class Collaborative Filtering Algorithm: Application to Target Gene Prediction of Transcription Factors. *IEEE/ACM transactions on computational biology and bioinformatics* 18, 1 (2020), 126–137. https://doi.org/10.1109/TCBB.2020.2968442

[57] Yankai Lin, Zhiyuan Liu, Maosong Sun, Yang Liu, and Xuan Zhu. 2015. Learning Entity and Relation Embeddings for Knowledge Graph Completion. In *Twenty-ninth AAAI conference on artificial intelligence*. https://doi.org/10.1609/aaai.v29i1.9491

[58] David Liu, Catherine Lu, and Karanveer Mohan. 2014. Pinterest analysis and recommendations. *Course report. Stanford University* (2014).

[59] Hui Liu, Yan Zhao, Lin Zhang, and Xing Chen. 2018. Anti-cancer Drug Response Prediction Using Neighbor-Based Collaborative Filtering with Global Effect Removal. *Molecular Therapy-Nucleic Acids* 13 (2018), 303–311. https://doi.org/10.1016/j.omtn.2018.09.011

[60] Alessandra Alaniz Macedo, Juliana Tarossi Pollettini, José Augusto Baranauskas, and Julia Carmona Almeida Chaves. 2016. A Health Surveillance Software Framework to deliver information on preventive healthcare strategies. *Journal of biomedical informatics* 62 (2016), 159–170. https://doi.org/10.1016/j.jbi.2016.06.002

[61] José Medina-Moreira, Oscar Apolinario, Harry Luna-Aveiga, Katty Lagos-Ortiz, Mario Andrés Paredes-Valverde, and Rafael Valencia-García. 2017. A Collaborative Filtering Based Recommender System for Disease Self-management. In *International Conference on Technologies and Innovation*. Springer, 60–71. https://doi.org/10.1007/978-3-319-67283-0_5

[62] Elnaz Mirzaei Mehrabad, Reza Hassanzadeh, and Changiz Eslahchi. 2018. PMLPR: A novel method for predicting subcellular localization based on recommender systems. *Scientific reports* 8, 1 (2018), 1–10. https://doi.org/10.1038/s41598-018-30394-w

[63] Peter Mika, Abraham Bernstein, Chris Welty, Craig Knoblock, Denny Vrandečić, Paul Groth, Natasha Noy, Krzysztof Janowicz, and Carole Goble. 2014. *The Semantic Web–ISWC 2014: 13th International Semantic Web Conference, Riva del Garda, Italy, October 19-23, 2014. Proceedings, Part I.* Vol. 8796. Springer.

[64] Sameh K Mohamed, Aayah Nounu, and Vít Nováček. 2019. Drug target discovery using knowledge graph embeddings. In *Proceedings of the 34th ACM/SIGAPP Symposium on Applied Computing*. 11–18. https://doi.org/10.1145/3297280.3297282

[65] Anam Mustaqeem, Syed Muhammad Anwar, Abdul Rashid Khan, and Muhammad Majid. 2017. A statistical analysis based recommender model for heart disease patients. *International journal of medical informatics* 108 (2017), 134–145. https://doi.org/10.1016/j.ijmedinf.2017.10.008

[66] Anam Mustaqeem, Syed Muhammad Anwar, and Muhammad Majid. 2020. A modular cluster based collaborative recommender system for cardiac patients. *Artificial intelligence in medicine* 102 (2020), 101761. https://doi.org/10.1016/j.artmed.2019.101761

[67] Maximilian Nickel, Lorenzo Rosasco, and Tomaso Poggio. 2016. Holographic embeddings of knowledge graphs. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 30. https://doi.org/10.1609/aaai.v30i1.10314

[68] Maximilian Nickel, Volker Tresp, and Hans-Peter Kriegel. 2011. A three-way model for collective learning on multi-relational data. In *Icml*.

[69] Rayan M Nouh, Hyun-Ho Lee, Won-Jin Lee, and Jae-Dong Lee. 2019. A smart recommender based on hybrid learning methods for personal well-being services. *Sensors* 19, 2 (2019), 431. https://doi.org/10.3390/s19020431

[70] Juan G Diaz Ochoa, Orsolya Csiszár, and Thomas Schimper. 2021. Medical recommender systems based on continuous-valued logic and multi-criteria decision operators, using interpretable neural networks. *BMC Medical Informatics and Decision Making* 21, 1 (2021), 1–15. https://doi.org/10.1186/s12911-021-01553-3

[71] Ilja Ormel, Charles C Onu, Mona Magalhaes, Terence Tang, John B Hughes, Susan Law, et al. 2021. Using a Mobile App–Based Video Recommender System of Patient Narratives to Prepare Women for Breast Cancer Surgery: Development and Usability Study Informed by Qualitative Data. *JMIR Formative Research* 5, 6 (2021), e22970. https://doi.org/10.2196/22970

[72] Makbule Guclin Ozsoy, Tansel Özyer, Faruk Polat, and Reda Alhajj. 2018. Realizing drug repositioning by adapting a recommendation system to handle the process. *BMC bioinformatics* 19, 1 (2018), 1–14. https://doi.org/10.1186/s12859-018-2142-1

[73] Matthew J Page and et al. 2021. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 372 (2021). https://doi.org/10.1016/j.ijsu.2021.105906

[74] Ladislav Peska, Krisztian Buza, and Júlia Koller. 2017. Drug-target interaction prediction: a Bayesian ranking approach. *Computer methods and programs in biomedicine* 152 (2017), 15–21. https://doi.org/10.1016/j.cmpb.2017.09.003

[75] Jhonny Pincay, Luis Terán, and Edy Portmann. 2019. Health recommender systems: a state-of-the-art review. In *2019 Sixth International Conference on eDemocracy & eGovernment (ICEDEG)*. IEEE, 47–55. https://doi.org/10.1109/ICEDEG.2019.8734362

[76] Evgenii Pustozerov, Polina Popova, Aleksandra Tkachuk, Yana Bolotko, Zafar Yuldashev, and Elena Grineva. 2018. Development and Evaluation of a Mobile Personalized Blood Glucose Prediction System for Patients With Gestational Diabetes Mellitus. *JMIR mHealth and uHealth* 6, 1 (2018), e6. https://doi.org/10.2196/mhealth.9236

[77] Zhiyun Ren, Bo Peng, Titus K Schleyer, and Xia Ning. 2021. Hybrid collaborative filtering methods for recommending search terms to clinicians. *Journal of Biomedical Informatics* 113 (2021), 103635. https://doi.org/10.1186/s13321-021-00495-2

[78] Francesco Ricci, Lior Rokach, and Bracha Shapira. 2015. Recommender Systems: Introduction and Challenges. In *Recommender systems handbook*. Springer, 1–34. https://doi.org/10.1007/978-1-4899-7637-6_1

[79] Darius A Rohani, Andrea Quemada Lopategui, Nanna Tuxen, Maria Faurholt-Jepsen, Lars V Kessing, and Jakob E Bardram. 2020. MUBS: A Personalized Recommender System for Behavioral Activation in Mental Health. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–13. https://doi.org/10.1145/3313831.3376879

[80] Shagahyegh Sadeghi, Jianguo Lu, and Alioune Ngom. 2021. A network-based drug repurposing method via non-negative matrix factorization. *Bioinformatics (Oxford, England)* (2021), btab826. https://doi.org/10.1093/bioinformatics/btab826

[81] Shengtian Sang, Zhihao Yang, Lei Wang, Xiaoxia Liu, Hongfei Lin, and Jian Wang. 2018. SemaTyP: a knowledge graph based literature mining method for drug discovery. *BMC bioinformatics* 19, 1 (2018), 1–11. https://doi.org/10.1186/s12859-018-2167-5

[82] Lynn M Schriml, James B Munro, Mike Schor, Dustin Olley, Carrie McCracken, Victor Felix, J Allen Baron, Rebecca Jackson, Susan M Bello, Cynthia Bearer, et al. 2022. The Human Disease Ontology 2022 update. *Nucleic acids research* 50, D1 (2022), D1255–D1261. https://doi.org/10.1093/nar/gkab1063

[83] Atsuto Seko, Hiroyuki Hayashi, and Isao Tanaka. 2018. Compositional descriptor-based recommender system for the materials discovery. *The Journal of chemical physics* 148, 24 (2018), 241719. https://doi.org/10.1063/1.5016210

[84] Guy Shani and Asela Gunawardana. 2011. Evaluating recommendation systems. In *Recommender systems handbook*. Springer, 257–297. https://doi.org/10.1007/978-0-387-85820-3_8

[85] Ekaterina A Sosnina, Sergey Sosnin, Anastasia A Nikitina, Ivan Nazarov, Dmitry I Osolodkin, and Maxim V Fedorov. 2020. Recommender Systems in Antiviral Drug Discovery. *ACS Omega* 5, 25 (2020), 15039–15051. https://doi.org/10.1021/acsomega.0c00857

[86] Benjamin Stark, Constanze Knahl, Mert Aydin, and Karim Elish. 2019. A literature review on medicine recommender systems. *International Journal of Advanced Computer Science and Applications (IJACSA)* 10, 8 (2019), 6–13.

[87] Chayaporn Suphavilai, Denis Bertrand, and Niranjan Nagarajan. 2018. Predicting Cancer Drug Response using a Recommender System. *Bioinformatics* 34, 22 (2018), 3907–3914. https://doi.org/10.1093/bioinformatics/bty452

[88] Ferran Torrent-Fontbona and Beatriz López. 2019. Personalized Adaptive CBR Bolus Recommender System for Type 1 Diabetes. *IEEE Journal of Biomedical and Health Informatics* 23 (2019), 387–394. https://doi.org/10.1109/JBHI.2018.2813424

[89] Wytze J Vlietstra, Ronald Zielman, Robin M van Dongen, Erik A Schultes, Floris Wiesman, Rein Vos, Erik M Van Mulligen, and Jan A Kors. 2017. Automated extraction of potential migraine biomarkers using a semantic graph. *Journal of biomedical informatics* 71 (2017), 178–189. https://doi.org/10.1016/j.jbi.2017.05.018

[90] Annie Wang, Hansaim Lim, Shu-Yuan Cheng, and Lei Xie. 2018. ANTENNA, a Multi-Rank, Multi-Layered Recommender System for Inferring Reliable Drug-Gene-Disease Associations: Repurposing Diazoxide as a Targeted Anti-Cancer Therapy. *IEEE/ACM transactions on computational biology and bioinformatics* 15, 6 (2018), 1960–1967. https://doi.org/10.1109/TCBB.2018.2812189

[91] Bin Wang, Yan Ding, Penghui Zhao, Wei Li, Ming Li, Jingbo Zhu, and Shuhong Ye. 2022. Systems pharmacology-based drug discovery and active mechanism of natural products for coronavirus pneumonia (COVID-19): An example using flavonoids. *Computers in Biology and Medicine* 143 (2022), 105241. https://doi.org/10.1016/j.compbiomed.2022.105241

[92] Hang Wang, Jianing Xi, Minghui Wang, and Ao Li. 2018. Dual-Layer Strengthened Collaborative Topic Regression Modeling for Predicting Drug Sensitivity. *IEEE/ACM transactions on computational biology and bioinformatics* 17, 2 (2018), 587–598. https://doi.org/10.1109/TCBB.2018.2864739

[93] Quan Wang, Zhendong Mao, Bin Wang, and Li Guo. 2017. Knowledge graph embedding: A survey of approaches and applications. *IEEE Transactions on Knowledge and Data Engineering* 29, 12 (2017), 2724–2743. https://doi.org/10.1109/TKDE.2017.2754499

[94] Zheyu Wang, Haoce Huang, Liping Cui, Juan Chen, Jiye An, Huilong Duan, Huiqing Ge, Ning Deng, et al. 2020. Using Natural Language Processing Techniques to Provide Personalized Educational Materials for Chronic Disease Patients in China: Development and Assessment of a Knowledge-Based Health Recommender System. *JMIR medical informatics* 8, 4 (2020), e17642. https://doi.org/10.2196/17642

[95] Zhen Wang, Jianwen Zhang, Jianlin Feng, and Zheng Chen. 2014. Knowledge Graph Embedding by Translating on Hyperplanes. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 28. https://doi.org/10.1609/aaai.v28i1.8870

[96] David S Wishart, Yannick D Feunang, An C Guo, Elvis J Lo, Ana Marcu, Jason R Grant, Tanvir Sajed, Daniel Johnson, Carin Li, Zinat Sayeeda, et al. 2018. DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic acids research* 46, D1 (2018), D1074–D1082. https://doi.org/10.1093/nar/gkx1037

[97] David S Wishart, Craig Knox, An Chi Guo, Dean Cheng, Savita Shrivastava, Dan Tzur, Bijaya Gautam, and Murtaza Hassanali. 2008. DrugBank: a knowledge base for drugs, drug actions and drug targets. *Nucleic acids research* 36, suppl_1 (2008), D901–D906. https://doi.org/10.1093/nar/gkm958

[98] Bishan Yang, Wen-tau Yih, Xiaodong He, Jianfeng Gao, and Li Deng. 2014. Embedding entities and relations for learning and inference in knowledge bases. *arXiv preprint arXiv:1412.6575* (2014). https://doi.org/10.48550/arXiv.1412.6575

[99] Longqi Yang, Cheng-Kang Hsieh, Hongjian Yang, John P Pollak, Nicola Dell, Serge Belongie, Curtis Cole, and Deborah Estrin. 2017. Yum-Me: A Personalized Nutrient-Based Meal Recommender System. *ACM Transactions on Information Systems (TOIS)* 36, 1 (2017), 1–31. https://doi.org/10.1145/3072614

[100] Shu Yang, Xiaoxi Liu, and Raymond T Ng. 2020. ProbeRating: a recommender system to infer binding profiles for nucleic acid-binding proteins. *Bioinformatics* 36, 18 (2020), 4797–4804. https://doi.org/10.1093/bioinformatics/btaa580

[101] Wanjuan Yang, Jorge Soares, Patricia Greninger, Elena J Edelman, Howard Lightfoot, Simon Forbes, Nidhi Bindal, Dave Beare, James A Smith, I Richard Thompson, et al. 2012. Genomics of Drug Sensitivity in Cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. *Nucleic acids research* 41, D1 (2012), D955–D961. https://doi.org/10.1093/nar/gks1111

[102] Nobuaki Yasuo, Yusuke Nakashima, and Masakazu Sekijima. 2018. CoDe-DTI: Collaborative Deep Learning-based Drug-Target Interaction Prediction. In *2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. IEEE, 792–797. https://doi.org/10.1109/BIBM.2018.8621368

[103] Wenbin Yue, Zidong Wang, Weibo Liu, Bo Tian, Stanislao Lauria, and Xiaohui Liu. 2021. An optimally weighted user-and item-based collaborative filtering approach to predicting baseline data for Friedreich's Ataxia patients. *Neurocomputing* 419 (2021), 287–294. https://doi.org/10.1016/j.neucom.2020.08.031

[104] Xiangxiang Zeng, Yinglai Lin, Yuying He, Linyuan Lü, Xiaoping Min, and Alfonso Rodríguez-Patón. 2019. Deep Collaborative Filtering for Prediction of Disease Genes. *IEEE/ACM transactions on computational biology and bioinformatics* 17, 5 (2019), 1639–1647. https://doi.org/10.1109/TCBB.2019.2907536

[105] Lin Zhang, Xing Chen, Na-Na Guan, Hui Liu, and Jian-Qiang Li. 2018. A Hybrid Interpolation Weighted Collaborative Filtering Method for Anti-cancer Drug Response Prediction. *Frontiers in pharmacology* 9 (2018), 1017. https://doi.org/10.3389/fphar.2018.01017

[106] Wen Zhang, Hua Zou, Longqiang Luo, Qianchao Liu, Weijian Wu, and Wenyi Xiao. 2016. Predicting potential side effects of drugs by recommender methods and ensemble learning. *Neurocomputing* 173 (2016), 979–987. https://doi.org/10.1016/j.neucom.2015.08.054

[107] Yin Zhang, Daqiang Zhang, Mohammad Mehedi Hassan, Atif Alamri, and Limei Peng. 2015. CADRE: Cloud-assisted drug recommendation service for online pharmacies. *Mobile Networks and Applications* 20, 3 (2015), 348–355. https://doi.org/10.1007/s11036-014-0537-4

[108] Zhi Zheng, Chao Wang, Tong Xu, Dazhong Shen, Penggang Qin, Baoxing Huai, Tongzhu Liu, and Enhong Chen. 2021. Drug Package Recommendation via Interaction-aware Graph Induction. In *Proceedings of the Web Conference 2021*. 1284–1295. https://doi.org/10.1145/3442381.3449962

[109] Xianyou Zhu, Xin He, Linai Kuang, Zhiping Chen, and Camara Lancine. 2021. A Novel Collaborative Filtering Model-Based Method for Identifying Essential Proteins. *Frontiers in Genetics* (2021), 1868. https://doi.org/10.3389/fgene.2021.763153