

“© 2015 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.”

# A framework of hybrid recommender system for personalized clinical prescription

Qian Zhang, Guangquan Zhang, Jie Lu, Dianshuang Wu

Decision Systems and e-Service Intelligence Laboratory,

Centre for Quantum Computation & Intelligent Systems,

Faculty of Engineering and Information Technology,

University of Technology Sydney, Australia

E-mail: qian.zhang-11@student.uts.edu.au, guangquan.zhang@uts.edu.au, jie.lu@uts.edu.au (J. Lu),

dianshuang.wu@uts.edu.au

**Abstract**—General practitioners are faced with a great challenge of clinical prescription owing to the increase of new drugs and their complex functions to different diseases. A personalized recommender system can help practitioners discover mass of medical knowledge hidden in history medical records to deal with information overload problem in prescription. To support practitioner’s decision making in prescription, this paper proposes a framework of a hybrid recommender system which integrates artificial neural network and case-based reasoning. Three issues are considered in this system framework: (1) to define a patient’s need by giving his/her symptom, (2) to mine features from free text in medical records and (3) to analyze temporal efficiency of drugs. The proposed recommender system is expected to help general practitioners to improve their efficiency and reduce risks of making errors in daily clinical consultation with patients.

*Keywords*—recommender systems, case-based reasoning, healthcare personalization, text mining

## I. INTRODUCTION

Due to the large number of patients general practitioners (GPs) are meeting every day, efficiency and accuracy of decision making are quite crucial in clinical medicine. Prescription is an important element in every consultation with patient. Because of the fast variation of diseases, numerous new drugs are developed to deal with different situations occurred to patients. It is sometimes difficult to keep up to date with plenty and complex drug information even for an experienced GP, let alone the ones in internship. Moreover, contraindications of many drugs may bring chronic harm to unsuitable patients or what’s worse may cause lethal effect. Therefore, GPs are facing a big challenge in the information explosion era with many new drugs flooded into market every year. To reduce mistakes and cut off the time consumed in encounter with patients, there is a requirement of medical E-services to help GPs search and select drugs when they are making decisions on prescription.

Some surveys show that GPs in clinics use an Electronic Medical Record System (EMRS) mainly for prescribing together with ordering laboratory tests and recording patient’s progress [1]. Wildly used EMRS lead to

fast accumulation of medical data where potential medical knowledge is to be extracted to guide GPs in prescription [2]. Prescription decision support to GPs is to supply them with personalized drug list based on patients’ medical symptoms, test results and other characteristics collected by EMRS. Similarly, to find appropriate item for specific customers in E-commerce is a problem solved by personalized recommendation [3].

Recommender systems (RS) aim to predict users’ potential interests on items they haven’t bought according to users’ preferences (explicitly as ratings on historical item and implicitly as social relations) thus making a personalized product list for customers to choose [4]. RS is broadly adopted to different areas both in academia and in industry in the past two decades [5]. Some RSs have been developed to support medical prescription. Collaborative filtering (CF) is used together with a trust-based relationship among clinicians using drug frequency in prescribing to represent its rating of a clinician [6]. Moreover, since clinicians often solve a new problem based on previous medical experience (solved case) with some adjustments implemented [7], application of knowledge based technique in RS or specifically, case-based reasoning (CBR) method is applied in this problem. CBR is refined by association rule mining and Bayesian reasoning to integrate general knowledge from relationship between clinical manifestation and prescribed medicine into cases [8, 9]. However, due to the patient privacy in medicine, implementation of recommender system techniques in drug recommendation has gained less attention [10].

Although these systems mentioned above filtered some unrelated drugs and achieved good performance on test, some questions are remained to be solved. First, unlike the items recommendation which depend on user’s preference in E-commerce, drugs prescribed by GP are corresponding to patient’s symptoms. How to define a patient’s need given a patient’s symptom is a primary problem in prescription recommendation. Existing method choose CBR to evade this issue or simplify it to a single one-to-one symptom-drug relationship. However, the relationship between symptom and drug is complex to deal with but seriously influence the final recommendation accuracy. Second, free text message in EMRS is not taken into consideration. Free text written by GP is the most flexible part containing latent characteristics of different patients, where distinguishable feature extracted

are valuable for personalization. Third, since numerous new drugs are developed each year to improve therapy performance, analyzing temporal efficiency of drugs will help GP grasp the trend in the leading edge. The three aspects mentioned above will affect the accuracy and practical utility of the prescription recommender system, which are not well developed in previous research.

To solve the identified issues, this study proposes a new framework of hybrid recommender system which uses artificial neural network and case-based reasoning to personalize prescriptions in a comprehensive way. This framework relates patient's need to different drug clusters, includes meticulous patient features in free text and mines up-to-date drug trends to support GP in drug decision making. The rest of the paper is organized as follows. It begins by review of related work in Section 2. Section 3 illustrates the architecture of prescription recommender system. Finally, conclusion and future studies are discussed in Section 4.

## II. RELATED WORKS

### A. Recommender system

Recommender system (RS) is first applied in E-commerce to solve information overload problem brought by Web 2.0 and quickly expanded to personalization of e-government, e-learning, e-tourism etc. [4]. Classification of RSs is different in literatures, commonly it is categorized in 4 types according to methods used in the system: content-based, collaborative filtering (CF)-based, knowledge based and hybrid based [11, 12]:

**Content-based RS:** Recommendations are based on the content of the item user preferred before. Usually heuristic method is applied to match item attribute based on keyword to user's profile. This method is to match user's preference to an item based on user's profile and item's attributes. Content-based technique is often used in webpage or book recommendation where semantic meaning of the content is quite important.

**CF-based RS:** Recommendations to a user are based on other users who have similar tastes. This method depends on user-item rating matrix and is the most famous method used in E-commerce. Although it suffers from limitations like sparsity of the rating matrix and scalability with large amount of data, it holds an irreplaceable position in the development of RS.

**Knowledge-based RS:** Recommendations were based on the relationship of user's need and item's function. It depends on the existed knowledge or rules where knowledge database management is needed.

**Hybrid-based RS:** Hybrid method combines two or more techniques to overcome the limitation of any individual one.

### B. Recommender systems in health care

In recent years, different data mining techniques of recommender system are applied in medical data to acquire medical knowledge to support diagnosis, treatment or

prognosis [13]. CF is adapted from e-commerce to predict risks in heart attack and achieves better performance compared with traditional classification methods like support vector machine or linear regression [14]. A system named iCARE based on CF and ensemble learning makes predictions of patient's disease risks according to their disease history [15]. An incremental CF algorithm, W-InCF, using Mahalanobis distance and fuzzy membership predicts the risk of pregnant women in delivery [16].

Besides, many systems are developed based on CBR. A recommendation system to predict treatment to different diabetes cases according to different features selected using rough set feature reduction [17]. A CBR method is also applied to determine radial dose in cancer treatment and weights of features are optimized by Bee Colony Optimization [18].

These works take lab test results in numerical form as patient feature to retrieve most similar case in predicting the unknown case. The prediction made in these cases are quite simple either a binary yes-or-no decision or a choice from less than 10 candidates. However, medical cases and treatment plans in clinical context can be quite complex and no simple and direct rules can be traced. Decision making in prescription is to be made among hundreds or thousands of drugs. Combinations and relationships between different drugs will increase the difficulty of the choice. Till now, no adequate system to support decision making of GP's prescription has been provided.

## III. A HYBRID RECOMMENDER SYSTEM FRAMEWORK

This section describes the component of the framework and how it is trained with data in EMRS and applied in GP's prescription.

TABLE I. PATIENT FEATURE AND LABEL DEMO

Category	Terms	Feature/Label
Demographic characteristics	Visit Date	patient demographic feature
	Age	
	Gender	
	BMI (Body Mass Index)	
	Allergy	
	Smoking	
Lab test result	Surgery history	patient symptom feature
	Blood glucose	
	Urine glucose	
	Blood hemoglobin	
	Alkaline phosphate	
Free text	AST	To be mined
Prescription	Drug1	drug label
	Drug2	
	Drug3	

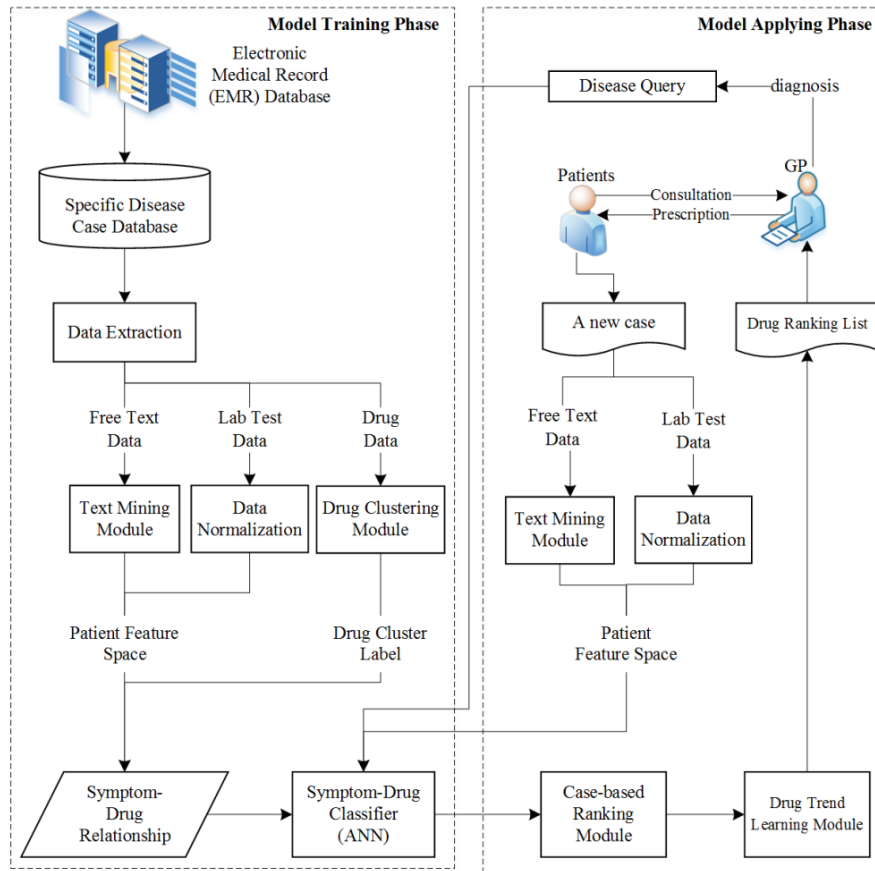


Figure 1. Framework of hybrid recommender system for personalized clinical prescription

### A. Framework overview

This part proposes a hybrid recommender system framework to recommend drugs to GP during the consultation of a patient.

The hybrid recommender system is relied on EMRS that GP used every day in clinics. As time goes by, electronic medical records (EMR) of various patients are collected and stored in EMR database. A patient's medical record is collected during treatment in clinic, hospital or other health organization. Basically, it covers demographic characteristics (like age, gender, weight, height, allergy, medical history), free text written by GP as notes during the consultation with patients, results of lab tests ordered by GP, and prescription or other treatment the patient received as demonstrated in Table 1.

Data discussed above contain various diseases from common disease like cold to chronic and complex disease like cancer. Since pathophysiological and pharmacological mechanisms are incomparable between different diseases, records or cases should not be mixed together. To mine medical knowledge that has practical value to offer a guide to a new case, history records should be about the same disease. Therefore, based on GP's diagnosis, medical records in EMR database should be separated to specific disease databases like diabetes database, heart attack database, skin cancer database, etc. The hybrid system is working on a

specific disease database. Actually, if a general system is to be set up for GP to deal with all kinds of diseases comprehensively, multiple models should be trained for each kind of disease.

As shown in Fig.1, the framework consists two parts: (1) model training phase which is done during the system designing and building process; (2) model applying phase which is an interactive part that directly connected with a GP interface.

In the model training phase, the system starts getting data from specific disease database to build a model to support GPs in prescription of this disease. These data are labeled training data containing symptom-drug relationship. After data pre-processing part (which will be introduced in next Section), patient feature space and corresponding drug cluster label are obtained to form the symptom-drug relationship. Based on this relationship, a symptom classifier is trained to determine which kind of drug to recommend to GPs. Before the system is put into practical use, models corresponding to each specific disease database are trained during the system development.

As to the model applying phase, the following is a scenario illustrating how the hybrid recommender system is applied in routine consultation of a GP. When a patient with morbidity comes for consultation, a GP may make inquiries and order lab tests for the patient. Information about this

patient is entered into the system as a new case during the process. The system will process the new data and extract patient features. Then, a GP makes a diagnosis based on the patient's problem. The diagnosis is matched to a specific disease category in the system, to determine which symptom-drug classifier to use. Patient features in the new case are put into the classifier to predict which drug cluster/clusters to choose for this patient. Drugs in each cluster will be ranked by the ranking module to form the final recommendation list. At last, a GP should dispense advice, prescription or other kind of treatment to patient to restore his/her health. This is where the recommender system supports GPs with drug recommendation list that is personalized for the specific patient.

### B. Definitions

In this study, each patient-GP encounter is a medical case stored in the EMR database. In a specific disease database, say diabetes as an example, set  $C = \{c_1, c_2, \dots, c_S\}$  represents all the cases related to diabetes, which is the training set.

In a case  $c_i$ , a GP gives a prescription to a patient based on his/her symptoms. A patient's information includes demographic characteristics and lab test results as structured data and free text description as unstructured data. Patient's demographic characteristics are important and valuable in decision of prescription, but it is not included in the symptom-drug relationship. This study asserts that symptoms are combination of features mined from unstructured free text written by GP and structured lab results. Patient feature space is a numeric representation of symptom description in patient's medical record. In a case  $c_i$ , a patient feature space is the union of text features and lab features.

On the other hand, a patient in case  $c_i$  took several drugs. Drugs are represented as  $D = \{d_1, d_2, \dots, d_K\}$  for a specific disease. Usually, a patient may take several drugs in case  $c_i$ . Unlike traditional two-class and multi-class classification problem, each instance in training set is associated with a set of labels. Given a training set, to predict what kinds of drugs to take according to the patient symptom is a multi-label classification problem. The problem is defined as follows:

Let  $\chi \subseteq \mathbb{R}^F$  denote the domain of patient features and let  $\mathcal{L} = \{1, 2, \dots, K\}$  be the finite set of drug labels. A case is denoted by  $c_i = (X_i, Y_i)$ , where  $X_i \in \chi$  and  $Y_i \subseteq \mathcal{L}$ . Given training set  $C = \{(X_i, Y_i)\}_{i=1}^S$ , the classification task is to train a multi-label classifier  $f: X \rightarrow 2^{|\mathcal{L}|}$ .

However, there are numerous drugs to deal with so  $|\mathcal{L}| = K$  may be quite large ( $K$  drugs will yield  $2^K$  classes). With the huge number of possible classes to deal with, the training set has to be large and very well distributed to avoid data sparseness. Since drugs are in different categories against different symptoms, they can be clustered into  $N$  drug clusters. Let  $\mathcal{L} = \{1, 2, \dots, N\}$  be finite set of drug labels, cardinals of  $\mathcal{L}$  will be reduced from  $K$  to  $N$ .

Therefore, the number of possible classes can be reduced significantly.

### C. Model training phase

Three key issues are disposed in the model training phase: to form a patient feature space, to form a drug label space and to train a classifier that is suitable to deal with the relationship between them.

#### 1) Patient Feature Space

To form a patient feature space  $X_i$  is to explore features from unstructured free text to form text feature space  $A_i$  and to obtain features from structured lab test data to form lab feature space  $B_i$ .

##### a) Text mining module

Free text is not a good format for the computer to deal with. Clinical text analysis is a branch in natural language processing [19] but seldom used in prescription recommendation or healthcare personalization. The text mining module of this system is to deal with free text in medical record as shown in Fig.2. In this module, medical record written by GP is processed in two steps:

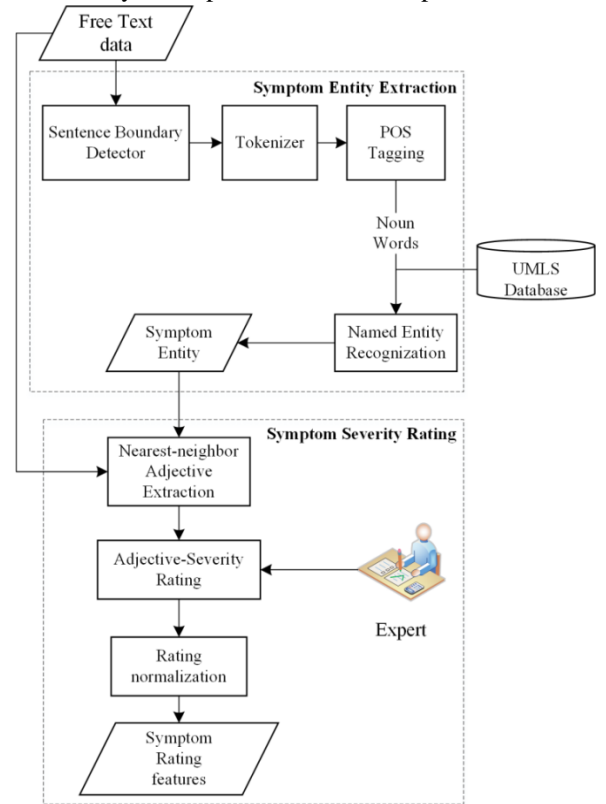


Figure 2. Architecture of text mining module

#### Step1: Symptom Entity Extraction

This step is to extract entities about symptom from free text as utilized in system like cTAKE[19]. These extracted entities are to be the feature of the patients of diabetes.

Text is firstly separated as sentences by period, question mark or exclamation mark. Then a tokenizer splits

the sentence into single words on space and punctuation. Each word is identified as nouns, verbs, adjectives, adverbs, etc. by POS (Part of Speech) tagging. The nouns words are extracted and mapped by a terminology dictionary in terms. This dictionary is a subset of UMLS (Unified Medical Language System) contains concepts of SNOMED CT (Systematized Nomenclature of Medicine, Clinical Terms). Each term in the dictionary belongs to diseases with symptoms, procedures and drugs. Finally, nouns describing symptoms are selected.

### Step2: Symptom Severity Rating

This step is to map the adjectives modifying symptom entities to corresponding ratings by its severity. Symptoms from step1 will be represented by ratings after this step.

In a sentence, the nearest neighbor adjective is to be the most probable one to modify the entity. Based on symptom entity acquired in step1, its nearest adjective can be extracted. Thus, the relationship of symptom and severity (i.e., the relationship of symptom entity and adjective) is acquired. To determine the severity of the symptom, a thesaurus of all the adjectives is set up and a severity-rating estimator under the supervision of domain expert like clinicians is used to rate these adjectives. After that, the adjective describing the severity of the symptom is corresponding to a rating number (range of 1-5).

Symptoms extracted from free text in different cases may not be the same. Text feature space  $A'_i = (a'_{i1}, a'_{i2}, \dots, a'_{ip})$  is the sum of all symptoms extracted in each case ( $A'_i$  represents  $A_i$  before normalization). For those cases who do not contain a specific symptom in its text, the rating is filled with 0. Thus, the codomain of symptom features acquired above is  $\{0,1,2,3,4,5\}$ . In order to standardize the range of features to  $(0,1)$ , a feature  $a'_{ij} \in A'_i$  is normalized by min-max scaling as follows:

$$a_{ij} = \frac{a'_{ij} - \min(A')}{\max(A') - \min(A')} \quad (1)$$

where  $a'_{ij}$  is the  $j$ th rating corresponding to the  $j$ th symptom in  $A'_i$ .  $\min(A')$  is the minimum rating and  $\max(A')$  is the maximum rating in  $A'$ ,  $A' = \{A'_i\}_{i=1}^N$ .

Finally, text mined symptom feature space  $A_i = (a_{i1}, a_{i2}, a_{i3}, \dots, a_{ip})$  is obtained. Its codomain is  $[0,1]$ .

#### b) Lab test normalization

This module is to normalize the lab results in different scope. The results of lab test are usually numeric values and evaluated by fixed clinical criteria. Generally, the criteria are a range to designate the normal state of a lab indicator. The result of test can be divided to a term set {normal, abnormal}. For instance, the normal blood glucose level is 3.9-6.1mmol/L. A test result below 3.9mmol/L or above 6.1mmol/L is defined as abnormal, and a test result falls right in the middle is defined as normal. This module is to map the result as a specific number to the term set above.

The term set can be simply represented as  $\{0,1\}$  corresponding to {normal, abnormal}. Lab test features are formed to  $B_i = (b_{i1}, b_{i2}, \dots, b_{iq})$ , and their codomain is  $\{0, 1\}$ .

Accordingly, a patient feature space  $X_i = (x_{i1}, x_{i2}, x_{i3}, \dots, x_{im})$  is a union of  $A_i$  and  $B_i$ ,  $m = p + q$ .

#### 2) Drug cluster label

To form a drug label space  $Y_i$  in case  $c_i$  as defined above, drugs are classified into different categories. Primarily, drugs are described by some attributes. Since drugs are in different categories against different symptoms, the patient feature space attributes are used to characterize a drug  $d$ . Patient feature space contains many symptom features both from free text and lab test. In one specific case  $c_i$ , many features must be 0 in its feature space  $X_i$ . The features in  $X_i$  that are not zero is the symptom related to the drugs prescribed in this case  $c_i$ .

Like term frequency/inverse document frequency (TF-IDF) used in content-based recommender system[12], a drug's attribute is measured by the frequency of concurrence of a drug  $d_j$  and symptom feature  $x_j$  in all cases. Measurement to symptom feature  $x_j$  is defined as follows:

$$freq_{d_j x_j} = \frac{n_{d_j x_j}}{n_{d_j}} \quad (2)$$

where  $freq_{d_j x_j}$  is the value corresponding to symptom feature  $x_j$ ,  $n_{d_j x_j}$  is the number of cases containing  $d_j$  and  $x_j$  is a non-zero value,  $n_{d_j}$  is the number of cases containing  $d_j$ . A drug  $d_j$  can be defined as follows:  $d_j = (freq_{d_j x_1}, freq_{d_j x_2}, \dots, freq_{d_j x_m})$ . As drugs are described by attributes mentioned above, a k-means cluster method is applied to cluster these drugs to  $n$  clusters  $\mathcal{L} = \{1, 2, \dots, n\}$ .

In a case  $c_i$ , drugs are prescribed by GP as shown in table1. For instance, drugs prescribed in  $c_i$  are  $d_1, d_3$  and  $d_5$ .  $d_1, d_3$  and  $d_5$  are in clusters 2, 3 and 6, then label space  $Y_i = (2, 3, 6)$ .

#### 3) Multi-label classification

Unlike traditional classification problem, each instance in training set  $C = \{(X_i, Y_i)\}_{i=1}^N$  is associated with a set of labels. The target is to output a label set with unknown size. So the classification task based on training set  $C$  is to train a multi-label classifier  $f: X \rightarrow 2^{|\mathcal{L}|}$ . There are many methods provided to solve this multi-label classification problem[19].

Here, artificial neural network (ANN), a classical and well-accepted machine learning method is used as a classifier. The weights in ANN are obtained during the training process. This model will be used in the next model applying phase to first decide which drug cluster should be chosen according to patient's symptoms.

### D. Model applying phase

#### 1) Case-based ranking module

When a new patient comes to consultation, it is a new case  $c_{new}$  to this system. A patient feature space is obtained as the same procedure in the model training phase, according to the free text data and lab test data typed in EMRS by GP. The patient feature space is put into the symptom-drug classifier from which categories of drugs the patient should take are exported. A list of drugs is then to be chosen in each drug cluster. Case-based ranking module is to give a ranking list in each drug cluster for the patient.

In one drug category, each drug is occurred in one or several old cases. These cases are retrieved to obtain attributes of each drug. In each case, the drug is related to a patient feature space. The similarity of an old case  $c_{old}$  and the new case  $c_{new}$  is a measurement of suitability of this drug to the new case. Similarity can be calculated by cosine similarity:

$$sim(c_{new}, c_{old}) = \frac{\sum_{i=1}^m X_{i,new} \times X_{i,old}}{\sqrt{\sum_{i=1}^m (X_{i,new})^2} \times \sqrt{\sum_{i=1}^m (X_{i,old})^2}} \quad (3)$$

where  $X_{i,new}$  and  $X_{i,old}$  are feature vectors of the new case and old case, respectively.

For a drug occurred in several old cases, a similarity is the average of these similarities to the new case  $c_{new}$ . Drugs in one category are ranked by their similarities to the new case and a recommendation list is formed. In the end, several recommendation lists are formed for the new patient to treat his/her different symptoms.

## 2) Drug trend learning module

A drug can be eliminated once patients develop resistance to a certain drug. The frequency of this drug prescribed by GP will be reduced since its clinical effect is not as good as new drugs. The curative effect of a drug is changing. This is time drifting data widely known as concept drift, which is also considered in e-commerce recommender system [20].

In medical environment, each drug is analyzed during a time period. Instead of using average value of similarities to the new case  $c_{new}$  for a drug occurred in several old cases, it is more reasonable to add a time-aware weight to adjust the final ranking.

$$r_{d_i} = \sum_{c_{old}} sim(c_{new}, c_i) \times w_i(t) \quad (4)$$

where  $C_{old}$  is the set of old cases a drug  $d_i$  occurred in, and  $w_i(t)$  is the time-aware weight learned from time-series medical records.

## IV. CONCLUSION AND FUTURE WORK

The fast development of human society will make people pay more attention to their health. Clinical data describing patient's symptom and treatment are accumulated every year. Exploring knowledge from them then improving human healthcare will keep being emphasized in the future. Thus, designing and implementing

healthcare recommender system will benefit both GPs and patients.

This paper proposes a hybrid recommender system to support GP in personalized clinical prescription by integrating ANN and CBR. In this model, information mined from text expands patient feature space which, in previous research, is usually restricted to lab test result or demographic characters. By clustering the drugs based on their remedy functions to symptoms, multiple choices of drugs can be narrowed to several clusters. An artificial neural network is used to deal with the complex relationship between symptom and drug that cannot be extracted simply as rules or linear relations. What's more, effectiveness of drug in a time period is analyzed to amend the recommendation list provided to GP. To sum up, this model overcomes the difficulty of mining the relationship between symptom and drug prescribed, fills the black of taking drug temporal effectiveness into consideration and better performed as more features are added to patient feature space.

Further study will focus on implementation of this framework and evaluation of this recommender system based on electronic medical record from hospitals or other health organizations. Because the system is related to healthcare, which requires domain knowledge of expert in this area, this system should be evaluated by experienced clinician before it comes to practical use.

## REFERENCES

- [1] D. K. McInnes, D. C. Saltman, and M. R. Kidd, "General practitioners' use of computers for prescribing and electronic health records: results from a national survey," *Medical Journal of Australia*, vol. 185, p. 88, 2006.
- [2] P. B. Jensen, L. J. Jensen, and S. Brunak, "Mining electronic health records: towards better research applications and clinical care," *Nature Reviews Genetics*, vol. 13, pp. 395-405, 2012.
- [3] J. Lu, "Personalized e-learning material recommender system," *International conference on information technology for application*, 2004, pp. 374-379.
- [4] J. Lu, D. Wu, M. Mao, W. Wang, and G. Zhang, "Recommender system application developments: a survey," *Decision Support Systems*, vol. 74, pp. 12-32, 2015.
- [5] Z. Huang, X. Lu, H. Duan, and C. Zhao, "Collaboration-based medical knowledge recommendation," *Artificial Intelligence in Medicine*, vol. 55, pp. 13-24, 2012.
- [6] S. Begum, M. U. Ahmed, P. Funk, N. Xiong, and M. Folke, "Case-based reasoning systems in the health sciences: a survey of recent trends and developments," *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews* vol. 41, ed, 2011, pp. 421-434.
- [7] S. Ting, W. M. Wang, S. K. Kwok, A. H. Tsang, and W. Lee, "RACER: Rule-Associated CasE-based Reasoning for supporting General Practitioners in prescription making," *Expert Systems with Applications*, vol. 37, pp. 8079-8089, 2010.
- [8] S. Ting, S. K. Kwok, A. H. Tsang, and W. Lee, "A hybrid knowledge-based approach to supporting the medical prescription for general practitioners: Real case in a Hong Kong medical center," *Knowledge-Based Systems*, vol. 24, pp. 444-456, 2011.

- [9] M. B. Sesen, M. D. Peake, R. Banares-Alcantara, D. Tse, T. Kadir, R. Stanley, et al., "Lung Cancer Assistant: a hybrid clinical decision support application for lung cancer care," *Journal of The Royal Society Interface*, vol. 11, p. 20140534, 2014.
- [10] F. Ricci, L. Rokach, and B. Shapira, "Introduction to recommender systems handbook," in *Recommender systems handbook*, F. Ricci, L. Rokach, B. Shapira, and P. B. Kantor, Eds. Berlin: Springer, 2011, pp. 1-35.
- [11] G. Adomavicius and A. Tuzhilin, "Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions," *IEEE Transactions on Knowledge and Data Engineering*, vol. 17, pp. 734-749, 2005.
- [12] N. Esfandiari, M. R. Babavalian, A.-M. E. Moghadam, and V. K. Tabar, "Knowledge discovery in medicine: current issue and future trend," *Expert Systems with Applications*, vol. 41, pp. 4434-4463, 2014.
- [13] S. Hassan and Z. Syed, "From netflix to heart attacks: collaborative filtering in medical datasets," in *Proceedings of the 1st ACM International Health Informatics Symposium*, 2010, pp. 128-134.
- [14] D. A. Davis, N. V. Chawla, N. A. Christakis, and A.-L. Barabási, "Time to CARE: a collaborative engine for practical disease prediction," *Data Mining and Knowledge Discovery*, vol. 20, pp. 388-415, 2010.
- [15] M. Komkhao, J. Lu, and L. Zhang, "Determining pattern similarity in a medical recommender system," in *Data and Knowledge Engineering*, Y. Xiang, M. Pathan, X. Tao, and H. Wang, Eds. Berlin: Springer, 2012, pp. 103-114.
- [16] X. Lu, Z. Huang, and H. Duan, "Supporting adaptive clinical treatment processes through recommendations," *Computer Methods and Programs in Biomedicine*, vol. 107, pp. 413-424, 2012.
- [17] D. Teodorović, M. Šelmić, and L. Mijatović-Teodorović, "Combining case-based reasoning with Bee Colony Optimization for dose planning in well differentiated thyroid cancer treatment," *Expert Systems with Applications*, vol. 40, pp. 2147-2155, 2013.
- [18] G. K. Savova, J. J. Masanz, P. V. Ogren, J. Zheng, S. Sohn, K. C. Kipper-Schuler, et al., "Mayo clinical Text Analysis and Knowledge Extraction System (cTAKES): architecture, component evaluation and applications," *Journal of the American Medical Informatics Association*, vol. 17, pp. 507-513, 2010.
- [19] M.-L. Zhang and Z.-H. Zhou, "Multilabel neural networks with applications to functional genomics and text categorization," *IEEE Transactions on Knowledge and Data Engineering*, vol. 18, pp. 1338-1351, 2006.
- [20] Y. Koren, "Collaborative filtering with temporal dynamics," *Communications of the ACM*, vol. 53, pp. 89-97, 2010.