An Intelligent Medicine Recommender System Framework

Youjun Bao
College of computer science, Zhejiang University
Hangzhou 310027, China
baoyoujun@zju.edu.cn

Xiaohong Jiang
College of computer science, Zhejiang University
Hangzhou 310027, China
jiangxh@zju.edu.cn

***Abstract*—More and more people are caring about the health and medical diagnosis problems. However, according to the administration's report, more than 200 thousand people in China, even 100 thousand in USA, die each year due to medication errors. More than 42% medication errors are caused by doctors because experts write the prescription according to their experiences which are quite limited. Technologies as data mining and recommender technologies provide possibilities to explore potential knowledge from diagnosis history records and help doctors to prescribe medication correctly to decrease medication error effectively. In this paper, we design and implement a universal medicine recommender system framework that applies data mining technologies to the recommendation system. The medicine recommender system consists of database system module, data preparation module, recommendation model module, model evaluation, and data visualization module. We investigate the medicine recommendation algorithms of the SVM (Support Vector Machine), BP neural network algorithm and ID3 decision tree algorithm based on the diagnosis data. Experiments are done to tune the parameters for each algorithm to get better performance. Finally, in the given open dataset, SVM recommendation model is selected for the medicine recommendation module to obtain a good trade-off among model accuracy, model efficiency, and model scalability. We also propose a mistake-check mechanism to ensure the diagnosis accuracy and service quality. Experimental results show our system can give medication recommendation with an excellent efficiency, accuracy and scalability.**

Keywords—Medicine; Recommender system; Data mining; SVM

I. INTRODUCTION

Health information is one of the most widely concerned topics on the Web. A survey in 2013 by the Pew Internet and American Life Project found that 59% of adults have looked online for health topics, and with 35% of respondents focusing on diagnosing a medical condition online [1]. Behind the data, we find that more and more people are caring about the health and medical diagnosis problem. However, there are still many people losing their lives due to medication errors. According to the administration's report, more than 200 thousand people in China, even 100 thousand in USA, die each year due to medication errors [2]. More than 42% medication errors are caused by doctors because experts write the prescription according to their experiences which are quite limited [3]. There are some facts that may lead to these issues: (i) many hospitals lack either doctors or medical experts for critical illness, (ii) expert diagnosis is mainly depended on the expert's experience, especially for those inexperienced novices, which are hard to avoid mistakes. Meanwhile, most diagnosis case data in hospitals is still kept untouched and has not been used for mining, so that the value behind the data cannot be explored.

Hospital Information System (HIS) generates massive data, how to discover potential and useful knowledge from the diagnosis case data is a big challenge. Data mining and recommender technologies represent a promising direction to solve these challenging problems. Since the mid-1990s, various recommender system techniques have been proposed, and many sorts of recommender system software have been developed recently for a variety of applications. Traditional recommender technologies conclude collaborative filtering (CF), content-based (CB), knowledge-based (KB) techniques and hybrid recommendation technologies, which have some limitations. CB has overspecialized recommendations, and CF has sparseness, scalability and cold-start problems. To solve these problems, we apply data mining technologies to recommendation systems and design our own medicine recommender system framework.

In this paper, we introduce a universal medicine recommender system framework that is designed and implemented to apply data mining technologies to recommendation systems that use the potential knowledge hiding in medical records to decrease medical errors. The medicine recommender system consists of database system module, data preparation module, recommendation model module, model evaluation model and data visualization module. Data visualization presents some valuable knowledge behind the diagnosis case data. We investigate the medicine recommendation algorithms of the SVM (Support Vector Machine), BP neural network algorithm and ID3 decision tree algorithm based on the diagnosis data. Finally, SVM is selected for the medicine recommendation model for its high accuracy, good efficiency and scalability. In consideration of the safety of the patient, we proposed a mistaken-check mechanism that ensures the safety and the quality of service. Experimental results show our system can give medication recommendation with an excellent efficiency, accuracy and scalability.

The structure of this paper is as follows. Section 2 gives the related work about recommender technologies and intelligent medical diagnosis. Section 3 gives a description of our universal medical recommender system framework. Section 4 focuses on preparation for model building. Section 5 presents the building and evaluation of the proposed medicine recommendation model. Section 6 gives the conclusion and future work.

978-1-4673-8644-9/16/$31.00 ©2016 IEEE

II. Related Work

Recommender systems aim to provide users with personalized products and service to deal with the increasing online information overload problem. Various recommender system techniques have been proposed since the mid-1990s, and many sorts of recommender system software have been developed recently for a variety of applications. Most of the recommender technologies are applied to the e-government area[4], e-business area[5], e-commerce/e-shopping area[6], e-learning area[7], e-tourism area[8] and so on. However, medicine area includes rare recommender technologies, and this paper focuses on the design of the medicine recommender system and mining knowledge from medical case data.

Commonly used recommendation techniques include collaborative filtering (CF)[9], content-based (CB)[10], knowledge-based (KB)[11] techniques and hybrid recommendation technologies[12]. Each recommendation technology has advantages and limitations: CB mainly generates recommendations by using traditional retrieval methods and machine learning methods, but CB has overspecialized recommendations; Collaborative filtering (CF)-based recommendation techniques help people to make choices based on the opinions of other people who share similar interests, while CF has sparseness, scalability and cold-start problems. Knowledge-based (KB) recommendation offers items to users based on knowledge about the users, items and/or their relationships. Usually, KB recommendations retain a functional knowledge base that describes how a particular item meets a specific user's need. To achieve higher performance and overcome the drawbacks of traditional recommendation techniques, a hybrid recommendation technique that combines the best features of two or more recommendation techniques into one hybrid technique has been proposed [12]. When it comes to a new application area, a new recommendation framework is necessary to solve these problem.

Intelligent medical diagnosis has get more and more concern. Some selected techniques for data mining in medicine are discussed in [13]. Data mining technology has been used to predict heart disease, and to diagnosis thyroid diseases [14][15].The workshop on health search and discovery [16] discusses several challenges and important topics about search and discovery in the medical domain, indicating that information retrieval, personalization, expertise modeling, data mining and privacy preservation are critical to enable advances in health and discovery. These approaches do solve some problems in medical diagnosis and apply new data-driven technologies to the medical field. However, there does not exist a universal model which can work well for all the data and conditions. When faced with different data and application scenarios, it is necessary to build a different model and conduct data analysis. In this paper, we introduce a universal medicine recommender system framework that is designed and implemented to apply data mining technologies to the recommender system that use the potential knowledge hiding in medical records to decrease medical errors.

III. Medicine Recommender System Framework

In this section, the characteristics of the recommender system will be discussed, and the detail of our medicine recommender system framework will be presented.

Recommender system has become a valuable research field as the development of artificial intelligent technologies. Unlike most current recommender systems focusing on e-business, book and movie recommendation, our system aims at providing a virtual experienced doctor for inexperienced novices and patients in using right drugs. Since high accuracy and efficiency is critical for such an online medicine recommender system, hence we evaluate some data mining approaches to obtain a good trade-off among the accuracy, efficiency and scalability.

Our medicine recommender system framework mainly consists of five modules, as shown in figure 1, which are database system module, data preparation module, recommendation model module, model evaluation model, and data visualization module.

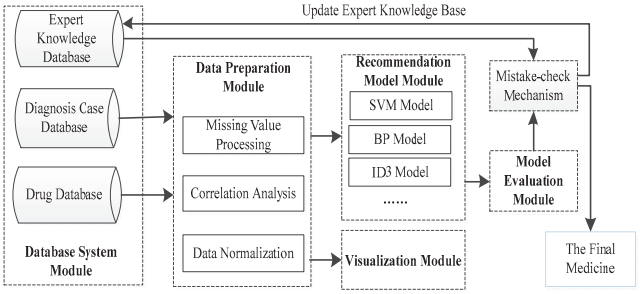

Fig. 1. The framework of the recommender system

Database system module provides data connections of modules, which contains diagnosis case database, drug database and expert knowledge database. Diagnosis database stores the diagnosis case data of the hospitals, and provides access to other modules. Drug database collects all of the drugs and the index is built for faster queries. Expert knowledge database stores the knowledge of experts, which is obtained by summarizing experts' experiences.

Data preparation module acts as a data-cleaner in our system. The real-world data is raw data which may be incomplete, noisy and dirty. Hence, data preparation is built to generate clean data. Data preparation module consists of missing value processing, correlation analysis and data normalization.

Recommendation model module is the kernel module of our system, which is built among SVM, BP neural network algorithm, ID3 decision tree algorithm and so on. In this paper, we mainly build our recommendation model among the three algorithms. We can also add new data mining algorithms to the module according to our demand. Visualization module mainly provides the visualization technology to present some valuable knowledge behind the diagnosis case data.

Model evaluation module evaluates the different recommendation models in a concrete dataset. When faced with different application scenarios, it is necessary to evaluate the models to obtain a good trade-off among model accuracy, model efficiency, and model scalability.

This universal medicine recommender system applies the data mining technologies to the medical diagnosis, which makes full use of diagnosis case data and experts' experience. Some recommendation models are built based on the diagnosis case data, and get the drug for the patient combined with the experts' experiences. What is more, a mistaken-check mechanism is proposed to make sure the diagnosis accuracy and service quality. That is, if the medicine recommended by SVM model is not matching to the medical experts' experience, the medical expert will diagnosis the patient and give more appropriate medicines, and update the expert experience database. Hence this strategy ensures the safety of the patient and improves the quality of the service.

IV. Preparation For Model Building

A. Business and data understanding

In business understanding step, we focus on what we use the diagnosis data to do or what problem we want to use data mining technologies to solve. This step is crucial to a successful data mining outcome. Given the diagnosis case data, we aim to build the recommendation model and recommend the appropriate medicine for inexperienced doctor, intern / medical student or the patient.

Data mining is a technology to find knowledge behind the data, and we should get a good understanding of our data. The data is from a scientific research website named Data Tang, which contains 1200 records [17]. Table 1 gives a presentation of the data attributes. As we can see, we use drugA,B,C,X,Y to denote the name of drug. Numeric attribute indicates that the attribute value is continuous, while the discrete attribute indicates the attribute value is discrete, and the value of the binary attribute takes only two values. The data contains the patient's normal diagnosis data, containing age, Sex, blood pressure, cholesterol level, K level, Na level and drug. Basically, a doctor can treat the patient and provide the medicine on the basis of the patient's diagnosis data.

Table 1. the attributes of the diagnosis case data

Attribute Name	Attribute Type	Attribute Value
Age	Numeric attribute	15~74
Sex	Binary attribute	F,M
BP(Blood pressure)	Discrete attribute	HIGH,NORMAL, LOW
Cholesterol	Discrete attribute	HIGH,NORMAL
Na	Numeric attribute	0.5005~0.8998
K	Numeric attribute	0.0202~0.0799
Drug	Discrete attribute	Drug A,B,C,X,Y

B. Data preparation

In practice, it has been generally found that data cleaning and preparation takes approximately 80% of the total data engineering effort [18]. So the data preparation plays an important role in the data mining system. There are three aspects that can contribute to this: (1) real-world data is raw data which may be incomplete, noisy and dirty. No good patterns will be found from this data; (2) quality data can improve the performance of mining system, since data preparation generates a smaller dataset than the original data; (3) quality data yields high-quality patterns. Data preparation in our medicine recommender system framework consists of missing value processing, correlation analysis and data nor

Real-world data is always incomplete and dirty, some data records lost some attribute value for some reasons. There are some approaches to address the issue: (1) delete the data record directly; (2) replace a missing value with the statistical value, such as mean value; (3) replace the missing value with the predicted value. In our system, we use the approach (2) to solve the missing value problem.

Data normalization is used to eliminate the influence of different scales, which transfer values with different scales to a uniform common scale. When faced the different scale data, each scale has a different meaning, hence we should make a transformation. In our recommender system framework, min-max method is applied to get the normalized data. We want to get each attribute data ranging from 0 to 1, we set y_min=0 and y_max=1. And normalized data y is denoted as below:

$$y = (y\_\max - y\_\min) * (x - x\_\min) / (x\_\max - x\_\min) + y\_\min$$

Where x, x_min and x_max are respectively denote the original data, the minimum and maximum of the original data.

Correlation analysis is an important step for data preparation, since correlations can indicate a predictive relationship that can be exploited in practice. In our system, there are both numerical and discrete attributes, and the target attribute is Drug which is discrete. So we use Chi-square test [19] and the method in [20] which is based on the comparison of variance to analysis the correlation between attributes. The correlation result in table 2 demonstrates that all attributes except Sex have correlation with Drug. So we delete the Sex attribute from the diagnosis case data. We build a set of attributes correlated with target attribute by correlation analysis.

Table 2. The result of correlation analysis

Attribute name	Target attribute	Correlation result
Age	Drug	Correlation value: 0.31, related
Na	Drug	Correlation value: 0.10, related
K	Drug	Correlation value: 0.58, strongly related
Sex	Drug	Chi-square value: 4.98, significant value: 9.49, irrelevant

BP	Drug	Chi-square value: 734.46, significant value: 15.51, strongly related
Cholesterol	Drug	Chi-square value: 140.56, significant value: 9.49, related

C. Data visualization

Data visualization is a technology which uses the visualization technology to present the data and find the pattern behind the data. Next we will show some patterns we find in the presentation of the diagnosis case data.

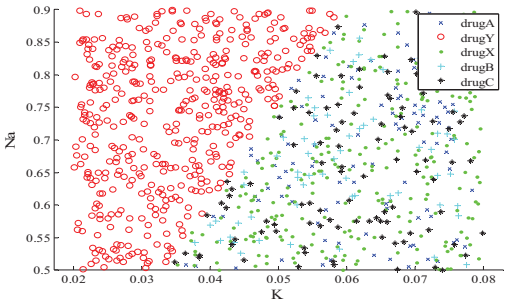

Fig. 2. K, Na and Drug visualization

Figure 2 shows the visualization of K, Na and Drug visualization. K is treated as x coordinate, and Na is treated as y coordinate, the scatter diagram is plotted shown as figure 2. The points are grouped by the different drugs, we use different colors and shapes to indicate five drugs, which are drugA, drugB, drugC, drugX and drugY. We can find that all the red points almost are in the left of the x coordinate, which indicates that the patient of low K should have drugY

When we regard the drug number as x coordinate, and regard age as y coordinate, the scatter diagram is plotted. The points are grouped by the sex, 0 indicates female, 1 indicates male. As shown in figure 3, the sex has no relation to the drug, because the female and male have a uniform distribution in every drug. It is quite clear that the drugA is appropriate for the patient less than 50 years old, and the drugB is appropriate for the patient more than 50 years old.

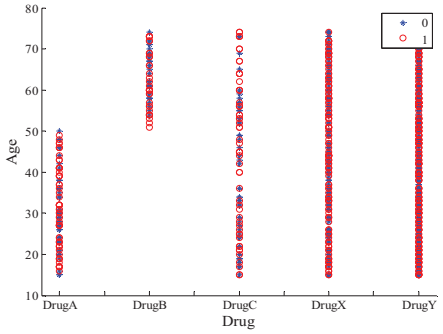

Fig. 3. Age, Drug and Sex visualization

V. Recommendation Model Building And Evaluation

This section will present some data mining algorithms, and our recommendation model will be built among the algorithms. Some experiments will be done to tune the parameters for each recommendation model to get better performance based on an open dataset. Finally we will evaluate the model to obtain a good trade-off among model accuracy, model efficiency and model scalability.

A. Support Vector Machine

In 1955, Cortes and Vapnik proposed the concept of Support Vector Network, which is a new learning machine for two-group classification problems [21]. That is the first edition of SVM. A support vector machine constructs a hyperplane or set of hyper-planes in a high-dimensional or infinite-dimensional space, which can be used for classification, regression, or other tasks. Intuitively, a good separation is achieved by the hyperplane that has the largest distance to the nearest training data point of any class (so-called functional margin), since in general the larger the margin the lower the generalization error of the classifier [22].

SVM is based on the basic idea that the input vector can be divided by a linear decision surface when mapped to a very high dimension feature space. SVM can only deal with the two classification problems initially, but CJ Lin etc. expended the model to deal with multi-class problems and regression problems [23]. SVM has a huge advantage for high dimensional data and nonlinear data. SVM overcomes the problem of overfitting and dimension disaster. The final optimized problem can be converted to a constrained QP (Quadratic Programming) problem [24]. When it comes to the large-scale data, SVM does not have a fast efficiency. There is also not a standard for the selection of kernel function. The work in [25] applies SVM model to recognize the complex human motion patterns and get a success.

SVM has a good scalability for multi dimension data and is fit for the problems of pattern recognition and regression, which is based on the statistics thoughts. Since SVM converts the task into an optimization problem, and two target attributes are contained in the typical SVM model. Given the training dataset $(x_i, y_i), i = 1,2,...l$, $x \in R^n, y \in \{\pm 1\}$, where 1 indicates the number of the training samples, n is the dimension of the dataset. And the final optimization problem is as follows [24]:

$$\max Q(a) = \sum_{j=1}^{l} \alpha_j - \frac{1}{2} \sum_{i=1}^{l} \sum_{j=1}^{l} \alpha_i \alpha_j y_i y_j (x_i \cdot x_j)$$

$$st. \sum_{j=1}^{l} \alpha_j y_j = 0, j = 1,2,...,l, \alpha_j \geq 0, i = 1,2,...,l$$

We can get the final hyperplane from the dual optimal problem. The classification function is as following:

$$f(x) = \operatorname{sgn}\{\sum_{i=1}^{l} \alpha_i^* y_i k(\vec{x} \bullet \vec{x_i}) + b^*\}$$

Where $x \in R^n$, and α_i^* , b^* is the optimized parameters. $k(\vec{x} \bullet \vec{x_i}) = \Phi(\vec{x}) \bullet \Phi(\vec{x_i})$, which is the definition of kernel function. We

can see that we only need to calculate the inner production of feature space, SVM introduces the kernel function to calculate the inner function. So we can deal with the nonlinear classification problems.

Due to these characteristics, SVM has a good scalability and performance. Next we will turn the parameters for SVM model. There are two important parameters which can extremely influence the performance of SVM, they are gamma parameter which is denoted as g in the kernel function and cost parameter denoted as c. And K-fold cross-validation method is applied to the experiments to get the appropriate g and c, in K-fold cross-validation, K-1 folds are used for training and the last fold is used for evaluation. This process is repeated K times, leaving one different fold for evaluation each time. In our experiments, we set K=5 and the number of examples is 1200. Figure 4 presents the optimization process and result of g and c. For simplicity of plotting, the logarithm of g is treated as y coordinate, and the logarithm of c is treated as x coordinate, the accuracy of SVM is regarded as z coordinate. The color of line from blue to red indicates that the accuracy ranges from low to high. And we can get that when g=0.125 and c=362.0387, SVM has a good accuracy of 95%.

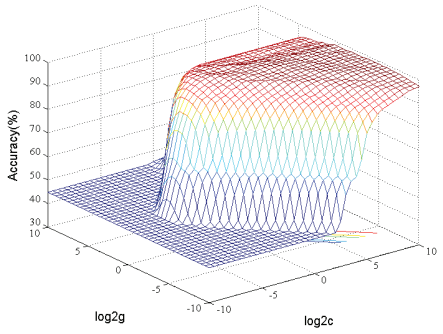

Fig. 4. the cross-validation result of *g* and *c*

B. BP neural network

Back-propagation (BP) is a well-known method to train an artificial neural network with an optimization method such as gradient descent. And the BP neural network is based on the Back-propagation which minimizes the sum of the squared errors be-tween the actual and the desired output values. The paper [26] demonstrates that artificial neural networks provide effective solutions to the problems encountered in building systems that emulate a physician's expertise and presents some neural network expert system in medical diagnosis. However, due to the method of gradient descend there exists some disadvantages for BP neural network, such as long training time and easily converging to local minimum. To address these issues, [27][28][29] propose some improved approaches.

Given the diagnosis data, we train our BP neural network model and do some experiments with different number of hidden nodes. We use Variable Learning Rate Backpropagation training function to train our model, figure 5 shows the error rate of testing and training error rate changed with the number of hidden nodes. We can see that the training error rate is decreasing with the increasing of the number of hidden nodes, while the testing error rate decreases first, then presents a trend of increasing. So we set number_hidden_nodes=8 to get an accuracy of 97%.

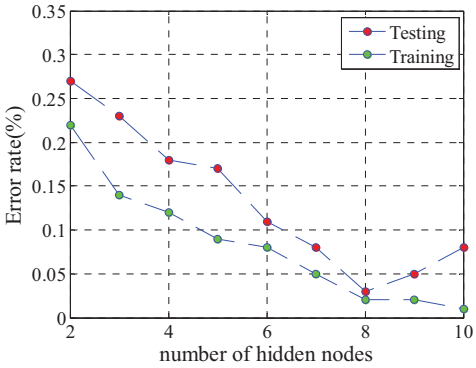

Fig. 5. The testing and training error rate evaluation

C. ID3 decision tree

ID3 is an algorithm proposed by Ross Quinlan used to create a decision tree based the data [30], and then we can make decision based the decision tree. The main challenge in the ID3 is the attribute selection, and the information gain is of great value. What's more, ID3 can help us make decision in an abstract and thoughtful way. Given these characteristics, coupled with the good performance we build the ID3 model. The final decision tree is presented in figure 6, and the numbers from one to five indicate the drugA,B,C,X,Y. We discretize all attributes and build the decision tree based on the information gain of each attribute. When coming a new patient, the ID3 recommendation model will first check the level of K, then find the leave node of the tree.

D. Model evaluation

We divide the diagnosis case data into two parts, which consist of 70% training data and 30% testing data. And ten group of experiments are conducted in each model to avoid accidental error. The model accuracy and running time is the mean of ten group of experiments. Figure 7 presents the experiment result.

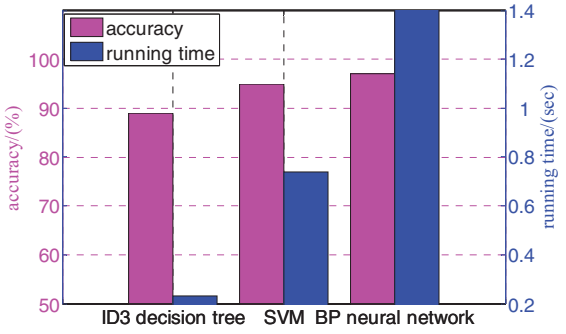

Fig. 6. Accuracy and running time of the three models

Figure 7 gives us a visualization of the accuracy and running time of models, the length of red bar indicates the accuracy of the model, and the length of blue bar shows the running time of the model. We can see from the figure 7, the accuracy gets higher and higher, and the efficiency changes apparently.

When a medicine recommender system gives a diagnosis result and recommend the medicine for a patient, the accuracy should be considered firstly. High accuracy results in high quality of service and high safety. But the efficiency is also an important factor for online system, we must get the diagnosis result as fast as possible. So we should get a good trade-off. ID3 decision tree model has the shortest time, but the accuracy is only 89%, and the ID3 has a bad scalability. When the attributes become larger, it is hard for ID3 to get the decision tree; the BP neural network has the highest accuracy, but the running time is 1.7s, it has a long training time and bad understanding of the result; SVM model has a good accuracy of 95% and 0.74 second running time. So we choose the SVM as the medicine recommendation model for its high accuracy, good efficiency and scalability in this open dataset.

VI. CONCLUSION

We devise a universal medicine recommender system framework that applies data mining technologies to the medical diagnosis, which consists of database system module, data preparation module, recommendation model module, model evaluation model, and data visualization module. And we give a concrete implementation of each module based on an open dataset. Experiments are done to evaluate the models, finally, SVM is selected for the medicine recommendation model for its high accuracy, good efficiency and scalability in this open dataset. In consideration of the safety of the patient, we also proposed a mistaken-check mechanism that ensures the safety and the quality of service. In future work, we plan to build our own recommendation model to improve the accuracy and efficiency of model further. We also plan to apply MapReduce parallel technologies to our medicine recommender system to enlarge the ability of processing big diagnosis data.

REFERENCES

1. S. Fox and M. Duggan. Health online 2013. Pew Internet and American Life Project. http://pewinternet.org/Reports/2013/Health-online.aspx, 2013.
2. Due to the medication errors there are 200 thousand people die every year in China, http://ln.qq.com/a/20140917/018143.htm, 2014.
3. Maorong Chen et al. The reason and prevention of hospital medication errors. Practical Journal of Clinical Medicine[J], 10(4): 84-87.
4. X. Guo, J. Lu, Intelligent e-government services with personalized recommendation techniques, International Journal of Intelligence Systems,2007,401–417
5. T. Lee, J. Chun, J. Shim, S.-g. Lee, An ontology-based product recommender system for B2B marketplaces, International Journal of Electronic Commerce, 2006,125–155.
6. J.B. Schafer, J. Konstan, J. Riedl, E-commerce recommendation applications, Applications of Data Mining to Electronic Commerce,Springer, US 2001, 115–153.
7. O.R. Zaiane, Building a recommender agent for e-learning systems, Proceedings of 2002 International Conference on Computers in Education, 2002, 55–59
8. T. Hung-Wen, S. Von-Wun, A personalized restaurant recommender agent for mobile e-service, 2004 IEEE International Conference on e-Technology, e-Commerce and e-Service. EEE, 2004, 259–262.
9. M. Deshpande, G. Karypis, Item-based top-N recommendation algorithms, ACM Transactions on Information Systems, 2004, 143–177.
10. M. Pazzani, D. Billsus, Content-based recommendation systems, The Adaptive Web, Springer, Berlin Heidelberg 2007, 325–341.
11. S. Middleton, D. Roure, N. Shadbolt, Ontology-based recommender systems, Handbook on Ontologies, Springer, Berlin Heidelberg 2009, 779–796.
12. R. Burke, Hybrid web recommender systems, , The Adaptive Web, Springer-Verlag, Berlin Heidelberg 2007, 377–408.
13. Lavrač N. Selected techniques for data mining in medicine[J]. Artificial intelligence in medicine, 1999, 16(1): 3-23.
14. Soni J, Ansari U, Sharma D, et al. Predictive data mining for medical diagnosis: An overview of heart disease prediction[J]. International Journal of Computer Applications, 2011, 17(8): 43-48.
15. Keleş A, Keleş A. ESTDD: Expert system for thyroid diseases diagnosis[J]. Expert Systems with Applications, 2008, 34(1): 242-246.
16. White R W, Yom-Tov E, Horvitz E, et al. Report on the SIGIR 2013 workshop on health search and discovery[C]//ACM SIGIR Forum. ACM, 2013, 47(2): 101-108.
17. http://www.datatang.com/data/44741
18. Zhang S, Zhang C, Yang Q. Data preparation for data mining[J]. Applied Artificial Intelligence, 2003, 17(5-6): 375-381.
19. Han J, Kamber M, Pei J. Data mining, southeast asia edition: Concepts and techniques[M]. Morgan kaufmann, 2006.
20. Ronghai Luo, etc. Correlation Calculation for Mixed Attribute [J]. Journal of Guangxi Normal University(Natural Science Edition), 2009, 27(1): 113-116.
21. Cortes C, Vapnik V. Support-vector networks[J]. Machine learning, 1995, 20(3): 273-297.
22. http://en.wikipedia.org/wiki/Support_vector_machine
23. Chang C C, Lin C J. LIBSVM: a library for support vector machines [J]. ACM Transactions on Intelligent Systems and Technology (TIST), 2011, 2(3): 27.
24. Platt J C. Using analytic QP and sparseness to speed training of support vector machines[J]. Advances in neural information processing systems, 1999: 557-563.
25. Schuldt C, Laptev I, Caputo B. Recognizing human actions: a local SVM approach[C]//Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on. IEEE, 2004, 3: 32-36.
26. Poli R, Cagnoni S, Livi R, et al. A neural network expert system for diagnosing and treating hypertension[J]. Computer, 1991, 24(3): 64-71.
27. Jin W, Li Z J, Wei L S, et al. The improvements of BP neural network learning algorithm[C]//Signal Processing Proceedings, 2000. WCCC-ICSP 2000. 5th International Conference on. IEEE, 2000, 3: 1647-1649.
28. Sun Y, Zhang S, Miao C, et al. Improved BP neural network for transformer fault diagnosis[J]. Journal of China University of Mining and Technology, 2007, 17(1): 138-142.
29. Kordylewski H, Graupe D, Liu K. A novel large-memory neural network as an aid in medical diagnosis applications[J]. Information Technology in Biomedicine, IEEE Transactions on, 2001, 5(3): 202-209.
30. Quinlan J R. Induction of decision trees [J]. Machine learning, 1986, 1(1): 81-106.
31. jcolibri2: A framework for building Case-based reasoning system