https://doi.org/10.1051/matecconf/201823202026

Traditional Chinese Medicine (TCM) Diagnosis Model Building Based on Multi-label Classification

Lu Zhou[1], Guang-geng Li[2], Yu-mei Zhou[1], Dan Yin[1], Yan Sun[1], Yan Zheng[2,a] and Yu-hang Li[1,a]
[1]School of Traditional Chinese Medicine, Beijing University of Chinese Medicine, 100029, Beijing China
[2]School of Computer Science, Beijing University of Posts and Telecommunications, 100876, Beijing China

Abstract. In the study, we propose a TCM diagnosis model that can be used for multi-label classification and give clear diagnosis, as well as the basis for diagnosis and differentiation when the symptoms correspond to multiple diseases or syndromes. The implementation of the model is divided into three steps. Firstly, choose the machine learning algorithm to train the TCM diagnosis model. The features of the training data are symptoms and the labels are diseases or syndromes. Secondly, give the number α ($\alpha>1$, $\alpha\in Z^{+}$) , the model will output the diagnoses with the top α highest probability according to the input symptoms as candidate diagnoses. Finally, the rules of differential diagnosis are designed to determine which candidate diagnoses should be reserved, thereby complete the multi-label classification. In our test dataset, by 10-fold cross-validation, the average accuracy of the single label classification was 0.882; the average precision was 0.974; the average recall was 1.000; the average f1 score was 0.967; the average accuracy of the multi-label classification was 0.706; the average micro precision was 0.934; the average micro recall was 0.941 and the average hamming loss was 0.060. Through the test we can know that this model had a good potential for auxiliary decision making in clinical diagnosis and treatment.

1 Introduction

In the traditional Chinese medicine(TCM) diagnosis, TCM practitioners analyze corresponding syndromes of patients and conduct differential diagnosis based on their information obtained through TCM four diagnostic methods to ensure the accuracy of diagnosis and avoid misdiagnosis and missed diagnosis.

For this process, an intelligent auxiliary model of TCM clinical diagnosis and treatment can be set up through machine learning, which can help practitioners use complex medical knowledge to deal with various medical problems more efficiently and quickly in the decision making of clinical diagnosis to avoid omissions and losses of important information and clues, so as to find more solutions for difficult miscellaneous diseases [1].

In recent years, the development of machine learning has provided many new methods for the auxiliary model of TCM clinical diagnosis. For example, neural network and random forest are used for modeling, which show high accuracy in clinical diagnosis of multi-class classification [2-4]. However, because of multi-class classification, which means only one diagnosis result is output, it is difficult for the model to deal with clinical symptoms corresponding to multiple diagnoses, and the diagnosis results output lack of understandable diagnostic evidence and differential diagnosis.

Therefore, in this study, according to the diagnosis and identification process based on symptoms of TCM, that is, inferring the corresponding TCM disease and syndrome according to the clinical symptoms and differential diagnosis, we put forward a TCM diagnosis model with multi-label classification. This model can provide multiple TCM diagnosis results that may be mapped according to the patient's clinical symptoms, and conduct differential diagnosis, remove the wrong results, make the model more consistent with the clinical diagnosis process.

2 Multi-label classification

In terms of the traditional classification problem, one instance (feature vector) is associated with one label [5]. That is, single instance and single label. For example, in the *Treatise on Febrile Diseases,* the disease corresponding to the following clinical symptoms including floating pulse, headache, stiff neck and aversion to cold is disease of the Taiyang channel.

But in reality, things themselves are complex, and one thing can associate multiple labels at the same time [5], namely single instance with multiple labels. For example, in the *Chinese Internal Medicine* [6], the disease corresponding to the following clinical symptoms including fever, little aversion to wind, sweating, headache, red face, cough, sputum, nasal congestion, a sharp tongue-red and rapid pulse is a cold with the syndrome of exterior attacked by wind-heat. That is, the corresponding diagnosis of this group of clinical symptoms is multi-label, which can be divided

[a] Corresponding author: yanzheng@bupt.edu.cn
[a] Corresponding author: liyuhang@bucm.edu.cn

into two levels to establish labels: the first level of labeling is the TCM disease level, i.e., "cold", and the second level of labeling is the syndrome level, i.e., "syndrome of exterior attacked by wind-heat". This example only takes two levels. In actual use, more levels can be built as needed. Compared with single label classification, TCM diagnosis model based on multi-label classification will have a wider scope of application.

3 Relative work

In last several years, great progress has been made in the study of clinical diagnosis and treatment models for multi-label classification. For example, reference [7] used BP neural network and chose the sigmoid function as the output function to train the TCM diagnosis and treatment model. The features of the training samples were symptoms, and the labels were traditional Chinese medicines. After the training was completed, the model can output a variety of traditional Chinese medicines according to the symptoms. Reference [8] used a recurrent neural network for multi-label diagnostic modeling, which can provide corresponding diagnoses based on the phenotyping features of dynamic changes. Since the training samples and the test samples were multi-label, the output function of this model was also sigmoid. Reference [9] used deep learning and one vs rest strategy for multi-label classification diagnosis modeling for the syndromes of TCM spleen and stomach diseases and achieved good results.

Compared with the above models, the model proposed in this paper is in accordance with the diagnosis and differential diagnosis process based on symptoms of TCM. Firstly, provide multiple possible candidate diagnoses for the input symptoms, then carry out differential diagnosis and exclude the candidate diagnoses that the model considers to be wrong, so as to achieve diagnostic multi-label classification.

4 Diagnosis model of TCM with multi-label classification

In traditional Chinese medicine, symptoms are the main basis for determining disease types and identifying syndromes, and thus are of important significance in the diagnosis of TCM [10]. Therefore, this model takes clinical symptoms as an important basis for identifying diseases and syndromes, and optimizes the decision-making process of the TCM diagnosis model with the use of single machine learning algorithm by means of differential diagnosis based on symptoms to achieve multi-label classification.

In short, the decision-making process of the model is as follows: firstly, output α ($\alpha > 1$, $\alpha \in Z^+$) diagnoses as candidate diagnoses based on the input clinical symptom. Secondly, extract the symptoms corresponding to each candidate diagnosis from the input symptoms. Finally, conduct differential diagnoses, namely, determine whether the symptoms corresponding to each candidate diagnosis simultaneously correspond to any possible diagnoses except the candidate diagnosis. If so, this candidate diagnosis can be excluded and if not, the diagnosis shall be reserved. Therefore, the main task of this model is to analyze the following issues:

- Which candidate diagnoses may be mapped by the currently input symptoms?
- Which current input symptoms are corresponding to each candidate diagnosis?
- Are the input symptoms corresponding to each candidate diagnosis corresponding to other possible diagnoses except the candidate diagnosis at the same time?

5 Methods

The model can be divided into two major parts. The first part is the TCM diagnosis model constructed by machine learning, and the second part is the differential diagnosis controlled by logic rules. The operation of this model includes three processing steps: multiple diagnostic outputs, reverse extraction and result identification.

In order to explain in detail the operation of the model, we use the Syndrome of Ephedra Decoction, Syndrome of Daqinglong Decoction, Syndrome of Puerariae Decoction and Syndrome of Gui-zhi Decoction in *Treatise on Febrile Diseases* for explanation.

5.1 Multiple diagnostic outputs

Multiple diagnostic outputs are completed in the machine learning, such as the diagnosis model with the use of the BP neural network. The features of the training set are symptoms, and the every label is disease or syndrome. For example, as for one sample, its label is "Syndrome of Gui-zhi Decoction" and its features include "fever, aversion to wind, headache, sweating, slower pulse". After the training is completed, the model will output the diagnoses with the highest probability of the top α according to the input symptoms as candidate diagnoses. The process after the model training is completed can be expressed as:

$$\mathrm{X} = \Phi(t) \tag{1}$$

$$\mathrm{y} = f(X) \tag{2}$$

$$C = max(\mathrm{y}, \alpha) \tag{3}$$

Equation (1) is a binarization function that converts a set of clinical symptoms to $\mathbf{X} \in \{0,1\}^{1\mathrm{xN}}$, where N is the number of symptoms. Equation (2) is an output function of a TCM diagnosis model constructed by a machine learning algorithm and can be used to output the probability of each diagnosis result according to **X**. Equation (3) indicates that the diagnosis results output by the model are sorted by probability, and the diagnoses with the highest probability of the top α are taken as the candidate diagnoses.

For example, input t as the main symptoms of Syndrome of Ephedra Decoction: [aversion to wind, fever, sweating, headache, rapid pulse], so **X** can be [1, 1, 0, 1, 0, 1, ..., 1, ..., 0]. If the model's 4 targets are classified as follows: 0: Syndrome of Ephedra Decoction;

1: Syndrome of Daqinglong Decoction; 2: Syndrome of Puerariae Decoction; 3: Syndrome of Gui-zhi Decoction. Then y can be [0.85, 0.13, 0.00, 0.00] (retain two significant figures after the decimal point). α is set to 2, which means that the first two results with high probability are taken out. The result can be [0, 1], $C_0 = 0$ for Syndrome of Ephedra Decoction, and C_1=1 for Syndrome of Daqinglong Decoction.

However, it should be noted that this process lowers the output threshold to output multiple diagnoses, so it usually contains wrong diagnosis, especially if there is only one correct result.

5.2 Reverse extraction

Reverse extraction is based on the multiple candidate diagnoses given in the previous step, extracting the symptoms supporting the candidate diagnoses from the input symptoms. The process can be expressed as:

$$Z_{\mathrm{i}} = T_{C_{\mathrm{i}}} \circ X \quad \mathrm{i} \in \{0,1,\ldots,\alpha\text{-}1\} \qquad (4)$$

Equation (4) can be used to calculate the intersection between the main symptoms of each candidate diagnosis and input symptoms for the extraction of symptoms Z_i that support the candidate diagnosis from the input symptoms. **T** refers to a binary-coded matrix of main symptoms of diseases and syndromes that can be provided by experts in the field. The order of **T** is the same as the target classification of the machine learning model. C_i represents the ith element in the candidate diagnoses C, then Tc_i represents the binary code of the candidate diagnosis C_i corresponding to the main symptoms. The process of combining (1)(2) is shown in Figure 1.

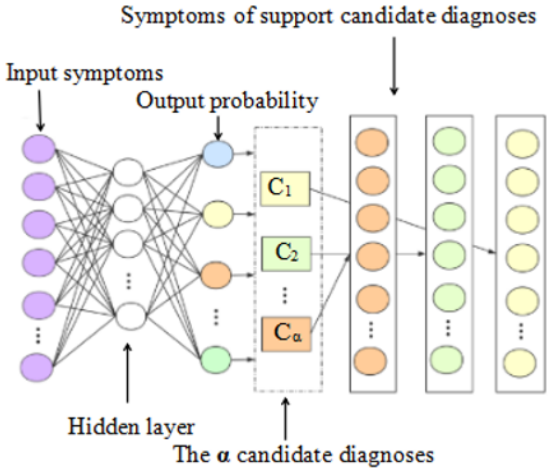
Symptoms of support candidate diagnoses
Input symptoms
Output probability
C1
C2
Cα
Hidden layer
The α candidate diagnoses

Figure 1. Output of multiple diagnostic outputs and the reverse extraction.

For example, the candidate diagnoses C calculated by Equation (3) in the above example is [0 1], then Tc_0 can be [1, 1, 0, 1, 0, 1, ..., 1 ,...,0], Z_0 can be [1, 1, 0, 1, 0, 1, ...,1, ...,0].

5.3 Result identification

The result identification shall return to the diagnosis model with the use of machine learning and conduct the second part, namely, the differential diagnosis. First, input Z_i into the diagnosis model to reacquire the corresponding probability of C_i, as shown in Figure 2. Next, activate the differential diagnosis shown in Figure 3, and judge whether there are other possible diagnoses corresponding to these symptoms Z_i at the level corresponding to C_i, so as to determine whether to reserve this candidate diagnosis C_i. For example C_0 is at the syndrome level of Syndrome of Ephedra Decoction, then the differential diagnosis is only executed between syndrome and syndrome. The main processing here is:

$$\mathrm{y} = f(Z_{\mathrm{i}}) \qquad (5)$$

$$P_{\mathrm{i}} = \mathrm{y}_{C_{\mathrm{i}}} \qquad (6)$$

$$H = max(\mathrm{y}, \alpha) \qquad (7)$$

$$D = T \circ Z_{\mathrm{i}} \qquad (8)$$

$$L_{\mathrm{i}} = G(P_{\mathrm{i}}, Z_{\mathrm{i}}, D, H) \qquad (9)$$

Equation (6) refers to update the probability that C_i corresponds based on y. Equation (8) can be used to obtain the intersection of Z_i and **T** and input the intersection into Equation (9) to determine whether Z_i corresponds to other diagnosis results except C_i. Equation (9) is a differential diagnosis rule as shown in Figure 3, and can be used to output the determined diagnosis; and the θ is a threshold of the output probability. In the differential diagnosis in Figure 3, the hamming distance between the clinical symptoms corresponding to Z_i and H_d can be determined. When the two are completely the same, the hamming distance is 0, which proves that the candidate diagnosis C_i has the same symptoms as other possible diagnoses. So the model will exclude candidate diagnosis C_i. Through differential diagnosis, the model can determine which candidate diagnoses can be reserved, so as to achieve multi-label classification of TCM diagnosis.

For example, substitute Z_0: [1, 1, 0, 1, 0, 1, ...,1, ...,0] in the above example, that is, the binary code of the corresponding symptoms of Syndrome of Ephedra Decoction to (5), the result y can be [0.85, 0.13, 0.00, 0.00]. Then the result P_0 of (6) is [0.85]. The result of (7) is still [0, 1], and the result of (8) is [[1, 1, 0, 1, 0, 1, ...,1, ...,0], [1 , 1, 0, 1, 0, 1, ...,1, ...,0], ...,[1, 1, 0, 1, 0, 0, ...,0, ...,1]]. Finally substitute these results into (9), then the output result can be 0, that is, to reserve the diagnosis of Syndrome of Ephedra Decoction.

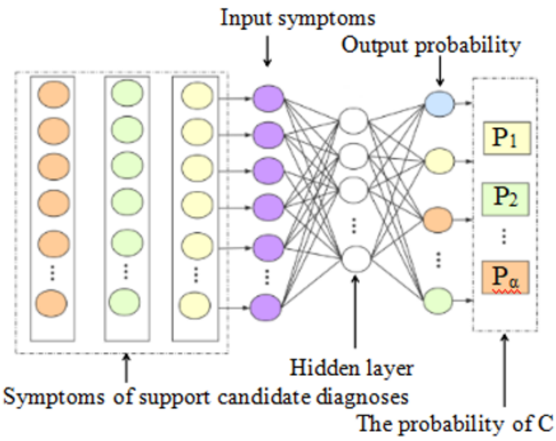

Figure 2. Update the corresponding probability.

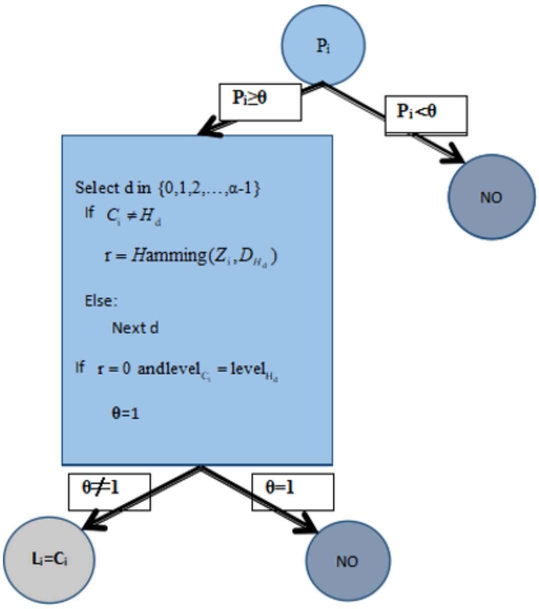

Figure 3. Differential diagnosis. Note: C_i refers to the current possible candidate diagnosis, P_i refers to the current probability of diagnosis, θ refers to the setting output threshold, H_d refers to the diagnosis results in need of differentiation, Hamming refers to the function to determine the hamming distance, and level refers to the corresponding level.

In summary, the TCM diagnosis model proposed in this study is to classify multiple labels into multiple single instance single label classification problems. It can not only output diagnosis results, but also have differential diagnosis and strict diagnostic basis control, allowing the physician to know how the model made identification and diagnose from multiple aspects. This model is not limited to BP neural networks, and other algorithms such as random forest can also be used. Modeling is shown in Table 1.

Table 1. Algorithm

Step 1: Build model
Input: Feature matrix **X**, label matrix **Y**
Select machine leaning algorithm and build model by input
Step 2: Predict
Input: **X**,α,**T**, θ
Evaluate probably results of C by (1-3)
Repeat
Select a C_i from C $i \in \{0, 1, 2, ..., \alpha-1\}$ Evaluate Z_i by (4) Evaluate probably results of H and probably of C_i by (5-6) Evaluate symptoms of H by (7) Evaluate final results by (8-9)
Until i = α

6 Experiment

For the experiment of model diagnosis performance, there is no publicly available TCM disease and syndrome dataset, so we use the dataset of 103 clinical samples collected from the experts of the Endocrinology Department of Nephropathy to train and test the model. The diseases included in the dataset are as follows: 8 cases of consumptive thirst, 7 cases of kidney water, and 6 cases of Guange (dysuria and vomiting). Syndromes contained: 11 cases of kidney-yang deficiency, 17 cases of water-rheum collecting internally, 21 cases of blood stasis, and 33 cases of deficiency of both qi and blood. The samples in the dataset include two levels, that is, disease and syndrome. Then establish label bases on above levels, and the labels of the level of disease are as follows: consumptive thirst, kidney water and Guange. The labels of the level of syndrome are as follows: kidney-yang deficiency, water-rheum collecting internally, blood stasis and deficiency of both qi and blood. The features of each sample in the dataset correspond to a set of clinical symptoms in the above labels, and the set of symptoms only correspond to one label. For example, the features are "ache of waist, nocturia, fatigue, white fur, sink pulse", then the label is "deficiency of kidney yang".

The samples in the dataset are single instance single label, and each sample corresponds to only one identified disease or syndrome label. Compared to the structure of one sample corresponding to multiple labels, this can make the sample's labels have clear symptoms as their basis, and because there is only one label, omission of labels or baseless labels can be avoided, ensuring the reliability of the dataset. This dataset is subject indexed by the *TCM Thesaurus* in the Chinese medicine prescription intelligent analysis system(CPIAS), which normalizes the dataset content and makes the features and labels with the same meaning expressed in the same way. The symptoms and number of the dataset, number of samples and category labels are shown in Table 2 and Table 3.

In this study, we use 10-fold cross-validation to test, clinical symptoms as input, and diagnosis results as output, and test results were evaluated to verify the manifestation of single label classification. The manifestation of multi-label classification was verified by collecting 15 samples with the label patterns of the disease + syndrome. For example, the multi-label: consumptive thirst + deficiency of kidney yang + blood stasis.

The models proposed in this study were modeled with the use of random forest and BP neural network (two machine learning algorithms used to train TCM diagnosis model). In addition, we used the same training set, adopted one vs rest strategy to establish a baseline model between random forest and BP neural network. The parameters of all models are same. After many times of filtering of the main parameters of modeling, a better performing set is shown in Table 4.

Table 2. The symptoms and number in the dataset.

Symptoms and number				
fatigue 43	poor sleep 23	ache of waist 22	lower limb swelling 19	dull tongue 19
swollen tongue 15	proteinuria 12	white fur 11	turbid urine 10	greasy fur 10
pulse is thin 8	dark red tongue 7	dark purple tongue 7	poor appetite 7	emaciati on 7
high serum creatinine 6	dry stool 6	light red tongue 6	diabetes history 6	sink pulse 5
polydipsia 5	nocturia 5	dark and gloomy face 5	teeth-print tongue 4	hypouro crinia 4
high IFG 4	urea nitrogen 4	edema 4	aversion to cold 3	tinnitus 3

Table 3 Statistics of diseases and syndromes.

Sample layer	Label	Sample size
Disease	consumptive thirst kidney water Guange	8 7 6
Syndrome	deficiency of kidney yang water-rheum collecting internally blood stasis deficiency of both qi and blood	11 17 21 33

Table 4. Modeling parameters.

Algorithm	Parameters	Output parameters

BP neural network	Activation=RELU
	Hidden Layer Sizes=20
	Learning Rate=0.01
	Max Iteration=300
	Solver=Adam
		θ=0.3
		α=7

random forest	Criterion=Gini
	Max Depth=20
	Max Leaf Nodes=30
	Estimators=100
		θ=0.3
		α=7

7 Results

The test results of the model single label classification were evaluated using accuracy, precision, recall and f1 score. For the multi-label classification, we used macro precision, macro recall, micro precision, micro recall and hamming loss to evaluate multi-label classification capabilities. All the tests use 10-fold cross-validation.

7.1 Single Instance single label classification

In single label classification, the average accuracy, average precision, average recall and average f1 score of 10-fold cross-validation are shown in Table 5.

Table 5. Single label classification evaluation.

Model	Average accuracy	Average precision	Average recall	Average f1 score
Our proposed model(BP)	0.882	0.974	1.000	0.967
Our proposed model(RF)	0.823	0.911	0.961	0.927
one vs rest model(BP)	0.882	0.974	0.971	0.968
one vs rest model(RF)	0.765	0.889	0.971	0.915

Note: BP is based on BP neural network; RF is based on random forest

7.2 Single Instance multi-label classification

The multi-label classification capabilities are evaluated by the average macro precision, average macro recall, average micro precision, average micro recall, and average hamming loss after 10-fold cross-validation. See Table 6 and Table 7.

Table 6. Multi-label classification evaluation of average accuracy, average macro precision and average macro recall.

Model	Average accuracy	Average macro precision	Average macro recall
our proposed model(BP)	0.706	0.934	0.946
our proposed model(RF)	0.690	0.937	0.939
one vs rest model(BP)	0.06	0.361	0.777
one vs rest model(RF)	0.306	0.754	0.944

Note: BP is based on BP neural network; RF is based on random forest

Table 7. Multi-label classification evaluation of average micro precision, average micro recall and average hamming loss.

Model	Average micro precision	Average micro recall	Average hamming loss
our proposed model(BP)	0.934	0.941	0.060
our proposed model(RF)	0.934	0.925	0.069
one vs rest model(BP)	0.320	0.914	0.358

one vs rest model(RF)	0.705	0.935	0.173

Note: BP is based on BP neural network; RF is based on random forest

8 Discussion

8.1 Advantages of multi-label classification of TCM diagnosis model

Multi-label classification allows a set of features to have multiple labels at the same time, and single label classification can be considered as a special case of only one label in a multi-label classification. In the TCM clinical diagnosis, a set of clinical symptoms often corresponds to multiple levels of diagnosis such as disease level and syndrome level. Therefore, multi-label classification is better for clinical diagnosis of TCM. It can be used to establish TCM diagnosis model of multi-label classification with the use of modeling strategies of multi-label classification including algorithm adaptation and provide results more consistent with the needs of clinical diagnosis of TCM according to the input TCM clinical symptoms.

8.2 Performance and analysis of the model

The one vs rest strategy can be used for multi-label classification. However, for samples collected in this study, every sample only corresponds to one label. This makes it difficult for the model to set the threshold of probability output when the input symptoms of the model correspond to multiple labels due to no similar sample in the training process. If the threshold set is too small, there will be too many output results, so there will be wrong diagnoses in the results; if the threshold set is too high, there will be omissions of diagnoses, so it will be difficult to exert the advantages of multi-label classification.

On the contrary, the model proposed in this paper does not depend entirely on the model's result probability to make the final diagnosis, but improves the judgment method of the model through algorithm adaptation. In the simulation, the model can narrow the diagnosis scope, identify the main symptoms of the syndrome and conduct differential diagnosis according to clinical symptoms. The model can select the corresponding disease or syndromes of the clinical symptoms, conduct differential diagnosis and exclude the results that the model considered to be wrong, so as to determine the number of labels, thereby realizing the learning of single instance single label sample and performing multi-label classification. In addition, when the sample is labeled in this study, we conducted attribute labeling of the level of TCM disease or TCM syndromes, so as to make manual labeling and model judgment have a clear level.

8.3 Problems and proposed solutions of models in multi-label classification

Although the TCM diagnosis model proposed in this paper has good performance in multi-label classification, there is also one problem, that is, heavy workload, because experts need to label the main symptoms corresponding to each diagnosis result.

For this problem, it may be further combined with the Gibbs sampling method used in the topic model [11] to quantify the importance of symptoms for the diagnosis of diseases or syndromes, assisting experts to label more easily, and promising to solve the heavy labeling workload of experts. In addition, there is another problem, that is, the training samples are too small and there is a risk of overfitting or underfitting. Therefore, we will further explore small samples learning method of one shot learning [12] in the future in order to solve the problem of small samples learning in TCM.

9 Conclusion

The intelligent TCM diagnosis model proposed in this paper not only can provide the probability of results, but also can realize the differential diagnosis and the extraction of syndrome evidence on the basis of multi-label classification, making the results more interpretable. In order to further establish a syndrome differentiation and treatment model based on the principle of "if there is manifestation of this syndrome, we can use its corresponding prescription", an effective method is provided to enable it to have the potential for auxiliary clinically diagnosis and treatment.

Acknowledgment

Yu-hang Li and Yan Zheng carried out the research and designed the experiment. Lu Zhou and Guang-geng Li wrote the manuscript. Lu Zhou, Guang-geng Li, Yu-mei Zhou, Dan Yin, contributed to experimental manipulation and data acquisition. Yan Sun contributed to experimental instruction and data analysis.

Funding

This study was funded by the National Key Research and Development Program of China (Grant No. 2017YFC1700303)

References

1. X. Chen, P. H. Liu, Y. Z. Sun, X. Shen, L. Zhang, X. Q. Wang. Research on Disease Prediction Models based on Imbalanced Medical Data Sets. Chinese Journal of Computers, 1-14 (2017)
2. J. Wang, R. Wu, X.Z. Zhou. Syndrome factors based on SVM from coronary heart disease treated by prominent TCM doctors. Journal of Beijing University of Traditional Chinese Medicine, **31**(08):540-543+560 (2008)

3. L. Xu, S.Q. Chen, J.H. Hou. W. X. Bi, F. Yuan. The Research on the Construction of the TCM Differentiation Model Based on BP Neural Network. World Chinese Medicine, **11**(02):335-338 (2016)
4. H.Z. Wang, X.Q. Hu. Intelligent diagnosis classification on TCM "five pathogens produced by five organs". Computer Engineer Application, **47**(06):156-160+163 (2011)
5. M. L. Zhang, Z. H. Zhou. A Review on Multi-Label Learning Algorithms. IEEE Transactions on Knowledge & Data Engineering, **26**(08):1819-1837 (2014)
6. M.H. Wu, X.Y. Wu. *Chinese Internal Medicine*, 72-73 (2012)
7. Q.W. Chen. A Model of TCM Syndromes and Treatment Based on Artificial Neural Network. Chinese Archives of Traditional Chinese Medicine, **27**(07):1517-1520 (2009)
8. Z.C. Lipton, D. C. Kale, R. C. Wetzell. Phenotyping of Clinical Time Series with LSTM Recurrent Neural Networks, arXiv preprint arXiv:1510.07641 (2015)
9. G.P. Liu, J. Yan, Y.Q. Wang, W. Zheng, T. Zhong, X. Lu, P. Qian. Deep learning based syndrome diagnosis of chronic gastritis. Computational and Mathematical Methods in Medicine, vol.2014, Article ID 938350, 8 pages (2014)
10. W.F. Zhu, *Diagnostics of Traditional Chinese Medicine*, 1(2007)
11. M. Rosen-Zvi, C. Chemudugunta, T.L. Griffiths, P. Smyth, M. Steyvers. Learning author-topic models from text corpora. Acm Transactions on Information Systems, 28(1):4 (2010)
12. F.F. Li, R. Fergus, P. Perona. One-Shot learning of object categories. IEEE Transactions on Pattern Analysis and Machine Intelligence, **28**(4):594-611(2006)