



OPEN ACCESS

EURASIA Journal of Mathematics Science and Technology Education  
ISSN: 1305-8223 (online) 1305-8215 (print)  
2017 13(8):5577-5587  
DOI: 10.12973/eurasia.2017.01011a



## Positive and Negative Association Rules Mining for Mental Health Analysis of College Students

Long Zhao

Qilu University of Technology, Jinan, China

Feng Hao

Qilu University of Technology, Jinan, China

Tiantian Xu

Qilu University of Technology, Jinan, China

XiangJun Dong

Qilu University of Technology, Jinan, China

Received 23 February 2017 • Revised 5 July 2017 • Accepted 14 August 2017

### ABSTRACT

The psychological problems of college students have aroused general concerns. A lot of college students are plagued by all kinds of psychological health problems. Psychological health problems brought a lot of negative effects to them. The psychological assessment data and the basic information collected from 6500 freshmen are used to analyze association rules and characteristics of college students' psychological factors in this paper. The symptom self-rating scale (SCL-90) was compiled by L. R. Derogatis in 1975, which contains 90 items. The SCL-90 has been used in a wide range of psychiatric symptoms. The SCL-90 includes ten factors, such as somatization, obsessive-compulsive symptoms, interpersonal sensitivity, depression, anxiety, hostility, terror, paranoia, psychosis and other factors. The PNARC model is introduced in this paper to mine the positive and negative association rules from real SCL-90 data set of one Chinese college. The support-confidence framework and the correlation test method are obtained to delete the contradict association rules and get the positive and negative association rules for correctly analyzing the potential relationship of SCL-90 factors. Mined positive and negative association rules are also helpful to analyze and coach the mental health of college students.

**Keywords:** psychological health problems, college students, SCL-90, the PNARC model, association rules

### INTRODUCTION

According to the relevant statistics, the mental health of college students is in a very unhealthy stage (Jena, S., & Tiwari, H. 2015). The 13% to 23% of students have mental health problem. About 9% of college students have serious mental disorders. This trend is increasing year by year. The research on college students' psychological analysis has become the focus of health education in universities. The symptom self-rating scale (SCL-90) was compiled by L. R. Derogatis in 1975, which contains 90 items (Herawan, T., Vitasari, P., & Abdullah, Z. 2012). Each

© **Authors.** Terms and conditions of Creative Commons Attribution 4.0 International (CC BY 4.0) apply.

**Correspondence:** Long Zhao, Lecturer, School of information, Qilu University of Technology, Jinan, China. Address to No. 3501, University Road, Changqing District, Jinan, China Tel: +86-531-89631258.

**Correspondence:** Tiantian Xu, Lecturer, School of information, Qilu University of Technology, Jinan, China. Address to No. 3501, University Road, Changqing District, Jinan, China Tel: +86-531-89631251.

✉ [zxcvbnm9515@163.com](mailto:zxcvbnm9515@163.com) & [xtt-ok@163.com](mailto:xtt-ok@163.com)

#### State of the literature

- The research on mental health of college students has been one of the most important issues that many researchers have focused on.
- For the large-scale properties and huge databases, traditional Apriori algorithm of association rules is difficult to achieve good effects.
- The commonly SCL-90 metrics are discretized into two values that is normal and abnormal. But the importance of the embodiment item does not come out.

#### Contribution of this paper to the literature

- The support-confidence framework and the correlation test method are obtained to delete the contradict association rules and get the positive and negative association rules for correctly analyzing the potential relationship of SCL-90 factors.
- In the mining process of this paper, the frequent item sets corresponding to the impact factor are extracted firstly, and then the PNARC model is put forward to mine the positive and negative association rules.
- This paper extracts the different association rules from psychometric properties of the 6500 freshmen samples. The results show that the mining association rules mined by proposed algorithm have practical guiding significance.

item has five grades. A wide range of psychiatric symptoms, such as feeling, emotion, thinking, consciousness, behavior guide habits, interpersonal relationships, diet and sleep, etc. are tested according to their actual situation. A person's symptoms are measured in a certain period of time (usually a week). The scales of SCL-90 are simple and easy to use. The SCL-90 has a good distinction for the potential psychological barrier. It also can apply and check the crowd who may have psychological barriers, which kind of psychological barriers and how serious. And it is always used as a reference in clinical diagnosis or as a primary screening tool.

The research on mental health of college students has been one of the most important issues that many researchers have focused on. In the university campus, in addition to the study of professional culture courses, try to live independently after leaving their parents, some problems will adversely affect the mental health of college students. In all the participating groups, there are about 16% to 30% of college students have depression, anxiety, obsessive-compulsive, interpersonal relationships, personality disorders and other psychological problems (Özyürek, P., & İbrahim Kılıç. 2015). Interpersonal relationships, academic problems, emotional problems and neurosis are the common psychological problems of college students (Zheng, G., et al. 2014). They cause a lot of trouble to the normal study and life. Psychological problems not only brought a lot of learning and life problems, but also brought harm to themselves. But for college students' psychological problems, most of the colleges and universities are simply to strengthen the theoretical education and management which can't create a protective effect (Burlaka, V, et al. 2014).

This paper extracts the different association rules from psychometric properties of the 6500 freshmen samples. Because Apriori algorithm design and implementation of multidimensional association rules mining algorithm, and study the relationship between psychometric properties of the scales are applied (Wang, D., & Psychology, S. O. 2015). The approach used can be applied to association rule mining and related fields (Nembhard, D. A., et al 2012). The analyses for college students' mental health have become the focus of domestic and foreign scholars' research. Mental health has a direct impact on the current college students' normal study and life. In order to deeply understand the main factors affecting the psychological health of college students and the correlation between psychological symptoms, the association rule mining is applied to college students' psychological health survey data. After preprocess the initial data of college students' psychological assessment information, a multidimensional association rule mining model is built (Qi, W. et al 2013).

In order to do better in guiding the education of college students' psychological health, this paper tries to find out the association rules between the various factors of SCL-90 data set. The research tests the psychological health of college students by using self-rating mental health clinical symptoms (Cheng, C. W, et al. 2014). In the mining process of this paper, the frequent item sets corresponding to the impact factor are extracted firstly, and

then the PNARC model is put forward to mine the positive and negative association rules. The hidden rules are obtained from the SCL-90 data set. The positive and negative association rules are confirmed. The results show that the mining association rules mined by proposed algorithm have practical guiding significance.

## FREQUENT ITEM SET EXTRACTION

J. Han proposed the Frequent-Pattern tree (FP-tree) algorithm for the defect of Apriori algorithm (Han, J. 2010). It isn't a generation of candidate frequent item sets method using divide rule to decompose data mining tasks into smaller tasks. The algorithm uses FP-tree structure to compress the database, avoiding the high cost of database scanning (Han, J., et al 2000). The FP growth algorithm adopted the strategy of divide and conquer to compress the database contained the frequent item sets in a FP-tree. A centralized information association letter is still preserved. The compressed database is then divided into a special projection database-the condition database. Each condition corresponding to a frequent item of each database is mined.

The main flow of FP-growth algorithm is: FP-tree → conditional pattern bases → conditional FP-tree → frequent patterns. Conditional pattern libraries are created for each node in the FP-tree. The conditions pattern bases are used to construct the corresponding conditional FP-tree. Recursively construct condition FP-tree at the same time increases the frequent sets. If the conditional FP-tree contains only one path, the generated frequent sets are directly generated. If the conditional FP-tree contains multiple paths, a hybrid approach is used. Specific steps are as follows:

Step 1: Create the root node of the tree, mark the node "null".

Step 2: Permute each item of transaction database in the descending support, and then according to the order of each transaction create a branch.

Step 3: when conceding add a branch for another transaction, each node on the corresponding prefix plus 1 along with the count of common affairs arrangement, creating new connection for the prefix nodes.

FP-growth algorithm maintains the complete information of frequent item set mining, without interrupting the long mode. The long frequent pattern problem is transferred to recursively find some short mode. The short mode and the suffix are connected. FP-growth algorithm uses the least frequent items as a suffix, which shows good selectivity. This method can save the search cost, and greatly improve the efficiency of the algorithm. The solutions is firstly dividing the database into a set of projection databases, then construct FP-tree on each projection database and dig it up. This can be used recursively for the projection database. For mining long and short frequent patterns, it is both efficient and scalable, and approximately one order of magnitude faster than the Apriori algorithm.

It is easy to calculate the degree of support and confidence in the positive association rules. However, it is difficult to directly calculate the support and confidence of the negative association rules. This method uses the known positive supports to calculate negative association rule's support and confidence by appropriate conversion. Here is the calculation method proposed in this paper. Transform the set operation of the item set to the set operation of the transaction set. This is more conducive to the application of some theorems and properties, but also easy to understand.

## ALGORITHM DESIGN

The FP\_growth algorithm and PNARC algorithm are used to dig out the corresponding association rules. Set  $A_s$  is an item set of transactions, the base  $|A_s|$  is the number of transactions in  $A_s$ . Similarly,  $B_s$  said the item set of  $B$  transaction. The base  $|B_s|$  is the number of transactions in the  $B_s$  database.  $D$  is the item set of all transactions in the database. The base  $|D|$  is all the number of transactions.

Input:  $D$  database; minimum support threshold (minsupport).

Output: the complete set of frequent patterns.

According to the following steps from the database to construct a frequent pattern tree FP-tree:

Step 1: Scan transaction database D once. Collect 1-frequent item L1 item set, record its support degree. The items in the L1 to support a descending sort order, get the descending frequent item table L1.

Step 2: Create a frequent pattern tree FP-tree root node, using the "null" to it as a tag. Re adjust the transaction database D, remove the non-frequent items in each transaction, and sort the remainder of the transaction in the L1 order.

Step 3: Mining condition pattern base from FP-tree tree and condition pattern tree. This procedure is implemented by calling FP growth. The process is as follows:

The Positive and Negative Association Rules on Correlation (PNARC) algorithm can determine the correlation between item sets and frequent item sets mining positive and negative association rules, it can detect and remove those independent item sets generated rules. In the algorithm, we assume that the frequent item sets are obtained and stored in the set. PNARC model used the support degree-confidence framework and the correlation test method to delete the contradictory association rules and get the correct positive and negative association rules (Dong, X. et al. 2007).

The connection of the two rules shows that there are contradictions between the two rules, which cannot exist at the same time in the result set. A and B is one item of data set. Such as  $A \Rightarrow B$ ,  $A \Rightarrow \neg B$  and  $A \Rightarrow B$  cannot exist at the same time. No connection between  $A \Rightarrow B$  and  $\neg A \Rightarrow \neg B$  means two rulers can exist at the same time. The application of statistical correlation can effectively solve this problem. According to the knowledge of probability theory, the support and confidence formula of positive and negative association rules are as follows:

$$s(\neg A) = 1 - s(A) \quad (1)$$

$$s(A \cup \neg B) = s(A) - s(A \cap B) \quad (2)$$

$$s(\neg A \cup B) = s(B) - s(A \cap B) \quad (3)$$

$$s(\neg A \cup \neg B) = 1 - s(A) - s(B) + s(A \cap B) \quad (4)$$

$$conf(A \Rightarrow B) = \frac{s(A \cap B)}{s(A)} \quad (5)$$

$$conf(A \Rightarrow \neg B) = \frac{s(A) - s(A \cap B)}{s(A)} = 1 - conf(A \Rightarrow B) \quad (6)$$

$$conf(\neg A \Rightarrow B) = \frac{s(B) - s(A \cap B)}{1 - s(A)} \quad (7)$$

$$conf(\neg A \Rightarrow \neg B) = \frac{1 - s(A) - s(B) + s(A \cap B)}{1 - s(A)} = 1 - conf(\neg A \Rightarrow B) \quad (8)$$

$$corr = \frac{s(A \cup B)}{s(A) s(B)} \quad (9)$$

PNARC algorithm:

Input: L: frequent item set; *minconf*: minimum confidence;  
 Output: PAR: positive association rule set; NAR: negative association rule set;  
 Step 1: PAR =  $\Phi$ ; NAR =  $\Phi$ ;  
 Step 2: The positive and negative association rules in the frequent item set L  
 For any frequent item set X in L do begin  
 For any item set  $A \cup B = X$  and  $A \cap B = \Phi$  do begin  
 If *corr* > 1 then begin  
 Step 3: Generate rules like  $A \Rightarrow B$  and  $\neg A \Rightarrow \neg B$   
 If  $conf(A \Rightarrow B) \geq minconf$  then  
     PAR = PAR  $\cup \{A \Rightarrow B\}$ ;  
 If  $conf(\neg A \Rightarrow \neg B) \geq minconf$  then  
     NAR = NAR  $\cup \{\neg A \Rightarrow \neg B\}$ ;  
 End;  
 If *corr* < 1 then begin  
 Step 4: Generate rules like  $A \Rightarrow \neg B$  and  $\neg A \Rightarrow B$   
 If  $conf(A \Rightarrow \neg B) \geq minconf$  then  
     NAR = NAR  $\cup \{A \Rightarrow \neg B\}$ ;  
 If  $conf(\neg A \Rightarrow B) \geq minconf$  then  
     NAR = NAR  $\cup \{\neg A \Rightarrow B\}$ ;  
 End;  
 End;  
 End;  
 Step 3: return PAR and NAR;

Obviously, the algorithm PNARC can be used for mining association rules in some kinds of forms, as long as the corresponding position in the algorithm to add the appropriate statement can be. Algorithm AR in the algorithm PNARC based on the addition of a part of the calculation and judgment statement, the two have the same time complexity.

## EXPERIMENTS AND ANALYSES

Experimental data sets are from the SCL-90 test of one college students. SCL-90 results were divided into ten factors, such as somatic symptoms, obsessive-compulsive symptoms, interpersonal sensitivity, depression, anxiety, hostility, terror, paranoia, psychosis, and other factors. We combined the ten factors and the gender of the subjects as the research objects. Data mining techniques are used to study the relationship between mental health factors and the health factors (Hodgins, D. C., et al 2015).

The cleaning of the experimental data, only ten factors and the subjects' gender are selected as mining object, remove other data records. Mental health clinical symptom checklist (Derogatis, L.R) assessment methods: according to the degree of mental health division of 1-5. Test questions are given to determine the answer, no, light, moderate, heavy and serious. There is no specific or rigid definition according to their own situation to experience. The time range of the answer is "the last week" or "now, the moment" true feeling. Ten influencing factors are analyzed in this paper somatization, Obsession, interpersonal relationship sensitiveness, depression, anxious, hostility, Phobic Anxiety, Paranoid, Psychoticism and Additional Items.

**Somatic factor :** It refers to the feelings and reactions of the human body. It includes basic discomfort, not physical illness. More is the subjective feeling. Such as gastrointestinal discomfort, difficulty breathing, pain unbearable, weakness, caused by climate events or anxiety, excessive excitement etc. The topic in the scale include 58, 56, 53, 52, 49, 48, 42, 40, 27, 12, 4, 1(numbers are reverse). In the volume table test topics include a total of 12 items of this category.

**Obsessive compulsive symptoms:** This is different from compulsive what we feel. It refers to we always cannot get rid of meaningless thoughts or behavior and we clearly know that is not necessary but do not believe in yourself, or it is perceived barriers, these all belong to the category of obsessive-compulsive symptoms. The measurement subjects are 65, 55, 51, 46, 45, 35, 8, 35, 28, 10, 9, and 3.

**Interpersonal sensitivity:** It reflected in the process of dealing with people, people often feel inferior and they don't get along well with people. They often exhibit an uncomfortable feeling. Other people do not want to close or along with him. This kind of symptom is more apparent when they compared to others. Test topics include 9 items, such as 73, 69, 61, 41, 37, 36, 34, 21, and 6.

**Depression:** It mainly reflects two concepts, one is the depression group, and another is the specific symptoms in clinic. Losing faith in life, lacking of confidence and energy, no contact or communication with the outside world or having suicidal thoughts when serious are all reflects of depression. Test topics include 13 items, such as 79, 71, 54, 31, 20, 29, 26, 22, 20, 15, 14 and 5.

**The anxiety factor:** It often refers to the tension, neuroticism or the physical signs which a person cannot control. It is the performance or experience of the significant association of the clinical manifestations and symptoms of anxiety. The main contents of this study include the uncertain mood which includes panic of attack or swim away from misty. Test the topics include ten items, such as 86, 80, 78, 72, 57, 39, 33, 23 and 17.

**Hostility factor:** It often refers to a bad manner when people face to something. They are not optimistic, often tired of things, and they argue, even breaking things, irrepressible impulse, outburst temper and so on. We can see these in the test of human emotions, thoughts and various unfriendly behavior. Test questions include 6 items, such as 11, 24, 63, 67, 74 and 81.

**Fear factor:** The subjects often panic mentality, often do not have a sense of security, especially in social phobia events, in addition to common travel, or riding public transport, these events will cause the fear of him, even the empty field, free people are not allowed to feel security. Test questions include 7 items such as 13, 25, 47, 50, 70 and 75.

**Paranoia factor:** It refers to that he individual often reflects projected thinking and exaggerated in thinking, such as too conceited, widespread suspicion, morbid jealousy, delusion, excessive hostility. The scale includes some of the basic contents. Test includes 6 items such as 8, 18, 43, 68, 76 and 83.

**Psychosis:** It refers to that the subjects showed less dynamic, eccentric, emotional about desert, rupture of thinking, thinking of relaxation, paranoia and other symptoms of schizophrenia, reaction of poor adaptability in the society, the social function is on the decline. Test items include ten items, such as 90, 88, 87, 85, 84, 77, 62, 35, 16 and 7.

**Other factors:** This is some additional questions, these questions are not included in any factor above, it mainly in order to make the total sum consistent with the factor scores. Through these test questions we can know about the standard of diet health status and sleep quality of personnel. To collect the freshmen's psychological test scale, to calculate the value of each impact factor, the processed results are shown in [Table 1](#).

**Table 1.** The partial data of the original data

somatization	Obsessive-Compulsive	Interpersonal Sensitivity	Depression	Anxiety	Hostility	Phobic Anxiety	Paranoid Ideation	Psychoticism	Additional Items
1.0	1.5	1.2	1.4	1.1	1.0	1.1	1.3	1.2	1.1
1.2	1.2	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
1.3	1.4	1.2	1.5	1.4	1.0	1.0	1.2	1.1	1.1
1.0	1.2	1.7	1.1	1.0	1.0	1.4	1.0	1.5	1.0
1.5	1.9	1.6	1.5	1.5	1.2	1.3	1.3	1.3	1.6
1.7	2.3	2.1	1.9	1.7	1.2	1.4	1.5	1.6	1.9
1.0	1.3	1.2	1.1	1.1	1.2	1.0	1.3	1.0	1.1
1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
1.1	1.8	1.4	1.2	1.4	1.3	1.1	1.8	1.4	1.4

The normal score of each factor is shown in table 1.

**Table 2.** The normal range of SCL-90 values

Column label	Item	normal range
C	Somatization	1.37 ± 0.48
D	Obsession	1.62±0.58
E	interpersonal relationship sensitiveness	1.65±0.61
F	Depression	1.5±0.59
G	Anxious	1.39±0.43
H	Hostility	1.46±0.55
I	Phobic Anxiety	1.46±0.55
J	Paranoid	1.43±0.57
K	Psychoticism	1.20±0.42

The commonly SCL-90 metrics are discretized into two values that is normal and abnormal. But the importance of the embodiment item does not come out.

**Table 3.** The partial converted data

somatization	Obsessive-Compulsive	Interpersonal Sensitivity	Depression	Anxiety	Hostility	Phobic Anxiety	Paranoid Ideation	Psychoticism	Additional Items
A1	B2	C1	D1	E1	F1	G1	H1	I1	J1
A1	B1	C1	D1	E1	F1	G1	H1	I1	J1
A1	B1	C1	D1	E1	F1	G1	H1	I1	J1
A1	B1	C2	D1	E1	F1	G1	H1	I2	J1
A2	B2	C2	D1	E2	F1	G1	H1	I1	J2
A2	B2	C2	D2	E2	F1	G1	H2	I2	J2
A1	B1	C1	D1	E1	F1	G1	H1	I1	J1
A1	B1	C1	D1	E1	F1	G1	H1	I1	J1
A1	B2	C1	D1	E1	F1	G1	H2	I1	J1

Because the value of each factor in the original data set is continuous, the data have to be discretized. The obtained factor scores have four degrees. Taking the body factor as an example, that A1 expresses no somatization symptoms, A2 expresses mild, A3 expresses moderate, and A4 expresses severe somatization. The other factors are converted like this, part of the transformed data is shown in **Table 3**.

With the increase of support degree, the number of association rules is changed as shown in **Figure 1**.

With the increase of confidence, the number of association rules is changed as shown in **Figure 2**.

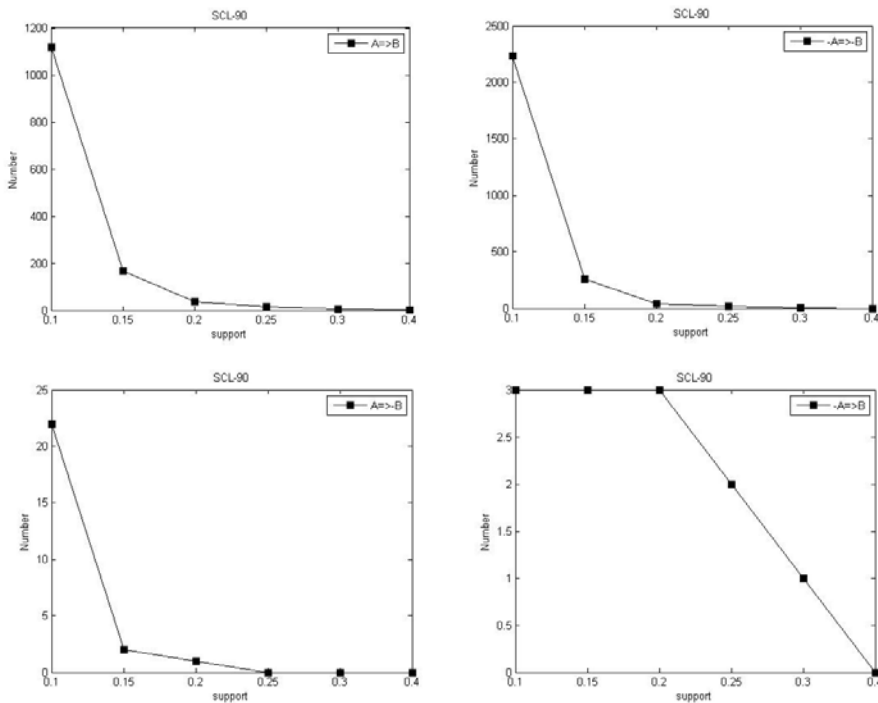


Figure 1. Number of association rules - trends in different supports

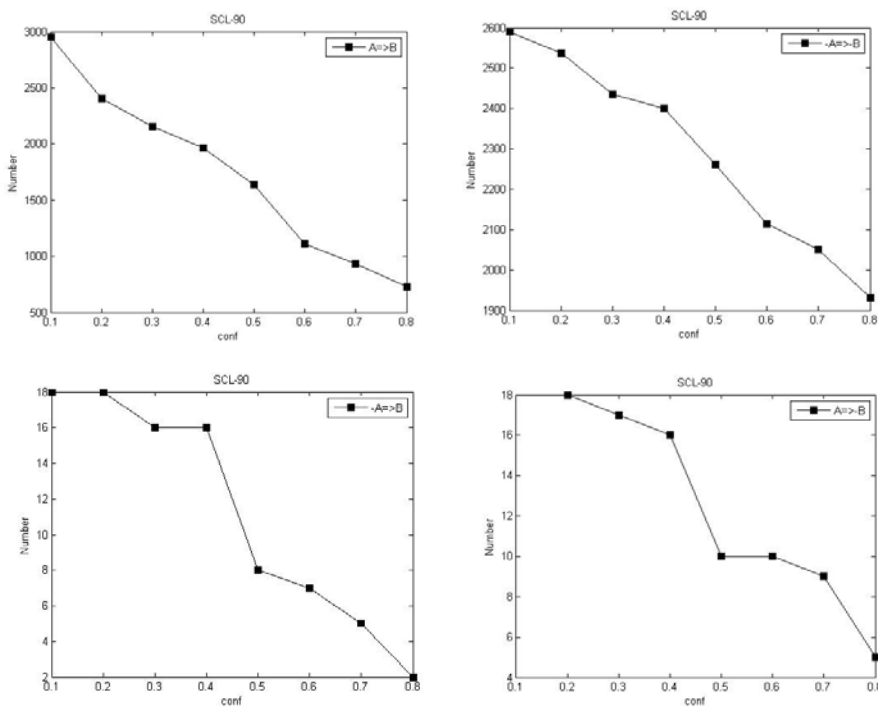


Figure 2. Number of association rules - trends in different confidences



**Table 4.** The support value of each factor

Item	Support
somatization	0.1267
Obsession	0.7600
interpersonal relationship sensitiveness	0.6180
depression	0.2635
anxious	0.3237
hostility	0.3093
Phobic Anxiety	0.2313
Paranoid	0.3744
Psychoticism	0.2418
Additional Items	0.2590

**Table 5.** The association rules(A⇒B)

Association rules(A⇒B)	Support	Confidence
interpersonal relationship sensitiveness=>Obsession	0.5032	0.8330
anxiety=>Obsession	0.4368	0.8076
depression=>anxiety	0.3236	0.8049
Obsession, depression=>interpersonal relationship sensitiveness	0.2676	0.8582
interpersonal relationship sensitiveness, depression=>Obsession	0.2676	0.8355
Obsession, anxiety=>interpersonal relationship sensitiveness	0.3602	0.8246
depression, hostility=>Paranoid	0.2092	0.8132
depression, anxiety=>Paranoid	0.2597	0.8025
interpersonal relationship sensitiveness, depression, anxiety=>Obsession	0.2232	0.8364
Obsession, depression, anxiety=>interpersonal relationship sensitiveness	0.2232	0.8802

In **Table 4** obsession and interpersonal relationship sensitiveness have higher supports mean two factors have higher frequencies. It shows that in the target population was tested. People with obsession and relationship sensitiveness interpersonal symptoms were up to.

**Table 5** shows the positive association rules between the factors. It can be seen that there is almost a correlation between each factor. Female is easy to suffer from male, and Obsession in dealing with interpersonal relationship is not as good as female. People with anxious, depression, Psychoticism are prone to suffer from Obsession. Relationship sensitiveness Psychoticism, interpersonal, Obsession between the three most closely linked. People with Psychoticism and the relationship between human and human sensitive symptoms are easy to suffer from Obsession. Psychological doctors in solving such mental health problems, you can focus on this relationship. Interpersonal sensitivity, anxiety and obsessive-compulsive disorder in the relationship between the most closely, and its confidence is more than 80%.

**Table 6** shows the association rules like  $\neg A \Rightarrow \neg B$ , which indicates support and confidence level when a mental state set is not normal, then the other mental state set is not normal. As is shown in the table, if the sample does not exist relationship sensitiveness interpersonal problem, then it may not be depression or anxious. If the sample isn't relationship sensitiveness interpersonal and not obsession, then it will not be hostility. If the sample isn't relationship sensitiveness interpersonal and not anxious, then it isn't depression.

**Table 6.** The association rules ( $\neg A \Rightarrow \neg B$ )

Association rules ( $\neg A \Rightarrow \neg B$ )	Support	Confidence
$\neg$ interpersonal relationship sensitiveness $\Rightarrow$ $\neg$ depression	0.3203	0.7751
$\neg$ interpersonal relationship sensitiveness $\Rightarrow$ $\neg$ anxious	0.4221	0.6733
$\neg$ (Obsession, interpersonal relationship sensitiveness) $\Rightarrow$ $\neg$ hostility	0.2823	0.6680
$\neg$ (interpersonal relationship sensitiveness, anxious) $\Rightarrow$ $\neg$ depression	0.2669	0.7661
$\neg$ (interpersonal relationship sensitiveness, depression) $\Rightarrow$ $\neg$ hostility	0.2112	0.6651
$\neg$ (interpersonal relationship sensitiveness, Paranoid) $\Rightarrow$ $\neg$ depression	0.2571	0.7468
$\neg$ (interpersonal relationship sensitiveness, Psychoticism) $\Rightarrow$ $\neg$ depression	0.2416	0.7496
$\neg$ (depression, hostility) $\Rightarrow$ $\neg$ Psychoticism	0.2053	0.6673
$\neg$ (anxious, hostility) $\Rightarrow$ $\neg$ Psychoticism	0.2361	0.6713

**Table 7.** The association rules ( $A \Rightarrow \neg B$ )

Association rules ( $A \Rightarrow \neg B$ )	Support	Confidence
female $\Rightarrow$ $\neg$ (Obsession, interpersonal relationship sensitiveness)	0.2313	0.5374
(female, Obsession) $\Rightarrow$ $\neg$ interpersonal relationship sensitiveness	0.2313	0.3936
female $\Rightarrow$ $\neg$ (Obsession, interpersonal relationship sensitiveness, Paranoid)	0.127	0.746
male $\Rightarrow$ $\neg$ anxious	0.147	0.7061
male $\Rightarrow$ $\neg$ (anxious, Paranoid)	0.1008	0.7985

There are some negative association rules as shown in the **Table 7**. From the table, it has been shown that female is not easy to suffer from obsession, relationship sensitiveness interpersonal, paranoid three kinds of symptoms. Male suffering from anxious and paranoid is also relatively low probability. Patients with obsession generally don't have depression. Depression and obsession have a very small chance of occurring simultaneously.

## DISCUSSION, CONCLUSION AND SUGGESTIONS

It has great practical significance to analyze and study the psychological problems of college students by using data mining knowledge. The obtained association rules have great practical significance to guide the students' thoughts and behaviors. In this paper, association rules mining algorithm is applied to analyze psychological relevance of college students, and then the potential valuable information is found. For college students' psychological correlation analysis, this paper use the SCL-90 symptoms self-rating scale to limit the related factors. At present, there is not much research on the application of data mining in the psychological analysis for college students. But the psychological problem of college students is directly related to the mental health, which also indirectly affect the physical health for college students. The PNARC association rules mining algorithm is combined with the psychological correlation analysis of college students to carry out information mining. Relevant experiments and analysis showed that there are correlations between the mental health factors. The discovery of association rules provides a basis for psychological counseling and evaluation of college students.

## ACKNOWLEDGEMENTS

This work was supported by the National Natural Science Foundation of China (71271125, 61502260) and Natural Science Foundation of Shandong Province, China (ZR2011FM028). Major project of soft science research program in Shandong Province (2015RZB01008)

## REFERENCES

- Burlaka, V., Churakova, I., Aavik, O. A., Staller, K. M., & Delva, J. (2014). Attitudes toward health-seeking behaviors of college students in ukraine. *International Journal of Mental Health & Addiction*, 12(5), 549-560. DOI:10.1007/s11469-014-9483-4

- Cheng, C. W., Martin, G. S., Wu, P. Y., & Wang, M. D. (2014). PHARM - Association Rule Mining for Predictive Health. *The International Conference on Health Informatics*. Springer International Publishing. DOI:10.1007/978-3-319-03005-0\_29
- Dong, X., Niu, Z., Shi, X., Zhang, X., & Zhu, D. (2007). Mining Both Positive and Negative Association Rules from Frequent and Infrequent Item sets. *International Conference on Advanced Data Mining and Applications*, 4632, 122-133. Springer-Verlag. DOI:10.1007/978-3-540-73871-8\_13
- Han, J. (2010). Fast mining of frequent item sets based on fp-tree. *Computer Applications & Software*.
- Han, J., Pei, J., & Yin, Y. (2000). Mining frequent patterns without candidate generation. *Acm Sigmod Record*, 29(2), 1-12. DOI:10.1145/342009.335372
- Herawan, T., Vitasari, P., & Abdullah, Z. (2012). Mining interesting association rules of students suffering study anxieties using slp-growth algorithm. *International Journal of Knowledge & Systems Science*, 3(2), 24-41. DOI:10.4018/jkss.2012040102
- Hodgins, D. C., Ranson, K. M. V., & Montpetit, C. R. (2015). Problem drinking, gambling and eating among undergraduate university students. What are the links? *International Journal of Mental Health & Addiction*, 1-19. DOI:10.1007/s11469-015-9598-2
- Jena, S., & Tiwari, H. (2015). Stress and mental health problems in 1st year medical students: a survey of two medical colleges in Kanpur, India. *New England Journal of Medicine*, 359(15), 228-232. DOI:10.5455/2320-6012.ijrms20150122
- Nembhard, D. A., Yip, K. K., & Stifter, C. A. (2012). Association rule mining in developmental psychology. *Data Mining Concepts Methodologies Tools & Applications*, 1(1), 23-37. DOI:10.4018/ijaie.2012010103
- Özyürek, P., & İbrahim Kılıç. (2015). The investigation of mental health problems and social phobia status of college students. *Journal of the Canadian Dental Association*, 79, d147-d147. DOI:10.15345/iojes.2015.04.016
- Qi, W., Yan, J., Huang, S., Guo, L., & Lu, R. (2013). The application of association rule mining in college students' mental health assessment system. *Journal of Hunan University of Technology*.
- Wang, D., & Psychology, S. O. (2015). Application of multidimensional association rules method in psychological measurement. *Intelligent Computer & Applications*.
- Zheng, G., Lan, X., Li, M., Ling, K., Lin, H., & Chen, L., Tao, J., Li, J., Zheng, X., Chen, B., & Fang, Q. (2014). The effectiveness of tai chi on the physical and psychological well-being of college students: a study protocol for a randomized controlled trial. *Trials*, 15(1), 1-9. DOI:10.1186/1745-6215-15-129

<http://www.ejmste.com>