Text Mining and Teaching Methods of College Teaching Resources Based on College Students' Mental Health

Hexia Yao[1,2,a,*], Elizabeth M. Sagubo[1,b]

[1]*College of Teacher Education, University of the Cordilleras, Baguio, Benguet, Philippines*
[2]*Student Affairs Department, Hebei University of Engineering, Handan, Hebei, China*
[a]*hexiayao73@163.com,* [b]*emsagubo@uc-bcf.edu.ph*
[*]*Corresponding author*

Abstract: College students' education and psychologic wholesome management is an important subject of ideological and political work in colleges and universities. The development of college students' psychologic wholesome needs to continuously strengthen and improve the education and management of college students' psychologic wholesome. After years of development, China's education system has deeply realized the influence and role of college students' psychologic wholesome education and management. This paper studied the text mining of teaching resources and teaching methods in colleges and universities, and explored new models of innovative management methods to solve current problems. This paper used various algorithms to investigate the content and teaching methods of text mining teaching resources under the psychologic wholesome of college students, and summarized and discussed the experiments. The research results show that the text mining model research and teaching methods established in this paper have improved the effectiveness of college students' psychologic wholesome by 9.31%.

1. Introduction

In recent years, with the intensification of social competition and the increase of employment pressure, the psychologic wholesome problems of college students are increasing day by day. In college education, college students' psychologic wholesome education course is a general course to popularize psychologic wholesome knowledge and improve students' psychological quality. The application of video teaching provides new ideas for college students' psychologic wholesome education class. Establishing an excellent and efficient management team is an important symbol of the current university's development capability and the basis for improving the university's basic competitiveness. Although in recent years the Ministry of Education has successively announced the management measures for college students' psychologic wholesome education and the policy of setting up students' psychologic wholesome education management groups, there are still many problems that need to be solved urgently in the implementation process.

China is entering the critical period of building a well-off society in an all-round way in the new century, and the psychologic wholesome education and management of college students has become a major problem that colleges and universities must solve. Based on this, many scholars have

conducted research on the psychologic wholesome of college students. Andersen R proposed that from an internal point of view, the psychologic wholesome problems of college students were individual psychological dysfunctions. The external manifestation is that the individual's social development is not good, and the psychologic wholesome of college students has a significant positive relationship with their sociality, and the two influence and restrict each other [1]. Baber M selected ordinary high school and college freshmen to conduct psychological characteristics survey, and analyzed the effective way to combine current psychologic wholesome education and political curriculum education according to the survey results [2]. According to Nesbitt A E, as an important part of the modern curriculum system, college students' psychologic wholesome education has played an important role in improving the physical structure of students and promoting the physical and psychologic wholesome of primary school students [3]. Ma Z R proposed three development policies based on the history and logic of the development of educational policies for political courses in secondary schools and universities [4]. Aiming at the shortcomings of traditional college students' psychologic wholesome early warning method with complex calculation and low accuracy, a psychologic wholesome early warning method based on SSA-ELM (Salp Swarm Algorithm-Extreme Learning Machine) was proposed by Medina A M [5]. Bas G believed that the network environment would have a significant impact on people's thinking and behavior, and the impact of the network environment on the physical and psychologic wholesome of college students was both positive and negative [6]. Wattick R A put forward corresponding countermeasures from the aspects of enhancing the public's awareness of public health emergencies, alleviating the negative emotions of college students, guiding college students to develop normal work and rest habits, and building a psychological care system of multi-party care and mutual assistance [7]. The psychologic wholesome of college students is hotly discussed in life, and also attracts attention in the academic circles.

In today's rapid economic development, the quality of high school teaching resources is improving day by day, and scholars pay more attention to the research on high school and university teaching resources and teaching methods. According to Yan C, data mining of educational texts was considered to have distinct advantages, the results of the study showed that the use of interdisciplinary research methods can effectively address the inherent problems of large volumes of Chinese texts, even in the fields of humanities and social sciences [8]. Bayrak T believed that under the general environment of quality education, the reorganization and integration of music teaching resources in high schools and universities should not only stimulate students' active learning ability, but also fully tap their potential, which was a basic premise for the realization of quality education [9]. Chao Y analyzed the background of ideological and political courses, proposed the creation of ideological and political courses, and dig deep into basic professional courses in ideological and political education based on hot events and special knowledge points [10]. Combined with the connotation of data mining technology, Liu H analyzed and discussed the teaching quality evaluation system in colleges and universities based on data mining [11]. Aiming at the current bottleneck problems in the teaching of ideological and political courses in colleges and universities, Bai J investigated from the aspects of systematicness, universality and operability, and proposed a general teaching method of "research + integration" of ideological and political course teaching [12]. Liu Z believed that in a limited classroom, with unlimited resources, scientifically and effectively extracting teaching resources inside and outside the classroom was one of the necessary teaching skills for any teacher [13]. Liu Y has fully discussed the theory and methods of art teaching, providing a new perspective and new experience for colleges and universities to carry out art practice teaching activities [14]. Based on the psychologic wholesome of college students, this paper studied the combination of text mining of college teaching resources and teaching methods.

To do a good job in psychologic wholesome education and management of college students, the relevant training of teachers is the top priority. Starting from the psychologic wholesome of college students, this paper built a good teaching environment, educational development and teacher-student communication, and on this basis, put forward feasible suggestions based on the research on college students' psychologic wholesome.

2. Text Mining of Teaching Resources in Colleges and Universities and Models of Teaching Methods

(1) Linear Support Vector Classifier

The support vector classifier uses the minimum value between the two data in the two spatial classes to determine the boundary of the class by searching for the optimal linear classifier based on the distribution of training samples. Different classifications of points in space are known as "generating vectors". In this way, a small number of training samples that are most effective for predicting behavior can be screened into large, high-dimensional data [15].

Given a set of training set data in the form:

$$D = \{(x_i, c_i) \mid x_i \in \Re^p, c_i \in \{-1,1\}\}(1 \le i \le n) \tag{1}$$

Now we need to find a hyperplane that can distinguish between data labeled 1 and -1. This hyperplane can be expressed as:

$$w \cdot x - b = 0 \tag{2}$$

w is the normal vector perpendicular to the hyperplane, and the parameter b represents the offset from the hyperplane along the normal vector. To maximize the gap, the gap between two parallel hyperplanes is maximized and different data types can be distinguished.

According to the data category representation method, the following formulas can be obtained to constrain these two parameters

$$\begin{cases} w_1 \cdot x - b = 1 \\ w_2 \cdot x - b = -1 \end{cases} \tag{3}$$

Because the training set data is linearly separable, there are no data points in the two selected hyperplanes, and the distance between them is the largest. From a geometrical point of view, the distance between these two hyperplanes is $\dfrac{2}{\|w\|}$, so w is the smallest. In order to avoid data points exceeding the range, the following constraints are imposed on each data vector x, that is, the deformation of formula (3).

$$\begin{cases} w_1 \cdot x_i - b \ge 1 \\ w_2 \cdot x_i - b \le -1 \end{cases} \tag{4}$$

For simplicity, formula (4) can be written as

$$c_i(w \cdot x_i - b) \ge 1, (1 \le i \le n) \tag{5}$$

(2) Naive Bayes

Naive Bayesian classification is a classification method based on Bayes' theorem [16]. The core content of the method is to use the known conditional probability and prior probability to classify the text to be classified, and classify it into a category with a high posterior probability. At the same

time, assuming that the occurrence probability of all texts is independent of the occurrence probability of each feature element, the calculation formula is as follows.

$$P(D|C_i) = \prod p(W_k|C_j)$$
(6)

The existence of this assumption effectively reduces the computational complexity of the Naive Bayes algorithm, but also reduces the classification accuracy. Suppose there are N categories of texts in the corpus, which are represented as $C_1, C_2..., C_i$; at the same time, D is any text, which is represented as a vector $D = \{W_1, W_2,..., W_k\}$. According to Yes's theorem, the posterior probability P(Ci|D) of the C_n class can be expressed as the following formula.

$$P(C_i|D) = \frac{P(c_i)P(D|C_I)}{P(D)}$$
(7)

Since P(D) is a constant for any class, the formula can be simplified to

$$P(C_i|D) = P(C_i)P(D|C_i)$$
(8)

$P(C_n)$ is easily obtained and equals the number of texts in class $P(C_n)$. The key of this algorithm is how to obtain $P(D|C_i)$, which is obtained by the following formula under the assumption that the elements of the feature are independent of each other.

$$P(D|C_i) = \prod_{j=1}^{k} P(W_i|C_i)$$
(9)

Among them, $P(W_j|C_i)$ represents the probability that the text containing the feature word w_m belongs to the class Ci. Arranging the calculation formula of $P(C_i|D)$, the following results are obtained.

$$P(C_i|D) = P(C_i)\prod_{j=1}^{k} P(W_j|C_i)$$
(10)

According to the calculated posterior probability result $P(C_i|D)$, the category to which the text D belongs can be determined, and the text D belongs to the category with the highest posterior probability. The algorithm has a high classification rate and is simple to implement. However, due to the independent calculation of text feature elements, the application scenarios are limited.

(3) Classification evaluation indicators

At the end of the text classification process, it is necessary to evaluate the performance of its classification. Usually, the evaluation indicators used to measure the quality of a sorting algorithm are as follows: Precision, Recall, F1 value and macro F1 value [17].

1) Accuracy Precise

The accuracy rate (precise rate) refers to the proportion of correctly classified texts among the texts identified as a certain category in the test set C_n, and the calculation formula is as follows.

$$P = \frac{TP}{TP + FP}$$
(11)

2) Recall rate

Recall rate refers to the ratio of the number of correctly classified texts in the texts identified as a

certain type of C_n in the test set to the actual number of texts belonging to the C_i type in the test set, and its calculation formula is as follows.

$$R = \frac{TP}{TP + FN}$$

(12)

The higher the precision and recall, the better the results, but in reality, there is some connection between the two methods. If the accuracy rate is high, it would decrease, and if the summon rate is high, it would decrease. In this context, the most common approach is to use the F-measure as a measure of the relationship.

3) F1 value

The F1 value is the most commonly used form of F-measure, and its mathematical expression is as follows.

$$F1 = \frac{2 \times P \times R}{P + R}$$

(13)

When the F1 value is higher, the classification algorithm in the experiment is more effective.

4) Macro F1 value

The macro F1 value is the average value of F1 of all categories, and its mathematical expressions are as follows.

$$MacroF1 = \frac{\sum_{k=1}^{n} F1k}{k}$$

(14)

$$MacroF- = \frac{\sum_{k=1}^{n} F-k}{k}$$

(15)

This paper mainly uses the F1 value and the macro F1 value to evaluate the classifier effect.

3. College Teaching Resources Text Mining and Teaching Method Factors of College Students' Psychologic Wholesome
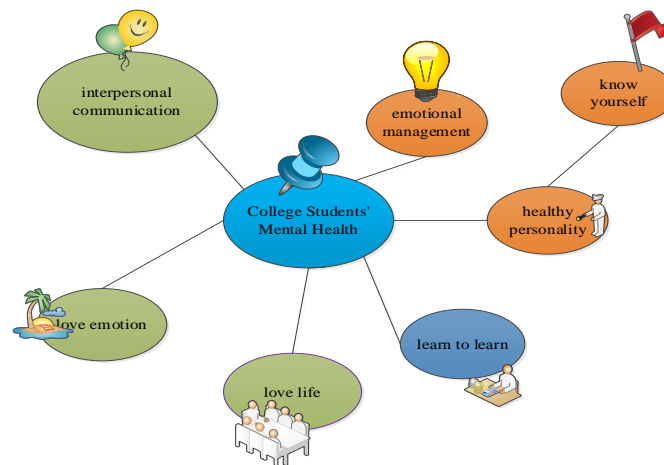

Figure 1: Analysis of the psychologic wholesome factors of college students

(1) Analysis of psychologic wholesome factors of college students

College student psychologic wholesome is an academic term in psychology. This means that the psychology of college students has many characteristics of middle-aged and young people, but because they are a special group, it is impossible for college students to be completely consistent with the youth in society [18]. Psychologic wholesome is usually measured on scales, and the standards are not fixed. Psychologic wholesome standards vary over time and cultural context, as shown in Figure 1:

From the analysis of college students' psychologic wholesome factors, the following 7 points are drawn:

1) Learning to learn: great pressure to study, lack of motivation to study, poor academic performance and academic problems are the four problems that have plagued college students for a long time.

2) Emotional management: Stable emotions and positive emotional reflection are important factors for students to become talents, and they are also issues worthy of attention in students' psychologic wholesome. Negative emotions of college students are greater than positive emotions, which deserves attention.

3) Interpersonal communication: Good interpersonal relationships play an important role in the growth and socialization of college students, and it is of great significance to maintain a good attitude.

4) Anxiety problem: The anxiety of college students is representative to some extent. Its root cause is not from the real threat, but from the inside, without a clear object and specific content. It is dominated by test-taking anxiety and individual anxiety.

5) Love emotions: Love, friendship and family are the three major elements of students' emotions.

6) Knowing the Self: The cognitive part of self-awareness mainly answers the question "Who am I?" It is the main component of self-perception and the psychological basis for controlling self-regulation, including self-perception, self-concept, self-observation, self-analysis and self-evaluation.

7) Learning to learn: It is an important part of college life for students to adapt to college life, realize the educational mission of transforming college students from "cultural person" to "social person", and socialize college students.
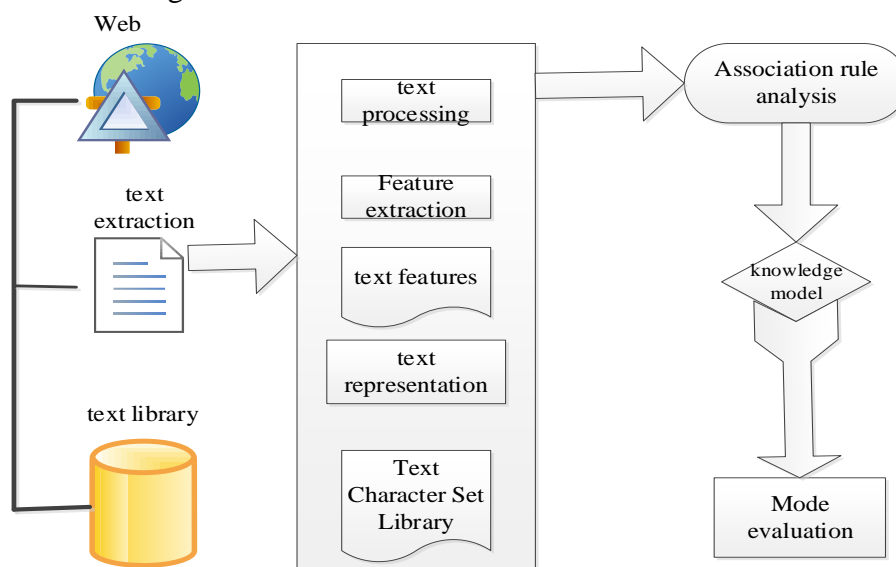

Figure 2: Web text data mining process

(2) Web-based text data mining model

In Web text mining, text feature-based mining is the foundation of data mining [19]. The use of document language to describe Web content is a natural language that is difficult for computers to understand, so it cannot be directly applied to existing data mining techniques. This requires preprocessing the text, extracting metadata that can represent its characteristics, and using it as an intermediate representation, and storing it as a structure, as shown in Figure 2:

Although the goals of Web-based text data mining may vary, the overall process can be described by the above method.

Text extraction: From the text data collected on the Internet, all text data related to the current operation can be extracted and stored in the database.

Word processing: Complete part-of-speech encoding, stop word deletion, word segmentation, etc., and obtain all text items in the text.

Feature extraction and text expression: Since the target of Web text data mining is often HTML (HyperText Mark-up Language) files, which lack data rules, document feature vectors are often used to achieve it. The model can reflect the content of the document, but in the current document expression method, the size of the feature vector is very large. Therefore, when mining Web text, a subset of features can be selected.

Text mining: After compressing the document feature vector dimension, it is mined by various data mining techniques such as association rule analysis.

Pattern evaluation: The mining results can be evaluated and output when they meet the requirements. On the contrary, it can go back to the previous stage, and after analysis and improvement, the next stage of mining can be carried out.

(3) Analysis of factors of psychologic wholesome teaching methods

Currently, psychological counseling in schools has become a hot topic [20]. School psychological counseling refers to the use of psychological theories and methods by school counselors to provide direct or indirect counseling for students' learning, adaptation, development, career choices and other psychological problems, as shown in Figure 3:
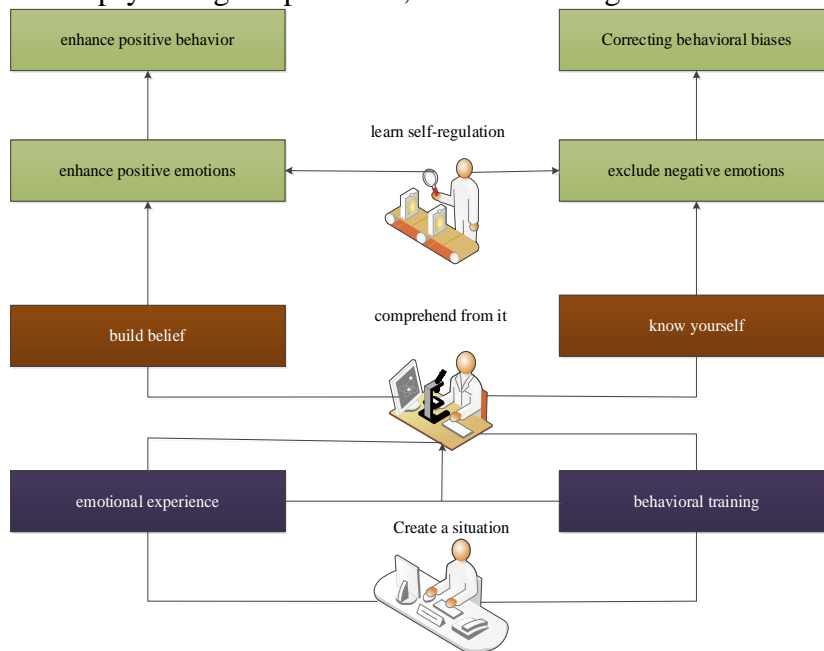

Figure 3: Factors analysis of psychologic wholesome teaching methods

The purpose of this study is to construct an internal mechanism of an intrinsic psychologic wholesome effect. In this model, there are also differences in the psychologic wholesome of people with different views. Ontological theory holds that an individual's personality traits do not change and are more likely to have negative emotional, behavioral, and physiological responses. Progressives believe that personality is malleable and can change and develop over time, making it easier to develop positive emotional, behavioral, and physical responses. Substantialism and gradualism can not only have a direct impact on an individual's emotions, behaviors, and physiological responses, but also regulate and control the individual's personality and motivational mechanisms, attribution methods, and cognitive mechanisms for beliefs.

This model shows that the functional mechanism of implicit theory is diverse and complex. These intermediaries themselves are independent of each other, but also penetrate each other. For example, success goals can explain both the role of implicit theory in psychologic wholesomeand the different attributions of individuals. At the same time, the personal factor cannot be regulated as an independent mechanism, it is also regulated by the related factors in the situation.

4. Text Mining of Teaching Resources and Teaching Methods Based on Psychologic Wholesome

(1) Self-assessment of psychologic wholesome

College students' depression is a common emotional and psychological disease, which can lead to interpersonal tension, emotional loss, etc. It is necessary to study its clinical manifestations, as shown in Figure 4:
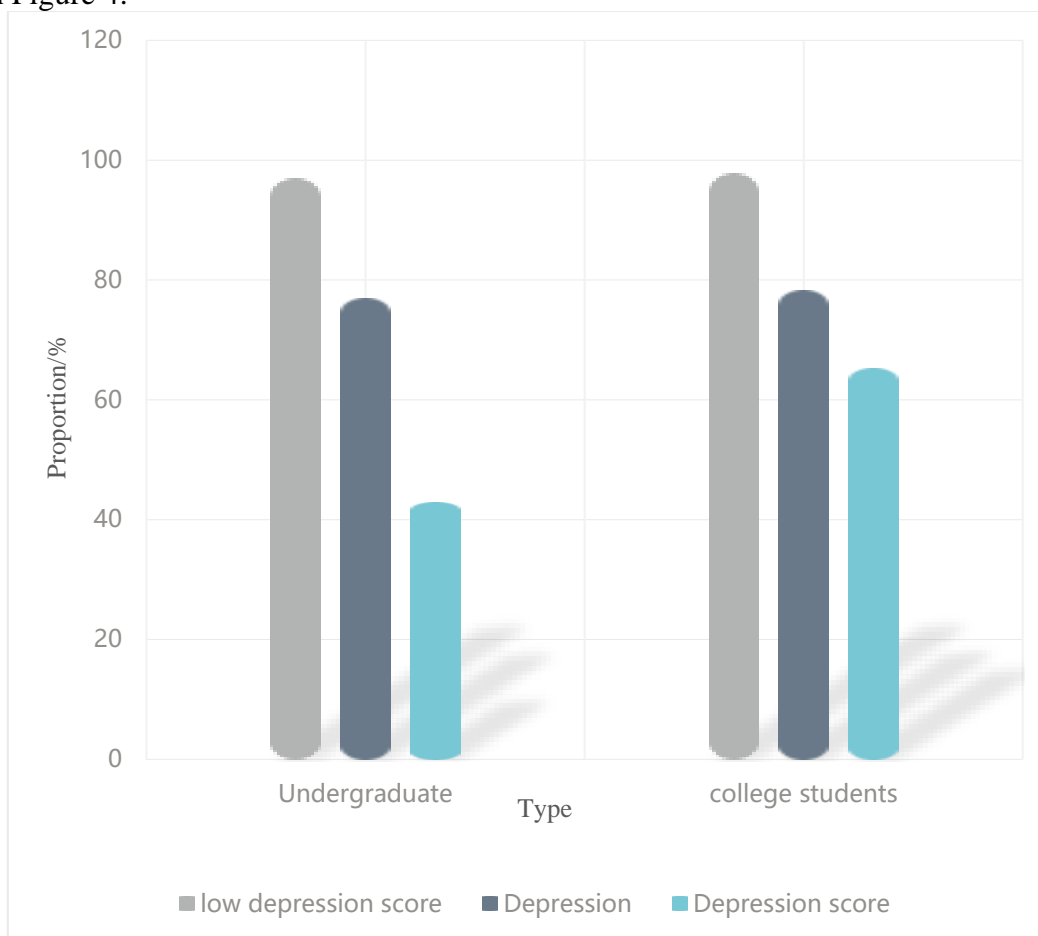

Figure 4: Analysis of depression scores between undergraduate students and college students

As can be seen from the above figure, among those with low scores on the depression scale, namely, no depression tendency, 96.5% of the college students believed that their psychologic wholesome was normal, and 97.7% of the college students believed that their psychologic wholesome was good. Good normal health reflects a more accurate self-assessment. Among those with a high score on the Depression Scale, that is, a high probability of depression, the proportion of those who considered themselves mentally healthy dropped significantly to 41.9% (undergraduate students) and 66.3% (college students). This reflects a tendency to overestimate one's own psychologic wholesome.

Table 1 shows the statistical distribution of the scores of depression tendency in psychologic wholesome among undergraduates, college students, men and women, and adults.

Table 1: Distribution of college students in each fraction segment of the Depression Scale

object of investigation	No tendency to depression (0-9 points)	Depressed tendencies (10-16 points)	High risk of depression (17-27 points)
Undergraduate boys	72.6%	20.8%	6.6%
Undergraduate girls	70%	22.4%	7.6%
college boys	90.2%	8.2%	1.6%
college girls	89.3%	8.8%	1.9%
National Adult	78.7%	16.7%	4.5%

It can be seen from the above table: according to the type of student source, undergraduates: 72, 6% of the male undergraduates have no tendency to depression. 20.8% of male undergraduate students scored 10 to 16 points (depressive tendencies), and 6.6% of male undergraduate students scored 17 points and above (high risk of depression). 70% of undergraduate girls had no depression tendency, 22.4% of undergraduate girls scored 10-16 points (depressive tendency), and 7.6% of undergraduate girls scored 17 points or above (high risk of depression). College students: 90.2% of college students had no depression tendency, 8.2% of college students scored 10-16 points (depressive tendency), and 1.6% of college students scored 17 points or above (high risk of depression). 89.3% of college students had no tendency to depression, 8.8% of college students scored 10-16 points (depressive tendency), and 1.9% of college students scored 17 points or above (high risk of depression).

(2) Text precision test

The precision rate in the classification evaluation index algorithm is used to further mine the text mining of college teaching resources, and the precision rate test of the two models is carried out. The higher the precision, the better the performance of the educational resource text classification model. The accuracy test was carried out on the use of the classification evaluation index algorithm to construct the education sub-language text data mining classification model and the traditional text data mining classification model, and the differences between the two models were observed, as shown in Figure 5.

It can be seen from Figure 5 that the accuracy rates of various text models are different, but the accuracy of the classification evaluation index algorithm to construct the educational text data processing classification model is higher than that of the traditional text data processing classification model. The accuracy rate is the lowest, and the F class has the highest accuracy rate. The higher the accuracy, the stronger the search ability of the text model. The more data about the category in the text model, the stronger the text model would be. In summary, the classification evaluation index algorithm can improve the accuracy of text mining classification models.
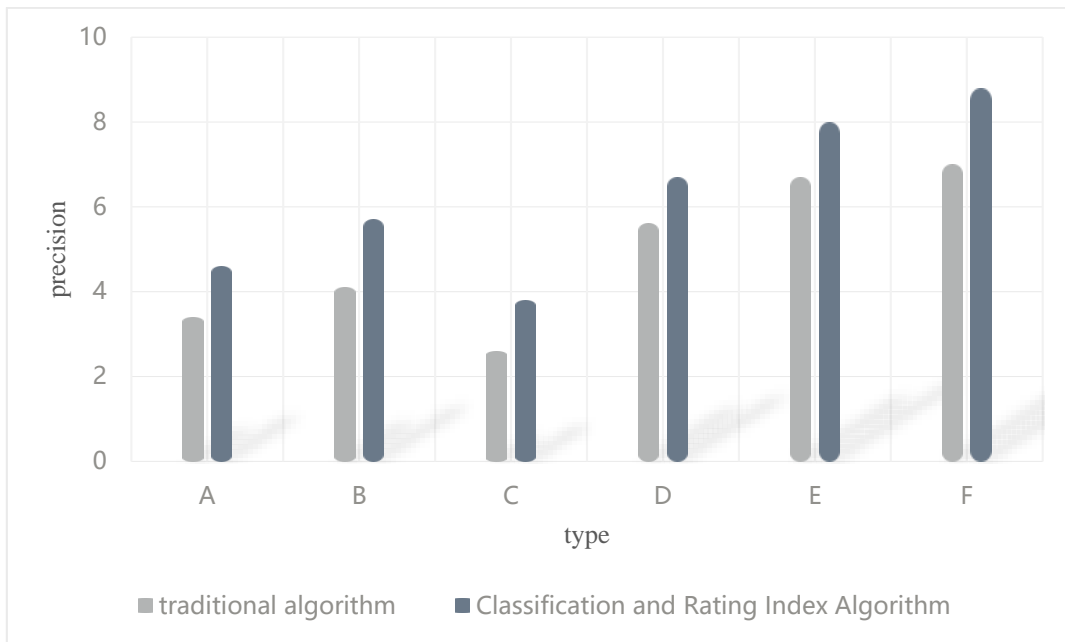
Figure 5: Classification accuracy test of text data mining of educational resources

(3) Text recall test

The recall rate in the classification evaluation index algorithm can be used to further mine the text mining of college teaching resources. The more data mined by the data mining technology, the higher the recall rate of the model and the better the text classification model. The recall test can be performed on different numbers of educational resource text classification models to observe the differences in the test results. The test data are 100, 200, 400, 800, and 1600, respectively. The test results are shown in Figure 6.
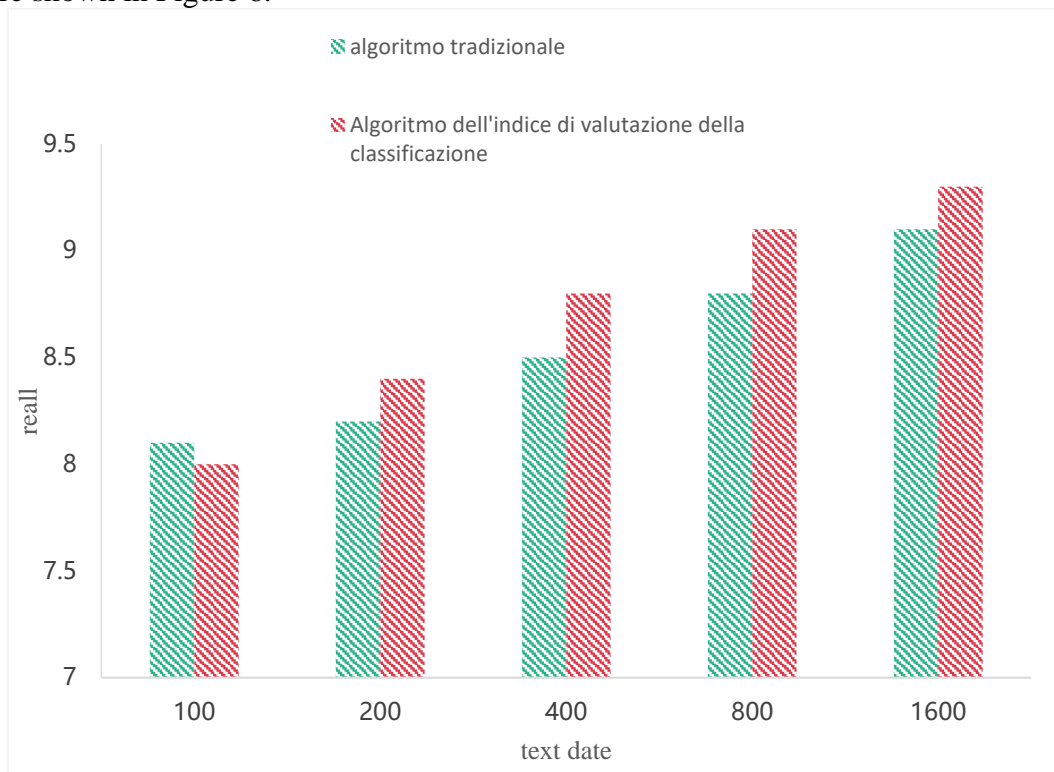

Figure 6: Educational resource text data mining classification recall test

As can be seen from the above figure, with the increase in the text data of educational resources, the recall rate of the text data mining classification model is also increasing. The growth rate of the classification evaluation index algorithm text data mining classification model is much faster than that of the traditional algorithm text data mining classification model. Among them, when the text data increases from 100 to 200, the model recall rate increases slowly. When the text data increases from 200 to 400, the model recall rate increases rapidly, especially when the classification model of educational text data mining is constructed based on the classification evaluation index algorithm. The classification evaluation index algorithm can make the model better analyze the educational text, so that the data mining can classify it reasonably, so that the more educational resource text data, the higher the recall rate of the text classification model.

(4) Analysis of college students' satisfaction with the teaching methods of psychologic wholesome courses

Satisfaction is a state of mind that reflects an individual's subjective assessment of the quality of a relationship. It refers to the pleasure generated after the customer's needs are satisfied, that is, the expected and actual feeling of the customer when using the product or service. If these emotions are measured with data, the data is called "satisfaction", and customer satisfaction is the basis of customer loyalty. Teaching satisfaction is the degree to which students' actual perceptions of education and teaching services compare with their expectations of education and teaching services. Figure 7 shows the satisfaction of college students with the way psychologic wholesome courses are taught:

Figure 7: Comparison of the satisfaction of college students in schools A and B with the teaching methods of psychologic wholesome courses

This paper analyzes the satisfaction of undergraduates in school A with the teaching methods of psychologic wholesome. The survey shows that the average total satisfaction of the undergraduates in School A to teaching reaches the average. Among them, 14.51% of college students chose "very satisfied", and 38.57% of them chose "relatively satisfied". 41.03% of college students chose "average", 4.13% of college students chose "not very satisfied", and 1.76% of college students chose "very dissatisfied". It can be seen that nearly 50% of college students have a degree of satisfaction with their studies that is "relatively satisfied" or above, excluding nearly 94.11% of college students who are very dissatisfied. It is not difficult to see from the above figure that School B is more satisfied with this teaching method. This shows that the teaching of this course has been

recognized by the students of School A. The model research of the constructed text mining and teaching methods improves the teaching effectiveness of college students' psychologic wholesome by 9.31%.

5. Conclusion

With the implementation of the national education reform plan, higher education outreach goals are being advanced in an orderly manner, the number of college students is increasing, and the number of students with psychologic wholesome problems is increasing. In addition, the current research also has the problems that the research system is not systematic enough, the research content is empty and the research object is marginalized. Faced with this situation, this paper uses the classification evaluation index model to study the text mining of college teaching resources and teaching methods for college students' psychologic wholesome, and through experimental research proves that it can fully improve the effectiveness of college teaching for college students' psychologic wholesome.

References

[1] Andersen R, Holm A. The student mental health crisis: Assessing psychiatric and developmental explanatory models. Journal of Adolescence, 2021, 86(7):101-114.

[2] Baber M, Bate W. Student mental health – a public health challenge? Perspectives in Public Health, 2021, 141(3):127-128.

[3] Nesbitt A E, Collins K J, Nalder E. Occupational Outcomes of a Physical Activity Intervention for Post-Secondary Student Mental Health: Canadian Journal of Occupational Therapy, 2021, 88(3):254-265.

[4] Ma Z R, Ma W H, Idris S. COVID-19 impact on high school student's education and mental health: A cohort survey in China. World Journal of Psychiatry, 2021, 11(6):10-17.

[5] Medina A M, Mead J. MS-Speech-Language Pathology Student Mental Health: Establishing a Virtual Safety Net during COVID-19. Journal of Indian Speech Language & Hearing Association, 2021, 35(1):22-25.

[6] Bas G. Relation between Student Mental Health and Academic Achievement Revisited: A Meta-Analysis. Journal of Physics: Conference Series, 2021, 2(9):4-10.

[7] Wattick R A, Hagedorn R L, MD Olfert. Impact of resilience on college student mental health during COVID-19. Journal of American College Health, 2021, 7(6):1-8.

[8] Yan C. Innovative Research on German Education Teaching Mode in Colleges and Universities from the Perspective of BD. Journal of Physics: Conference Series, 2021, 1852(3):6-32.

[9] Bayrak T. A content analysis of top-ranked universities' mission statements from five global regions. International Journal of Educational Development, 2020, 2(7):102-130.

[10] Chao Y, Guo X, Wang R. Exploration of Teaching and Education: Integration of Sports and Medicine in Physical Education in Colleges and Universities. Contemporary Education Research, 2022, 6(1):63-67.

[11] Liu H. Application and Strengthening Strategies of Network Resources in the Construction of Teaching Platform. Journal of Physics: Conference Series, 2021, 1915(4):42-56.

[12] Bai J. A Brief Talk on the Computer Multimedia Network as a Powerful Aid to the Teaching of Guzheng in Colleges and Universities. Journal of Physics Conference Series, 2020, 15(33):16-22.

[13] Liu Z. Discussion on Network Classroom Teaching of the Meteorology and Climatology in Common Colleges and Universities. Meteorological and Environmental Research: English Edition, 2020, 11(5):2-3.

[14] Liu Y, Li J. Study on the application and teaching strategy of multimedia technology in music teaching in colleges and universities. Journal of Intelligent and Fuzzy Systems, 2021, 7(1):1-8.

[15] Yang Feng. Application Research of Text Classification Based on Linear Support Vector Machine. Information Technology and Informatization, 2020, 42(3):3-4.

[16] Zhu Baojun, Xian Shili, Fan Huifang. Research on Transformer Fault Diagnosis Technology Based on WRSR and Improved Naive Bayesian Fusion. Power System Protection and Control, 2021, 49(20):9-32.

[17] Niu Guiqin, Chen Xiaoping. Construction of a classification evaluation index system for young scientific and technological talents. Future and Development, 2020, 44(9):8-17.

[18] Xie Jun. Analysis of mental health factors of college students. Psychological Exploration, 2019, 21(1):5-6.

[19] Li Rui, Huang Yukun, Dong Lutong. Research and application of machine learning algorithm based on multi-text data mining of power grid regulation business. Electronic Design Engineering, 2021, 29(9):51-62.

[20] Qin Jing. Research on the optimization of online teaching evaluation indicators of mental health education courses under the background of curriculum ideology and politics—taking Shanghai Electronic Information Vocational and Technical College as an example. Educational Research, 2021, 4(9):107-109.