Research Article

Zhang He*

Prediction mechanism of depression tendency among college students under computer intelligent systems

https://doi.org/10.1515/jisys-2023-0209
received October 25, 2023; accepted December 13, 2023

Abstract: In response to the existing problems in the prediction accuracy, data collection, real-time monitoring, and consideration of factors leading to depression in the current mechanism for predicting depression among college students, this article used computer intelligent systems to study the prediction mechanism of depression among college students. This article conducted a survey on students at University A using a survey questionnaire to understand the main reasons that affect their tendency to develop depression. It processed and analyzed the data using the Beck Depression Scale and Statistical Product and Service Solution 21.0 (SPSS 21.0). Meanwhile, natural language processing techniques in computer intelligence systems can be utilized. This article combines emotional dictionaries and word frequency-inverse document frequency to construct a prediction mechanism model for depression tendencies among college students, improving the accuracy of predicting student depression tendencies. The experiment shows that the average accuracy of the depression tendency prediction mechanism model constructed based on Natural Language Processing technology after 50 experiments was 97.02%, which was 5.33% higher than the model constructed based on neural network calculations. Overall, research on the prediction mechanism of depression tendency among college students based on computer intelligence systems can provide more effective mental health support and intervention measures for schools, helping students improve their psychological state, academic achievement, and quality of life.

Keywords: depression tendencies, predictive mechanisms, computer intelligent systems, natural language processing, emotional dictionary, neural network

1 Introduction

In the twenty-first century, with the acceleration of the social rhythm, people have to face many hardships. The incidence rate of depression is also rising year by year, and college students are high-risk groups for depression. Academic stress in college students is an important cause of depression. Many students face expectations from school, family, and society. These expectations may exceed their ability to bear, leading to emotional stress and depression; second, communication problems with parents and classmates are also an important factor in student depression. The increasing incidence of depression among college students not only affects their learning and quality of life but also leads to the occurrence of extreme behaviors. Many scientific studies have confirmed that depression patients are more likely to engage in violent behavior, with the majority of depression patients engaging in suicidal behavior, and about half of all suicide cases are caused by depression patients. The number of college students suffering from depression accounts for about 20% of the total

*** Corresponding author: Zhang He,** School of Marxism, Suzhou Vocational University, Suzhou, 215001, China,
e-mail: zhanghe@jssvc.edu.cn

number. Therefore, it is necessary and urgent to conduct research on depression tendencies among college students. However, existing methods for predicting depression tendencies among college students have problems with low prediction accuracy and inability to conduct real-time monitoring. Therefore, this paper uses computer intelligent systems to study the prediction mechanism of depression tendency among college students, accurately predicting and intervening in a timely manner for this depression group in universities. This can effectively prevent students from transitioning from depressive tendencies to depression, hoping to prevent tragedies and provide a scientific basis for research and intervention in university psychology.

Depression is a mental illness that often manifests as persistent symptoms such as low mood, irritability, insomnia, poor appetite, fatigue, and inferiority complex [1,2]. Depression tendency is the beginning of depression, exhibiting certain symptoms of depression but not yet meeting the diagnostic requirements for depression [3,4]. To investigate the role of social skills in depression among college students through the mediation of loneliness, Fauziyyah and Ampuni [5] and Ozimek and Bierhoff [6] used the Social Skills Scale to measure social skills and the Beck Depression Scale II to measure depression and conducted regression analysis using a simple mediation model [5,6]. Zhengjie et al. believed that depression was one of the most common mental illnesses in the world. Its characteristics are sustained low emotions, slow thinking, and decreased energy. In recent years, predicting the risk of depression through physiological techniques has become a hot topic in the field of neuroscience [7,8]. Yano et al. believed that sensory-processing sensitivity (SPS) is a genetically determined trait characterized by sensitivity and responsiveness to environmental stimuli. Therefore, they conducted a study on the relationship between life skills based on individual differences in SPS and depressive tendencies [9]. Qin et al. conducted a study on the correlation between depression and depressive symptoms in the adult population of China. They found that women, elderly people, and people living in the Midwest and rural areas are more prone to depression [10,11]. Nurwulan and Selamaj studied the impact of improving students' dual roles as students and workers on the depression rate of working college students and found that working during the study period does not necessarily affect students' academic performance [12]. In summary, scholars have conducted in-depth research on depression tendencies and student depression, but there is not much research on depression prediction. It has problems such as low prediction accuracy and incomplete data collection, and computer intelligent systems have great advantages in improving prediction accuracy.

Research based on computer intelligence systems refers to the use of technologies such as artificial intelligence and machine learning. It endows computers with intelligent abilities similar to or even surpassing those of humans to complete various complex tasks and problem-solving and has significant advantages in improving prediction accuracy. He et al. proposed an end-to-end trainable intelligent system that can effectively simulate potential depression patterns and achieve better performance in most video-based depression recognition methods [13]. Bohu et al. [14] and Pimenov et al. [15] proposed the meaning of modeling and simulation technology for new artificial intelligence systems. They explored the impact of new artificial intelligence system simulations on modeling and simulation techniques, as well as the ability to improve the accuracy of system predictions [14,15]. The principle of system intelligent prediction proposed by Zhao [16] and Ye et al. [17] is to perform reasonable structured clustering and prediction modeling based on existing data in the system. Li et al. [18] and Ayon et al. [19] proposed an asynchronous joint optimization algorithm for convolutional neural networks (NNs) based on joint learning to improve communication costs and convergence rate. The results indicate that the proposed method can effectively protect user privacy while ensuring prediction accuracy [18,19]. Qureshi et al. [20] and Yang et al. [21] proposed a new deep NN model based on attention for multitasking learning. It promotes the fusion of various patterns and uses this network to regress and classify the level of depression, improving the accuracy of predicting depression [20,21]. In summary, computer intelligence systems have a wide range of applications and use a wide range of technologies, which can greatly help improve the accuracy of system predictions. It can be applied to the research on the prediction mechanism of depression tendency in university learning, which can effectively solve the problem of inaccurate prediction and improve the accuracy of prediction.

With the implementation of expanded enrollment policies in higher education institutions, many schools have experienced the phenomenon of relaxed admission and strict withdrawal, which has led to the emergence of some negative phenomena and brought great difficulties to student education and management. To address the management issues of students with depression or depressive tendencies, a comprehensive analysis of the causes and manifestations of depression is necessary. Depression generally has endogenous pathogenic factors, that is, physiological or hereditary causes, which are usually the consequences of changes in the body's biochemical factors. This kind of depression has certain psychotic symptoms. Patients are unable to respond appropriately to the individual's internal and external environment, and they may exhibit unconventional behaviors that are already very abnormal. Therefore, this article first conducts a survey of University A students through a survey questionnaire. This can help understand the reasons that lead to students' tendency towards depression. It conducted a specific survey on 3,620 students. This article studies the indicators that affect depressive tendencies; next, it combines natural language processing technology in computer intelligent systems with emotion dictionaries and word frequency-inverse document frequency. This can construct a prediction mechanism model for depression tendency among college students, helping schools guide students to adapt to school life and grow healthily using scientific and effective methods.

The novelty of the research on the prediction mechanism of depression tendency of college students based on computer Intelligent Systems is mainly reflected in the following aspects:

1. Data-driven predictive model: Through computer intelligent systems technology, a large amount of college student data can be collected, including social media data, academic performance, physiological indicators, etc., and the predictive model can be trained through machine learning algorithms to accurately predict students' depressive tendencies.
2. Personalized treatment plan: Through computer intelligent systems technology, a personalized treatment plan can be formulated according to the specific situation of each student to improve the treatment effect, and at the same time, it can better meet the needs of students.

2 Data on the prediction mechanism of depression tendency among college students

Depression is often manifested as an emotional and psychiatric disorder, which is a clinical symptom based on personal experiences of depressive emotional states [22,23]. Depression is a symptom of pathological depression or emotional depression. According to research, depression can be divided into three types: emotional depression, neurological depression, and psychological depression. A strict definition of depression is undoubtedly beneficial for the diagnosis of depression. At the same time, it often overlooks those who suffer from depression but do not reach the essence, which is the so-called depression-prone population. Depression tendency is between the treatment of normal depression and depression neurosis. This group does not have existing diagnostic criteria as pathological. If they miss the opportunity for active intervention, their physical and psychological discomfort also has an impact on the lives of the corresponding group, causing the group to suffer from depression and have nowhere to go. More importantly, mild depression, moderate depression, and severe depression are different stages of emotional change from good to bad. A person's emotions can go from bad to good or from good to bad. Depression patients' emotions are weak, persistent, and diffuse [24,25]. This emotional way of thinking can make people think that things around them also feel the same way, which can lead to a more difficult emotional state. This can lead to depression-prone individuals gradually becoming depression patients and may even become high-risk groups for depression.

2.1 Reasons for depression tendencies among college students

With the changes in society, social competition has become increasingly fierce, which has had a very negative impact on the mental health of college students. For example, because they are not sure about whether they can find a job with higher income and social status in the future, they have no sense of control over future events, which leads to the increased psychological pressure on some students. This also makes people pay close attention to the mental health of college students, and depression is a very common negative emotion [26,27].

There are many reasons that cause depression among college students. A survey was conducted on 1,000 students from University A, and two main causes of depression were identified (multiple choices are available). The specific investigation is shown in Table 1.

Table 1: Causes of depression among college students

Serial number	Cause classification	Explanation of meaning	Number of people
1	Personal reasons	Personal personality, self-care ability, and personal emotional regulation ability are all closely related to whether depression can occur	300
2	Family reasons	Problems such as the discordant family atmosphere, sudden and major family changes, and difficulties in maintaining family livelihoods have increased the psychological pressure and psychological burden on students in the past	220
3	Social reasons	The development of the economy has improved people's living standards, but it has also caused many undesirable social problems, among which social indifference is a major killer of depression among college students	100
4	Network reasons	As a group of Internet users, college students may receive some bad information, which in turn produces negative emotions and cause psychological depression	140
5	Physiological reasons	People's condition manifests as depression, which is mainly because of a disorder in the chemicals released by their brain nerves. This disorder is related to human genes	100
6	Psychological reasons	The person is noticed and sympathized because of depression, which satisfies the individual's need to be noticed and makes up for some of the individual's emotional needs, so the state of depression will be strengthened	250
7	Economic difficulties	Students may need to independently bear financial responsibilities such as tuition and living expenses. For college students with financial troubles, financial pressure may cause a mental burden	130
8	Academic pressure	Higher education often has high requirements for students, and these pressures may bring frustration and depression to students	150
9	Interpersonal reasons	College students are facing new social environments and interpersonal challenges, and these difficulties may make them feel tired, anxious, and depressed	165

As shown in Table 1, a survey was conducted on 1,000 selected students, and it was found that they have nine main causes of depression among college students. Among them, the number of students who choose personal and psychological reasons is the highest, and most students believe that a person's depression is mainly due to internal factors. For external factors, students believe that family reasons are the main external cause of depression, as well as interpersonal and academic pressures, which can also lead to depression in students.

2.2 Experimental materials

2.2.1 Beck depression inventory (BDI)

BDI is an indicator evaluation scale that measures depressive symptoms and emotions in self-reports over the past 2 weeks [28]. The scale has a total of 21 items, with different selection scores for each item. The total score for the 21 items would be between 0 and 63 points. The total score was no depression in the range of 0–13, mild depression from 14 to 19, moderate depression from 20 to 28, and severe depression from 29 to 63.

2.2.2 Self-rating depression scale (SDS)

The SDS is a common psychological assessment tool and is mainly used to assess the degree of depression in individuals. By answering the questions on the scale, individuals can assess their depression. SDS can be used to measure the severity of depression, with 20 items divided into 4 levels. It mainly measures the frequency of symptoms, and the higher the frequency, the more severe the depression situation of the surveyed people [29,30]. "1" indicates no or little time has; "2" means sometimes has; "3" means most time has; "4" means the majority of time. The main statistical indicator of SDS is score, which is the final total score obtained by adding the scores of 20 items and multiplying the total score by 1.25 to obtain the standard score.

2.3 Data collection

The purpose of this study is to understand the current state of depression among students at University A and identify relevant influencing factors. This article proposes solutions based on influencing factors to reduce the overall depression rate of University A students, improve their learning quality, and ensure their mental health. Based on the 1,000 students selected in Table 1, this article expands the sample data. It randomly selected 3,750 students of all grades from University A for scale testing. A total of 3,620 questionnaires were collected, with a recovery rate of 96.53%. The specific situation is shown in Table 2.

Table 2: Specific situation of students selected by University A

Variable	Category	Number of people	Proportion (%)
Gender	Male	1,815	50.14
	Female	1,805	49.86
Grade	Freshman year	1,000	27.62
	Sophomore	820	22.65
	Junior year	930	25.69
	Senior year	870	24.04
Professional	Medical major	560	15.47
	Accounting major	1,020	28.18
	Business administration major	960	26.52
	Civil engineering	710	19.61
	Art major	200	5.52
	Semester education major	170	4.7
Only child	Yes	1,620	44.75
	No	2,000	55.25
Source	Urban	1,567	43.29
	Rural	2,053	56.71

As shown in Table 2, a total of 3,620 valid questionnaires were selected. Among them, the proportion of boys and girls is very close, ensuring the rationality of the research data in terms of gender and avoiding large-scale differences. In terms of grade, the number of freshmen is the highest, with 1,000 students accounting for 27.62%, while the number of sophomores is the lowest, with only 820 students accounting for 22.65%. For the selected student sources, the number of students from rural areas is slightly higher than that from urban areas. For the majors of students, the largest number is the accounting major, accounting for 28.18%, while the number of art majors and preschool education majors is relatively small, with a total of only 370 people. Regarding whether students are only children, the number of non only child students is 380 more.

The research object of this research is a student of University A. First, for the criteria of depression tendency and health, students can be divided into those with depression tendency and those with health through the established criteria. Based on the statistical results obtained from the school's mental health survey scale, individuals who meet the requirements can be selected, and the experimental content and needs can be summarized to these individuals. Next, it would inquire about the students' willingness to participate in the investigation, whether there is a reason to participate. If they are willing, they can sign an informed consent form and basic personal information registration form. At the same time, they would fill out BDI and SDS according to the requirements and finally conduct the experiment. The informed consent form is an important legal document in the medical process. It records the process by which the doctor informs the patient of medical information when the patient receives medical services, and the patient decides whether to receive medical services. Throughout the entire experiment, the subjects were alone in a quiet place using professional writing equipment. If there were any questions, the experimenter could help answer them. The specific framework is shown in Figure 1.

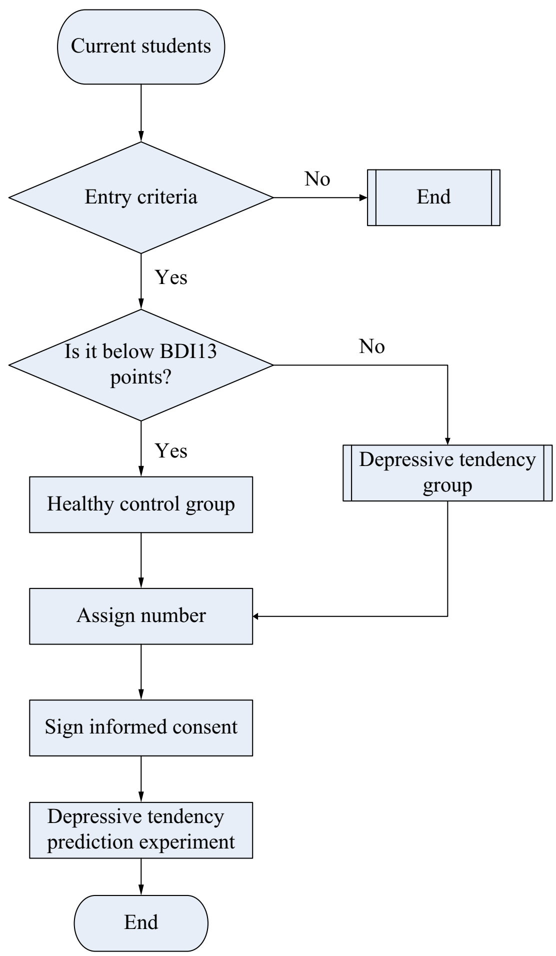

Figure 1: Data collection framework diagram.

As shown in Figure 1, this article collected data from students at University A and screened the learning criteria for inclusion. If it does not meet the criteria, it ends. If it does, it is determined whether the criteria for compliance are below BDI 13. If it is, it is divided into a healthy group; if it is not, it is divided into a depression-prone group. Next, these two groups can be numbered, and eligible students can sign informed consent forms to conduct depression tendency prediction experiments and obtain experimental results.

2.4 Feature extraction

Emotional features can be extracted, and then sentiment analysis models can be used to determine the sentiment of student scale texts. Emotion can be used to study students' depressive tendencies, which takes into account students' emotional demands and is not limited to identifying students' language patterns. Considering that depression patients have negative emotions and symptoms, they are also more inclined to express negative emotions in social settings. Common ones include anxiety and tension, avoidance and withdrawal, low self-esteem and self-doubt, and anger and hostility. Therefore, emotions can be used to identify whether students suffer from depression. Therefore, emotional features are used to extract the depressive emotional features of University A students [31,32]. Using an emotion dictionary can effectively capture users' emotional characteristics and judge students' emotions. It is currently widely used in emotion analysis and has achieved good results. Based on the above considerations, sentiment distribution is used to analyze users' language and text features, and sentiment features are evaluated based on sentiment dictionaries. In the feature quantification stage, emotional features Senta are added. The language structure of users is disassembled and refined, and the user's emotion value, word frequency weight, and other information are calculated through feature quantification. The calculation formula is as follows:

$$\text{Senta} = \frac{n}{N}, \tag{1}$$

where n represents the number of negative sentiment Weibo posts detected by the sentiment classification model, and N is the total number of Weibo posts posted by users.

2.5 Data

This study used a simple random sampling method for questionnaire distribution and reviewed the collected questionnaires to eliminate those that were not effectively utilized. The valid questionnaires obtained can be used for statistical calculations using Statistical Product and Service Solution 17.0 (SPSS17.0). At the same time, a dataset can be created and transferred to SPSS21.0 software for preliminary data preprocessing and calculation. The Cronbach's alpha method was used to study the consistency of questionnaire data. In this study, the relevant techniques included in different levels of depression were clearly described through frequency and percentage.

The methods used in data analysis can be used to study the detection of depression symptoms among students of different grades and majors. It mainly uses BDI and SDS for detection, and these two measurement scales are used for depression monitoring. Scores greater than 50 indicate obvious depressive symptoms, scores between 30 and 49 indicate probable depressive symptoms, while scores below 29 indicate no depressive symptoms, as shown in Tables 3 and 4.

Table 3: Detection of depressive symptoms among students of different grades

Symptoms		Freshman year	Sophomore	Junior year	Senior year
No depressive symptoms	Number of people	620	400	445	365
	Composition ratio (%)	62	48.78	47.85	41.95
May have depressive symptoms	Number of people	220	250	265	240
	Composition ratio (%)	22	30.49	28.49	27.59
Obvious depressive symptoms	Number of people	160	170	220	265
	Composition ratio (%)	16	20.73	23.66	30.46
Total	Number of people	1,000	820	930	870
	Composition ratio (%)	100	100	100	100

Table 4: Detection of depressive symptoms among students of different majors

Symptoms		Medical major	Accounting major	Business administration major	Civil engineering	Art major	Semester education major
No depressive symptoms	Number of people	60	364	500	368	40	70
	Composition ratio (%)	10.71	35.69	52.08	51.83	20	41.18
May have depressive symptoms	Number of people	170	376	271	220	100	60
	Composition ratio (%)	30.36	36.86	28.23	30.99	50	35.29
Obvious depressive symptoms	Number of people	330	280	189	122	60	40
	Composition ratio (%)	58.93	27.45	19.69	17.18	30	23.53
Total	Number of people	560	1,020	960	710	200	170
	Composition ratio (%)	100	100	100	100	100	100

As shown in Table 3, the survey on depression among students of different grades should be divided into three types. It can be found that the number and proportion of students without depressive symptoms are the highest in their freshman year, while the number and proportion of students with obvious depressive symptoms are the lowest. As freshmen, they have just started school and are full of freshness in everything. At the same time, their academic pressure is relatively low, and they do not have to face financial pressure. Everyone has just started school, and interpersonal communication is also relatively less stressful. The number of students without depressive symptoms in their senior year is the lowest, and the proportion is also the lowest. At the same time, the number of students with obvious depressive symptoms is the highest, and the proportion is also the highest. Because the senior year is the graduation season, students need to face internships, enter society, and begin to bear pressure from society. At the same time, their academic focus on papers may also exacerbate stress and lead to depression among students. The proportion of sophomores who may have depressive symptoms is the highest, with 30.49%, while the number of people who may have depressive symptoms in their third year is the highest, with 265 people.

As shown in Table 4, a study on the depression situation of students in different majors reveals that the proportion of medical students without depression symptoms is low, only 10.71%. The proportion of people with obvious depressive symptoms is as high as 58.93%, which is also the highest. This is because medical students face high academic pressure and also face various experiments, which can easily lead to depression. The proportion of students majoring in business administration without depressive symptoms is the highest, with 52.08%. The proportion of civil engineering students with obvious depressive symptoms is the lowest, only 17.18%. The proportion of art students who may have depressive symptoms is the highest, with 50%. Students majoring in accounting may have the highest number of depression symptoms, with 376 people, while students majoring in education during the semester have the lowest number of obvious depression symptoms, with only 40 people.

3 Construction of a prediction mechanism model for depression tendency in university learning

3.1 Model design

Many methods are involved in computer intelligent systems, and Natural Language Processing (NLP) technology is chosen to study the prediction of depression tendency in college learning [33,34]. By using NLP to analyze students' emotions and text content mining, students' current emotional state and thinking mode can be clearly understood. This information can help predict whether students are depressed and providing interventions for school teachers to help avoid increased depressive symptoms. NLP technology can be used to explore important features in student language and improve the effectiveness of stress assessment. First, pre-trained text can be used to create word embedding vectors, and then the vector representation of words can be modified based on semantic content. Then, based on the concept of words and the characteristics of phonetics, the information of words can be embedded into an emotion dictionary to obtain fine-grained feature representations of words. This can detect the feature information with the highest correlation to the model, thereby obtaining the weights of words and phrases, thereby improving the accuracy of search results. The effectiveness of this method can be demonstrated through a large number of experiments in the end.

The first step is to convert the text into a sequence of words. Assuming H represents a text $H = \{v_1, v_2, v_3, \cdots, v_m\}$ in the data, where $v_i(1 \le i \le m)$ represents the ith word in the text. For sequence H, for each word v_i, its corresponding word vector is represented as w_i^v:

$$w_i^v = fv_i. \tag{2}$$

The input text data can be transformed into vector representations of words.

The foundation of NLP technology is to use contextual information to predict the appearance of target words. The formula is

$$Q = (M_t \mid M_{t-j}, \cdots, M_{t-2}, M_{t-1}, M_{t+1}, M_{t+2} \cdots, M_{t+j}), \tag{3}$$

where M_t represents the word to be predicted, and j represents the window size.

Multiply the initial vector Y with the corresponding matrix M to obtain the result of the hidden vector:

$$f_0 = \frac{1}{2j}\sum_{i=1} Y_i^T \times M_{a\times b}. \tag{4}$$

Then multiply the hidden layer f_0 by the hidden layer matrix M' to obtain the hidden layer output:

$$R = f_0 \times M'_{a\times b}. \tag{5}$$

During the training process, the training error is calculated through the loss function, and the matrices M and M' are iterated and optimized through backpropagation.

After preprocessing and word segmentation of text data, a depression candidate sentiment word set would be generated. At the same time, it is necessary to remove irrelevant useless words related to depression and stress and only retain useful words related to them. To this end, the term frequency-inverse document frequency (TF-IDF) weighting algorithm was used for word frequency weight analysis, resulting in a depression seed word set [35,36]. The formulas for calculating the frequency index used in TF-IDF are as follows:

$$\mathrm{TF} - \mathrm{IDF} = tf \times idf, \tag{6}$$

$$tf(i) = \frac{di}{d}, \tag{7}$$

$$\mathrm{i}df(i) = \log\frac{m}{m(i) + 1}. \tag{8}$$

In the above formulas, the feature word i belongs to the candidate emotion word set, and $tf(i)$ represents the frequency of the feature word i in the total word count d of the text. $idf(i)$ represents the frequency of the inverse document, which is the degree of domain generalization of the word to the text. m represents the total number of text. Overall, the $tf - \mathrm{i}df$ value of the ith word is determined by the frequency of its occurrence in the text and the domain generalization of the text.

By using TF-IDF to process the weighted algorithm of college students' texts and extending the calculation of cosine similarity between words, the calculation formula for cosine similarity is as follows:

$$\cos(\theta) = \frac{\sum_{i=1}^{m}(a_i \times b_i)}{\sqrt{\sum_{i=1}^{m}(a_i)^2} \times \sqrt{\sum_{i=1}^{m}(b_i)^2}}, \tag{9}$$

where θ refers to the angle between two words after vectorization; a_i is the numerical value of the word i on the dimension a after vectorization representation; the cos value of the angle between two words indicates their similarity.

The seed word set filtered by the TF-IDF weighted algorithm belongs to a set of depression emotion dictionaries. Due to the incomplete generalization ability of subset domains, the seed word set also needs to be expanded. In order to improve the quality and efficiency of the dictionary, this article uses a basic sentiment dictionary to expand and support the depression seed word set on the basis of promises. Figure 2 shows the extension of the depression dictionary:

By fusing TF-IDF word frequency features with text sentiment features, a text feature vector is obtained, and the input vector h_i of this part is represented as:

$$h_i = w_{\mathrm{tfidf}} : w_{\mathrm{rule}}, \tag{10}$$

where w_{tfidf} represents the TF-IDF word frequency feature, and w_{rule} represents the emotional value feature of the text.

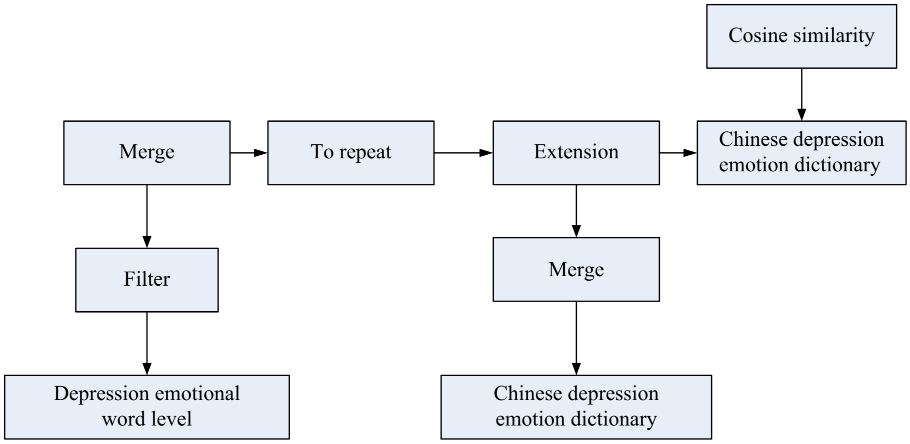

Figure 2: Expansion process of depression emotion dictionary.

The depression tendency detection module would fuse the W_i and H_i feature vectors obtained, represented as formulas (11) and (12).

$$\text{Mixed}_{\text{Feature}} = \text{concatenate}(W_i, H_i), \tag{11}$$

$$\text{Fea} = \text{BiNLP}(\text{Mixed}_{\text{Feature}}), \tag{12}$$

where $\text{Mixed}_{\text{Feature}}$ represents the fused feature vector, and Fea represents the output vector obtained from the fused feature vector.

Then, input Fea into the Sigmoid layer for classification and obtain the final prediction Result, represented as:

$$\text{Result} = \text{Sigmoid}(\text{Fea}). \tag{13}$$

3.2 Model construction and training

In order to evaluate the performance of the model, precision rate, recall rate, and $F1$ score values are mainly used for evaluation. Among them, the text of patients with depression is the positive sample size, while the normal text is the negative sample size. In general, accuracy and recall affect the overall performance of a model and are positively correlated with the overall performance of the model. The specific formulas are as follows:

$$\text{Precision} = \frac{T_{\text{P}}}{T_{\text{P}} + F_{\text{P}}}, \tag{14}$$

$$\text{Recall} = \frac{T_{\text{P}}}{T_{\text{P}} + F_{\text{M}}}, \tag{15}$$

$$F_1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{(\text{Precision} + \text{Recall})}, \tag{16}$$

where T_{P} represents the number of correctly classified positive samples, F_{P} represents the number of negative samples classified as positive samples, and F_{M} represents the number of positive samples classified as negative samples.

Based on the collected data, a predictive model can be constructed, with depression tendency as the target variable and other features as input variables. Many methods are involved in computer intelligent systems for conducting experiments on extracted sample data using different methods. In addition to NLP, there are also NN, support vector machine (SVM), and other methods that use text to predict depression tendencies, such as recurrent neural network (RNN), long short term memory (LSTM), transformer model, etc. The results of the depression tendency prediction models constructed using these different methods can be compared with the results obtained from the research methods in this article. It is mainly compared from three aspects: precision rate, recall rate, and *F*1 score value, as shown in Table 5.

Table 5: Comparison of experimental results of different models

Model	Precision (%)	Recall (%)	*F*1-score (%)
NLP	98.67	97.92	98.06
NN	90.47	91.76	89.47
SVM	91.68	90.33	88.96
RNN	91.24	89.75	91.08
LSTM	89.86	90.72	90.76
Transformer	90.33	88.69	91.33

As shown in Table 5, regardless of the precision rate, recall rate, or *F*1 score value, the student depression tendency prediction model constructed based on NLP technology is higher than other student depression tendency prediction models. Among them, the model built based on the LSTM algorithm has the lowest precision rate, while the model based on the Transformer has the lowest recall rate, and the model based on the SVM algorithm has the lowest *F*1 score value. The model precision rate based on NLP technology is 8.2, 6.99, 7.43, 8.81, and 8.34% higher than that based on NN, SVM, RNN, LSTM, and Transformer, respectively. The recall rate of the model based on NLP technology is 6.16% higher than that of the model constructed based on the NN algorithm. In summary, the model constructed based on NLP technology has more accurate performance results in predicting students' depressive tendencies through text.

4 Experiment of the predictive mechanism of depression tendency among college students on NLP

Studying the predictive mechanism of depression tendencies among college students can help prevent their depression psychology in advance. To study the student data collected in Table 2, this article utilizes NLP technology in computer intelligent systems to construct a prediction model and perform predictive analysis on the student data in Table 2. In order to make the experiment more scientific and reasonable, multiple experiments were conducted. The accuracy of each prediction obtained can be compared with the experimental results of prediction models constructed based on NN, SVM, and RNN algorithms. The specific comparison results are shown in Figure 3.

In Figure 3, the *x*-axis represents the number of experiments, with a total of 50 conducted. The *y*-axis represents the accuracy of experimental predictions. As shown in Figure 3, 50 experiments were conducted, but the prediction accuracy of the depression tendency prediction mechanism model constructed based on NLP technology was much higher in each experiment than in the other three prediction models. Among them, the predictive accuracy of the depression tendency prediction mechanism model based on NLP technology was over 95.68%. The prediction models constructed based on NN, SVM, and RNN algorithms were below 94.21%, below 93.86%, and below 93.74%, respectively. At the same time, the average prediction accuracy of the NLP-based depression tendency prediction model in 50 experiments was 97.02%, which was 5.33, 4.79, and 5.03% higher than the prediction average based on NN, SVM, and RNN algorithms, respectively. The experimental results indicate that the depression tendency mechanism model based on NLP technology improves the accuracy of prediction through data mining on language and text.

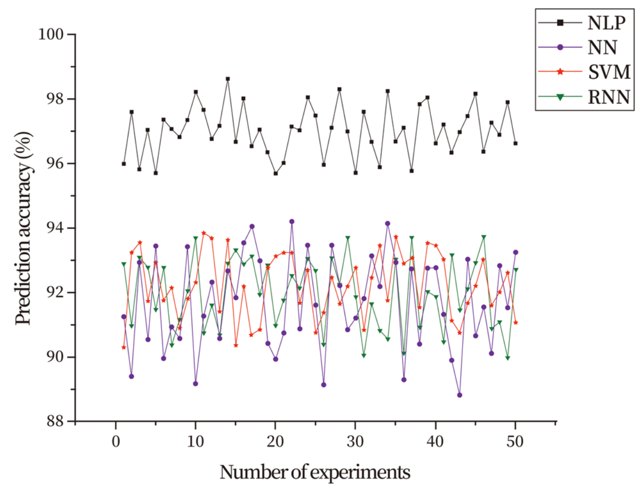

Figure 3: Comparison of prediction accuracy of different depression tendency prediction models.

It predicts depression tendencies among students from University A selected in Table 2. These students can collect a lot of data through surveys, and the faster the analysis is completed, the better it can help predict the school's depression tendency. Based on the sample data of 3,620 people in Table 2, this article expands the data and increases the number of people in the dataset. It uses the NLP technology-based prediction mechanism for college students' depression tendency to predict and analyze these student data and would obtain the analysis time. The comparison of the time required for prediction analysis with models built based on NN, SVM, and RNN algorithms is shown in Figure 4.

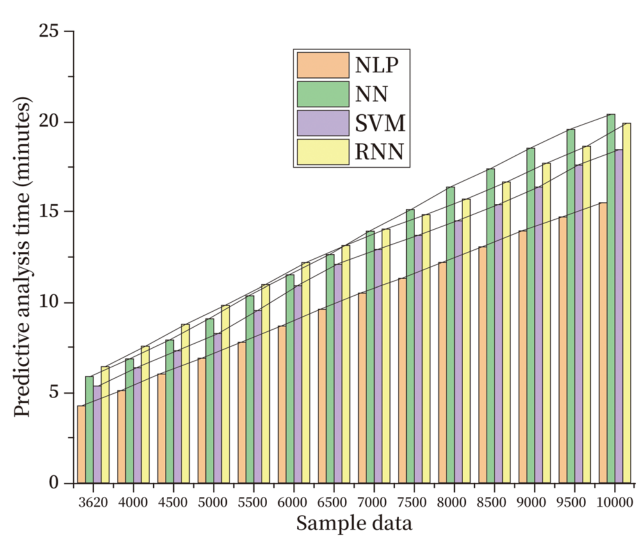

Figure 4: Comparison of time for student data processing using different prediction models.

In Figure 4, the *x*-axis represents the sample data. The *y*-axis represents the time of predictive analysis, and the unit is minute. As shown in Figure 4, based on the student data in Table 2, the number of students has been expanded. Different numbers of students would generate different amounts of data, and the time required for predicting and analyzing school depression tendencies also varies. The prediction mechanism model for college students' depression tendency based on NLP technology takes much longer to predict and analyze students' depression tendency than other prediction models. Among them, for the predictive model of SVM, it takes longer to analyze the depression tendency of different numbers of students than the model based on NLP technology but lower than the model constructed by NN and RNN algorithms. When the number of students was below 7,000, the time required for predicting depression tendency based on the NN algorithm model was higher than that based on NLP and SVM models, but it was lower than that based on RNN models. When the number of students exceeded 7,500, the time required for predicting and analyzing depression tendencies using the model constructed based on the NN algorithm was higher than the other three models. When the number of students was 3,620, the time required for the predictive analysis of depression tendency based on NLP technology was 4.29 min. It took 1.59, 1.1, and 2.13 min less than models built based on NN, SVM, and RNN algorithms. When the number of students was 10,000, the time required for the predictive analysis of depression tendency based on NLP technology was 15.54 min. It took 4.87, 2.91, and 4.35 min less than models built based on NN, SVM, and RNN algorithms.

This article numbers the nine reasons in Table 1 that may cause students to develop depressive tendencies or depression. The accuracy of predicting and identifying these nine reasons using the method studied in this article is very high. In order to further demonstrate the superiority of the research method in this article, the experimental results are compared with models constructed based on NN, SVM, RNN, and LSTM algorithms. The comparison of the recognition accuracy of these three models for the nine causes of depression tendencies in students in Table 1 is shown in Figure 5.

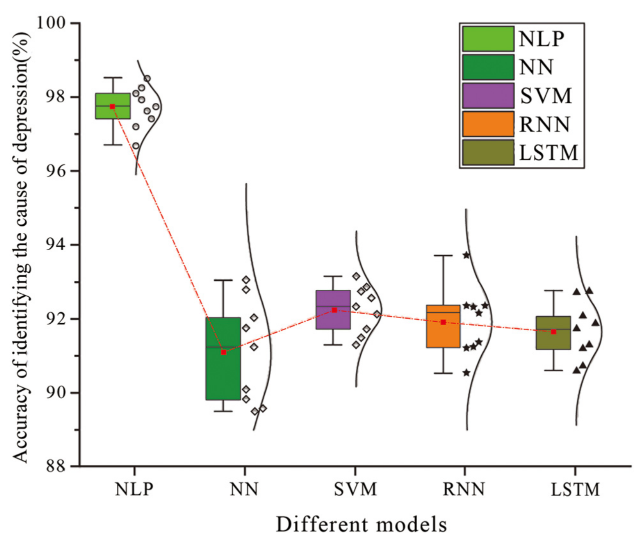

Figure 5: Comparison of recognition accuracy of different models for students' depression tendencies.

In Figure 5, the *x*-axis represents different models for predicting student depression tendencies. Five models are constructed based on different technologies or algorithms to predict the cause of depression, and the *y*-axis represents the accuracy of identifying the cause of depression. The various elements in the figure represent the accuracy of identifying each of the nine causes. The position of the red square is the mean of the accuracy of identifying nine causes, and the red connecting line is the mean connecting line of each algorithm. As shown in Figure 5, the prediction mechanism model constructed based on NLP technology had an accuracy of over 96.7% for identifying nine causes. The recognition accuracy of models constructed based on NN, SVM, RNN, and LSTM algorithms was below 93.07, 93.16, 93.73, and 92.78%, respectively. Among them, the average accuracy of the research method in this article for identifying nine causes was 97.74%, which was 6.63, 5.47, 5.81, and 6.06% higher than the average recognition accuracy of models constructed based on NN, SVM, RNN, and LSTM algorithms, respectively.

5 Conclusions

Depression is a psychological problem, and depressive tendencies are early symptoms of depression. If not intervened in a timely manner, it is easy to develop into depression. The main symptoms of depression among college students include disappointment in academic and daily life, decreased academic performance, and social difficulties. Based on the preliminary analysis results, regression analysis was conducted using the influencing factors of the research survey as the independent variable and depression tendency as the dependent variable. The results indicate that students' family reasons, academic pressure, and interpersonal relationships are all important factors that affect their tendency to develop depression. For this reason, as well as the current mechanism for predicting depression tendencies among college students, there are problems in terms of prediction accuracy, data collection, real-time monitoring, and consideration of factors leading to depression. The NLP algorithm can be further used to construct a depression tendency prediction model. Comparing this model with other models in experiments, it can be found that this model can accurately predict the depression tendency of college students, and the speed of prediction analysis for students is also faster. It can analyze a large amount of student data in a very short period of time, helping to lay the foundation for subsequent predictions. It can also effectively identify the reasons that lead to students' tendency towards depression and can provide targeted medication and psychological treatment for students based on different reasons. Through the research on the depression tendency of college students, it can be helpful to predict the depression tendency of college students in advance, and targeted intervention measures can be taken to reduce the incidence rate of depression. It should be noted that this research field involves personal privacy and ethical issues and that students' informed consent and privacy protection are ensured in the process of data collection and use. In addition, the accuracy and interpretability of the prediction model are also issues that need to be considered to ensure the credibility and practicability of the model.

Funding Information: This study did not receive any funding in any form.

Author contributions: Z. H. was responsible for the research question, determining the research method and experimental protocol, and interpreting and discussing the experimental results, and the writing of the article.

Conflict of interest: The author(s) declare(s) that there is no conflict of interest regarding the publication of this article.

Data availability statement: The data used to support the findings of this study are available from the corresponding author upon request.

References

[1] Khandaker GM, Zuber V, Rees JM, Carvalho L, Mason AM, Foley CN, et al. Shared mechanisms between coronary heart disease and depression: findings from a large UK general population-based cohort. Mol Psychiatry. 2020;25(7):1477–86. doi: 10.1038/s41380-019-0395-3.

[2] Cookson C, Luzon O, Newland J, Kingston J. Examining the role of cognitive fusion and experiential avoidance in predicting anxiety and depression. Psychol Psychotherapy: Theory Res Pract. 2020;93(3):456–73. doi: 10.1111/papt.12233.

[3] Yaofeng H, Yuting Y, Ya F. Research on the relationship between depressive tendency and cognitive function change trajectory of elderly people in rural areas of our country. Chin J Epidemiol. 2022;43(9):1491–6. doi: 10.3760/cma.j.cn112338-20220415-00302.

[4] Zhen Y, Huijun Z. Construction of a predictive model for the risk of nursing depression among elderly caregivers of elderly patients with dementia. J Nurs. 2021;28(21):7–11. doi: 10.16460/j.issn1008-9969.2021.21.007.

[5] Fauziyyah A, Ampuni S. Depression tendencies, social skills, and loneliness among college students in Yogyakarta. J Psikol. 2018;45(2):98–106. doi: 10.22146/jpsi.36324.

[6] Ozimek P, Bierhoff HW. All my online-friends are better than me–three studies about ability-based comparative social media use, self-esteem, and depressive tendencies. Behav Inf Technol. 2020;39(10):1110–23. doi: 10.1080/0144929X.2019.1642385.

[7] Zhengjie L, Xiaojuan Z, Qingxuan L, Ziyang Z, Hong L, Zhongming Z. The combination of reward positive wave and late positive potential predicts depressive tendency: A study based on ERP. Psychol Sci. 2023;46(4):980. doi: 10.16719/j.cnki.1671-6981.202304027.

[8] Nan J, Ping W. The relationship between life satisfaction of medical students and depressive tendencies. J China Med Univ. 2018;47(9):838–41. doi: 10.12007/j.issn.0258-4646.2018.09.016.
[9] Yano K, Kase T, Oishi K. Sensory processing sensitivity moderates the relationships between life skills and depressive tendencies in university students 1. Jpn Psychol Res. 2021;63(3):152–63. doi: 10.1111/jpr.12289.
[10] Qin X, Wang S, Hsieh CR. The prevalence of depression and depressive symptoms among adults in China: estimation based on a National Household Survey. China Economic Rev. 2018;51(1):271–82. doi: 10.1016/j.chieco.2016.04.001.
[11] Sullivan MD. Depression effects on long-term prescription opioid use, abuse, and addiction. Clin J Pain. 2018;34(9):878–84. doi: 10.1097/AJP.0000000000000603.
[12] Nurwulan NR, Selamaj G. Working university students in riau archipelago: dual role and depression. J Educative: J Educ Stud. 2020;5(2):123–35. doi: 10.30983/educative.v5i2.3330.
[13] He L, Guo C, Tiwari P, Pandey HM, Dang W. Intelligent system for depression scale estimation with facial expressions and case study in industrial intelligence. Int J Intell Syst. 2022;37(12):10140–56. doi: 10.1002/int.22426.
[14] Bohu L, Xudong C, Lin Z, Tan L, Duzheng Q, Tingyu L, et al. Preliminary research on modeling and simulation technology for new artificial intelligence systems. J Syst Simul. 2018;30(2):349. doi: 10.16182/j.issn1004731x.joss.201802001.
[15] Pimenov DY, Bustillo A, Wojciechowski S, Sharma VS, Gupta MK, Kuntoglu M. Artificial intelligence systems for tool condition monitoring in machining: Analysis and critical review. J Intell Manuf. 2023;34(5):2079–121. doi: 10.1007/s10845-022-01923-2.
[16] Zhao K. Review of the application of set pair analysis in system intelligent prediction. J Intell Syst. 2021;17(2):233–47. doi: 10.11992/tis.202103023.
[17] Ye Y, Yong Z, Han D. The key technologies of industrial artificial intelligence and their application status in predictive maintenance. J Autom. 2020;46(10):2013–30. doi: 10.16383/j.aas.c200333.
[18] Li J, Jiang M, Qin Y, Zhang R, Ling SH. Intelligent depression detection with asynchronous federated optimization. Complex Intell Syst. 2023;9(1):115–31. doi: 10.1007/s40747-022-00729-2.
[19] Ayon SI, Islam MM, Hossain MR. Coronary artery heart disease prediction: a comparative study of computational intelligence techniques. IETE J Res. 2022;68(4):2488–507. doi: 10.1080/03772063.2020.1713916.
[20] Qureshi SA, Saha S, Hasanuzzaman M, Dias G. Multitask representation learning for multimodal estimation of depression level. IEEE Intell Syst. 2019;34(5):45–52. doi: 10.1109/MIS.2019.2925204.
[21] Yang S, Zhou P, Duan K, Hossain MS, Alhamid MF. emHealth: towards emotion health through depression prediction and intelligent health recommender system. Mob Netw Appl. 2018;2(3):216–26. doi: 10.1007/s11036-017-0929-3.
[22] Shorey S, Ng ED, Wong CH. Global prevalence of depression and elevated depressive symptoms among adolescents: A systematic review and meta-analysis. Br J Clin Psychol. 2022;61(2):287–305. doi: 10.1111/bjc.12333.
[23] Aalbers G, McNally RJ, Heeren A, De Wit S, Fried EI. Social media and depression symptoms: A network perspective. J Exp Psychol: Gen. 2019;148(8):1454. doi: 10.1037/xge0000528.
[24] Twenge JM, Thomas EJ. US Census Bureau-assessed prevalence of anxiety and depressive symptoms in 2019 and during the 2020 COVID-19 pandemic. Depression Anxiety. 2020;37(10):954–6. doi: 10.1002/da.23077.
[25] Acharya L, Jin L, Collins W. College life is stressful today–Emerging stressors and depressive symptoms in college students. J Am Coll Health. 2018;66(7):655–64. doi: 10.1080/07448481.2018.1451869.
[26] Peng W, Luo W, Zhou R. HRV evidence that working memory refresh training improves the emotional regulation ability of college students with depressive tendencies. J Psychol. 2019;51(6):648. doi: 10.3724/SP.J.1041.2019.00648.
[27] Ji D, Qin W, Minlei Q, Xiaojian Y, Huipan W, Yuanyuan M, et al. The correlation between adolescent family environment and depressive symptoms. Chin Sch Hyg. 2023;44(5):677–81. doi: 10.16835/j.cnki.1000-9817.2023.05.009.
[28] Haisha X, Xiaopu L, Xinyun G, Wenjing T, Dongling Z, Yu L, et al. Pay attention to the Meta-analysis of the intervention effect of bias correction therapy on depressed patients. Chin Gen Med. 2022;25(17):2145. doi: 10.12114/j.issn.1007-9572.2022.0182.
[29] Jokelainen J, Timonen M, Keinanen-Kiukaanniemi S, Harkonen P, Jurvelin H, Suija K. Validation of the Zung self-rating depression scale (SDS) in older adults. Scand J Prim Health Care. 2019;37(3):353–7. doi: 10.1080/02813432.2019.1639923.
[30] Komiyama M, Ozaki Y, Miyazaki Y, Yasoda A, Wada H, Yamakage H, et al. Short-term changes in self-rating depression scale scores after smoking cessation in neurotic patients. Intern Med. 2021;60(8):1175–81. doi: 10.2169/internalmedicine.4868-20.
[31] Haiqi W, Ran C, Shiqing W, Li G, Tao F. Use the geographic and emotional characteristics of Chinese Weibo data to mine. J Wuhan Univ (Inf Sci Ed). 2020;45(5):699–708. doi: 10.13203/j.whugis20180138.
[32] Park Y, Seo DG, Park J, Kim B, Choi J. The influence of behavioral and emotional characteristics on academic achievement of middle school students: A growth modeling approach. Sch Psychol Int. 2019;40(5):433–55. doi: 10.1177/0143034319853010.
[33] Khurana D, Koli A, Khatter K, Singh S. Natural language processing: State of the art, current trends and challenges. Multimed Tools Appl. 2023;82(3):3713–44. doi: 10.1007/s11042-022-13428-4.
[34] Otter DW, Julian RM, Jugal KK. A survey of the usages of deep learning for natural language processing. IEEE Trans Neural Netw Learn Syst. 2020;32(2):604–24. doi: 10.1109/TNNLS.2020.2979670.
[35] Larasati UI, Muslim MA, Arifudin R, Alamsyah A. Improve the accuracy of support vector machine using chi square statistic and term frequency inverse document frequency on movie review sentiment analysis. Sci J Inform. 2019;6(1):138–49. doi: 10.15294/sji.v6i1.14244.
[36] Falasari A, Muslim MA. Optimize naive bayes classifier using chi square and term frequency inverse document frequency for amazon review sentiment analysis. J Soft Comput Explor. 2022;3(1):31–6. doi: 10.52465/joscex.v3i1.68.