

Research Article

Research on Psychological Emotion Recognition of College Students Based on Deep Learning

Yuebo Li¹ and Yatong Zhou ²

¹*School of Marxism, Hebei Vocational University of Industry and Technology, Shijiazhuang 050091, China*

²*Department of Marxism Teaching, Hebei GEO University, Shijiazhuang 050030, China*

Correspondence should be addressed to Yatong Zhou; zhouyatong@hgu.edu.cn

Received 21 March 2022; Revised 6 April 2022; Accepted 13 April 2022; Published 12 May 2022

Academic Editor: Jie Liu

Copyright © 2022 Yuebo Li and Yatong Zhou. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

College students' anxiety, depression, inferiority complex, interpersonal sensitivity, and a series of mental health problems have a very serious negative impact on individuals, families, and society. In order to obtain better psychological emotion recognition effect of college students, this paper proposes a psychological emotion recognition algorithm based on multisource data. One-dimensional convolutional neural network (1D-CNN) was used to mine students' online patterns from online behavior sequences. According to the consumption data of students in the canteen, abnormal scores are calculated to depict the dietary differences among students. At the same time, the students' psychological state data provided by the psychological center are used as labels to improve the shortcomings of the questionnaire. Five kinds of common classification algorithms are trained by training set, and the classifier with the best effect is selected through evaluation of verification set, which is used to identify students with mental health problems in the test set. Experimental results show that precision, recall, and *F1*-measure reach 0.68, 0.56, and 0.67, respectively. 75% of students with mental health problems are identified. The psychological and emotional recognition system of college students based on deep learning provides systematic method and theoretical support for the school to find students with psychological problems in time and provide intervention.

1. Introduction

With the rapid development of knowledge economy and the increasing popularity of higher education, the number of college students is increasing day by day, and there are more and more students with psychological problems [1]. College students, as a part of high-level talents, have always been regarded as the outstanding generation of the society, with strong psychological quality by default, but the fact is disappointing. All kinds of pressures such as study, social relationship, and employment make college students mentally exhausted, and timely detection of abnormal students has become one of the most concerned and intractable problems in universities [2].

People's psychological characteristics can be expressed through daily living behaviors and work-and-rest rules. Researchers have mined information that can reflect the

mental health state from the daily behavior data of college students. Research shows that there is a close relationship between mental health and online behavior. Jorm proposed an algorithm to predict mental health problems through online use behavior. Based on web usage behavior, a computational model for predicting scores of SCL-90 dimensions was established [3]. Nie [4] proposed a new method to predict the future mental illness severity (MIS) of users sharing posts on Instagram in 2014. Paola [5] discussed the universality and treatment of college students' psychology and psychosis, proposed the uniqueness of college students' development stage and environment in terms of psychological abnormalities, and summarized the influence of college students' psychological problems [6]. Zanganeh [7] (2018) explored the role of emotional factors in doctoral students' online information retrieval. Domestic scholars have conducted relevant research on this topic [8]. Tang [9]

used SCL-90 scale to measure the mental health of graduate students, established a prediction model through BP neural network to analyze and predict the mental health status of students, and put the model into practical application [10]. Wang Yinmei et al. [11] used symptom measurement scale to investigate 532 college students and found that college students' mental health varies in different grades, mainly influenced by interpersonal relationship, diet and sleep in college, and students of different majors have different psychological conditions [12]. Zhu [13] (2019) proposed a new method to detect depression through time-frequency analysis of network behaviors. A classification model was established to distinguish the high SDS group from the low SDS group, and a more accurate prediction model was established to identify the psychological state of the depression group [14].

To sum up, this paper carries out an analysis and research on students' mental health state through the data of students' education in school and excavates the relationship between college students' behaviors and their mental health state, which plays an important role and shows significance for university student managers to carry out early intervention for students with psychological abnormalities.

2. Principle of Deep Learning Algorithm

Deep learning is the process of using computers to imitate human learning behaviors. In the process of learning, new knowledge and experience are constantly accumulated, and their own knowledge structure is established according to this knowledge so as to improve their learning ability [15]. The research in this paper is a binary classification problem, and five commonly used classification algorithms are used in the experimental process, including decision tree, random forest, naive Bayes, gradient ascending tree, and BP neural network. In the feature extraction process, in order to reflect the dining situation of students, k -means clustering is used to cluster the dietary data of students.

2.1. k -Means Clustering. The main idea of k -means clustering algorithm: firstly, select K sample points randomly from the sample set and take them as the center of the cluster. Then, according to the distance between each sample and the k centroids, it is divided into the nearest cluster, and then the centroids of each cluster are recalculated [16]. Euclidean distance is used in this paper, and the calculation formula is shown as

$$d(x, C_i) = \sqrt{\sum_{j=1}^m (x_j - c_{ij})^2}, \quad (1)$$

where x is the data object, C_i is the i th cluster center, m is the dimension of the data object, x_j , C_{ij} are the j attribute values

of x and C_i . The calculation formula of error square and SSE of the whole data set is shown as

$$SSE = \sum_{i=1}^k \sum_{x \in C_i} |d(x, C_i)|^2. \quad (2)$$

2.2. Decision Tree Algorithm. The different decision tree algorithms according to different split attribute algorithms are as follows. ID3 algorithm selects the optimal attribute for splitting by calculating the information gain, and the attribute with the maximum information gain after partitioning is the optimal attribute [17]. The information gain is calculated on the basis of information entropy, which is shown as

$$\text{Entropy}(D) = - \sum_{i=1}^m p_i \log_2 p_i. \quad (3)$$

Among them, m is the number of categories in a data set D , and p_i is the first i category of probability. The smaller the value of (D) is, the higher the purity of D is.

If attribute a in the data set is used, the possible values of a are $v = \{a^1, a^2, a^3, \dots, a^v\}$, which can be calculated according to formula (4) to attribute a classified information entropy D^v , because the various values of a sample are different, add weight to each branch node $|D^v|/|D|$, when dividing the data set with property a , can according to the formula to calculate the sample set of type D information gain.

$$\text{Gain}(D, a) = \text{Entropy}(D) - \sum_{v=1}^v \text{Entropy}(D^v). \quad (4)$$

However, the information gain tends to favor attributes with more values. In order to improve this defect, the famous C4.5 algorithm is proposed [7], which uses the information gain rate when selecting attributes, as shown in formula (5). The detailed attribute selection process is as follows: firstly, the information gain of the attribute is calculated, and the attribute whose information gain is higher than the average value is selected, and then the highest gain rate is selected from these attributes for classification.

$$\text{Gain_ratio}(D, a) = \frac{\text{Gain}(D, a)}{\text{SI}(a)}. \quad (5)$$

$$\text{SI}(a) = - \sum_{v=1}^v \frac{|D^v|}{D} \log_2 \frac{|D^v|}{D}. \quad (6)$$

CART algorithm uses Gini index to select the optimal splitting attribute, and the calculation speed is faster than the information gain rate [13]. The purity of dataset D can be measured by Gini index, as shown in (7). The smaller (D) is, the higher the purity of the data set is. Therefore, which

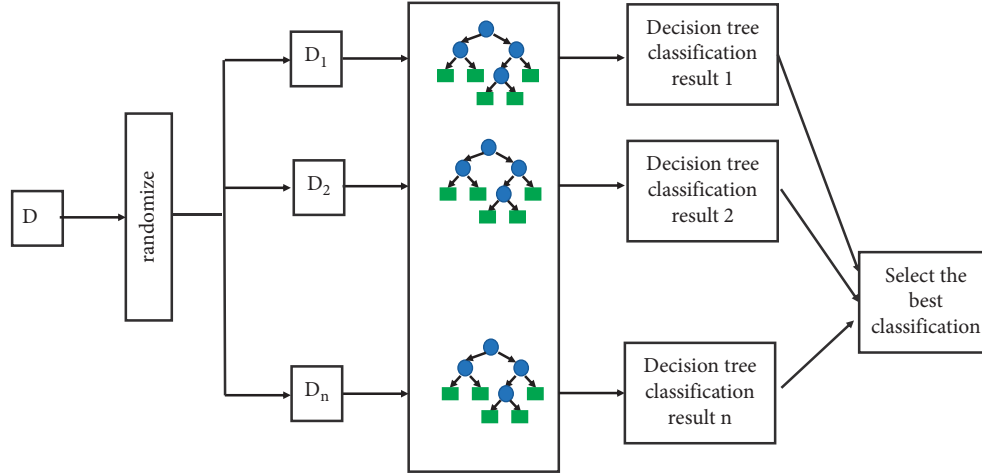


FIGURE 1: Random forest workflow.

attribute is the optimal splitting attribute can be judged according to the size of Gini index.

$$\text{Gini}(D) = 1 - \sum_{i=1}^n p_i^2. \quad (7)$$

The idea of decision tree algorithm is simple, easy to understand, and easy to extract rules. Insensitivity to data can process both nominal and numerical data and can process unrelated features. However, it also has some disadvantages. First, it is easy to produce overfitting, and second, it is easy to ignore the correlation between attributes.

2.3. Random Forest Algorithm. The random forest algorithm takes decision tree-based classifier and combines Bagging construction with random attribute selection. Its workflow is shown in Figure 1. Specifically, random forest determines the category of samples by voting, and each base classifier has one vote to determine the category of samples according to the principle of majority rule. The randomness of the random forest in random attribute selection, for each node base classifier, randomly selected from the set from the current node attributes first k attribute a subset, and then according to the attribute selection algorithm from subset to calculate the optimal split attribute, rather than the traditional decision tree from the current node, find out the best in all of the attributes of the attribute. The parameter k is introduced to reflect the randomness of the random forest Algorithm 1.

2.4. Naive Bayes. Naive Bayes algorithm takes Bayes algorithm as the premise and introduces the “independence” hypothesis, whose assumed attributes are independent of each other [18]. Under this assumption of independence, naive Bayes algorithm is especially suitable for solving multiattribute problems. Naive Bayes algorithm is currently

recognized as a simple and effective probability classification method, and its classification process is as follows:

- (1) Assume that each sample has n attributes. Each sample is represented by an n -dimensional eigen-vector $d = \{d_1, d_2, d_3, \dots, d_n\}$.
- (2) Suppose there are m categories $c = \{c_1, c_2, c_3, \dots, c_m\}$. Given an unknown data sample d , the classification predicts that d belongs to the class with the maximum posterior probability, that is, if and only if $p(c|d) > p(c|d)1 \leq j \leq m$ and $j \neq i$. The Bayesian classification assigns the sample to be classified to class c_i , and the value of $p(c_i|d)$ can be calculated by the conditional probability formula in probability theory, as shown in the following formula.

$$P(c_i|d) = \frac{P(d|c_i)P(c_i)}{P(d)}. \quad (8)$$

- (3) Since the sample set is known and the value of $P(d)$ can be calculated, the maximum value of $P(c_i|d)$ can only be calculated by the maximum value of $p(d|c_i)p(c_i)$. Without knowing the prior probabilities of the classes, assume that each class is equally likely, so you just maximize $(d|c_i)$. Otherwise, maximize $(d|c_i)(c_i)$.
- (4) In the case of multiple attributes, the overhead of calculating $(d|c_i)$ may be very large. In this case, we can make the naive assumption of class condition independence to reduce the overhead of calculating $(d|c_i)$. The classes of a given sample are labeled, assuming that the attributes are independent of each other.

$$P(d|c_i) = \prod_{j=1}^a P(d_j|c_i), \quad (9)$$

- (i) Input: data set $T = \{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\}$, loss function L
- (ii) Output: Additive model $f(x) = f_M(x)$
- (1) Initialize the base classifier $f_0(x) = \arg \min_{\gamma} \sum_{i=1}^N L(y_i, \gamma)$
 - (2) Each base classifier needs to perform the following operations:
 - ① According to the loss function, the negative gradient of each sample in the current model is calculated and used as an estimate of the residual $r_{im} = -[\partial L(y_i, f(x_i))/\partial, f(x_i)]$.
 - ② For a given Γ_{im} , the leaf node R_{mj} , R_{mj} , $j = 1, 2, \dots, J$, J is the number of leaf nodes
 - ③ For leaf nodes, calculate $c_{mj} = \arg \min_c \sum_{x_i \in R_{mj}} L(y_i, f_{m-1}(x_i) + c)$.
 - ④ Update regression tree $f_m(x) = f_{m-1}(x) + \sum_{j=1}^J c_{mj} I(x \in R_{mj})$.
 - (3) Output gradient lifting tree $\hat{f}(x) = f_M(x)$.

ALGORITHM 1

where $(d_1|ci)$, $(d_2|ci)$, \dots , $P(dj|ci)$ can be computed from the sample set.

- (5) When classifying classified sample d , calculate $(d|ci)(ci)$ for each class c_i in turn. If and only if $(d|ci)(ci) > P(d|cj)P(j)$, $1 \leq j \leq m$, $j \neq i$, sample d to be classified is divided into class c_j .

2.5. BP Neural Network. BP neural network is a multilayer feedforward neural network trained according to error reverse propagation algorithm. The training process of BP neural network is divided into two parts: one is signal forward propagation, and the other is error reverse propagation, as shown in Figure 2. Among them, BP neural network has many advantages such as strong nonlinear mapping ability, generalization ability, and fault tolerance ability due to its error back propagation learning process.

BP neural network consists of three parts: input layer, hidden layer, and output layer. Among them, the number of training sample attributes determines the number of neurons in the input layer, and the number of target classification determines the number of neurons in the output layer. For example, for a binary classification problem, there are only two neurons in the output layer. Users can freely define the number of hidden layers and the number of neurons in each hidden layer. The lines between each layer represent the weights.

BP neural network has the following advantages: first, it has strong adaptability and self-learning ability, which can continuously learn and correct path weight and save learning content during training. Second, the nonlinear mapping and fault tolerance ability can better solve the internal mechanism of the more complex nonlinear mapping. However, BP neural network has some defects, such as slow training speed, no corresponding standard for hidden layers, and number of neurons.

2.6. Pearson's Correlation Coefficient. Pearson's correlation coefficient is used to measure the linear correlation between two variables in most cases. It is defined as the quotient of covariance and standard deviation between two variables, and its calculation is shown in the following formula.

$$\rho_{X,Y} = \frac{\text{cov}(X, Y)}{\sigma_{X,Y}} = \frac{E(XY) - E(X)E(Y)}{\sqrt{E(X^2) - E^2(X)}\sqrt{E(Y^2) - E^2(Y)}} \quad (10)$$

where X and Y represent two random variables, $\text{cov}(X, Y)$ represents the covariance between X and Y , $\sigma_{X,Y}$ represents the standard deviation between X and Y , and E represents statistical mathematical expectation.

Pearson's correlation coefficient ranges from -1 to 1 . A positive value means a positive correlation between two variables (1 means a complete positive correlation), a negative value means a negative correlation between two variables (-1 means a complete negative correlation), and a 0 value means there is no linear correlation between two variables. When comparing the correlation between feature and target problem, the larger the absolute value of correlation coefficient is, the more relevant the feature is to target problem.

3. Student Mental Health Problem Identification Algorithm Based on Multisource Data

The experimental process of identifying students with psychological and emotional problems based on multisource data includes data acquisition and preprocessing, feature extraction, model training, and recognition. In the experiment, students' consumption data, access control data, network data, historical performance data, and psychological center data were collected.

3.1. Algorithm Framework. The overall framework of the student mental health problem recognition algorithm based on multisource data is shown in Figure 2. The whole algorithm process is divided into three parts, including data acquisition and preprocessing, feature extraction, and model training and recognition. In the stage of data acquisition and preprocessing, we obtained four data sources, namely, network log, access control data, achievement data, and consumption data. In the feature extraction stage, the relevant features such as students' online pattern and consumption abnormal score are extracted from the four data

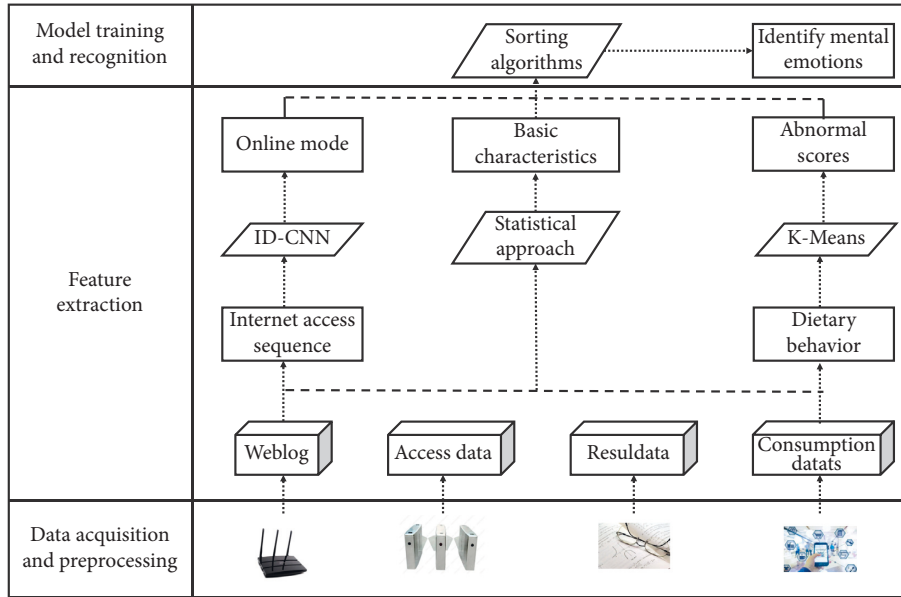


FIGURE 2: Multisource data-based student mental health problem recognition algorithm framework.

sources. The optimal classifier is selected from five common classification algorithms.

3.1.1. Data Acquisition and Preprocessing. After applying to relevant departments and obtaining students’ informed consent, this study obtained a variety of behavioral data of students, including students’ consumption data, historical performance data, network logs, access control and card swiping data, and psychological state data.

(1) Consumption Data. There will be a record every time students swipe the campus card. There are three tables of consumption data, namely, consumption flow chart, student account information table, and shop account information table. According to the shop account, connect the shop account information table with the consumption table and the student account information table, and the results are shown in Table 1.

(2) Access Control Data. At present, college students can enter and leave dormitories only after swiping their card through the access control system using a Metro Card. Every time a student swipes the card, an entry-and-exit record will be generated, and these records are stored in the access control information table, as shown in Table 2.

The access control information not only reflects the time trajectory of students entering and leaving the dormitory, but also calculates the duration of students in the dormitory according to their entering and leaving states. Students usually go in and out of dormitories with friends, and digging in this direction may also reveal interesting information, such as friend relationships.

(3) Students’ Network Log Records. Table 3 shows the network log information. Network logs record students’ online activities, including online time and browsing

content. How to dig out the patterns and characteristics of students’ online behavior from these records is a problem we need to solve.

With the approval of the school network center and the informed consent of the students, we obtained the log files of the students’ Internet access. We understand and analyze the network log; the format of the network log looks very simple, but the information it contains is very complex. The URL is divided into seven types, respectively: (1) comprehensive, (2) adult, (3) entertainment, (4) news, (5) publicity, (6) shopping, and (7) learning.

(4) Students’ Historical Performance Data. There is a significant negative correlation between mild, moderate, and severe mood disorders and academic performance [19], and students’ historical performance data reflect their past learning status and process. With the approval of the faculty and the informed consent of the student, this table details the students for the first time to attend this exam year students (XN), the first time to attend the exam of semester exam (XQ), students take the exam last time last year (QDXN), students take part in this exam semester (QDXQ), students, student id (XH), course name (KCMC), test scores (KS CJ), credits (XF), GRADE points (JD), grade points (XFJD), and credit acquisition methods (QDFS), including preliminary courses, retakes, and make-up exams, as shown in Table 4. Each record in the student history grade table reflects the student’s learning of a certain course, and the student’s comprehensive learning situation can be understood through the student number.

In the analysis and understanding of the student history score table, we found that the scoring method of the examination score (KSCJ) is not uniform; some courses use the hundred-mark system, while some use the five-level system [19]. For the convenience of calculation and analysis, we change the five-grade system into a percentage system, and

TABLE 1: The three tables join in the consumption data.

Student ID	Consumption time	Charge time	Transaction amount	Shop's account	Shop's name
2021****	20210421	2021/4/21 8:09:59	-5	1050229	Noodle section on the first floor of the third restaurant
2020****	20210417	2021/4/17 13:46:11	-3	1574702	Service building food street
2019****	20210425	2021/4/25 18:07:22	-2.5	1050253	The third restaurant, first floor malatang

TABLE 2: Access control information table.

Date of credit card	Charge time	Student ID	State of charge	Professional	Location	Dormitory building	Dormitory room no.
20210417	9:11:27	2020****	3	Mechanical engineering	9	3	309
20210606	10:52:0''	2019****	3	Transportation process	3	1	116
20210425	21:55:9	2018****	1	Materials and engineering	7	1	127

TABLE 3: Network log original table.

Time online	URL	Student ID	ULONG source IP	Source port	ULONG destination IP	Destination port
2021/4/21 8:16:54	vcheck.f.360cn/checker	2020****	172.24.8.118	80	221.228.204.185	34532
2021/4/21 7:02:44	rq.wh.cmam.111com/res''	2019****	172.20.52.152	443	203.248.210.232	56547
2021/4/25 17:32:25	https://www.imooc.com/wenda	2018****	1172.24.8.238	80	117.121.10141-	75362

TABLE 4: Student history score table.

XN	XQ	QD XN	QD XQ	XH	KCMC	XF	KS CJ	JD	XF JD	QDFS
2019	0	2019	0	2019****	Advanced mathematics	1	Pass	1	1	Early repair
2020	0	2021	0	2020****	C+ language	0	0	0	0	Early repair
2018	1	2019	1	2018****	Linear algebra	0	54	0	0	Rebuild
2019	1	2019	1	2019****	Linear algebra	1	65	1	2	Supplementary examination

the corresponding relationship is as follows: excellent (95), good (85), medium (75), pass (60), and fail (50).

It is found that there are missing values in the grade point column of the student's history grade table. Since credits are the product of grade points multiplied by course credits, a loss of grade points results in loss of corresponding credits. By observing the existing GPA, find out the calculation rule of GPA, where $CGPA_i$ stands for GPA of the course i , $Score_i$ stands for GPA of the course i , and the corresponding relationship between the calculated GPA and GPA is shown in Table 5. Next, according to the formula to calculate the credit grade point of the problem, $CGPA_i$ stands for the first i course credit grade point and $Credit_i$ stands for the first i course credits.

$$GPA_i = \frac{(Score_i - 50)}{10}, \quad (11)$$

$$CGPA_i = Credit_i \times GPA_i.$$

TABLE 5: Conversion relationship between grade points.

Designation	Conversion relationship						
Results	>90	89	88	61	60	<
Grade point	4	3.9	3.8	1.1	1	0

(5) *Psychological State Data*. With the approval of the psychological center and the informed consent of the students, the psychological state table of the students was obtained. Each record in the table contains the student's student ID, gender, department, grade, class, attention level, and update time. The concern level includes mild, moderate, and severe. "Mild" means suffering from mild mental health problems; "moderate" means having moderate mental health problems; and "severe" means having severe mental health problems [11].

In this study, we targeted a dichotomous problem, namely, students with mental health problems and normal

students. Students whose concern level is light, medium, and heavy in the psychological state table of students were taken as positive samples, and other students who were not reported were taken as normal students.

3.1.2. Feature Extraction. This section mainly describes the process of extracting student behavior characteristics from data sources. The behavior of students on campus is rich and colorful. We divide the behavior characteristics of students into four types: consumption characteristics, access control characteristics, Internet access characteristics, and performance characteristics [9].

(1) Consumption Characteristics. Some studies have found that people with mental health problems will have eating disorders, especially depression [19]. Based on this fact, the consumption records of students in the canteen were extracted through the shop names to analyze their eating patterns. Pay special attention to students' breakfast/lunch/dinner routines. Set breakfast from 6 to 9, lunch from 11 to 13, and dinner from 17 to 19:30. Since there are often multiple records for each meal, use the time of the first swipe as the meal time. For example, if there are three records in a breakfast, the time of occurrence is 7:20, 7:21, and 7:22, use 7:20 as the breakfast time. The regularity of a behavior can be considered repeatable and will be measured by entropy of the probability that the behavior occurs at a particular time interval. Assuming that there are n time intervals $T = \{t_1, t_2, t_3, \dots, t_n\}$, for any given student, the probability $p_v(T = t_i)$ of behavior $v \in V = \{\text{breakfast, lunch, dinner}\}$ occurring within time interval t_i is calculated according to formula (12), where $n_v(t_i)$ represents the frequency of behavior v occurring within time interval t_i . Then, calculate the entropy, which is calculated by formula (13). Assume that for the three behaviors of morning/afternoon/dinner, each time interval span is half an hour.

$$P_v(T = t_i) = \frac{n_v(t_i)}{\sum_{i=1}^n n_v(t_i)}, \quad (12)$$

$$E_v = - \sum_{i=1}^n P_v(T = t_i) \log P_v(T = t_i). \quad (13)$$

The $b_entropy$, $l_entropy$, and $d_entropy$ values for breakfast/lunch/dinner can be calculated from formulas 12 and 13. Meanwhile, according to formula (12), the smaller the entropy value of a behavior is, the more concentrated its probability distribution is over time, and the higher its regularity is. If one student only goes to the canteen occasionally during a concentrated period of time, while the other student goes to the canteen frequently, although the regularity of the two students' dining in the canteen is not the same, they will also have similar entropy values. In order to better distinguish the two types of students, the number of breakfasts/luncheons/dinners for each student was calculated separately. k -means clustering algorithm is used to cluster entropy and meal times. It is considered that students

have three common dining behaviors: less frequent dining, poor dining regularity, and good dining regularity. Therefore, the number of clustering k is set as 3. The clustering results are shown in Figure 3.

The goal was to find students with abnormal eating habits. It is assumed that the smaller the cluster a student is in, and the further away from the center of the cluster, the higher the abnormal score (20) according to the formula to calculate the abnormal scores for students, which is c_i student of n in the cluster centroid of C_i , said all the students, $|c_i|$ clustering C_i number of students, $|C|$ represents all the students, students n the distance to the center of mass of $(n, c_i)c_i$ said. In this experiment, Euclidean distance is used to calculate the distance.

$$AS = D(n, c_i) * \left(1 - \frac{|C_i|}{C}\right). \quad (14)$$

(2) Characteristics of Access Control. It is found that the more difficult the interpersonal communication is, the worse the mental health level is, and the stronger the interpersonal communication ability is, the better the mental health level is [5]. The dormitories of undergraduate students in colleges and universities are usually divided according to classes, and students in the same dormitory generally have the same entry and exit schedule. The interpersonal relationship of students in the dormitory can be reflected by calculating the times of students and roommates entering and leaving the dormitory together, which is called card swiping in front and back foot. The rule is as follows: if two students are roommates, the difference in credit card time is less than 20 seconds, and both students enter or leave the dormitory at the same time. Further analysis of the card swiping data from front and back showed that some students' values were less than 10 or even 0.

(3) Characteristics of Internet Access The overall framework of the one-dimensional convolutional neural network designed in this study consists of five neural network layers. In the training of the model, the input of each student is a network sequence, that is, a one-dimensional vector. Since the length of the network sequence is inconsistent, 0 is used to fill it. The operation of the convolution layer is shown in formula (15), where x_j^l is the feature graph of the j convolution operator of the l layer, x_i^{l-1} represents the input, k_{ij}^l represents the j convolution kernel of the l layer, b_j^l represents the bias of the j convolution operator, $f(\cdot)$ represents the activation function, and $*$ represents the multiplication operation.

$$x_j^l = f \left(\sum_{i \in M_j} x_i^{l-1} * k_{ij}^l + b_j^l \right). \quad (15)$$

The second and fourth neural network layers adopt the pooling layer, and the pooling mode is divided into average pooling and maximum pooling. In this study, the pooling mode is maximum pooling. The operation of the pooling

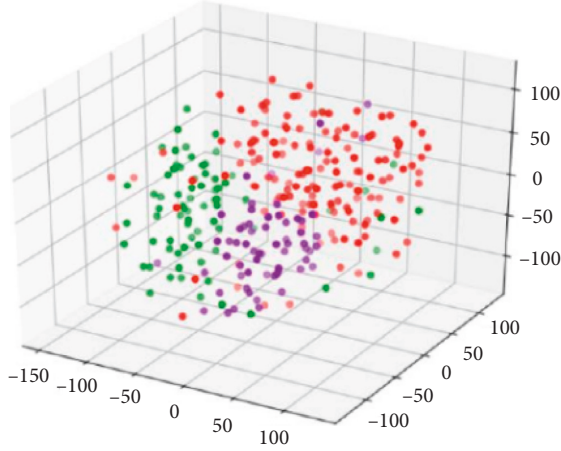


FIGURE 3: Clustering results based on entropy and frequency of breakfast, lunch, dinner.

layer is shown in formula (16), where x_j^i , w_j^i , b_j^i , $\text{down}x_j^i$ represents the input, weight matrix, bias, and downsampling function, respectively.

$$\hat{x}_j^l = f(w_j^l) \text{down}(x_j^l) + b_j^l. \quad (16)$$

The fifth layer uses a full connection layer, and the operation of the full connection layer is shown as follows:

$$s^{\text{pat}} = f(w^{(5)} * x^4 + b^{(5)}). \quad (17)$$

For parameter setting, the number of convolution kernels of the two convolution layers is 16 and 32, respectively. ReLU was used as the activation function and Adam as the optimization algorithm. Additionally, to prevent overfitting, we use three dropout layers with parameters 0.15, 0.15, and 0.5. In the model training stage, 70 positive samples and 70 negative samples were used to train the model. In the feature extraction stage, all the experimental samples were input into the trained 1D-CNN model, and finally the results of the full connection layer spat were output as the online mode.

(4) *Performance Characteristics.* (1) Grade point average students' mental health problems are often caused by external factors, and grade pressure is one of the external factors, so it is necessary to extract grade-related characteristics to reflect students' performance. GPA plays a decisive role in the school's evaluation and award, and the calculation formula is as follows:

$$\text{GPA}_{\text{avg}} = \frac{\sum_{i=1}^n \text{CGPA}_i}{\sum_{i=1}^n \text{Sore}_i}. \quad (18)$$

(2) Failing credits for college students: failing credits are the best characteristic of performance pressure. Failing credits are counted by traversing a student's history score sheet and counting the corresponding credits as failing credits if the QDFS value is a make-up exam. Two characteristics, GPA and failing credits, are extracted from historical achievement data. Finally, a total of 69 features

were extracted from consumption data, access control data, and network logs and historical performance data.

3.1.3. *Model Training and Recognition.* The research objective of this paper is a binary classification problem. Five common classification algorithms are selected as candidate algorithms, including random forest (RT), gradient ascending tree (GBDT), Naive Bayes (NB), neural network (NN), and decision tree (DT). We hope to select a classification algorithm with the best performance as the classifier of the algorithm.

The data set is divided into training set, verification set, and test set, and each classification algorithm is trained with the training set, and then each classification algorithm is evaluated with the verification set. Obfuscation matrix is an index to evaluate model performance, which is mainly used to judge the performance of classifier. The confusion matrix has four basic elements, which are True Positive (TP), False Negative (FN), False Positive (FP), and True Negative (TN). The meanings of these four elements will be introduced in detail below.

However, the obfuscation matrix is only a statistical quantity. When the sample number is very large, it is difficult to evaluate the performance of the model only by calculating the quantity. Therefore, the confusion matrix extends some evaluation indexes based on the statistical quantity. In educational data research, there are five commonly used evaluation indicators, including accuracy, precision, recall, F1-measure, and AUC (area under curve) [19]. In this experiment, precision, recall, and F1-measure were selected as evaluation indexes.

Precision refers to the proportion of students who are truly positive samples among the students who are predicted to be positive samples, and its calculation formula is shown as follows:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}. \quad (19)$$

Recall refers to the proportion of students who are predicted to be positive samples among students who are real and positive samples, and its calculation formula is shown as follows:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}. \quad (20)$$

Precision and recall are two important indicators. However, when evaluating the performance of several models, precision of one model is often higher than that of other models, but recall is lower than that of other models. Therefore, the optimal model cannot be selected. A comprehensive index F -measure was introduced, and precision and recall were weighted and averaged. The calculation formula is shown as (21). The larger the F -measure is, the better the model performance is. When α is equal to 1, it is commonly used $F1$ -measure, and $F1$ -measure is used in this study.

$$F - \text{Measure} = \frac{(\alpha^{2+1})\text{Precision} * \text{Recall}}{\alpha^2 * \text{Precision} + \text{Recall}} \quad (21)$$

The training process of the model is the process of parameter selection; how to choose the optimal parameter is a key problem. In this algorithm, the way of permutation and combination is used to select the optimal parameter combination. For each classification algorithm, we adjust only a few important parameters and use the default values for the rest parameters. After trying different parameter combinations, the optimal parameter combinations of each candidate classification algorithm are selected.

4. Experiment and Result Analysis

4.1. Experimental Data. Among the 280 undergraduates who participated in the survey, 70 students were assessed as students with mental health problems by experts at the university's psychological center in May and June 2019. Their mental health levels were classified as mild, moderate, or severe. In addition, 210 undergraduates were randomly selected from the entire school. Finally, the proportion of students at each mental state level is shown in Figure 4.

Students were divided into two groups, with a positive sample for mild, moderate, and severe cases, and a negative sample for students without mental health problems. The label has high credibility and avoids the drawbacks of concealing facts in questionnaires. In order to protect students' privacy, the student id was encrypted throughout the experiment. As the psychological state is a short-term state, generally lasting from half a month to one month, only the consumption data access control data and network logs in April 2021 are used.

4.2. Training Set and Test Set. The essence of deep learning is the process of model selection. However, only parameters of the model are unknown. By using various tuning methods, an optimal value can be obtained, which enables the model to map independent variables to dependent variables well. In practice, it is not always easy to sample data evenly. So, the actual process is random. In this experiment, the data set sample is 280 students, among which 70 students are positive samples and 210 students are negative samples; that is to say, the ratio of positive samples to negative samples is 1 : 3. In order to make the sample ratio of the training set and the test set 1 : 3, 49 and 147 students were randomly selected from the positive and negative samples of the data set, respectively, as the training set, and 25% of the data were used as the verification set, and the positive and negative sample ratio was 1 : 3. Meanwhile, the ratio of positive and negative samples in the test set was 1 : 3.

4.3. Comparison and Analysis of Experimental Results. The extracted features are input into the classification algorithm for model training and prediction. Five common machine learning classification algorithms, including random forest (RT), gradient ascending tree (GBDT), Naive

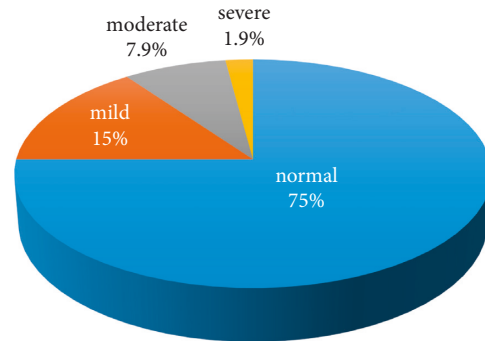


FIGURE 4: Proportion of each mental state in the data set.

Bayes (NB), neural network (NN), and decision tree (DT), are compared. The experimental results are shown in Table 6.

First of all, it can be obviously observed that in all models, precision is higher than recall, and recall is only 0.58 at its highest, indicating that few samples can be correctly identified. Secondly, decision tree has the best comprehensive performance, especially recall. Therefore, we choose decision tree as the classifier of mental health recognition algorithm based on multisource data.

To verify the generalization ability of the algorithm, we input the test data set into our algorithm and obtain the following experimental results: precision 0.68, recall 0.56, and $F1$ -measure 0.67. Based on the test set experiment results, our algorithm was able to identify 56 percent of students with mental health problems. Table 7 lists 20 characteristics with high Pearson's correlations.

It can be seen from the table that 7 of the first 8 positive correlation characteristics are related to early rise. In general, students' courses are arranged by week, and students have the same class time every week. The larger standard deviation or average value of early rise may be caused by lateness or truancy. There was a positive correlation with mental health problems. Working day and day of rest time to stay in the dormitory all show the negative correlation. There is a negative correlation between the time spent in the dormitory on weekdays and rest days. The students spend more time in the dormitory and less time in club activities, student work, and dinner parties. These students have poor interpersonal communication and are more likely to have mental health problems, consistent with people's intuitive feelings. This kind of students have generally poor interpersonal, namely, interpersonal communication; the worse the student, the more possible mental health problems, which is consistent and intuitive feeling of people. Students who spend more money in the canteen on weekdays and holidays may seldom participate in dinner parties, which may also reflect poor interpersonal communication. In addition, the correlation coefficient of students' swiping times is high and negative. On the one hand, students who swiped cards less often participated in fewer activities and had poorer interpersonal communication. On the other hand, students who swipe their credit cards only a few times a month should live off-campus and be the focus of attention. The insomnia rate (correlation coefficient 0.1058), the number of takeout orders (correlation coefficient 0.1024), GPA_avg, and failed

TABLE 6: Results of five classification algorithms to identify students with psychological problems.

Sorting algorithms	Precision	Recall	F1-measure
RT	0.68	0.53	0.65
GBDT	0.65	0.52	0.62
NB	0.62	0.50	0.60
NN	0.65	0.50	0.63
DT	0.68	0.58	0.69

TABLE 7: The top 20 most relevant features.

Characteristics of the category	Characteristic	Signalment	Coefficient of association
Entrance guard	Getup std 1	Standard deviation of Monday morning	0.2627
Entrance guard	Reststayroom	Time spent in the dormitory on days off	-0.2290
Entrance guard	Swipe_num	Credit card number	-0.2159
Entrance guard	getup_mean6	Saturday morning average	0.2014
Entrance guard	last_count5	Number of Friday nights	0.1990
Entrance guard	getup_mean1	Monday morning average	0.1915
Entrance guard	getup_count7	Number of early Sundays	0.1830
Entrance guard	getup_mean4	Thursday morning average	0.1780
Entrance guard	getup_mean3	Wednesday morning average	0.1766
Entrance guard	getup_std3	Standard deviation of Wednesday morning	0.1761
Consume	lunch_mean	The average amount spent on lunch	0.1825
Entrance guard	Work stay room	Time spent in the dormitory on weekdays	-0.1564
Entrance guard	last_std6	Saturday night is standard deviation	0.1540
Entrance guard	Feetcard	Swipe cards front and back	0.1440
Consume	rest_consume	Rest day consumption amount	-0.1424
Consume	Work consume	Consumption amount on working days	-0.1314
Score	GPA avg	Grade point average	-0.1244
Score	guake_count	Failing credit	0.1218
Consume	abnormal_score	Abnormal scores	0.1209
Network	cnn	The internet features	-0.1149

credits were all weakly correlated, but still valuable for analysis.

Through the above analysis, some characteristics related to mental health problems are obtained, and the impact on mental health is in line with common sense. Moreover, these related features come from multiple data sources, which shows that the use of multisource data can improve the recognition effect from one aspect.

5. Conclusion

This paper proposes a mental health problem recognition algorithm based on multisource data, trains five common classification algorithms, evaluates through the verification set, and establishes a recognition system to judge the psychological emotions of college students, and analyzes the results of the system verification test. The following conclusions can be drawn:

- (1) Analysis of student behavior data can be used to identify students with mental health problems.
- (2) Using a variety of behavioral data can improve the identification effect of students' mental health problems and can identify 75% of students with mental health problems.

To establish college students' mental emotion recognition system, this paper through the prediction results can

identify the students with risk of mental health problems and counselors for school administrators to better understand the students' psychological condition, giving psychological help to the student as soon as possible and targeted intervention method provides the theory support, able to prevent the further deterioration of the psychological condition of students.

Data Availability

The dataset can be accessed upon request to the corresponding author.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the publication of this paper.

Acknowledgments

This research was supported by Distinguished Teachers' Studio of Ideological and Political Course of Hebei Province.

References

- [1] A. M. Sylvia, C. O. Desmond, M. Rucha, and Z. Jamil, "Empathy and well-being correlate with centrality in different

- social networks[J],” *Proc Natl Acad Sci U S A*, vol. 114, no. 37, pp. 9843–9847, 2017.
- [2] A. Mc and J. B. Suo, “Analysis and research on mental health of college students based on cognitive computing[J],” *Cognitive Systems Research*, vol. 56, pp. 151–158, 2019.
- [3] A. F. Jorm, “Mental health literacy: e,” *American Psychologist*, vol. 67, no. 3, pp. 231–243, 2012.
- [4] D. Nie, A. Li, Z. Guan, and T. Zhu, “Your search behavior and your personality[J],” *Pervasive Computing and the Networked World*, vol. 8351, pp. 459–470, 2014.
- [5] P. Paola, N. Maren, Y. Albert, Z. Courtney, and W. Timothy, “Colleg students: mental health problems and treatment considerations[J],” *Academic Psychiatry*, vol. 39, no. 5, pp. 503–511, 2015.
- [6] N. Min, Y. Lei, J. Sun et al., “Advanced forecasting of career choices for college students based on campus big data[J],” *Frontiers of Computer Science*, vol. 12, no. 3, pp. 494–503, 2018.
- [7] M. Yari Zanganeh and N. Hariri, “The role of emotional aspects in the information retrieval from the web,” *Online Information Review*, vol. 42, no. 4, pp. 520–534, 2018.
- [8] Z. Lee and C. Lee, “A parallel intelligent algorithm applied to predict students dropping out of university[J],” *The Journal of Supercomputing*, vol. 76, no. 2, pp. 1049–1062, 2020.
- [9] J. L. Tang, D. Wang, Z. G. Zhang, U. He, J. Xin, and Y. Xu, “Weed identification based on K-Means feature learning combined with convolutional neural network[J],” *Computers and Electronics in Agriculture*, vol. 135, pp. 63–70, 2017.
- [10] B. Rupali, B. Machhindra, W. Pragati, M. V. Raut, and H. S. Habeebullah, “Tracking and predicting student performance using machine learning[J],” *SSRN Electronic Journal*, vol. 6, no. 2, pp. 439–442, 2019.
- [11] Y. M. Wang, C. He, and S. J. Tong, “The present situation and educational enlightenment of college students’ mental health development[J],” *Modern educational science (higher education research)*, vol. 2, pp. 65–71, 2018.
- [12] L. Q. Hao, “Analysis of data mining method based on related rules[J],” *Journal of Taiyuan University*, vol. 38, no. 03, pp. 42–45, 2020.
- [13] C. Y. Zhu, B. B. Li, and A. Li, “Predicting depression from internet behaviors by time-frequency features[J],” *Web Intelligence*, vol. 17, no. 3, pp. 199–208, 2019.
- [14] M. Y. Zanganeh and N. Hariri, “The role of emotional aspects in the information retrieval from the web [Online, 2018 web[J],” *Online Information Review*, vol. 42, no. 4, pp. 520–534, 2018.
- [15] Y. M. Shi and Y. J. Luo, “The attentional bias of optimistic and pessimistic attributive style college students to emotional information[J],” *Chinese Journal of Mental Health*, vol. 5, pp. 395–399, 2017.
- [16] X. F. Chen, M. Tong, C. Shi et al., “Visual analysis of multi-source college students’ mental health questionnaire data[J],” *Journal of Computer Aided Design and Graphics*, vol. 2, no. 32, pp. 181–193, 2020.
- [17] H. Yao, D. Lian, and Y. Cao, “Predicting academic performance for college students: ACampus behavior perspective [J],” *ACM Transactions on Information Systems*, vol. 10, no. 3, pp. 24.1–24.21, 2019.
- [18] P. R. Selvaraj and C. S. Bhat, “Predicting the mental health of college students with psychological capital[J],” *Journal of Mental Health*, vol. 27, no. 3, pp. 279–287, 2018.
- [19] C. Shu, J. Ma, and J. Zhu, “Comparison of UPI and SCL-90 on mental health survey of freshman in higher vocational college [J],” *Chinese Journal of School Health*, vol. 35, no. 4, pp. 523–526, 2014.