# Data Collection and Analysis of Psychological Health Signs of College Students Based on Artificial Intelligence Technology

Yongming Mao[1*], Xiaoxiao Ding[2]
[1]Media College, Zhengzhou University of Technology, Zhengzhou 450044, China
[2]School of Marxism, Henan University, KaiFeng 475001, China
E-mail: 13526772342@163.com, 1392323005@qq.com
[*]Corresponding author

**Keywords:** mental health, CNN, GRU, knowledge graph, depression

**Received:** Septembr 25, 2025

*The problem of mental health concerns among university students is growing more and more noticeable, and early identification of mental health signs is crucial for intervention. Current assessment methods rely on questionnaire data, which suffer from inefficiencies in data statistics. Taking Rednote as a case study, this research employs web crawling technology to gather diverse user information. By utilizing data mining techniques, linguistic, emotional, and behavioral features are extracted, and a depression knowledge graph is constructed to model the deep-level associations among these features. Additionally, the study selects a Convolutional Neural Network (CNN) and a Gated Recurrent Unit (GRU) to develop a predictive model that integrates textual features with user characteristics, ultimately yielding mental health predictions. The results indicate that the predictive model achieves an accuracy of 87.3% and an F1 score of 86.5%. Without the knowledge graph, the accuracy drops to 81.2%, representing a 6.1% decrease compared to the predictive model, demonstrating that the knowledge graph can effectively identify key depression-related pathways. The F1 score for the combined CNN and GRU model reaches 98.5%, showing a 6.3% improvement over the GRU alone. This research provides a feasible approach for social media data-driven mental health monitoring, with the application of the knowledge graph enhancing model interpretability and aiding in the development of precise psychological intervention strategies. It offers an automated screening tool for college psychological counseling services and provides data support for depression prevention research in the realm of public health.*

*Povzetek: Predstavljen je celovit, podatkovno podprt sistem za zgodnje zaznavanje depresije pri študentih, ki z analizo družbenih omrežij združuje graf znanja o depresiji (DKG) in model CNN-GRU.*

## 1 Introduction

With the development of society, the mental health of college students is increasingly being valued. Depression, as a common mental health problem, is quite common among college students. However, severe depression poses great harm and can lead to students' mental exhaustion, affecting their daily learning and life [1]. Early identification of psychological health signs is of great significance for preventing psychological problems such as depression among college students. Depression, as a frequently encountered mental health issue, often occurs and develops with subtle changes in language patterns, social behavior, and other aspects. These changes are often intuitively reflected in daily social media interactions. Such interactions provide new data dimensions for AI-based psychological health analysis [2]. The current research on mental health mainly focuses on the analysis of explicit emotional expressions in the use of social media data, while systematic mining of the broader concept of mental health signs is still insufficient [3]. Psychological health signs not only include typical symptoms of depression, but also encompass multiple indicators such as sleep quality, social frequency, and self-awareness patterns, which together form a panoramic picture of the psychological state. Most of the existing analysis methods adopt a single mode data processing method, which is difficult to capture the synergistic mechanism between different signs. Additionally, there is a clear explanatory gap between traditional psychological measurement tools and real social behavior data, which limits the generalization ability of prediction models [4]. Convolutional Neural Network (CNN) automatically learns local features in text and extracts global semantic information through multi-layer convolution operations.

Many scholars have adopted various machine learning algorithms to assess the mental health of college students. For example, Krishnan et al. [5]collected data from an online questionnaire using the Hospital Anxiety and Depression Scale (HADS) and applied Support Vector Machine (SVM) to analyze the differences in mental health status between engineering and medical college students during the COVID-19 pandemic. This study relied entirely on online questionnaires published on social media, and the results showed that the prediction accuracy of this method reached 100%. Zhai [6] identified high-risk populations for anxiety and depression disorders among American college students and developed a prediction model using eXtreme Gradient Boosting (XGBoost), Random Forest (RF), decision tree, and logistic regression. The results showed that the Area Under Curve (AUC) of the model reached 0.77. Cai et al. [7] proposed a particle swarm optimization based Light Gradient Boosting Machine algorithm to analyze the mental health assessment data of college students, and evaluated the actual situation based on the symptom checklist 90 (SCL-90). The results showed that the F1 score of the algorithm was 97.90%, and the AUC was 99.45%. Tyulepberdinova et al. [8] used RF algorithm to analyze physiological indicators such as blood pressure and body mass index, as well as various psychological scales, for predicting the physical and mental health status of college students. The findings demonstrated that the accuracy of the algorithm reached 99.4%, and the F1 score reached 98.7%. Bhavani and Naveen [9] built an RF prediction model to analyze the factors influencing the mental health of international students. The model was used to analyze their living conditions and mental health data, and the findings demonstrated that the prediction accuracy of the model reached 80%.

With the development of deep learning technology, there are increasingly significant achievements in its innovative applications in the field of mental health. Ji [10] designed a student mental health system that combines backpropagation neural network and deep fuzzy neural network to analyze mental health status and predict crises, addressing the complexity and uncertainty of mental health data. The results showed that the system had accuracy and effectiveness. Ming [11] proposed a text processing algorithm based on metaphorical attention mechanism using CNN and Bidirectional Long Short-Term Memory (BiLSTM) to improve the accuracy of assessing mental health issues among college students. The results showed that the algorithm improved F1 score and recall rate by 4.04% and 6.52%, respectively. Li [12] addressed the issue of capturing dynamic changes in students' psychological states by combining a Transformer based multimodal feature extraction network and an improved variational autoencoder with an LSTM model. The results showed that the model had high efficiency and accuracy. Marco et al. [13] used decision tree algorithm to analyze gene metabolism status in patients with severe depression in order to predict drug treatment efficacy. The findings demonstrated that the model had a prediction accuracy of up to 70%. Zhang [14] addressed the problem of insufficient early screening for global adolescent mental health issues by integrating electronic health records and neuroimaging data from over 50000 adolescents. They used a hybrid architecture combining CNN spatial feature extraction and LSTM time series modeling, and the results showed that the accuracy and AUC of the architecture were 95% and 97%, respectively. Although complex CNN and LSTM can capture subtle patterns, they are also more prone to overfitting to the noise of training data, leading to performance degradation on independent external validation sets.

In summary, existing research has laid an important foundation for the assessment of college students' mental health based on AI technology. However, these studies rely on single data, and traditional machine learning methods have limited ability to handle serialized, unstructured social media text. Meanwhile, existing deep learning models have a high level of complexity. Therefore, the study innovatively used the Rednote platform as data support and designed a Depression Knowledge Graph (DKG) from the perspective of college students' depression tendency. Additionally, a Mental Health Prediction (MHP) model based on CNN and Gated Recurrent Unit (CNN-GRU) was developed to analyze the psychological health signs of college students and provide psychological counseling and intervention. The research aims to improve MHP efficiency and provide an effective technical tool for university psychology departments.

# 2 Methods and materials

## 2.1 Design of DKG for collecting data on psychological health signs of college students

The study selected the platform Rednote as the dataset to obtain and analyze the psychological health signs of college students from user posts. Rednote, also known as China's Instagram, is a social media and e-commerce platform that serves young people. All data collected by this research came from users' public posts on the Rednote platform, and no technical means have been taken to obtain non-public or privacy protected content. To maximize the protection of user privacy, all data were thoroughly anonymized immediately after collection. Specific measures include: removing all identifiable personal information such as usernames, nicknames, avatars, device IDs, precise geographic locations, etc., and conducting strict cleaning and aggregation analysis on text data. At any stage of research, it is impossible to trace back to specific individuals. All data usage activities have not affected the normal operation of any data source platform, nor have they infringed upon the personal rights of any users.

The research focuses on depression and selects various data. The study selected 600 users from social media platforms and obtained 125000 personal profiles and text data on Rednote through web crawling technology. After data preprocessing, 67500 valid Rednote posts were obtained [15]. To implement MHP analysis for college students, the DKG was designed with a focus on depression. The basic composition diagram of DKG is shown in Figure 1.

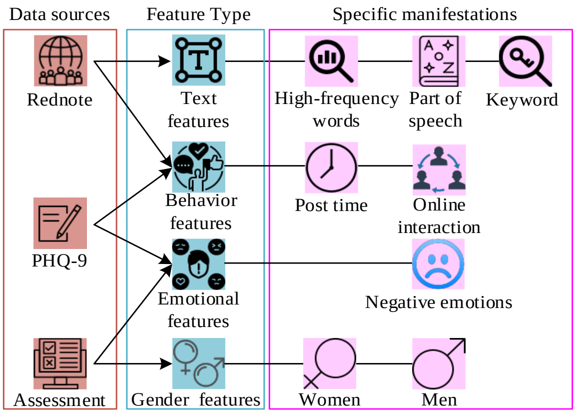



Figure 1: Basic composition diagram of DKG.

In Figure 1, DKG has three levels, namely data sources, feature types, and their specific manifestations. The data sources include the Rednote social media platform, the Patient Health Questionnaire-9 items (PHQ-9) scale, and professional psychological assessment institutions. Social media platforms provide authentic and dynamic records of user behavior, capturing daily language expressions and interaction patterns, and compensating for the lag and subjective bias of the PHQ-9 scale. The PHQ-9 scale, as a standardized depression screening tool, has structured data that can validate the clinical relevance of social media features [16]. Professional psychological assessment institutions provide diagnostic level labels to enhance the reliability and validity of the model. From raw data such as text and scales to abstract features such as emotions and behaviors, and then to specific expressions such as high-frequency words and posting time, an interpretable semantic network DKG is constructed.

The entity used by DKG is the user, and the ontology consists of four main feature types. This pattern defines the asymmetric correlation between features. There are four types of features: text features, behavioral features, emotional features, and gender features. Among them, the specific manifestations of text features are high-frequency words, part of speech, and keywords. The language patterns of users with depressive tendencies are specific, such as an increase in first person pronouns and a higher frequency of negative vocabulary. Behavioral features denote the observable and quantifiable behavioral patterns of users on the platform, which are believed to have potential associations with depressive tendencies. Its specific manifestations include posting time and online interaction. This is because depressed patients often exhibit nighttime activity, so the behavior of posting time can be used to determine the user's depression status. Due to depression and nocturnal activity, the circadian rhythm is disrupted, resulting in a sudden decrease in the frequency of personal posting behavior and a reduction in online interaction behavior [17]. Emotional features refer to the emotional tendencies and expression patterns extracted from the text content posted by users through natural language processing technology. Its specific manifestation is negative emotions. The core symptom of depression is persistent low mood, and sentiment analysis of social media texts is significantly correlated with PHQ-9 clinical depression scores. The specific manifestations of gender characteristics are female and male. According to data from the official website of the World Health Organization (WHO), the prevalence of depression among women worldwide is approximately 1.5-2 times that of men, with women accounting for about 5.1% and men accounting for about 3.6% [18].

## 2.2 Multi dimensional feature processing and vectorization based on DKG

After designing DKG, due to the existence of multidimensional features, the study processed different features separately. According to the feature information extracted from DKG, college students with depression tendencies in DKG often use negative vocabulary in text features, and the frequency of using first person is much higher than that of normal users. The flowchart of multidimensional feature processing based on DKG is shown in Figure 2.

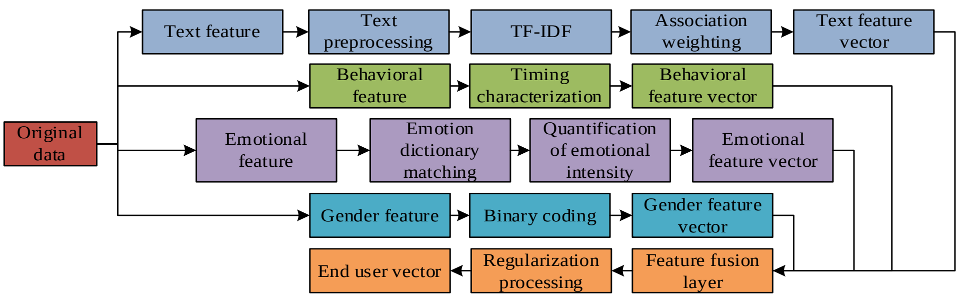



Figure 2: Flowchart of multi dimensional feature processing based on DKG.

In Figure 2, the raw data obtains four types of features: text features, behavioral features, emotional features, and gender features. After preprocessing the text features for word segmentation, the process encodes the semantic information of the text in vector form using Term Frequency Inverse Document Frequency (TF-IDF). Afterwards, feature weighting is applied to the word vector, specifically by assigning different weights based on the degree of association between the word and depression to construct the user's text feature vector. TF-IDF is a weighting technique commonly used in information retrieval and natural language processing, which combines two metrics: Term Frequency (TF) and Inverse Document Frequency (IDF) to assess the significance of a word within a document [19]. The expressions for TF, IDF, and TF-IDF are shown in equation (1).

$$\begin{cases} \mathrm{TF}_{kn} = \dfrac{c_{k,n}}{\sum_j c_{j,n}} \\ \mathrm{IDF}_{kn} = \log \dfrac{|F|}{1+|n : e_k \in f_n|} \\ \text{TF-IDF}=\mathrm{TF}_{kn} \cdot \mathrm{IDF}_{kn} \end{cases} \tag{1}$$

In equation (1), $\mathrm{TF}_{kn}$ is the frequency of the entry $e_k$ appearing in the file $f_n$. $\mathrm{IDF}_{kn}$ means that if there are fewer documents containing entries and the IDF is larger, it means that the entries have a good ability to distinguish categories. $c_{k,n}$ is the number of occurrences of entry $e_k$ in file $f_n$, and $\sum_j c_{j,n}$ is the sum of the number of occurrences of all words in file $f_n$. $|F|$ is the total number of files in the corpus. $|n : e_k \in f_n|$ indicates the number of documents containing entries [20]. The text feature vector of a certain Rednote user is $T_k$, $T_k = (q_1 d_1, q_2 d_2, ..., q_u d_u)$. Among them, $q_u$ represents the feature weight of TF-IDF and $d_u$ represents the $u$ th dimension of the vector. For behavioral features, the study generates a feature behavior vector after sequential feature processing and normalization. It then processes emotional features using a Chinese sentiment lexicon for matching, followed by labeling of emotional polarity, with negative assigned as -1, neutral as 0, and positive as +1. The emotional intensity is quantified to generate an emotional feature vector [21]. Based on the processed text feature vectors, the study conducts emotional feature vectors, behavioral feature vectors, and gender feature vectors, feature fusion at the feature fusion layer. After regularization, the final user vector $Q_k$ is obtained, and the representation of $Q_k$ is shown in Equation (2). [22].

$$\begin{cases} Q_k = (q_s s_k + q_p p_k + q_a a_k + q_b b_k) \\ q_s + q_p + q_a + q_b = 1 \end{cases} \tag{2}$$

In equation (2), $q_s$ is the weight of the user's gender characteristics, $s_k$ is the user's gender, 0 for female and 1 for male. $p_k$ and $q_p$ are the frequency and weight of users' posting at night. $a_k$ and $q_a$ are online interactive data and their weights, $b_k$ and $q_b$ are user emotion categories and their weights.

### 2.3 Construction of MHP model based on CNN and GRU

DKG acts as a feature engineering and association framework, concatenating all processed multidimensional features (text, behavior, emotion, gender) in the feature fusion layer to form the final user vector representation, which serves as the input for the CNN-GRU model. This method effectively injects prior knowledge into feature representations without the need for complex graph embedding operations. By using users' text data and personal characteristics on Rednote, it is possible to predict whether they have a tendency towards depression, thereby achieving psychological health assessment. To ensure the full exploration of the richness of text information in Rednote, the study develops a MHP model combining CNN and GRU to better utilize text sequence relationships for feature extraction and classification. The schematic diagram of the structure of CNN and GRU is shown in Figure 3.

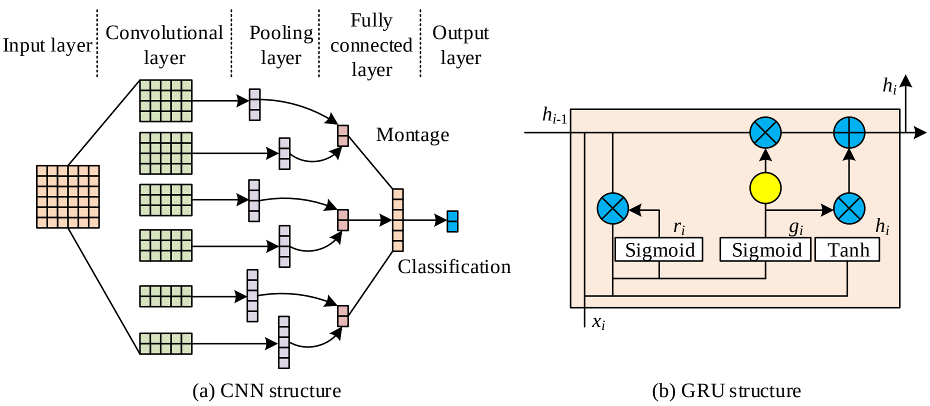



Figure 3: Schematic diagram of the structure of CNN and GRU.

In Figure 3 (a), the input layer of CNN consists of four features of the user in DKG, and the output variables are the MHP results of that user. The CNN designed in MHP only uses local connections, combined with weight sharing to compress and purify text features, to ensure the generalization of the training model, reduce the computational complexity of training, and accelerate the convergence of the entire network. In Figure 3 (b), the

GRU function of the MHP model is to further obtain hidden information from the text sequence and improve the quality of feature extraction. It can overcome the local field of view limitations of CNN and understand the logical connections across sentences. GRU consists of update gate $g_i$ and reset gate $r_i$, which dynamically control the flow of information through gating mechanisms. The update gate controls the fusion ratio of new and old information, while the reset gate determines how much historical information needs to be discarded.The expression of $r_i$ and $g_i$ is shown in equation (3) [23].

$$\begin{cases} r_i = \sigma\left\{q_r\left[h_{i-1}, x_i\right]\right\} \\ g_i = \sigma\left\{q_g\left[h_{i-1}, x_i\right]\right\} \end{cases} \quad (3)$$

In equation (3), $\sigma$ is the Sigmoid function. $q_r$ and $q_g$ are the weights of reset door and update door respectively. $x_i$ is the input at time $i$, and $h_{i-1}$ is the hidden state at the previous time $i-1$. The expression of hidden state $h_i$ at time $i$ is shown in equation (4).

$$h_i = (1-g_i) \odot h_{i-1} + g_i \odot h^* \quad (4)$$

In equation (4), $h^*$ is the candidate state. The reset gate determines the proportion of the incoming candidate state $h^*$ from $h_{i-1}$, and gets the reset data through the reset gate after obtaining the gating signal. The candidate state generates the temporary memory of the current step, and the expression of $h^*$ is shown in equation (5) [24].

$$h^* = \mathrm{Tanh}\left\{q\left[r_i \odot h_{i-1}, x_i\right]\right\} \quad (5)$$

In equation (5), $r_i \odot h_{i-1}$ represents the data after reset, $\left[r_i \odot h_{i-1}, x_i\right]$ indicates the concatenation operation with $x_i$ . Tanh is the activation function. On the basis of DKG, a MHP model structure was constructed by combining CNN and GRU, as shown in Figure 4. In Figure 4, the MHP model primarily consists of a DKG, a CNN, a GRU, and a Fully Connected Layer (FC). Initially, the textual features from the DKG are transformed into continuous vector representations and then fed into the input layer of the CNN for spatial convolution. The CNN functions by extracting local features and performing spatial hierarchical processing to capture key semantic patterns indicative of depression from the textual data. It employs convolutional kernels of varying sizes to scan the word vector sequences, detect specific phrase combinations, and retain prominent features while suppressing noise through max pooling. The semantics processed by the CNN are subsequently modeled sequentially in the GRU, which can handle the sequential dependencies in the text. Compared to LSTM, the GRU has a simpler structure, making it suitable for medium-sized datasets and helping to avoid overfitting. The model uses a CNN to predict depressive tendencies among Rednote users, with input variables comprising DKG-based user features and binary output labels (0/1) indicating depression status. The role of textual feature extraction is to identify depression-related keywords, phrases, and their contextual patterns. Pre-trained Tencent word vectors are employed to convert the text into dense vectors, capturing semantic associations. The convolutional layer uses filters of different kernel sizes to extract local textual features. Max pooling retains the most significant features from each convolutional window, reducing dimensionality and enhancing robustness. The FC in the MHP model is an independent network layer positioned after the joint CNN-GRU feature extraction and before the final classification. This approach integrates textual features (obtained from the CNN-GRU) with user-level features into a composite multidimensional vector. The vector is then transformed into a lower-dimensional representation using a weight matrix, ensuring retention of key information. The ReLU activation function is used to learn complex interactions among features [25]. The dimensionality of the final output matches the number of categories, specifically representing two-dimensional information: depressed and non-depressed.

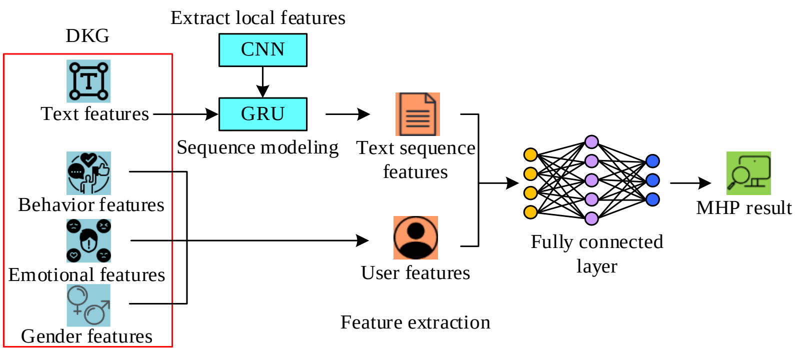



Figure 4: MHP model structure.

An automated warning and intervention system is designed based on MHP results. When the MHP model analyzes a student's social media data and their depression risk probability exceeds the preset threshold for one consecutive week, the system will automatically trigger a warning. This warning will be sent in real-time to authorized psychological counselors through an encrypted internal messaging platform linked to the campus counseling center. Upon receiving a system alert, the counselor will initiate proactive outreach by dispatching standardized care messages through the designated messaging platform. Within three working days, the counselor will extend psychological counseling appointment invitations to targeted students, guided by objective data indicators provided by the system. Subsequent counseling content follows standard psychological counseling ethics and procedures.

# 3 Results

## 3.1 Testing and analysis of DKG and CNN-GRU models

The hardware configuration for experimental testing was NVIDIA Tesla-A100-80G GPU, 16 core CPU, and 128GB DDR4 memory. The crawler framework selected was Scrapy+Selenium, and the deep learning framework Pytorch 2.0.1. The hyperparameters and training configurations of the model are shown in Table 1.

Table 1: Model hyperparameters and training configuration.

| Parameter category | Parameter name | Parameter value/description |
|---|---|---|
| CNN architecture | Convolution kernel size | [2, 3, 4] |
| | Number of convolution kernels per size | 128 |
| | CNN activation function | ReLU |
| | Pooling method | Maximum global pooling |
| GRU architecture | GRU number of plies | 2 |
| | Number of hidden cells per layer | 128 |
| Fully connected layer | Number of hidden layer neurons | 128 |
| | Number of output layer neurons | 2 |
| | Activation function | ReLU / Softmax (output layer) |
| Training parameters | Optimizer | Adam |
| | Learning rate | 0.001 |
| | Batch size | 32 |
| | Training rounds | 200 |
| | Dropout rate | 0.5 |
| | Loss function | Cross-entropy loss |

To ensure the statistical validity of the research results, independent sample *t*-test and repeated measures analysis of variance were used in this study. Cross validation calculations were conducted through 5 repeated experiments to evaluate the performance indicators of the model. All model performance metrics, such as accuracy, F1 score, etc., represented the performance of the model on an independent test set. To analyze the rationality of the selection of specific behavioral characteristics in DKG, the study analyzed the posting time and online interaction changes of depressed and normal users on the Rednote social media platform, in order to analyze the behavioral characteristics of different users. On the platform, the indicators of user followers, fans, likes, comments, and shares were important criteria for evaluating user activity and influence. To investigate the differences in online interaction between depressed users and normal users, the crawled user posts were statistically analyzed, and their number of followers and fans were recorded. The posting time and online interaction changes of the two types of users are shown in Figure 5.

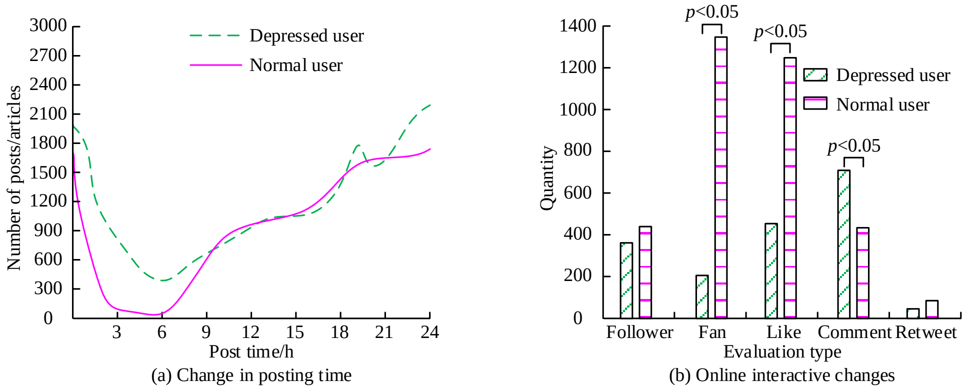



Figure 5: Posting time and online interaction changes.

In Figure 5(a), depressed users posted more messages than normal users between 0:00-7:00 and 22:00-24:00, while the number of posts by the two types of users showed little difference and no significant discrepancy during the 7:00-21:00 period. Specifically, depressed users posted an average of 350 more messages than normal users between 0:00-7:00. In Figure 5(b), depressed users had 216 followers and received 461 likes, whereas normal users had 1,348 followers and received 1,254 likes. The independent sample *t*-test results showed that the number of fans and likes of depressed users were significantly lower than those of normal users ($t$=3.21, $p$<0.05). Depressed users made 719 comments, reflecting an increased need for emotional catharsis, and their number of comments was significantly higher than that of normal users ($p$<0.05). However, there were no significant differences in the number of follows and reposts between the two types of users.

To evaluate the generalization ability of the MHP model, this study used a dataset collected and constructed from the Rednote platform as the unique training set to train the final MHP model. The data from Weibo, Zhihu, and campus forum platforms was only used as an independent test set and did not participate in any training or parameter adjustment of the model. The study examined the emotional distribution of datasets obtained from four platforms: Rednote, Weibo, Zhihu, and campus forums. Specifically, 26,000 data entries were collected from Weibo, 20,000 from Zhihu, and 12,500 from campus forums. The emotional distribution of datasets from different platforms is shown in Table 2.

Table 2: Emotion distribution in datasets from different platforms.

| Platform type | Dataset partitioning | Positive | Neutral | Negative |
|---|---|---|---|---|
| Rednote | Training set | 24000 | 18000 | 12000 |
| | Test set | 6000 | 4500 | 3000 |
| Weibo | Training set | 9000 | 6000 | 5000 |
| | Test set | 3000 | 2000 | 1000 |
| Zhihu | Training set | 6000 | 4500 | 4500 |
| | Test set | 2000 | 1500 | 1500 |
| Campus forum | Training set | 4800 | 3200 | 2000 |
| | Test set | 1200 | 800 | 500 |

Based on the obtained text data, the emotional categories of users were analyzed, and the proportion of basic emotions between depressed users and normal users was compared. Emotions were divided into happiness, sadness, anger, fear, disgust, surprise, and shame. The proportion of different emotional intensities is shown in Figure 6.

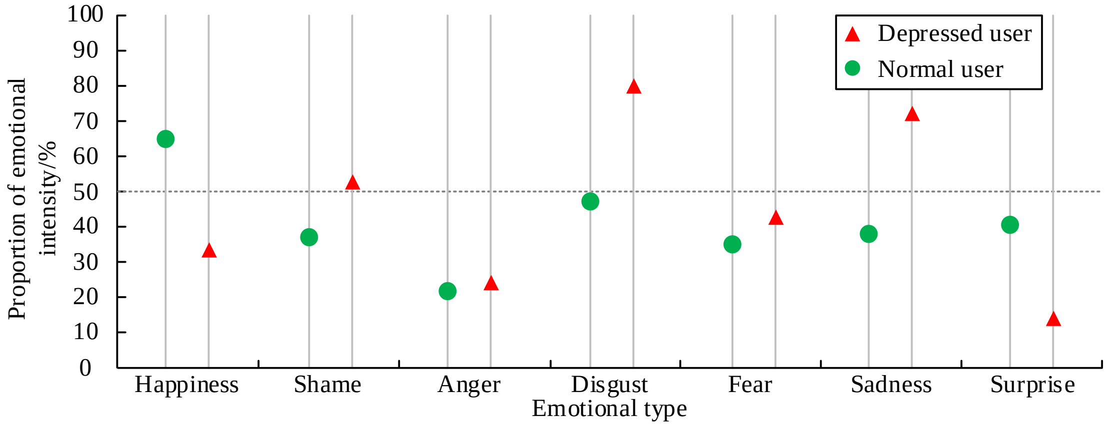



Figure 6: Proportion of intensity of different emotions.

In Figure 6, the proportion of sadness, disgust, and shame among depressed users exceeded 50%, accounting for 72.4%, 80.3%, and 53.1% respectively. These emotions were more consistent with the signature depressive emotions in the scale. In contrast, the proportion of sadness, anger, fear, disgust, surprise, and shame emotions among normal users was less than 50%, while happiness accounts for 65.4%. Normal users had a certain level of healthy emotional regulation.

The experimental parameters were set with a batch size of 32, a training iteration of 200, convolution kernel sizes of [2-4], a learning rate of 0.001, and a Dropout of 0.5. The study compared and analyzed GRU, Transformer, and CNN-GRU, and the results of their accuracy, recall, and F1 score changes are shown in Figure 7.

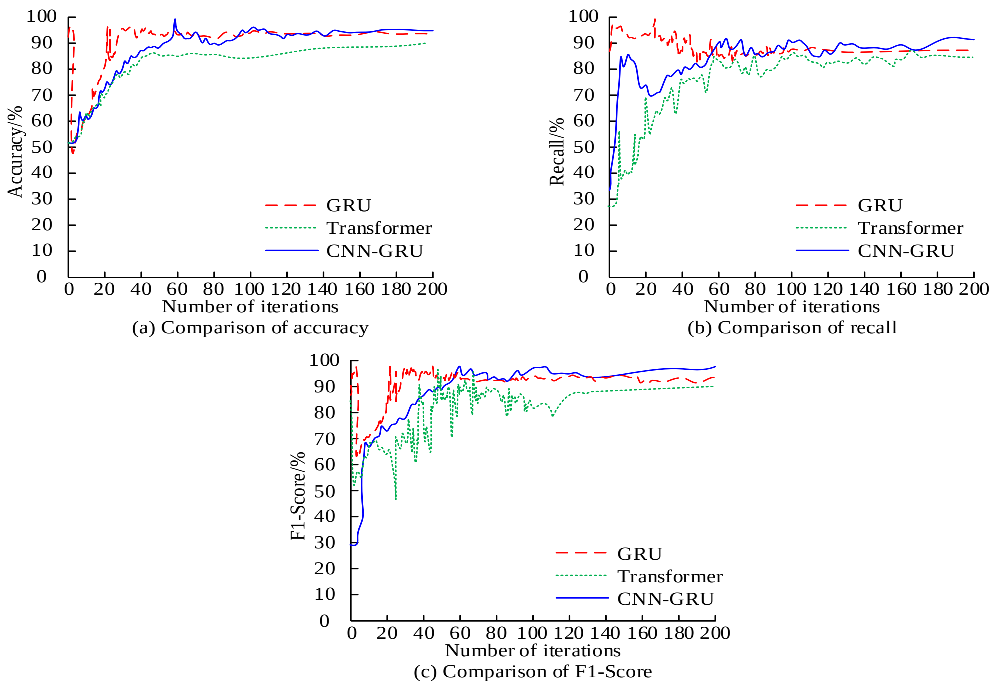


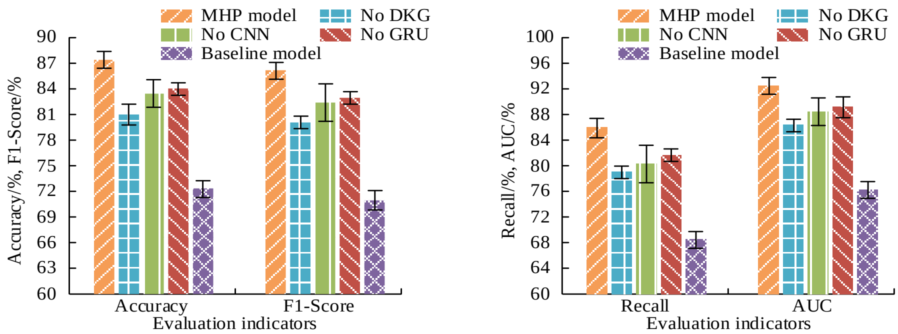

Figure 7: Accuracy, recall, and F1 score changes of different models.

In Figure 7 (a), the three models converged after approximately 110 iterations, and the overall GRU was lower than CNN-GRU, with an accuracy of about 92.8%. GRU had strong modeling ability for long sequences, but was susceptible to interference from irrelevant words. The accuracy of CNN-GRU increased rapidly in 200 iterations and remained stable at a high level of approximately 95.6% in the later stages. The accuracy of Transformer was only 89.3%. In contrast, CNN-GRU had stronger classification accuracy and recognition ability. In Figure 7 (b), the overall change in GRU recall rate was a fluctuating decrease, ultimately stabilizing at 87.7%. The recall rates of Transformer and CNN-GRU increased with the number of iterations, and the final recall rates were 85.1% and 90.4%, respectively. The advantages of CNN-GRU in local extraction and temporal features enabled better analysis of emotions, thus enabling MHP implementation. In Figure 7 (c), after 0-60 iterations, the F1 score of GRU was higher than that of CNN-GRU, but its variation curve fluctuated more significantly. After 60-200 iterations, the

F1 score of CNN-GRU was higher than that of GRU, reaching 98.5%, which was 6.3% higher than GRU's 92.2%. The F1 score of Transformer fluctuated in the first 110 iterations and ultimately reached about 90.1%. It had good global dependency capture ability, but required a large amount of training data to obtain good detection results.

## 3.2 MHP model performance analysis and application results

To verify the effectiveness of different modules in the MHP model, ablation experiments were designed. Each model configuration (including MHP and its variants) in the ablation experiment underwent 5 random initialization repetitions of training and testing to evaluate the robustness of the model. The MHP model included DKG+CNN-GRU+FC. The module without DKG included CNN-GRU+FC, the module without CNN included DKG+GRU, the module without GRU included DKG+CNN, and the baseline model was pure FC. The evaluation results of different models on the same dataset are shown in Figure 8.

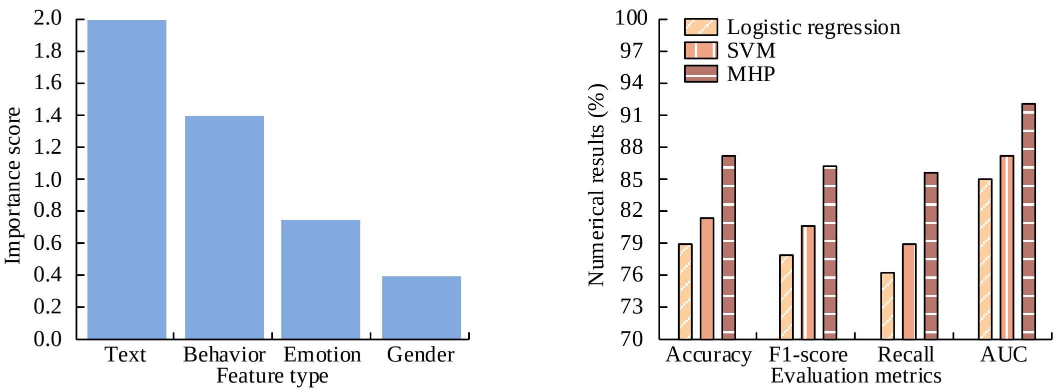


(a) Comparison of accuracy and F1-Score
(b) Comparison of recall and AUC

Figure 8: Results of ablation experiment.

In Figure 8 (a), the accuracy and F1 score of MHP model were 87.3% and 86.5%, respectively. Compared with MHP model, the baseline model showed the greatest decrease in accuracy and F1 score, with a decrease of 14.9% and 15.7%, respectively. The accuracy and F1 score without DKG decreased by 6.1% and 6.4% respectively compared to MHP model, indicating a significant contribution of DKG to feature association modeling. The F1 scores without CNN and GRU decreased by 4.1% and 3.5% respectively compared to MHP model. CNN in MHP model could detect negative phrases in text information, while GRU could analyze emotional evolution in text information. In Figure 8 (b), the recall rate and AUC of MHP model were 85.8% and 92.3%, respectively. Compared to MHP model, the recall rate of the baseline model decreased by 17.5%. The AUC without DKG was 86.1%. The AUC without GRU was 89.1%, which was 0.9% higher than the 88.2% without CNN.

To analyze the contribution of four types of features in DKG to MHP model decision-making, this study conducted feature importance analysis and trained the MGP model with logistic regression and SVM models on the same multi-dimensional feature vectors of DKG. The results of the two experimental analyses are shown in Figure 9.



(a) Comparison of Feature Importance Scores
(b) Model performance comparison

Figure 9: Feature importance analysis and model performance comparison.

In Figure 9 (a), the importance of text features was the highest, with a score of 2.0. Its ranking importance score was five times that of gender characteristics (0.4 points). This confirmed that language patterns were the strongest signals for identifying depressive tendencies on social media. The importance of behavioral characteristics was secondary, with a score of 1.4. Its importance was significantly higher than emotional foundation and gender characteristics, which validated the necessity of incorporating behavioral patterns into DKG. Emotional features also showed some importance, but their contribution was smaller than that of textual and behavioral features. The importance of gender characteristics was relatively low, but as prior knowledge of risk statistics, it still played an auxiliary adjustment role in the model. This analysis not only quantified the contributions of different features in DKG, but also made the decision-making process of the model more transparent. In Figure 9 (b), the proposed MHP model significantly outperformed logistic regression and SVM in accuracy, F1 score, and AUC. Compared with logistic regression (78.0%), the F1 score of MHP model (86.5) increased by about 8.5%. This confirmed the advantage of deep learning models in capturing complex nonlinear patterns in text sequences. Compared with SVM (79.2%), MHP model performed better in recall rate (85.8%) and could more effectively identify users with depressive tendencies. This comparison proved that introducing the CNN-GRU architecture to handle such problems could bring substantial performance improvements compared to classical methods that only use feature engineering.

The study selected a 6-month post sequence of 100 depressed college students from the Rednote platform. The changes in the proportion of nighttime posts, online interaction attenuation rate, negative word growth rate, and PHQ-9 growth rate of users are shown in Table 3.

Table 3: Changes in the temporal characteristics of users.

| Time (month) | Night post proportion (%) | Online interaction attenuation rate (%) | Negative word growth rate (%/week) | PHQ-9 increase |
|---|---|---|---|---|
| 1~2 | 12→18 | 0.05 | 1.2 | 1.5 |
| 3~4 | 18→34 | 0.12 | 3.5 | 4.8 |
| 5~6 | 34→51 | 0.21 | 6.7 | 9.2 |

In Table 3, during the observed months of January and February, the proportion of nighttime posts by this type of user increased by 6% ($p<0.05$), by 16% ($p<0.01$) from March to April, and by 17% ($p<0.1$) from May to June. Within 6 months, the online interaction attenuation rate of users gradually increased, the growth rate of negative words gradually accelerated, and the PHQ-9 growth rate increased.

To verify the direct cross platform generalization ability of the MHP model without domain adaptation, this study applied the final model trained on Rednote directly to independent test sets from Weibo, Zhihu, and campus forums for evaluation. There were significant differences between these test sets and training sets in terms of user groups, language styles, and platform culture. The results are shown in Table 4.

Table 4: Application testing results on different platforms.

| Platform type | Data volume | Accuracy (%) | F1 score (%) | Recall (%) |
|---|---|---|---|---|
| Rednote | 67500 | 87.3 | 86.5 | 85.8 |
| Weibo | 26000 | 87.2 | 85.6 | 83.9 |
| Zhihu | 20000 | 81.3 | 79.7 | 77.5 |
| Campus Forum | 12500 | 89.5 | 88.1 | 86.3 |

In Table 4, the MHP model had an F1 score of 85.6% on the Weibo platform, and the text structures published on both platforms were similar. However, the F1 score on the Zhihu platform was relatively low, only 79.7%, which was due to the fact that the Zhihu platform had a more professional academic background, so its performance was not as good as the other two platforms. The campus forum platform had specialized vocabulary and focused on different topics. The complexity of model analysis was not high, and the recognition of text information was high, with an F1 score of 88.1%. The evaluation results on Weibo and campus forums showed that the accuracy of the MHP model was 87.2% and 89.5%, respectively, demonstrating high performance. This indicated that the features learned by the MHP model from Rednote had good transferability on these platforms. However, the accuracy on the Zhihu platform was 81.3%, indicating a decrease in performance. This indicated that there were indeed distribution differences between different platforms, and also emphasized the necessity of introducing domain adaptation technology in the future to further enhance the universality of the model. Overall, cross platform testing results demonstrated that the MHP model had a certain degree of inherent robustness and practical potential.

To verify the effectiveness of the proposed automated warning and intervention system, the study selected 400 students from Beijing to conduct real-time intervention effect tests on the MHP model, and the participants' PHQ-9 scale results were all greater than or equal to 10 points. This study was approved by the ethics review committee and informed consent was obtained. The participants were divided into two groups, A and B, with 200 people in each group. Group A adopted the intervention method of MHP model warning and consultation, while Group B only conducted routine paper questionnaire screening, with a trial evaluation period of 4 months. The intervention results are shown in Table 5, using the change in PHQ-9 score, consultation conversion rate, and number of crisis events as measurement indicators.

Table 5: Comparison of intervention results.

| Measurement indicator | A group | B group | Cohen's d/OR | *p* value |
|---|---|---|---|---|
| PHQ-9 score change value (points) | -4.2±1.8 | -1.1±2.3 | 1.57 | <0.001 |
| Consultation conversion rate (%) | 63 | 22 | / | 0.008 |
| Number of incidents of crisis events (cases) | 2 | 11 | 0.17 | 0.013 |

In Table 5, the PHQ-9 score of Group A decreased by an average of 4.2 points after the end of the experiment, while Cohen's d was 1.57, indicating that the intervention method of Group A was highly effective. The counseling conversion rate of Group A was 63%, and the MHP model warning of this group could effectively improve the conversion rate of psychological counseling. Comparing the number of crisis events between the two groups, the Odds Ratio (OR) was 0.17, which meant that the risk of crisis events in Group A was reduced by 83% compared to Group B.

# 4 Discussion and conclusion

There are problems with insufficient multidimensional data fusion and weak real-time detection in the current assessment of college students' mental health. To achieve efficient processing of college students' mental health sign analysis, a MHP model based on DKG and CNN-GRU was proposed. The findings demonstrated that the accuracy and F1 score without DKG were 81.2% and 80.1%, respectively, which were 6.1% and 6.4% lower than MHP model, respectively. This indicated that DKG contributed significantly to feature association modeling. Reference [5] used the HADS scale combined with SVM, but did not take into account daily behavioral characteristics. In contrast, the design of this study was more efficient in integrating multidimensional data on social media platform, PHQ-9 scale and DKG. The F1 score of the CNN-GRU designed in the study was 98.5%, with an accuracy of 95.6%, which was better than the CNN and BiLSTM performance in reference [11]. The approach proposed in reference [14] considered electronic health records and neural impact data, but the economic cost of actual use was relatively high. In contrast, this study used crawler technology to obtain multidimensional data, and the cost was lower. This study not only achieved the fusion of multi-source data, but also improved the accuracy of MHP, providing a reliable intervention tool for college students' psychology. Although this study achieved good preliminary results, the current research relied on a single data source, which did limit the full validation of the model's universality. Just as adaptive fuzzy control [26] and adaptive backstepping control [27] methods respond to unmodeled dynamics and external disturbances in dynamic systems by adjusting parameters online, a truly robust MHP model also needed to be able to adapt to the uncertainty brought by various data distributions, such as those from different social media platforms and student groups. The current work can be seen as validating the effectiveness of the core architecture under relatively controlled initial conditions, laying a solid foundation for subsequent adaptive scaling. Inspired by the search for dynamic optimal solutions in complex systems using nonlinear optimal control [28], this study planed to analyze the dynamic update mechanism to enable the MHP model to make online adjustments over time and changes in user data distribution, to maintain its prediction accuracy and achieve continuous tracking of the evolution of user psychological states. Future research will expand the scope of the study from ordinary college students to include more diverse groups such as international students and graduate students, in order to evaluate the performance of the model in different subgroups and ensure the broad applicability of its intervention strategies.

# References

[1] Li H .Multicultural data assistance mining analysis for ideological and political education in smart education platforms using artificial intelligence.Wireless Networks, 2025, 31(1):567-581.DOI:10.1007/s11276-024-03772-8.

[2] Gao Y. The impact and application of artificial intelligence technology on mental health counseling services for college students. Journal of Computational Methods in Sciences and Engineering, 2025, 25(2): 1686-1701. DOI: 10.1177/14727978241302641

[3] Peng X , Nie R , Tong S .Correction to: Research on the selfefficacy and resilience of female graduate students in the era of artificial intelligence: analysis of the mechanism of mobile phone dependence, anxiety and mentoring relationship.Archives of Women's Mental Health, 2025, 28(4):963-963.DOI:10.1007/s00737-025-01556-3.

[4] Lamontagne L G , Doty J L , Diehl D C ,Nesbit, T. S. , Gage, N. A. , & Kumbkarni, N. , et al.Acceptability, usage, and efficacy of mindfulness apps for college student mental health: A systematic review and meta-analysis of RCTs.Journal of Affective Disorders, 2024, 367(000):951-971.DOI: 10.1016/j.jad.2024.09.014.

[5] Krishnan R, Kumari S, Badi A A, Jeba, S., & James, M. Predictive machine learning model for mental health issues in higher education students due to COVID-19 using HADS assessment. Arab Gulf Journal of Scientific Research, 2024, 42(4): 1472-1490. DOI: 10.1108/AGJSR-01-2023-0034.

[6] Zhai Y, Zhang Y, Chu Z, Geng, B., Almaawali, M., Fulmer, R., Du, X. et al. Machine learning predictive models to guide prevention and intervention allocation for anxiety and depressive disorders among college students. Journal of Counseling & Development, 2025, 103(1): 110-125. DOI: 10.1002/jcad.12543.

[7] Cai Y, Lin D, Lu Q. Comparison of Different Machine Learning Algorithms in the Mental Health Assessment of College Students. Journal of ICT Standardization, 2024, 12(4): 409-427. DOI: 10.13052/jicts2245-800X.1243.
[8] Tyulepberdinova G, Mansurova M, Sarsembayeva T, Issabayeva, S., & Issabayeva, D. The physical, social, and mental conditions of machine learning in student health evaluation. Journal of Computer Assisted Learning, 2024, 40(5): 2020-2030. DOI: /10.1111/jcal.12999.
[9] Bhavani B H, Naveen N C. An approach to determine and categorize mental health condition using machine learning and deep learning models. Engineering, Technology & Applied Science Research, 2024, 14(2): 13780-13786. DOI:10.48084/etasr.7162.
[10] Ji X. Research on Mental Health Assessment and Intervention Methods for College Students based on Big Data Analysis. Scalable Computing: Practice and Experience, 2024, 25(6): 4702-4711. DOI: 10.12694/scpe.v25i6.3285.
[11] Ming Q. An evaluation model for college students' mental health based on machine learning algorithm. International Journal of Wireless and Mobile Computing, 2025, 28(1): 58-67. DOI: 10.1504/IJWMC.2025.143017.
[12] Li J. Dynamic evaluation of college students' psychological state based on multimodal physiological signal fusion and deep generation model. International Journal of Information and Communication Technology, 2025, 26(17): 133-146. DOI: 10.1504/IJICT.2025.146669.
[13] Marco Calabrò, Fabbri C , Serretti A ,et al.A machine learning approach to predict treatment efficacy and adverse effects in major depression using CYP2C19 and clinical-environmental predictors.Psychiatric Genetics, 2025, 35(2):17-25. DOI: 10.1097/YPG.0000000000000379.
[14] Zhang Z. Early warning model of adolescent mental health based on big data and machine learning. Soft Computing, 2024, 28(1): 811-828. DOI: 10.1007/s00500-023-09422-z.
[15] Gao Q, Wei Y. Understanding the cultivation mechanism for mental health education of college students in campus culture construction from the perspective of deep learning. Current Psychology, 2024, 43(2): 1715-1732. DOI: 10.1007/s12144-023-04320-x
[16] Zeng Z , He J , Xiang T , Wang, N. , Chen, B. , & Guo, S. Cognitive Tracing Data Trails: Auditing Data Provenance in Discriminative Language Models Using Accumulated Discrepancy Score.Cognitive Computation, 2024, 16(6):3379-3390. DOI:10.1007/s12559-024-10315-y.
[17] Cherradi M, El Haddadi A. Enhancing Data Lake Management Systems with LDA Approach. Journal of Data Science and Intelligent Systems, 2025, 3(1): 58-66. DOI: 10.47852/bonviewJDSIS42022312.
[18] Agbesi V K, Chen W, Ukwuoma C C, Kuadey, N. A., Agbesi, C. C. M., Ejiyi, C. J., Kuadey, A. M. et al. Multichannel 2D-CNN attention-based BiLSTM method for low-resource ewe sentiment analysis. Journal of Data Science and Intelligent Systems, 2025, 3(1): 67-77. DOI: 10.47852/bonviewJDSIS32021512.
[19] Singh H, Kaur B, Sharma A, & Singh, A. Framework for suggesting corrective actions to help students intended at risk of low performance based on experimental study of college students using explainable machine learning model. Education and Information Technologies, 2024, 29(7): 7997-8034. DOI: 10.1007/s10639-023-12072-1.
[20] Norouzi F , Machado B L M S , Nematzadeh S .Schizophrenia Diagnosis and Predictionwith Machine Learning Models. International Journal of Scientific and Academic Research, 2024, 04(09):113-122. DOI: 10.54756/ijsar.2024.26.
[21] Pei J, Amanvermez Y, Vigo D, Puyat, J., Kessler, R. C., Mortier, P., Cuijpers, P. et al. Sociodemographic correlates of mental health treatment seeking among college students: A systematic review and meta-analysis. Psychiatric services, 2024, 75(6): 556-569. DOI: 10.1176/appi.ps.2023041.
[22] Wedyan M, Saeidi-Rizi F. Assessing the impact of urban environments on mental health and perception using deep learning: A review and text mining analysis. Journal of Urban Health, 2024, 101(2): 327-343. DOI: 10.1007/s11524-024-00830-6.
[23] Huda A M, Shidik G F, Praskatama V. COMPARATIVE ANALYSIS OF LSTM, BILSTM, GRU, CNN, AND RNN FOR DEPRESSION DETECTION IN SOCIAL MEDIA. Jurnal Teknik Informatika (Jutif), 2024, 5(6): 1723-1735. DOI: 10.52436/1.jutif.2024.5.6.4060.
[24] Taylor M E , Liu M , Abelson S , Eisenberg, D. , Lipson, S. K. , & Schueller, S. M.The Reach, Effectiveness, Adoption, Implementation, and Maintenance of Digital Mental Health Interventions for College Students: A Systematic Review.Current Psychiatry Reports, 2024, 26(12):683-693.DOI: 10.1007/s11920-024-01545-w.
[25] Magalhães E. Dual-factor models of mental health: A systematic review of empirical evidence. Psychosocial intervention, 2024, 33(2): 89-102. DOI: 10.5093/pi2024a6.
[26] Boubellouta; A B Z .Adaptive fuzzy control for practical fixed-time synchronization of fractional-order chaotic systems. Sage Publications[2025-10-20]. DOI: 10.1177/10775463251320258.
[27] Boulkroune A , Hamel S , Zouari F ,et al.Output-Feedback Controller Based Projective Lag-Synchronization of Uncertain Chaotic Systems in the Presence of Input Nonlinearities.Mathematical Problems in Engineering, 2017, 2017:1-12. DOI: 10.1155/2017/8045803.
[28] Rigatos G, Abbaszadeh M, Sari B, et al. Nonlinear optimal control for a gas compressor driven by an induction motor. Results in Control and Optimization, 2023, 11: 100226. DOI: 10.1016/j.rico.2023.100226.