RESEARCH ARTICLE

Construction of college students' psychological intervention system based on stacking model

Yanmei Gao, Yuanyuan Wang*, Junzhi Zhao

Department of Ideological and Political Theory Teaching and Research, Shijiazhuang Institute of Railway Technology, Shijiazhuang, Hebei, China.

Received: December 5, 2025; **accepted:** April 6, 2026.

College students nowadays face multiple pressures from academic studies, social interactions, and future development, making mental health issues increasingly prominent. Traditional mental health service systems mostly rely on students' voluntary help-seeking and regular scale screening, revealing specific limitations such as fragmented information, delayed intervention timing, and vulnerability to subjective evaluation bias. To break through these technical and model bottlenecks, this study constructed an intelligent and closed-loop mental health intervention system that integrated explainable artificial intelligence (XAI). The core early warning module of the proposed system employed a stacked ensemble learning framework with the first layer composed of heterogeneous models of support vector machine (SVM), random forest (RF), and extreme gradient boosting (XGBoost) to learn the depth features of multidimensional data and the second layer applied the meta-learner chooses logistic regression (LR) for fusion discrimination to ensure the accuracy and robustness of early warning. To overcome the black box limitation inherent in conventional predictive models, this study incorporated the Shapley Additive exPlanations (SHAP) framework to conduct attribution analysis on outputs generated by the stacking model. This approach quantified the contribution of each individual indicator to mental health risk prediction, enabling a critical interpretive shift from understanding what was predicted to explaining why such predictions occurred. Based on multi-source data, the study described the psychological state of college students and found that the average of academic stress, anxiety, depression, and overall stress was about 21 points. Students had mild and moderate risk accounting for over 80% and high-risk accounting for 11.2%, showing a high-pressure concentration distribution. The stacking ensemble model achieved an overall accuracy of 0.91, outperforming all individual baseline models. In the intervention experiment, anxiety was reduced by 18.7%, and happiness was increased by 21.5%. The closed loop of "high-precision early warning + personalized intervention" was realized, which provided an executable paradigm for the intelligent mental health support system.

Keywords: mental health; college students; stacking model; explainable artificial intelligence; psychological intervention; early warning.

***Corresponding author:** Yuanyuan Wang, Department of Ideological and Political Theory Teaching and Research, Shijiazhuang Institute of Railway Technology, Shijiazhuang, Hebei 050000, China. Email: xiaomann1992@126.com.

Introduction

Currently, college students have emerged as a high-risk population for mental health disorders influenced by multiple stressors including academic demands, social adaptation, and future career development. The current mental health service system in colleges and universities mostly relies on students' active help-seeking, regular scale screening, and counselors' daily observation. These traditional models have exposed obvious lag and passivity in practice [1, 2]. Interventions are often implemented only after psychological crises have become evident, resulting in the loss of optimal windows for early prevention and support. Information collected through routine psychological assessments is static and fragmented, making it difficult to track the dynamic progression of students' mental health status. Moreover, the reliability of self-reported data can be substantially compromised by subjective biases [3].

The development of artificial intelligence (AI) and big data technology provides a new perspective for understanding the complex psychological and behavioral patterns of human beings. It is possible to build an active and intelligent mental health early warning system by fusing and analyzing the multi-source heterogeneous data generated by students every day [4, 5]. Tamanna *et al*. revealed the regional differences of mental health of college students in Bangladesh through comprehensive factor analysis and predictive modeling and constructed an effective predictive model [6]. Kirlic *et al*. used machine learning method to deeply analyze the risk and protective factors of college students' suicidal thoughts and behaviors [7]. Their research not only focused on predicting but also on discovering the key variables affecting mental health, which provided directional guidance for subsequent intervention. As a powerful ensemble method, stacking model shows great potential in the field of educational data mining. Scientists successfully applied the stacking ensemble learning method to identify students at risk of dropping out of school at an early stage, which proved the superiority of this model in dealing with complex student data [8]. Oz *et al*. also used stacking method to analyze the data of program for international student assessment (PISA) and accurately predicted the academic performance of students, which further verified the robustness and effectiveness of stacking model in students' related prediction tasks [9]. These studies have laid a solid methodological foundation for applying stacking model to the accurate early warning of college students' mental health. Meanwhile, Zhang *et al*. put forward an AI-driven framework by integrating predictive modeling and intervention strategies to comprehensively improve college students' mental health education [10]. Zhao *et al*. combined AI algorithm with positive psychological intervention strategy to deal with the phenomenon of "lying flat" of college students, showing the great application value of the combination of technology and psychological theory [11]. Cheng *et al*. implemented psychotherapy by flipping the classroom teaching reform, which effectively relieved the psychological anxiety of college students [12], which confirmed the importance of innovative and active intervention measures in improving students' mental health. Nowadays, mental health issues among college students are becoming increasingly prominent, and the limitations of traditional intervention models and existing AI research are becoming more and more apparent. Some studies have attempted to introduce AI technology into the field of mental health, but most of them only focus on improving prediction accuracy, lacking systematic explanations for the reasons behind the prediction results, and failing to form a complete closed loop from early warning to intervention. In the assessment of mental states, a single model often suffers from insufficient generalization ability due to problems such as data noise and complex feature interactions, making it difficult to play a stable role in practical applications. Existing mental health early warning systems mostly rely on static scale data, ignoring the potential information contained in multi-dimensional dynamic indicators such as students' academic behaviors, social interactions, and sleep rhythms, which limits the sensitivity and timeliness of early identification. How to ensure the accuracy of early warnings, enhance the interpretability of models, and effectively convert early warning results into personalized intervention strategies has become a core problem that needs to be solved urgently in building intelligent mental support systems.

This study proposed an intelligent mental health intervention system for college students by connecting the complete link from data collection, risk early warning, cause analysis to intervention suggestion generation and realizing a business closed loop of accurate early warning, transparent attribution, and personalized intervention. The system adopted the stacking ensemble learning framework as the core model to ensure the accuracy and robustness of early warning, which deeply integrated the advantages of multiple heterogeneous learners through a two-layer learning structure including support vector machine (SVM), random forest (RF), extreme gradient boosting (XGBoost) as the base layer and logistic regression (LR) for secondary learning to fully mine complex patterns in multi-source heterogeneous data as the meta-learning layer. To break the black box dilemma in the application of traditional machine learning models, this study introduced explainable artificial intelligence (XAI) technology and used Shapley Additive exPlanations (SHAP) framework to conduct attribution analysis on the prediction results of the stacking model to quantify the contribution of each dynamic indicator to mental state early warning. This study identified the key driving factors affecting the mental health early warning of each student, built a solid data-driven bridge between precise early warning and personalized intervention to overcome the shortcomings of weak generalization ability of single models and information lag of static data, and achieved a cognitive leap from knowing "what" to knowing "why". It provided mental health practitioners with understandable and highly practical decision support and offered an implementable systematic paradigm for the scientific community to explore and implement intelligent mental health service systems.

Materials and methods

Framework of college students' psychological intervention system with XAI

The core goal of the framework was to build a complete closed loop from data collection, risk early warning, cause analysis to intervention suggestion generation to ensure that the predictive ability of the technical model could effectively serve the practical needs of mental health education. The workflow of the system started from the data input layer, which was responsible for integrating multi-source heterogeneous data from college students to form a comprehensive feature vector set (Figure 1). Let the feature vectors of individual students be $X = \{x_1, x_2, \cdots, x_n\}$, where x_i was a specific feature such as academic stress score, sleep quality index or social interaction frequency, and n was the total number of features. The core function of the system could be formally expressed as a mapping function F_{sys} from feature space to intervention suggestion space as below.

$$A = F_{sys}(X) \tag{1}$$

where X was the input student characteristic data. A was the set of personalized intervention suggestions output by the system. This mapping process was not completed in one step but was realized by several subsequent modules [13]. The data first entered the intelligent early warning layer, which was the "perception engine" of the system, and its core was a pre-trained mental health state early warning model M. In this study, the model was constructed by stacking integrated learning strategy to analyze the input feature vector X and output a prediction label Y_{pred} about the student's current mental health status as follows.

$$Y_{pred} = M(X) \tag{2}$$

The value of Y_{pred} represented different levels of mental health such as "good state", "potential risk", or "high-risk warning". When the model outputted the warning results, the system would not end directly. To break through the "black box" limitation of the traditional machine learning model, the early warning result Y_{pred} was sent to the interpretable attribution layer together with the original feature data X. This layer was the key of this system framework and

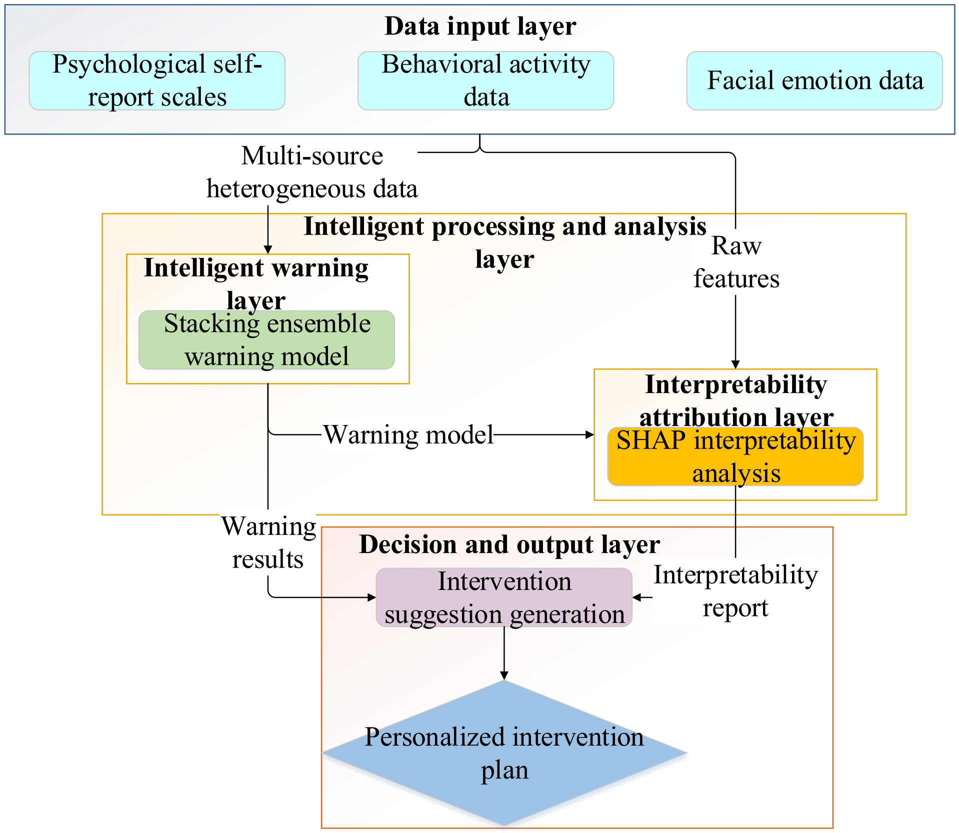

Figure 1. Framework of college students' psychological intervention system with XAI.

introduced XAI technology [14]. SHAP was used as its core algorithm E to analyze the internal decision logic of model M. The function of this module was to generate an explanation for each prediction and quantify the contribution ϕ_i of each feature x_i to the final prediction result Y_{pred}, as follows.

$$E(M,X) = \{\phi_1, \phi_2, \cdots, \phi_n\} \tag{3}$$

where ϕ_i was the SHAP value of the feature x_i. These attribution values constituted a detailed interpretable report, which clearly revealed the key driving factors for the model to make specific judgments. The intervention suggestion generation layer received a clear warning level Y_{pred} and a detailed attribution report $E(M,X)$. This module embedded a set of expert rule base or a lightweight decision-making model G, which matched and generated highly targeted and operable personalized intervention advice A from the knowledge base according to the risk level of students and the key negative factors that led to the risk, i.e. the characteristics with high negative contribution.

Early warning model of mental health based on stacking ensemble learning

To fully explore the complex patterns in multi-source heterogeneous data and improve the generalization ability of the model, this study adopted stacking ensemble learning strategy [15]. Stacking combined multiple different learning algorithms through a hierarchical structure with the core idea of letting multiple models learn different aspects of data and then learn how to optimally combine the prediction results of these models through a meta-learner. In the training process of stacking model, the original training dataset $D_{train} = \{(X_i, y_i)\}_{i=1}^{N}$ was subjected to K -fold cross-validation processing, where N was the total number of

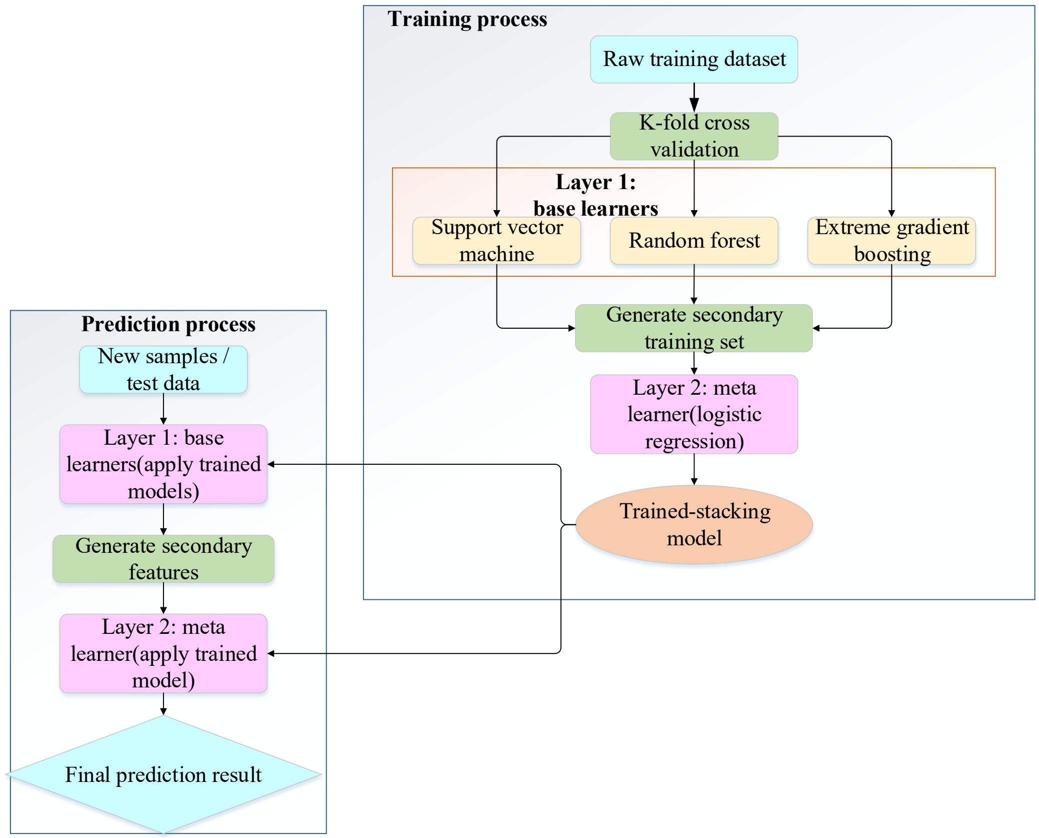

Figure 2. Stacking ensemble learning model structure.

training samples [16]. For each basic learner $M_j(j = 1,2,\cdots,J)$ in the first layer with $J = 3$, the model was trained on K-1 folded data and then predicted on the remaining folded data. This process was repeated k times until each sample had been predicted once as a verification set. In this way, for the j-th base learner, a prediction vector $Z_j = [p_{1j}, p_{2j}, \cdots, p_{Nj}]^T$ with the same size as the original training set was obtained, where p_{ij} was the prediction result of the base learner M_j on the i-th training sample (Figure 2). In this study, three heterogeneous models with different learning mechanisms were carefully selected in the basic learner layer including SVM, RF, and XGBoost. SVM classified in high-dimensional space by finding an optimal hyperplane [17], and the expression of its decision boundary was shown below.

$$f(x) = sign(\sum_{i=1}^{N} \alpha_i \, y_i K(x_i, x) + b) \quad (4)$$

where $K(x_i, x)$ was a kernel function. α_i was a Lagrange multiplier. RF was an integration method based on decision trees, and the stability of the model was improved by building multiple decision trees and voting the results [18]. For a RF containing B trees, the final prediction result was shown below.

$$\hat{y}_r f = argmax_c \sum_{b=1}^{B} I(\, \hat{y}_b(x) = c) \quad (5)$$

where $\hat{y}_b(x)$ was the prediction result of the b tree. $I(\cdot)$ was indicator function. As an advanced implementation of gradient lifting decision tree, XGBoost iteratively added new trees to fit the residual of the previous model, and its objective function included loss function and regularization term as shown below [19].

$$Obj^{(t)} = \sum_{i=1}^{N} L(y_i \, , \hat{y}_i^{(t-1)}) + f_t(x_i) + \Omega(f_t) \quad (6)$$

where L was the loss function. $\Omega(f_t)$ was the regularization term of the t-tree, which was used to control the model complexity. When all the base learners had completed the prediction of the training set through K-fold cross-validation [20], their prediction results $Z_1, Z_2, \cdots, Z_J$ were merged in columns to form a secondary training set D'_{train} as follows.

$$D'_{train} = \{([p_{i1}, p_{i2}, \cdots, p_{ij}], y_i)\}_{i=1}^{N} \quad (7)$$

This new dataset was used to train the second-level meta-learner M_{meta} to learn how to make a final and more accurate judgment according to the predictions of different base learners [21]. In this study, the meta-learner chose LR model because it was simple, efficient, and the results could be explained. The LR model mapped the input of linear combination to (0, 1) interval through Sigmoid function, indicating the probability that the sample belonged to a certain category as follows.

$$P(y = 1|Z') = \frac{1}{1+e^{-(\omega^T Z' + b_0)}} \quad (8)$$

where $Z' = [p_{i1}, p_{i2}, \cdots, p_{ij}]$ was a sample in the secondary training set. ω and b_0 were the weights and biases that the meta-learner needed to learn. When the model was deployed for prediction, a brand-new sample X_{new} would first pass through all the trained base learners M_j to generate a new prediction vector $Z'_{new} = [M_1(X_{new}), M_2(X_{new}), \cdots, M_J(X_{new})]$. This vector was input into the trained meta-learner M_{meta} , and the final mental health state prediction result Y_{pred} was obtained below.

$$Y_{pred} = M_{meta}(Z'_{new}) = M_{meta}([M_1(X_{new}), M_2(X_{new}), \cdots, M_J(X_{new})]) \quad (9)$$

Through this two-layer learning structure, the stacking model could effectively integrate the "wisdom" of different models. The basic learner focused on extracting different feature patterns from the original data, while the meta-learner focused on how to optimize the combination of this information to obtain more powerful prediction performance than any single model.

Model explainable analysis method based on SHAP

Although the stacking integrated model was excellent in improving the prediction accuracy, its complex multi-layer structure intensified the indeterminacy of the model, making it a typical "black box" model. In the application scene of psychological intervention for college students, which was highly dependent on trust and professional judgment, it was not enough to provide only an early warning result. It was equally important to understand the reasons for making decisions by the model. To open this "black box", this study introduced the SHAP method based on game theory and made an in-depth explanatory analysis of the built stacking early warning model. SHAP is a model-independent additive feature attribution method, which can decompose the prediction output of any complex model into the sum of the contribution values of each feature, thus clearly revealing the influence of each feature on the prediction results. For a model f with n input features, its predicted value $f(x)$ for a single sample x could be approximated by an additive explanatory model g as follows.

$$f(x) = g(x') = \phi_0 + \sum_{i=1}^{n} \phi_i x'_i \quad (10)$$

where x' was a simplified binary input vector, indicating whether a feature existed. n was the number of input features. ϕ_0 was the benchmark value to explain the model, which represented the average prediction output of the model without any feature information. ϕ_i was the SHAP value of the i -th feature, which represented the contribution of this feature to the prediction result from the benchmark value to the final prediction value. The calculation of SHAP value ϕ_i followed three ideal properties of local accuracy, missing, and consistency. Local accuracy ensured that the sum of the SHAP values of all features plus the reference value was exactly equal to the original prediction value of the model. Missing required that a feature missing in simplified input had a SHAP value of zero. Consistency ensured that if the marginal contribution of a feature increased or remained unchanged after a model was changed, the SHAP value of the feature should also increase or remain unchanged. The only solution that satisfied these three properties was the calculation of Shapley value as below.

$$\phi_i = \sum_{S \subseteq N\{i\}} \frac{|S|!(n-|S|-1)!}{n!} [f_{S\cup\{i\}}(x_{S\cup\{i\}}) - f_S(x_S)] \quad (11)$$

This equation calculated the weighted average marginal contribution of feature i on all possible feature subsets S (excluding i), where N was the set of all features, and $f_S(x_S)$ was the predicted output of the model when only the values of features in the feature subset s were known. Because the accurate calculation of SHAPley value faced the problem of combinatorial explosion when there were many features, SHAP framework provided a series of efficient approximation algorithms such as Kernel SHAP and Tree SHAP, which made it possible to apply it to large-scale datasets and complex models. In this study, SHAP analysis was carried out from the global and local levels to provide comprehensive and in-depth insight into the model. At the interpretation level, SHAP framework provided multi-level analysis from global and local dimensions. Global interpretation identified key factors that greatly influenced the mental health status of all students through SHAP summary plots and presented the positive and negative correlations between feature values and prediction results in the form of scatter plots. Combined with the feature importance ranking plot, features were sorted in descending order based on the mean of absolute SHAP values across all samples, intuitively showing the overall importance of each feature. Local interpretation targeted a single high-risk student sample and used single-sample prediction plots or waterfall plots to show in detail how each feature gradually pushed the prediction result from the baseline value to the final warning level, quantifying the specific contribution of each factor. Dependence plots revealed the inherent relationship between specific features and model prediction output, explored the interaction effects between the feature and other features, and helped mine deep-seated risk patterns. The above attribution analysis used the open-source SHAP analysis library in the Python environment (https://github.com/slundberg/shap).

Dataset description and verification strategy

This study introduced a public college student mental health assessment dataset (https://www.kaggle.com/datasets/zoya77/college-student-mental-health-assessment-dataset) with fully anonymized open-source public records involving no substantive recruitment of human subjects or privacy collection. The dataset contained multi-dimensional records of 125 college students aged from 18 to 25 years old with gender composition covering male, female, and other gender identity groups and integrated students' subjective psychological questionnaire scores with objective behavioral and emotional observation results including multiple core features of academic stress score, anxiety score, depression level, overall stress level, sleep quality index, behavioral activity level, social interaction frequency, and facial emotion labels. Mental health status was ultimately classified into four grades including healthy (0), mild risk (1), moderate risk (2), and severe risk (3). In the model training and validation phase, the 10-fold cross-validation method was adopted to fully test the stacking model and each base learner, ensuring a stable evaluation of the generalization ability.

Comparison model and evaluation index

To comprehensively verify the effectiveness of the constructed stacking framework, a parallel comparative test was conducted between the framework and the individual base learners for SVM, RF, and XGBoost comparison. The implementation of these comparison algorithms relied on open-source machine learning frameworks within the Python ecosystem. The SVM and RF algorithms were invoked from the scikit-learn library (https://scikit-learn.org), and the XGBoost algorithm was built *via* its official independent dependency library (https://xgboost.ai). For the classification and early warning task of mental health status, four key indicators of accuracy, precision, recall, and F1-score were introduced to evaluate the recognition performance of each model. Accuracy reflected the proportion of samples correctly classified by the model overall. Precision focused on the degree to which samples predicted as a specific risk level actually belonged to that level. Recall quantified the proportion of students actually in a high-risk state who were successfully identified by the model, which had extremely high practical significance in preventing psychological crises. As the harmonic mean of precision and recall, the F1-score provided a more objective and balanced comprehensive performance evaluation when the sample distribution of each risk level was unbalanced.

Results and discussion

Demographic characteristics of samples and the distribution of psychological indicators

The analysis of the psychological and behavioral characteristics of 125 college students revealed obvious clustering patterns in the core stress indicators. The academic stress score was 21.9 ± 11.97 ranging from 1 to 40. The scores of anxiety and depression were 21.18 ± 12.16 and 21.34 ± 11.55, respectively. The overall stress score reached 21.32 ± 11.72. The results indicated that the sample generally endured a heavy psychological burden with significant individual differences. The sleep quality score was 5.1 ± 2.97, reflecting generally poor sleep quality. Behavioral activity level showed large individual dispersion with a range from 21 to 139 and a mean of 70.72. The social interaction frequency was 10.63 ± 3.63. These statistical characteristics portrayed a realistic profile of contemporary college students, whose stressors were highly concentrated during their self-development stage, while their emotional regulation ability and daily routine regularity remained unstable.

Logical exploration of correlation of psychological core variables

An in-depth investigation of the correlation strength among various psychological variables showed that there was an extremely strong positive correlation among academic stress, anxiety, depression, and overall stress. The results of the Pearson correlation coefficient (PCC) indicated that the correlation between anxiety and depression was as high as 0.94, and the correlation coefficient between academic stress and overall stress reached 0.91. The correlations between academic stress and anxiety, depression remained at high levels of 0.9 and 0.88, respectively. These high-correlation results suggested a significant chain effect in the psychological stress process of college students that heavy academic burden often induced intense emotional fluctuations. In contrast, the correlations between sleep quality and indicators of anxiety and depression were in the low-to-moderate range of 0.24 to 0.32. This difference suggested that although sleep problems coexisted with perceived stress, they were not the single core factor leading to emotional deterioration, reflecting the complex interaction logic between the common late-night habits and psychological burden among contemporary college students.

Risk grade distribution and demographic characteristics difference

The mental health status distribution of the sample population showed an obvious sub-health tendency. Healthy individuals accounted for only 4.00% (5 students), while the mild-risk, moderate-risk, and high-risk groups accounted for 40.80% (51 students), 44.00% (55 students), and 11.20% (14 students), respectively. This "inverted pyramid" distribution structure revealed that college students generally suffered from psychological pressure, and most students were in a critical range that required early warning and attention. There were characteristic differences in risk performance among different gender groups. Among female participants, 7 students were at high risk, accounting for a relatively high proportion within this gender.

Table 1. Performance comparison of mental health early warning models.

Model	Accuracy	Precision	Recall	F1 score
SVM	0.82	0.79	0.81	0.80
RF	0.84	0.83	0.84	0.84
XGBoost	0.86	0.85	0.86	0.86
Stacking	0.91	0.90	0.91	0.91

Male participants were mostly concentrated in the mild and moderate risk categories with 19 and 18 individuals, respectively. The group labels as “other” in gender classification, specifically referring to non-binary gender or individuals self-identifying beyond traditional male and female categories, was particularly densely distributed in the moderate risk range with 28 individuals. This phenomenon suggested that students with diverse gender identities might face more complex psychological adjustment pressures in campus environments. Overall, the data depicted a health profile of the sample that high-risk groups were highly concentrated, and gender background was associated with the distribution of mental health status, providing a basis for subsequent differentiated psychological interventions.

Quantitative comparison of recognition efficiency of early warning model

The recognition performance of the mental health early warning model was comprehensively evaluated by using four key metrics of accuracy, precision, recall, and F1-score through the performance comparison of the SVM, RF, XGBoost, and stacking ensemble learning framework. The results showed that the proposed stacking ensemble model achieved scores above 0.9 across all evaluation metrics. Compared with the best performing single model XGBoost with the accuracy of 0.86, it improved recognition accuracy by approximately 5% (Table 1). Single models were susceptible to local optima or noisy data when processing multi-dimensional nonlinear features, leading to limitations in balancing recall and precision. By integrating the learning advantages of different underlying algorithms, the stacking framework effectively reduced the probability of false positives and false negatives, demonstrating stronger robustness in identifying high-risk student groups. This multi-layer learning structure fully absorbed the predictive strengths of heterogeneous models and could more accurately capture the correlation features hidden in academic, behavioral, and emotional data. These test results provided solid data support for the practical deployment of intelligent early warning systems in campus settings.

Empirical effect evaluation of intelligent intervention scheme

By analyzing the changes in multi-dimensional indicators before and after the intervention, the optimizing effect of the intelligent intervention strategy on students’ mental state could be directly observed. The anxiety score (AS), depression score (DS), sleep quality index (SQI), and subjective well-being (SWB) all exhibited significant improvement trends with the results showing that the anxiety score decreased from 27.1 to 22.0 with an improvement rate of 18.7%. The depression score also dropped by 14.3%. In terms of quality of life, the sleep quality score increased from 4.8 to 5.4 with an increase of 12.1%, while the subjective well-being score rose significantly from 5.1 to 6.2 with an improvement rate as high as 21.5% (Table 2). The optimization trajectory of this series of data confirmed the practical effectiveness of psychological counseling suggestions and behavioral guidance programs in relieving emotional load. The targeted strategies implemented for high-risk students induced positive psychological changes in the short term, indicating that transforming transparent model attribution results into

Table 2. Evaluation of intervention effect of core psychology and happiness index.

Evaluation indicator	Before intervention	After intervention	Improvement rate
AS	27.1	22.0	- 18.70%
DS	24.5	21.0	- 14.30%
SQI	4.8	5.4	+ 12.10%
SWB	5.1	6.2	+ 21.50%

personalized intervention plans could effectively enhance students' self-regulation ability and emotional efficacy. This result provided a practical basis for the long-term promotion of intelligent psychological support systems.

Conclusion

Based on the data analysis and model verification of college students' mental health, this study revealed that the group generally bore significant psychological pressure. Academic and emotional indicators were concentrated at relatively high levels with discrepancies between emotional burden and living conditions, demonstrating an overall structural pattern characterized by "high pressure, moderate sleep quality, and strong fluctuation". The stress chain was very obvious among study, anxiety, and depression, which showed the driving effect of study pressure on emotional state, indicating that the core of campus psychological issues focused on development anxiety, goal anxiety, and internal evaluation conflict. Gender groups showed differentiated psychological performance, and non-dual gender students were in a higher risk gradient, reflecting that emotional safety and identity experience of multi-groups should be paid more attention to in the supply of psychological support. Stacking fusion model showed high recognition ability. The intervention experimental results showed that targeted strategies could bring obvious improvement, and intelligent recognition and behavior guidance had application potential. Limited by the sample range and intervention period, this study failed to cover different cultural backgrounds and long-term evolution process, and the intervention effect needed to be verified on a larger scale. The future research should expand multi-school samples, integrate more real-time behavioral data and physiological signals, increase the vertical tracking of emotional fluctuation mechanism, expand psychological support scenarios, build a sustainable intelligent psychological service system, enabling technology to become gentler and more humanistic in the educational field.

References

1. Roy K, Easin M, Reno S, Sarker AM, Shamsi R. 2025. Leveraging a stacking ensemble model for accurate depression prediction and diagnosis across all ages. Engineer Rep. 7(10):e70416.
2. Mastour H, Dehghani T, Moradi E, Eslami S. 2025. Explainable artificial intelligence for predicting medical students' performance in comprehensive assessments. Sci Rep. 15(1):23752.
3. Yan L, Liu Y. 2020. An ensemble prediction model for potential student recommendation using machine learning. Symmetry. 12(5):728.
4. Powell SR, Berry KA, Fall AM, Roberts G, Fuchs LS, Barnes MA. 2021. Alternative paths to improved word-problem performance: An advantage for embedding prealgebraic reasoning instruction within word-problem intervention. J Educ Psychol. 113(5):898.
5. Sepanloo K, Shevelev D, Son YJ, Aras S, Hinton JE. 2025. Assessing physiological stress responses in student nurses using mixed reality training. Sensors. 25(10):3222.
6. Tamanna T, Barua E, Najmul Kabir M, Ahmed Z. 2025. Regional mental health disparities among university students in Bangladesh: A comprehensive factor analysis and predictive modeling approach. SN Social Sciences. 5(5):63.
7. Kirlic N, Akeman E, DeVille DC, Yeh HW, Cosgrove KT, McDermott TJ, *et al*. 2023. A machine learning analysis of risk and protective factors of suicidal thoughts and behaviors in college students. J Am Coll Health. 71(6):1863-1872.
8. Talamás-Carvajal JA, Ceballos HG. 2023. A stacking ensemble machine learning method for early identification of students at risk of dropout. Educ Inf Technol. 28(9):12169-12189.

9. Öz E, Bulut O, Cellat ZF, Yürekli H. 2025. Stacking: An ensemble learning approach to predict student performance in PISA 2022. Educ Inf Technol. 30(6):7753-7779.
10. Zhang X, Shi W. 2025. An AI-driven framework integrating predictive modeling and intervention strategies to enhance psychological health education among college students. Discov Appl Sci. 7(10):1095.
11. Zhao K, Gao M, Lian L, Jiang T. 2024. Positive psychology integration intervention strategies for the psychological phenomenon of "laying flat" in college students—An exploration and practice based on artificial intelligence algorithms. Studies in Psychological Science. 2(4):53-72.
12. Cheng P, Chao S. 2023. Flipped classroom-based psychological therapy teaching reform on college students' psychological anxiety. CNS Spectrums. 28(S2):S122.
13. De Souza VB, Nobre JC, Becker K. 2022. DAC stacking: A deep learning ensemble to classify anxiety, depression, and their comorbidity from Reddit texts. IEEE J Biomed Health. 26(7):3303-3311.
14. Smirani LK, Yamani HA, Menzli LJ, Boulahia JA. 2022. Using ensemble learning algorithms to predict student failure and enabling customized educational paths. Sci Programming. 2022(1):3805235.
15. Basysyar FM, Dwilestari G, Purnamasari AI. 2024. Analysis student emotions and mental health on cumulative GPA using machine learning and SMOTE. Jurnal Ilmu Pengetahuan dan Teknologi Komputer. 10(2):361-368.
16. Bekkouche NS, Schmid RF, Carliner S. 2022. "Simmering pressure": How systemic stress impacts graduate student mental health. Perform Improv Q. 34(4):547-572.
17. Biswas A. 2024. Modelling an innovative machine learning model for student stress forecasting. Global Perspectives in Management. 2(2):22-30.
18. Wang J. 2024. Optimization method based on machine learning for college students' psychological control source propensity classification. J Test Eval. 52(3):1714-1727.
19. Alghamdi S, Soh B, Li A. 2025. ISELDP: An enhanced dropout prediction model using a stacked ensemble approach for in-session learning platforms. Electronics. 14(13):2568.
20. Althaqafi T, Saleem F, Al-Ghamdi AALM. 2025. Enhancing student performance prediction: The role of class imbalance handling in machine learning models. Discov Comput. 28(1):79.
21. Nafea AA, Mishlish M, Al-Ani MM, Alheeti KMA, Mohammed HJ. 2023. Enhancing student's performance classification using ensemble modeling. Iraqi Journal for Computer Science and Mathematics. 4(4):16.