Building Machine Learning Predictive Models for Adolescent Internet Addiction: Key Findings on Self-Esteem and Resilience Interaction

Rongmei Liu
Central China Fuwai Hospital, Central China Fuwai Hospital of Zhengzhou University
Saiyi Wang
Zhengzhou University
Clifford Silver Tarimo
Zhengzhou University
Quanman Li
Zhengzhou University
Yifei Feng
Zhengzhou University
Lipei Zhao
Zhengzhou University
Shuaibin Liu
Zhengzhou Economic Development District
Xinghan Chen
Henan University of Science and Technology
Jian Wu
wujian@zzu.edu.cn

Zhengzhou University
Qiuping Zhao
Central China Fuwai Hospital, Central China Fuwai Hospital of Zhengzhou University

---

Article

Keywords: adolescents, internet addiction, machine learning, SHAP, self-esteem, psychological resilience

Posted Date: January 9th, 2025

DOI: https://doi.org/10.21203/rs.3.rs-5606509/v1

Additional Declarations: No competing interests reported.

Abstract

Objective: Internet addiction (IA) is a significant mental health concern among adolescents. This study aimed to develop machine learning (ML)-based predictive models to identify and explain key risk factors for IA.

Method: A total of 8176 junior high school students from Henan Province were surveyed from April to May 2023. The dataset was randomly divided into training and test sets in an 8:2 ratio. Four ML algorithms were used to predict IA, and feature importance was determined using SHapley Additive exPlanations (SHAP). The XGBoost model, which achieved the highest area under the curve (AUC), was selected for detailed analysis and individualized prediction explanations.

Results: The five most important predictors of IA were negative life events, self-esteem, school connectedness, parent-adolescent cohesion, and psychological resilience. Importantly, an interaction effect was found between self-esteem and psychological resilience: as self-esteem increased, the influence of low resilience transitioned from being a risk factor to a protective factor against IA.

Conclusion: This study demonstrates the power of ML models combined with SHAP for predicting IA and identifying its psychosocial determinants. The findings highlight the critical interplay of self-esteem and psychological resilience, offering valuable insights for clinicians and educators in addressing IA among adolescents.

Highlights

● This study developed machine learning (ML)-based predictive models to identify key risk factors for adolescent internet addiction (IA).

● Negative life events, self-esteem, school connectedness, parent-adolescent cohesion, and psychological resilience were identified as the top five predictors of IA.

● SHAP analysis highlighted a strong positive association between negative life events and IA risk. An interaction effect between self-esteem and psychological resilience showed that higher self-esteem transformed low resilience from a risk factor to a protective factor against IA.

● The integration of ML and SHAP provided clear and interpretable insights into IA risk factors, offering valuable guidance for prevention strategies.

1. Introduction

Internet addiction(IA), is defined as an increased desire to use the internet and out-of-control behavior, which is characterized by an intensified desire to reuse the internet (Tams, Legoux (1).The internet into the daily lives of teenagers has not only brought convenience but also introduced risks. The prevalence estimates for IA exhibit significant variability, current reports in China have shown that there is persistent rise in the prevalence of IA among adolescents, indicating the incidence ranging from 2.2–21.5%), which surpasses that in the United States and several European countries(2). In HK and Macau, about 1.9% of the college students were found to fulfil the criteria of IA(3). More recent studies had found that about 12% of the young adults in South Korea are addicted to the internet(4). Improper internet use is emerging as a public health issue that seriously influences the physical and mental health of adolescents(5). While currently there are no evidence-based treatments for IA, evidence suggests that IA leads to severe social and psychological damage(6). Hence, there is an emphasis on the importance of accurate projections and early intervention.

Adolescence constitutes a pivotal phase in human life, marked by the formation of habits, behaviors, and social confidence, during which individual undergoes various physical and mental transformation(7, 8). Adolescents experiencing IA are commonly engaged in diverse online activities including gaming, social media usage, and online pornography(9). IA in adolescents has been linked to diverse psychological factors(10), encompassing personality traits, negative emotions, self-esteem and impulsivity(11). Prior studies have demonstrated a correlation between elevated IA levels and increased aggressive behavior, including thoughts of suicide and self-harm among adolescents(12). Despite the noted negative consequences of IA among the adolescent population in China, there is a lack of exploration into the potential burden, prevalence and predictors of IA in research studies. Accurate factors predictions for IA in adolescents and a comprehensive understanding of the underlying factors are crucial for designing timely and targeted interventions.

Previous studies have employed statistical models, which are known for their stringent assumptions of a linear relationship between the outcome and explanatory variables, in modeling the predictors of IA across diverse populations. Machine Learning (ML) algorithms offer researchers powerful tools. Emphasizing algorithmic prediction in studies within this field is considered promising because ML algorithms are known for their reliability. ML algorithms provide researchers with robust tools utilized across various medical domains, including diagnosis, outcome prediction, treatment, and interpretation of medical images(13). Nevertheless, there remains a deficiency in research concerning ML for the risk prediction of IA. This gap is due to limited evidence supporting real-world clinical applications and the development of interpretable risk prediction models.

To address these limitations, the current study used ML algorithms (Logistic Regression (LR), Random Forest (RF), Extreme Gradient Boosting Machine (XGBoost), Support Vector Machines (SVM)), combined with SHapley Additive exPlanations (SHAP), to investigate the predictions of IA in adolescents. These methodologies can assess the interpretability of the model in guiding decisions related to interventions.

2. Methods

2.1 Participants

Participants were selected between April and May 2023 using a stratified random cluster sampling approach from six junior high schools in Henan Province, China. Within each school, 8 to 10 classes were randomly chosen from each grade. A total of 8,176 valid questionnaires were collected. The protocol received approval from Zhengzhou University, and all participants provided informed consent.

2.2 Measures

2.2.1 Internet addiction

A 20-item Internet Addiction Test was used to assess Internet addiction (14). It was evaluated by a Likert 5-point scale (ranging from 1 = rarely to 5 = always). Each question was summed up to calculate the total IA score. A total score above 50 was considered indicative of internet addiction (15) (Cronbach's α coefficient = 0.91).

2.2.2 Negative life events

The Adolescent Self-rating Life Events Checklist was adopted to assess the frequency of negative life events (☒☒☒ (16), where in a higher score denotes a greater prevalence of negative life events. The Cronbach's α coefficient in the study was 0.91.

2.2.3 Parent-adolescent cohesion

The 10-item Parent-Adolescent Cohesion Questionnaire was employed to measure parent-child and mother-child cohesion levels(17). For questions 3, 4, 8, and 9, reverse coding was applied. The Cronbach's α coefficient in this questionnaire was 0.87.

2.2.4 School connectedness

Based on prior studies (18), school connectedness was evaluated using 10 questions, encompassing three dimensions: teacher support (items 1, 5, and 8), school belonging (items 3, 6, and 9) and classmate support (items 2, 4, 7, and 10) (19). Response options ranged from 1 to 5 (1 = strongly disagree, 2 = strongly disagree, 3 = uncertain, 4 = strongly agree, 5 = strongly agree). Items 1 and 10 were reverse coded. The higher scores indicated higher levels of student connection (Cronbach's α coefficient = 0.85).

2.2.5 Psychological resilience

Psychological resilience was measured by the Connor-Davidson Resilience Scale(20) (21). 10 items were assessed using a 5-point Likert scale (0 = never, 1 = rarely, 2 = sometimes, 3 = often, 4 = always). The composite score ranged from 0 to 40, with higher scores indicating better psychological resilience. The Cronbach's α coefficient of the scale was 0.90.

2.2.6 Self-esteem

Self-esteem was assessed using the Self-esteem Scale developed by Rosenberg(22). 10 items were assessed on a 5-point Likert scale. Items 1, 2, 4, 6, and 7 were reverse coded, and higher scores indicate higher self-esteem (Cronbach's α coefficient = 0.90).

2.2.7 Control variables

Control variables for this study included gender, residence (rural, urban), grade (7th, 8th, 9th), family structure (intact family, others), maternal educational level (primary school and below, junior high school, senior high school, university and above), economic level (poor, moderate, good), study burden (light, moderate, heavy) and academic performance (poor, moderate, good).

2.3 Statistical analysis

Statistical analysis and data visualization were conducted using Python version 3.7. Categorical variables were presented as numbers and proportions, while continuous variables were reported in Mean ± SD. Binary logistic regression analysis was employed to explore factors related to IA. Additionally, the data were split into training and test sets with an 8:2 ratio, where 80% of the data were used for training the models and 20% for testing the models employing four algorithms, including Logistic LR, RF, XGBoost, and SVM. The LR model forecasts the probability of the binary dependent variable through the application of maximum likelihood estimation for ascertaining the regression coefficient. Both RF and XGBoost are algorithms grounded in tree-based learning. SVM serves as a generalized linear classifier within the framework of supervised learning. SHAP values were utilized to assess the contribution of each feature within each prediction model(23). P<0.05 were considered statistically significant.

3. Results

3.1 Sample characteristics

In total, 8176 adolescents (average age 14.42 years; 53.1% men) were included in this study. The prevalence of IA was observed in 1584 (19.4%) participants. Significant statistical differences were observed in gender, age, grade, family structure, maternal educational levels, family economic status, residence, study burden, academic performance, negative life events, psychological resilience, school connectedness, parent-adolescent cohesion, and self-esteem to IA. (See Table 1).

Table 1
Descriptive statistics of the sample (n = 8176)

Variables	Internet addiction			
	No (%)/Mean ± SD	Yes (%)/Mean ± SD	Total	*P*
Gender				< 0.001
Men	3576(82.3)	769(17.7)	4345(53.1)	
Women	3016(78.7)	815(21.3)	3831(46.9)	
Grade				0.017
7th	2528(82.9)	523(17.1)	3051(37.3)	
8th	2290(78.3)	635(21.7)	2925(35.8)	
9th	1774(80.6)	426(19.4)	2200(26.9)	
Residence				0.014
Rural	1964(79.0)	522(21.0)	2586(30.4)	
Urban	4628(81.3)	1062(18.7)	5690(69.6)	
Only-child				0.422
No	6031(80.7)	1439(19.3)	7470(91.4)	
Yes	561(79.5)	145(20.5)	706(8.6)	
Family structure				< 0.001
Intact family	6116(81.3)	1410(18.7)	7526(92.0)	
Others	476(73.2)	174(26.8)	650(8.0)	
Maternal educational levels				< 0.001
Primary school and below	517(74.8)	174(25.2)	691(8.5)	
Junior high school	2991(81.1)	695(18.9)	3686(45.1)	
Senior high school	1440(80.3)	353(19.7)	1793(21.9)	
Un University and above	1644(82.0)	362(18.0)	2006(24.5)	
Family economic status				< 0.001
Poor	363(72.5)	138(27.5)	501(6.1)	
Moderate	5062(81.4)	1160(18.6)	6222(76.1)	
Good	1167(80.3)	286(19.7)	1453(17.8)	

Variables	Internet addiction			
	No (%)/Mean ± SD	Yes (%)/Mean ± SD	Total	*P*
Study burden				< 0.001
Light	399(83.1)	81(16.9)	480(5.9)	
Moderate	3792(85.6)	636(14.4)	4428(54.1)	
Heavy	2401(73.5)	867(26.5)	3268(40.0)	
Academic performance				< 0.001
Poor	1528(72.8)	572(27.2)	2100(25.7)	
Moderate	3600(83.4)	716(16.6)	4316(52.8)	
Good	1464(83.2)	296(16.8)	1760(21.5)	
Age(years)	14.40 ± 0.94	14.46 ± 0.91	14.42 ± 0.94	0.020
Negative life events	39.25 ± 11.76	52.56 ± 16.39	41.83 ± 13.83	< 0.001
Psychological resilience	24.23 ± 8.15	19.50 ± 8.03	23.32 ± 8.34	< 0.001
School connectedness	38.00 ± 6.77	32.86 ± 7.42	37.01 ± 7.19	< 0.001
Parent-adolescent cohesion	36.29 ± 8.32	31.47 ± 8.40	35.36 ± 8.55	< 0.001
Self-esteem	29.95 ± 5.38	25.98 ± 6.15	29.18 ± 5.75	< 0.001

3.2 Model evaluation

Four ML models, LR, RF, XGBoost, and SVM were utilized to predict the occurrence of IA in adolescents (See Table 2). XGBoost had the highest AUC (Area under the curve,0.790) and precision (0.795), and its accuracy (0.822), recall (0.608), and F1(0.792) were second high among the four ML models. Figure 1 shows the ROC (Receiver operating characteristic) curves for all models in test sets, while the ROC curves in the training set are supplied as Supplemental Fig. 1. Therefore, we selected XGBoost for further analysis.

Table 2
Evaluation of the machine learning model performance.

Algorithm	AUC	Accuracy	Precision	Recall	F1	P^a
LR	0.782	0.828	0.749	0.618	0.799	< 0.001
RF	0.773	0.811	0.765	0.525	0.739	< 0.001
XGBoost	0.790	0.822	0.795	0.608	0.792	< 0.001
SVM	0.724	0.816	0.760	0.549	0.757	< 0.001

Note: AUC: Area under the curve. a: *P* value is the result of one-way analysis of variance for the AUC of the five models. LR: logistic. RF: Random Forest. XGBoost: Extreme Gradient Boosting. SVM: Support Vector Machine

3.3 SHAP model interpretation

The SHAP value with less than 0 indicates a negative contribution, equal to 0 indicates no contribution, and greater than 0 indicates a positive contribution. The top five features are negative life events, self-esteem, school connectedness, parent-adolescent cohesion, and psychological resilience. The higher the SHAP value of a feature, the higher the probability of developing IA. Adolescents with elevated levels of negative life events (depicted as red dots) were more prone to developing IA compared to those with lower levels (depicted as blue dots). Conversely, adolescents with low levels of self-esteem, school connectedness, parent-adolescent cohesion, and psychological resilience were more likely to develop IA (see Fig. 2).

Supplemental Fig. 2 depicts the top five variables of feature importance on the model output, revealing a nearly monotonic increase in local SHAP values for negative life events. The SHAP interaction plot (see Fig. 3) demonstrates the interaction effects between self-esteem and psychological resilience. A low value for psychological resilience (depicted as blue dots) poses a risk factor. However, as self-esteem improves, a low value for psychological resilience undergoes a transition from being a risk to exhibiting a somewhat protective effect.

3.4 SHAP values of individual prediction for interpretation

The force plot illustrates predictions for two randomly selected adolescents No. 5 and adolescents No. 6692, explaining the individual predictions in this study. The function f(x) represents the model output, indicating the predicted probability for each adolescent, while E[f(X)] means the average of the model predictions. Adolescents No. 5, a boy from rural, demonstrates a low risk of IA (-0.023) attributed to protective factors, including self-esteem (10), psychological resilience (21), grade (9), school connectedness (12), rural residence, and moderate family economic status (see Fig. 4(a)).

In contrast, Adolescent No. 6692, a girl diagnosed with IA in the study, exhibits a high probability of IA (0.442) due to risk factors such as negative life events (55), parent-adolescent cohesion (33), and good academic performance. (see Fig. 4(b)).

4. Discussion

In this study, we proposed a prediction model developed based on ML algorithms to accurately identify IA in adolescents. Our study provides a meaningful explanation based on the SHAP model. As previously described, the incidence of IA varied significantly under the influence of different social, cultural, and economic backgrounds(24). After the COVID-19 pandemic, IA may become a common problem for society, particularly affecting adolescents(25). Hence, reducing prevalence has become a crucial goal in the management of IA. In the present investigation, comprising 8716 adolescents, and ML prediction models were constructed utilizing 13 distinct features. The significance of a prediction is contingent upon its precision, thereby providing substantial contributions to clinical application.

Our study suggests that compared with other features, psychological features play a more significant role in the ML prediction of IA in adolescents. Specifically, the top five features in the prediction models are negative life events, self-esteem, school connectedness, parent-adolescent cohesion and psychological resilience. The ranking of variable importance based on SHAP values revealed that negative life events were the most significant factor influencing IA, a finding that has been rarely reported in previous literature on IA. More specifically, compared to adolescents with lower negative life events, adolescents with higher negative life events were more likely to experience social media addiction. These results indicated that negative life events may play an important role in IA(26).

However, our results indicate that lower self-esteem, school connectedness, and parent-adolescent cohesion are associated with the likelihood of IA occurrence. Studies have shown an association between IA and lower self-esteem(27), aligning with our findings. Additionally, lower school connectedness is related to increased IA has been reported(28). It seems plausible that the phenomenon is linked to the promotion of school connectedness, which enhances children's sense of belonging within the school environment. Improving positive interactions between teachers and students may reduce the risks of IA among adolescents. Furthermore, a study has reported parent–adolescent cohesion association with IA, which is consistent with our research. Worthy to pay attention, these points toward different psychological factors are closely related to IA in adolescents(10). The research found that adolescents with psychosocial problems are more prone to IA. A study from Turkey showed that the risk of psychosocial problems in adolescents was 19.8% and 18.8%, respectively(29). It was worth noting that the current findings were aligned with existing evidence suggesting that participants with poor mental health were more likely to cause IA.

More importantly, the SHAP model can be used to investigate the individual effects of risk factors and their interaction effects. The SHAP dependence plot depicts the top five variables of feature importance on the model output is presented in Supplementary Fig. 2, revealing a discernible positive correlation between negative life events and IA. Moreover, our study proves the observed SHAP interaction between self-esteem and psychological resilience in adolescents. As self-esteem improves, a lower value for psychological resilience undergoes a shift from being a risk factor to displaying a slightly protective effect.

To provide a detailed explanation and interpretation of IA prediction, we presented an individualized explanation of the model predictions through a force plot, illustrating predictions for two randomly selected adolescents. Remarkably, our ranking of variable importance closely aligns with observed differences in variables between subjects with and without IA. For example, adolescents without IA tend to exhibit higher levels of self-esteem, psychological resilience, and school connectedness. Conversely, Adolescent No. 6692 is associated with a high probability of IA risk (f(x) = 0.442) due to factors that elevate the prediction, including higher negative life events, lower parent-adolescent cohesion, good academic performance, and a tendency to be in grade 9 and reside in a rural area.

The study contains the following advantages. To date, no predictive models for IA in adolescents have been developed using ML algorithms and SHAP models. Unlike previous research in China, which predominantly relies on traditional regression models, our study represents an innovative approach to addressing the challenges associated with predicting IA in adolescents (30). The utilization of advanced ML models holds the potential to significantly enhance prediction accuracy. The application of ML algorithms to the medical field is a new trend, demonstrated in studies exploring connections between posttraumatic stress disorder (PTSD) and emotion regulation(31), predicting alcohol use(32) in adolescents, and other clinical applications(33). Additionally, the interaction analysis between self-esteem and psychological resilience is a novel approach to examining the interplay between variables for IA.

However, it is imperative to acknowledge the limitations of this study, when interpreting its findings. Firstly, Firstly, adolescents answered self-report questionnaires have recall bias. Secondly, analysis of cross-sectional data sets cannot be used to draw arbitrary conclusions. Finally, other factors affecting IA in adolescents, such as genetic factors(34), were not taken into account in the current study. Future research endeavors may incorporate genetic factors and other potential molecular level predictors, providing insights into the role of genetic mechanisms in IA.

5. Conclusion

In summary, our study combined the ML models and the explanation model to reliably predict the risk of IA in adolescents. This approach could assist physicians in intuitively understanding the influence of key features and detecting IA risks early by observing signs of psychological health in individuals. The findings have the potential to promote the identification of IA factors and the development of subsequent strategies for the follow-up care of adolescents.

Declarations

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Human Ethics and Consent to Participate declarations

Not applicable.

Ethical approval

The protocol received approval from Zhengzhou University.

Informed consent

All participants provided informed consent.

Data availability statement

Due to the confidentiality of the data, our data will not be publicly released. Our data will be provided by the corresponding author if required.

Funding Statement

Collaborative Innovation System Research on Drug Intervention & Non-drug Intervention in Proactive Health Context (20220518A); Research on Cardiovascular Disease Screening and Healthy Lifestyle Intervention in Children and Adolescents(20230014B) and Platform for Dynamic Monitoring and Comprehensive Evaluation of Healthy Central Plains Action (20220134B).

Acknowledgements

The authors wish to express their gratitude to all investigators and all participants and thank all their colleagues for their valuable inputs to the study design and data collection.

References

1. Tams S, Legoux R, Leger PM. Smartphone withdrawal creates stress: A moderated mediation model of nomophobia, social threat, and phone withdrawal context. Computers in Human Behavior 2018;81:1-9.
2. Mihara S, Higuchi S. Cross-sectional and longitudinal epidemiological studies of Internet gaming disorder: A systematic review of the literature. Psychiatry Clin Neurosci 2017;71:425-444.
3. Ding YJ, Lau CH, Sou KL, et al. Association between internet addiction and high-risk sexual attitudes in Chinese university students from Hong Kong and Macau. Public health 2016;132:60-63.
4. Na E, Lee H, Choi I, et al. Comorbidity of Internet gaming disorder and alcohol use disorder: A focus on clinical characteristics and gaming patterns. The American journal on addictions 2017;26:326-334.
5. Christakis DA. Internet addiction: a 21st century epidemic? BMC Medicine 2010;8.
6. Boer M, Stevens G, Finkenauer C, et al. Attention Deficit Hyperactivity Disorder-Symptoms, Social Media Use Intensity, and Social Media Use Problems in Adolescents: Investigating Directionality. Child development 2020;91:e853-e865.

7. Karaer Y, Akdemir D. Parenting styles, perceived social support and emotion regulation in adolescents with internet addiction. Comprehensive psychiatry 2019;92:22-27.
8. Seider S, Jayawickreme E, Lerner RM. Theoretical and Empirical Bases of Character Development in Adolescence: A View of the Issues. Journal of youth and adolescence 2017;46:1149-1152.
9. Griffiths MD, van Rooij AJ, Kardefelt-Winther D, et al. Working towards an international consensus on criteria for assessing internet gaming disorder: a critical commentary on Petry et al. (2014). Addiction (Abingdon, England) 2016;111:167-175.
10. Zhang W, Pu J, He R, et al. Demographic characteristics, family environment and psychosocial factors affecting internet addiction in Chinese adolescents. Journal of affective disorders 2022;315:130-138.
11. Gao YX, Wang JY, Dong GH. The prevalence and possible risk factors of internet gaming disorder among adolescents and young adults: Systematic reviews and meta-analyses. Journal of psychiatric research 2022;154:35-43.
12. Kuang L, Wang W, Huang Y, et al. Relationship between Internet addiction, susceptible personality traits, and suicidal and self-harm ideation in Chinese adolescent students. Journal of behavioral addictions 2020;9:676-685.
13. Rajkomar A, Dean J, Kohane I. Machine Learning in Medicine. The New England journal of medicine 2019;380:1347-1358.
14. Young KS. Internet Addiction: The Emergence of a New Clinical Disorder. Mary Ann Liebert, Inc 1998.
15. Zhang J H, Zhang X Q, Lu X Y, et al. The mediating role of depression in the relationship between childhood abuse and Internet addiction in adolescents. Chinese Journal of Health Psychology 2022;30:6.
16. Liu Xianchen. Development and reliability and validity test of adolescent life Events Scale. Shandong Psychiatry 1997;10: 5.
17. Zhang W X, Wang M P, Andrew, et al. Adolescents' expectation of autonomy, attitude towards parental authority and parent-child conflict and affinity. Journal of Psychology 2006.
18. Mcneely CA, Nonnemaker JM, Blum RW. Promoting School Connectedness: Evidence from the National Longitudinal Study of Adolescent Health. (Research Papers). Journal of School Health 2002; 72.
19. Yu C F, Zhang W, Zeng Y Y, et al. The relationship between gratitude and problem behavior in adolescents: the mediating role of school bonding. Psychological Development and Education 2011;27;9.
20. Campbell-Sills L, Stein MB. Psychometric analysis and refinement of the Connor-davidson Resilience Scale (CD-RISC): Validation of a 10-item measure of resilience. Journal of Traumatic Stress 2010;20:1019-1028.
21. Connor KM, Davidson JRT. Development of a new resilience scale: The Connor-Davidson Resilience Scale (CD-RISC). Depression and Anxiety 2003;18.

22. Rosenberg M. Rosenberg Self-Esteem Scale (RSES). . APA PsycTests 1965.
23. Lundberg S, Lee SI. A Unified Approach to Interpreting Model Predictions. Nips; 2017; 2017.
24. Twh C, Smy S, Mwl C. Adolescent Internet Addiction in Hong Kong: Prevalence, Psychosocial Correlates, and Prevention. The Journal of adolescent health : official publication of the Society for Adolescent Medicine 2019;64:S34.
25. Li YY, Sun Y, Meng SQ, et al. Internet Addiction Increases in the General Population During COVID-19: Evidence From China. The American journal on addictions 2021;30:389-397.
26. Wang X, Ding T, Lai X, et al. Negative Life Events, Negative Copying Style, and Internet Addiction in Middle School Students: A Large Two-year Follow-up Study. International journal of mental health and addiction 2023:1-11.
27. Sevelko K, Bischof G, Bischof A, et al. The role of self-esteem in Internet addiction within the context of comorbid mental disorders: Findings from a general population-based sample. Journal of behavioral addictions 2018;7:976-984.
28. Liu S, Yu C, Conner BT, et al. Autistic traits and internet gaming addiction in Chinese children: The mediating effect of emotion regulation and school connectedness. Research in developmental disabilities 2017;68:122-130.
29. Ozturk F, Ayaz-Alkaya S. Internet addiction and psychosocial problems among adolescents during the COVID-19 pandemic: A cross-sectional study. Archives of psychiatric nursing 2021;35:595-601.
30. Zhang X, Zhang J, Zhang K, et al. Effects of different interventions on internet addiction: A meta-analysis of random controlled trials. Journal of affective disorders 2022;313:56-71.
31. Christ NM, Elhai JD, Forbes CN, et al. A machine learning approach to modeling PTSD and difficulties in emotion regulation. Psychiatry research 2021;297:113712.
32. Afzali MH, Sunderland M, Stewart S, et al. Machine-learning prediction of adolescent alcohol use: a cross-study, cross-cultural validation. Addiction (Abingdon, England) 2019;114:662-671.
33. Li W, Wang J, Liu W, et al. Machine Learning Applications for the Prediction of Bone Cement Leakage in Percutaneous Vertebroplasty. Frontiers in public health 2021;9:812023.
34. Tereshchenko S, Kasparov E. Neurobiological Risk Factors for the Development of Internet Addiction in Adolescents. Behavioral sciences (Basel, Switzerland) 2019;9.

Figures

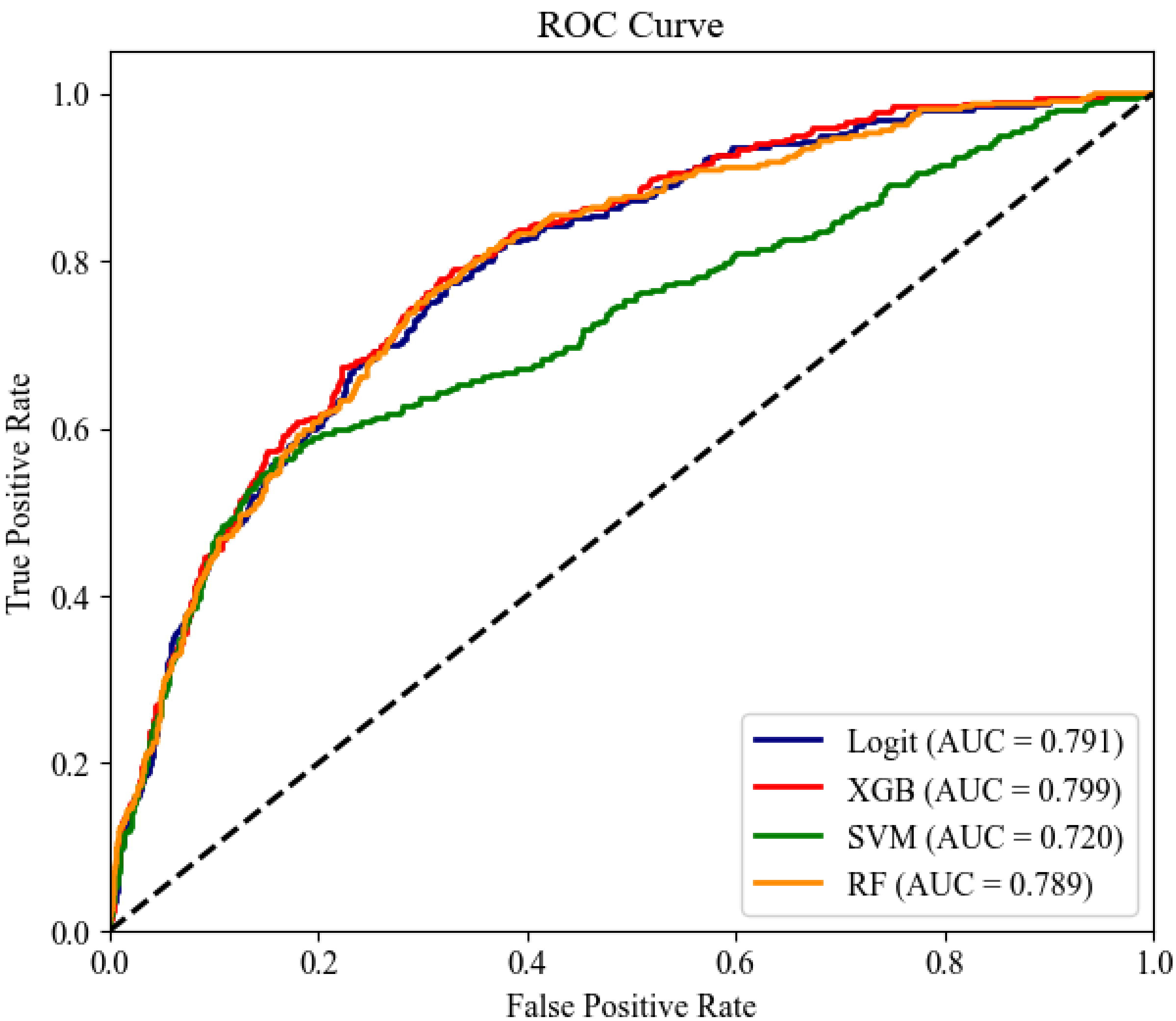

Figure 1

ROC curves for all models in test sets.

Notes: ROC: Receiver operating characteristic.

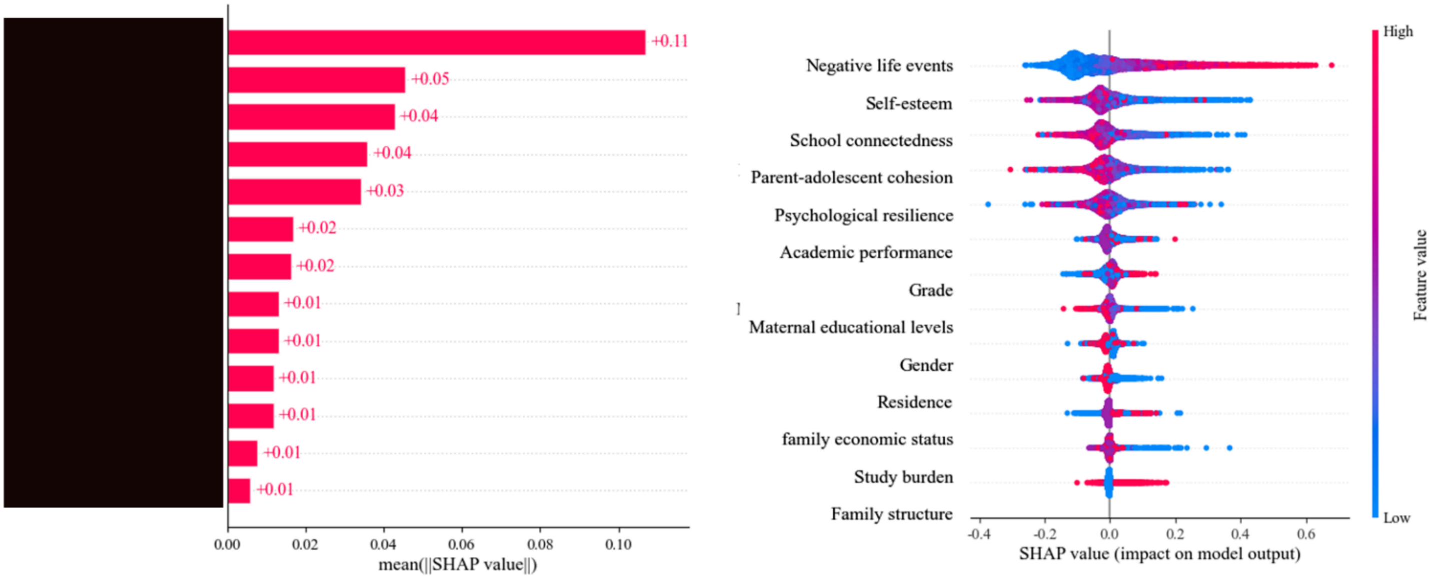

Figure 2

Feature importance ranking based on SHapley Additive exPlanations (SHAP) values in XGBoost.

The left figure: importance ranking. The right figure: SHAP summary diagram. Each row is a variable, each dot represents a sample, and the color represents the value of the variable in order of highest absolute importance. The point to the left of the X-axis is the protection point, and the point to the right of the X-axis is the danger point in the right graph.

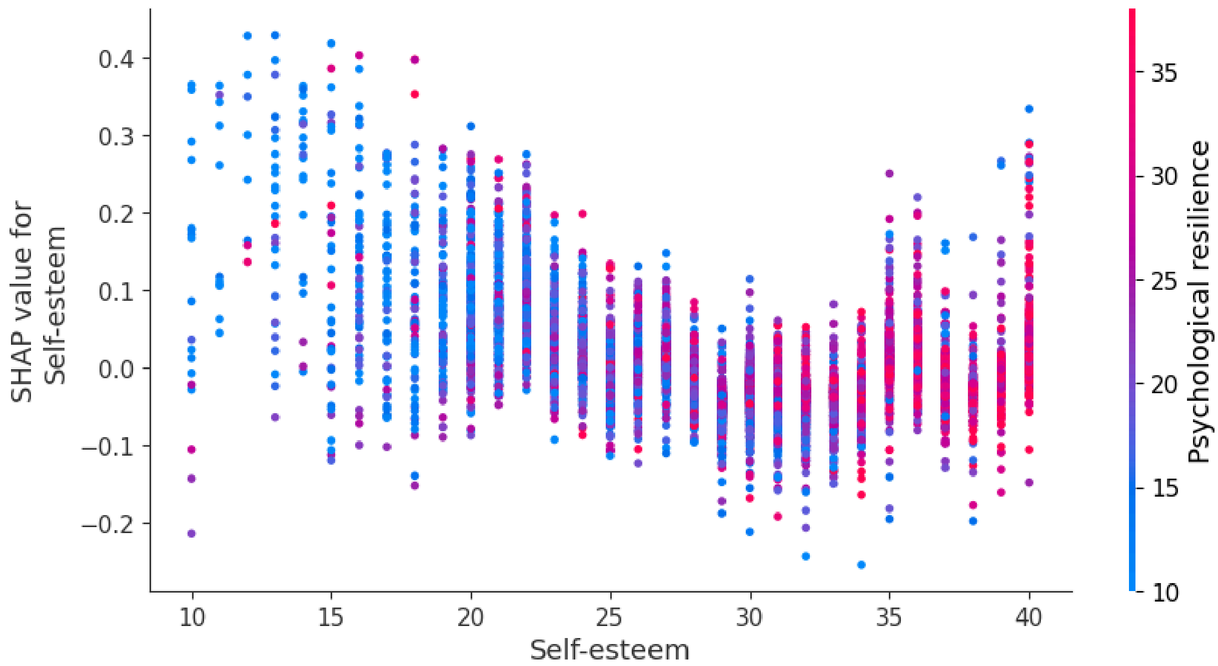

Figure 3

SHAP interaction plots between self-esteem and psychological resilience.

The SHAP dependence plots combine the main and interaction effects of a variable, where the interaction effects, can be interpreted as how two variables affect the model output simultaneously. The x-axis represents the value for self-esteem, while the y-axis is the SHAP values of their interaction. The color of the point represents the value of psychological resilience.

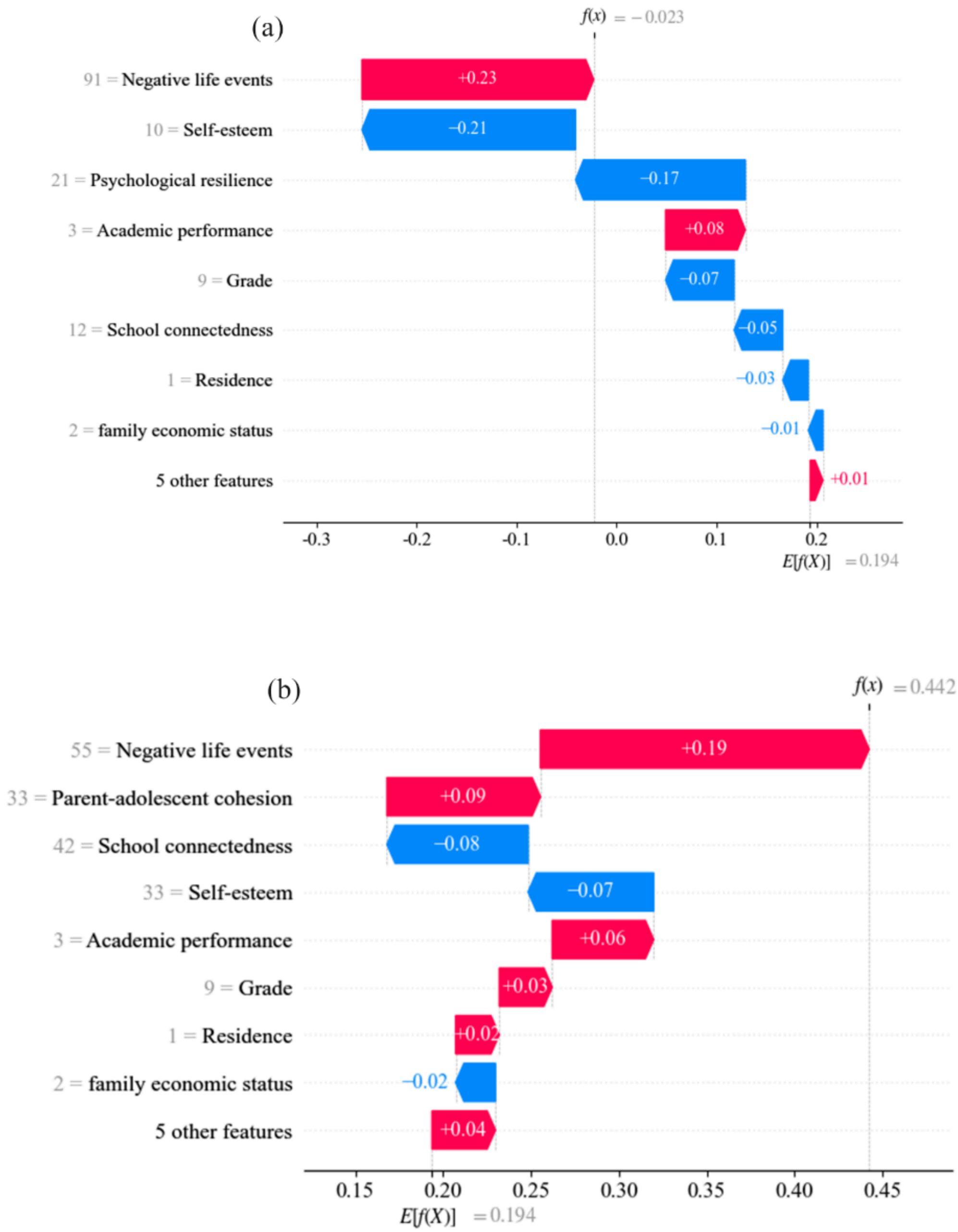

Figure 4

The force plot of two adolescents in the XGBoost model.

The force plot depicts the contribution of each feature to the process of moving the value of the decision score from the base value to the value predicted by the classifier. Red denotes features that make the model score higher, blue denotes features that make the model score lower. The longer the arrow length, the more influence the features.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- supplementary.docx