DYNA
http://dyna.medellin.unal.edu.co/

UNIVERSIDAD NACIONAL DE COLOMBIA

Predicting Internet addiction in college students using a 1D-CNN model: analysis of influencing factors

Xi Wang [a], Enyou Zhang [a], Yingjun Cui [b], Jie Huang [b] & Meng Cheng*[b]

[a] School of Educational Science, Ludong University, Yantai, China. xw259259@163.com, 13503120556@163.com
[b] School of Agricultural Engineering and Food Science, Shandong University of Technology, Zibo, China. cyj522819@163.com, huangjie0306@126.com, chengmeng0110@163.com

Received: February 4th, 2024. Received in revised form: July 22nd, 2024. Accepted: July 27th, 2024.

Abstract
This study constructs a deep learning-based model to predict internet addiction among college students and analyzes significant influencing factors. A random survey of 4,895 students from a university in Shandong Province was conducted using questionnaires on general information, internet addiction (CIAS-R), personality (CBF-PI-B), psychological traits (SDS, SAS), parenting styles (EMBU), behavioral issues (SAS-C), and social support (ASSRS) to establish a database. A predictive model was developed using a 1D Convolutional Neural Network (1D-CNN), extracting key influencing factors of internet addiction. The model showed 92.77% accuracy, with high precision and recall rates for predicting normal users and addicts. The gradient calculation indicates that in second-year students, negative and withdrawal behaviors, depression, over-interfering families, and anxiety significantly contribute to Internet addiction, with factors exceeding 0.5. The 1D-CNN model offers robust performance and accuracy in predicting internet addiction, identifying significant factors for early prevention and potential integration with apps for real-time monitoring.

Keywords: internet addiction; 1D-CNN; predicting; college students; model.

Predicción de la adicción a Internet en estudiantes universitarios mediante un modelo 1D-CNN: análisis de los factores influyentes

Resumen
Este estudio construye un modelo basado en el aprendizaje profundo para predecir la adicción a Internet entre los estudiantes universitarios y analiza los factores influyentes significativos. Se realizó una encuesta aleatoria a 4.895 estudiantes de una universidad de la provincia de Shandong mediante cuestionarios sobre información general, adicción a internet (CIAS-R), personalidad (CBF-PI-B), rasgos psicológicos (SDS, SAS), estilos parentales (EMBU), problemas de conducta (SAS-C) y apoyo social (ASSRS) para establecer una base de datos. Se desarrolló un modelo predictivo utilizando una red neuronal convolucional 1D (1D-CNN), extrayendo los factores clave que influyen en la adicción a Internet. El modelo mostró una exactitud del 92,77%, con altos índices de precisión y recuerdo para predecir usuarios normales y adictos. El cálculo del gradiente indica que, en los estudiantes de segundo curso, los comportamientos negativos y de retraimiento, la depresión, el exceso de interferencia familiar y la ansiedad contribuyen significativamente a la adicción a Internet, con factores superiores a 0,5. El modelo 1D-CNN ofrece un rendimiento y una precisión robustos en la predicción de la adicción a Internet, identificando factores significativos para la prevención temprana y la integración potencial con apps para la monitorización en tiempo real.

Palabras clave: adicción a Internet; red neuronal profunda; predicción; estudiantes universitarios; modelo.

1 Introduction

With the revolution of science and technology, Internet technology has become an indispensable tool in people's daily lives. However, this also brings about a brand new problem, the abuse of the Internet and the unrestrained use of modern electronic devices, causing a large number of teenagers to have Internet Addiction Disorder (IAD) [1, 2].

Goldberg et al. [3] first proposed the concept of Internet addiction, and Young. [4] further studied the influencing

How to cite: Wang, X., Zhang, E., Cui, Y., Huang, J., and Cheng, M., Predicting internet addiction in college students using a 1D-CNN model: analysis of influencing factors. DYNA, 91(233), pp. 66-74, July - September, 2024.

factors of Internet addiction based on the Diagnostic and Statistical Manual of Mental Disorders (DSM) [5, 6]. The study found that excessive Internet use harms physical health, causes interpersonal difficulties, decreases academic performance and interferes with normal work. Therefore, in order to prevent college students' Internet addiction in advance, it is urgent to build an Internet addiction prediction model and extract significant influencing factors of Internet addiction, so as to provide a research basis for deploying Internet addiction detection software to grasp the Internet addiction status of college students in real time.

Early research on Internet addiction mainly used methods based on statistical analysis. Yang. [7] conducted a questionnaire survey on 1,216 junior high school students in a nine-year compulsory education school in Jinan City, Shandong Province. After T-test analysis, the results show that the incidence rate of Internet overuse among junior high school students is 9.5%, and they have obvious characteristics such as anxiety, depression, psychoticism, and neuroticism. Chemnad et al. [8] conducted a survey among 479 teenagers in Qatar. After K-Prototype clustering, the results showed that family environment and school environment were negative and significant predictors of adolescent Internet addiction, with a prevalence rate of 29.64%. Lai et al. [9] conducted a psychological test on 844 Chinese adolescents in Hong Kong using demographics, the Internet Addiction Test Scale (IAT), and the Internet Addiction Scale-Revised (CIAS-R).

After validation analysis in statistical methods, the results show that “the IAT is a valid and reliable scale for screening Internet addiction among Chinese adolescents”. The above research methods usually require complex data processing and calculations before obtaining results through analysis, and their reliability and accuracy need to be further improved. Considering the limitations of the number of psychologists and self-answer questionnaires, it is difficult to effectively screen out Internet addicted subjects on a large scale.

With the rise of Internet technology and the development of science and technology, using machine learning and deep learning algorithms to process massive data and identify data patterns has become a trend and even a necessary method. This technology is widely used in manufacturing, finance, life sciences and healthcare industry [10,11]. Di et al. [12] used multiple personality questionnaire data and support vector machine algorithms to detect Internet addiction symptoms among middle school students. They classified the samples and optimized parameters through support vector machines and grid search algorithms, and their Internet addiction detection performance reached 96.32%. Chaudhury et al. [13] used naive Bayes, support vector machine and radial basis function neural network to analyze the impact of smartphones on the academic performance of 222 college students. The results showed that there is a correlation between smartphone addiction and academic performance sex. Shae et al. [14] constructed an unsupervised artificial intelligence (AI) model (XGBOOST) for internet addiction prediction using an auto encoder AI model (GAN) to generate training data. This model preliminarily discusses the prediction mechanism of Internet addiction based on deep learning technology, and provides a direction for research on using deep learning technology to detect Internet addiction. The limitations of the above-mentioned studies are: ① There will be bias in predictions for subjects in different geographical and cultural backgrounds ② The generalization of machine learning models needs to be improved ③ The research on Internet addiction and its main influencing factors is not systematic and comprehensive enough.

In view of this, this study randomly surveyed 4895 college students in a university in Shandong Province. First, one basic questionnaire and seven assessment scales were used to investigate Internet addiction, and then statistics and analysis of factors related to Internet addiction were conducted. On this basis, an Internet addiction database was established, an Internet addiction prediction model was constructed based on a one-dimensional convolutional neural network, and influencing factors with an impact factor greater than 0.5 were extracted in order to prevent Internet addiction among college students in Shandong Province, China. Thereby providing a research foundation.

2 Materials and methods

2.1 Data collection and processing

2.1.1 Data collection

This study employs seven different questionnaires to thoroughly investigate the factors affecting Internet addiction, capturing a wide spectrum of personal attributes. The general information questionnaire collects basic demographic, grades, majors, and family data. Internet addiction questionnaire are measured by the Revised Chinese Internet Addiction Scale (CIAS-R), with a scoring system from normal to addiction levels and a high reliability (Cronbach's $\alpha = 0.930$) [15]. In the present study, Cronbach’s $\alpha = 0.90$. Personality traits are assessed through the Chinese Big Five Personality Inventory Brief (CBF-PI-B), covering five key dimensions with an overall Cronbach's α range of 0.76 to 0.81 [16]. In the present study, Cronbach’s $\alpha = 0.92$. Mental health is evaluated using the Self-Rating Depression Scale (SDS) and the Self-Rating Anxiety Scale (SAS), each with their own scoring criteria and demonstrated reliability (Cronbach's α of 0.751 for SDS and 0.82 for SAS) [17,18]. In the present study, Cronbach’s $\alpha = 0.91$ and 0.90 respectively. The Evaluation of Parenting Styles Scale (EMBU) analyzes parenting attitudes and behaviors, with a reliability range of 0.58-0.922 [19]. In the present study, Cronbach’s $\alpha = 0.98$. Smartphone addiction is assessed using the SAS-C, which features a 5-point scale and a Cronbach’s α of 0.88 [20]. In the present study, Cronbach’s $\alpha = 089$. Finally, the Adolescent Social Support Rating Scale (ASSRS) examines various dimensions of social support, with a reliability range of 0.60-0.76 [21]. In the present study, Cronbach’s $\alpha = 0.93$. Each questionnaire is meticulously validated for accuracy and reliability, making them effective tools for measuring various aspects pertinent to internet addiction in this deep learning-based study.

2.1.2 Participants

The research was initiated by a university in Shandong Province, aiming to gain a comprehensive understanding of the different characteristics of the student body at the university. The research covered four grades (from freshman

Table 1.
The characteristics of college students in Shandong.

Survey type	Relevant factor	Classification description	Internet addiction，number/% Normal n=3882 (79.29)	Addiction n=1013 (20.71)	Sum N=4895	P value
General information	1. Age (Mean, SD)		20.45±3.65	19.87±5.63	21.75±2.31	0.587
	Gender	2. Male	1947 (79.47)	503 (20.53)	2450 (50.05)	0.939
		3. Female	1935 (79.14)	510 (20.86)	2445 (49.95)	
	Class	4. First grade	1368 (76.99)	409 (23.01)	1777 (36.30)	0.009
		5. Second grade	1206 (79.24)	316 (20.76)	1522 (31.09)	
		6. Third grade	784 (79.35)	204 (20.65)	988 (20.18)	
		7. Fourth grade	616 (88.00)	84 (12.00)	608 (12.42)	
	Only children	8. Yes	2522 (80.55)	609 (19.45)	3131 (63.96)	0.446
		9. No	1360 (77.10)	404 (22.90)	1940 (36.04)	
	Experience of left-behind children	10. Yes	2004 (75.20)	661 (24.80)	2665 (54.44)	<0.001
		11. No	1878 (84.22)	352 (15.78)	2230 (45.56)	
	Single parent	12. Yes	1075 (61.08)	685 (38.92)	1760 (35.96)	0.002
		13. No	2807 (89.54)	328 (10.46)	3135 (64.04)	
	Family economic conditions	14. Poor	1171 (84.43)	216 (15.57)	1387 (28.34)	0.002
		15. Normal	1139 (64.39)	630 (35.61)	1769 (36.14)	
		16. Rich	1572 (90.40)	167 (9.60)	1739 (35.53)	
Psychological characteristics	Depression	17. No	944 (85.12)	165 (14.88)	1109 (22.66)	0.018
		18. Mild	1005 (85.17)	175 (14.83)	1180 (24.11)	
		19. Moderate	923 (79.36)	240 (20.64)	1163 (23.76)	
		20. Severe	1010 (69.99)	433 (30.01)	1443 (29.48)	
	Anxiety	21. No	904 (85.28)	156 (14.72)	1060 (21.65)	0.017
		22. Mild	1004 (83.74)	195 (16.26)	1199 (24.49)	
		23. Moderate	932 (82.04)	204 (17.96)	1136 (23.21)	
		24. Severe	1042 (69.47)	458 (30.53)	1500 (30.64)	
Big five personality	25. Extraversion		770 (79.14)	203 (20.86)	973 (19.88)	<0.001
	26. Agreeableness		788 (79.36)	205 (20.64)	993 (20.29)	
	27. Conscientiousness		767 (79.32)	200 (20.68)	967 (19.75)	
	28. Neuroticism		870 (73.11)	320 (26.89)	1190 (24.31)	
	29. Openness		687 (88.99)	85 (11.01)	772 (15.77)	
Parenting	Paternal	31. caring and understanding	611 (82.97)	127 (17.21)	738 (15.08)	<0.001
		31. Harsh punishment	657 (78.97)	175 (21.03)	832 (17.00)	
		32. overly intrusive	669 (79.36)	174 (20.64)	843 (17.22)	
		33. Favoring the subject	611 (77.64)	176 (22.36)	787 (16.08)	
		34. Denial	656 (79.52)	169 (20.48)	825 (16.85)	
		35. Overprotection	678 (77.93)	192 (22.07)	870 (17.77)	
	Maternal	36. caring and understanding	781 (84.71)	141 (15.29)	922 (18.84)	<0.001
		37. overly intrusive	775 (79.16)	204 (20.84)	979 (20.00)	
		38. Denial	806 (77.50)	234 (22.50)	1040 (21.25)	
		39. Harsh punishment	768 (78.13)	215 (21.87)	983 (20.08)	
		40. Favoring the subject	752 (77.45)	219 (22.55)	971 (19.84)	
Behavioral problems	Smartphone addiction	41. Social reassurance	641 (83.68)	125 (16.32)	766 (15.65)	0.001
		42. Negative behavior	660 (76.04)	208 (23.96)	868 (17.73)	
		43. Withdrawal behavior	605 (76.20)	189 (23.80)	794 (16.22)	
		44. Highlight behavior	655 (77.33)	192 (22.67)	847 (17.30)	
		45. APP use	667 (79.40)	173 (20.60)	840 (17.16)	
		46. APP update	654 (83.85)	126 (16.15)	780 (15.93)	
Social support	47. Strong		979 (82.97)	201 (17.03)	1180 (24.11)	<0.001
	48. Good		938 (80.72)	224 (19.28)	1162 (23.74)	
	49. Fair		990 (81.15)	230 (18.85)	1220 (24.92)	
	50. Poor		975 (73.14)	358 (26.86)	1333 (27.23)	

Source: The authors.

to senior) and 175 classes at the university, involving nine colleges. To collect the data, the research team used the Tencent questionnaire tool, which converted the base questionnaire and seven assessment scales into an online format. The questionnaires were distributed by the counsellors of each class through a link.

Regarding the gathering and validity of questionnaires, a sum of 5,247 were amassed, with 4,895 qualifying as valid. The gender distribution among student participants was nearly equal, comprising 2,450 males and 2,445 females, accounting for 50.05% and 49.95% of the participants respectively. Participants' ages varied from 17 to 24 years, averaging at 21.75 years with a standard deviation of 2.31.

In terms of college distribution, the College of Mechanical Engineering had the highest number of participants with 887, followed by the College of Vehicle Engineering (695) and the College of Agricultural Engineering (700). The College of Electrical Engineering had 589 participants, the College of Computer Science had 548, the College of Mathematics and Statistics had 446, the College of Law had 346, the College of Management had 401, and the College of Foreign Languages had 283. The percentage of each college is 18.14%, 14.21%, 14.32%, 12.04%, 11.21%, 9.01%, 7.07%, 8.20%, and 5.80% respectively.

Statistical analysis revealed that the rate of internet addiction was 20.71% and the rate of no internet addiction was 79.28%. Among them, there was no statistical difference ($x^2 = 0.43$, $P > 0.05$) in internet addiction among age, gender, and whether they were only child, whereas there was a statistical difference ($x^2 = 15.43$, $P < 0.05$) in internet addiction among grade level, whether they were single parent, family economic condition, depression condition, anxiety condition, and mobile phone addiction condition. Big 5 personality situation, parenting style, and social support were statistically significantly different, $P < 0.001$.

The detailed findings are listed in Table 1. This research provides a comprehensive and detailed understanding of the student body of a university in Shandong Province.

This study will ensure that participants clearly understand the purpose of the study and participate on a voluntary basis during the data collection phase. The project has been approved by the University's student work department and is strongly supported by the relevant faculty members to ensure compliance and ethicality of the study.

2.1.3 Data processing

The study's dataset, encompassing both normal Internet users and those with Internet addiction, was strategically divided into three subsets for developing and evaluating a predictive model. This division was done randomly in an 8:1:1 ratio, resulting in a training set of 3916 subjects, a validation set, and a test set, each containing 489 subjects.

The training set's primary role was to facilitate the initial training of the predictive model. The validation set was crucial for testing the model's predictive accuracy and for fine-tuning key parameters such as the learning rate and regularization coefficient. This process was vital to enhance the model's performance and to mitigate overfitting. The test set's function was to conduct the final assessment of the model's prediction capabilities. For effective model training and evaluation, participants were distinctly labeled as "0: normal" for regular Internet users and "1: addiction" for those exhibiting addictive behaviors. This labeling was integral to the model's accuracy in distinguishing between the two groups.

2.2 1D-CNN prediction model

A CNN is a network model that can simulate human vision and recognize images, and was inspired by research on biological visual systems [22, 23]. The CNN can be used to process data such as characters, text, audio, images, and videos, and the 1D-CNN can also be used to deal with traditional data classification problems to achieve the fast learning of the pattern relationships between data [24].

Internet addiction influences are one-dimensional vectors, a multivariate sequential data, involving 20 aspects of influences, containing a total of 50 potential independent variables, and 2 dependent variables. Drawing inspiration from the classical Convolutional Neural Network (CNN) model, AlexNet, a series of experiments were conducted to refine the structure and hyper-parameters of the CNN, culminating in the development of a predictive model for Internet addiction.

This proposed model is architecturally composed of several layers: an input layer for initial data reception, a convolutional layer for feature extraction, a normalization layer to stabilize learning, an activation function layer to introduce non-linearity, a pooling layer for dimensionality reduction, a fully connected layer to synthesize learned features, and an output layer for final prediction. The structural layout of this model is illustrated in Fig. 1.

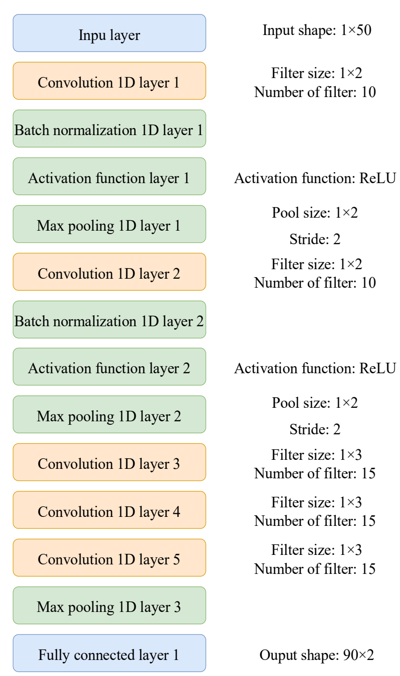

Figure 1. Extraction results of significant influencing factors of Internet addiction.
Source: The authors.

In Fig. 1, he input dimension of the input layer is 1×50, representing the 50 influencing factors of Internet addiction. After data normalization, the characteristics of Internet addiction in the input group with a size of t are $x_t = \{x_{1t}, \dots x_{2t}, \dots x_{50t}\}$. Each group of data corresponds to an addiction sample, and the sample labels are $\{y_t\}_{t=0}^{N}$, $y_t = \{0,1\}$, where y_0 corresponds to the normal Internet use sample, y_1 corresponds to the Internet addiction sample, and N denotes the total number of samples.

In a convolutional neural network model, the inter-layer relationships primarily facilitate data extraction. The concept of a weight matrix is central to this process. It embodies a subset of locally connected neurons from one layer to the next within the network, thereby generating a feature map. The convolutional layer's multiple feature maps, characterized by their respective weight matrices, are commonly referred to as convolution kernels. The mathematical representation of this convolution process is as follows [25]:

$$y^j = f\left(\sum_i k^{ij} * x^i + b^j\right) \tag{1}$$

In this formula, the symbol $*$ denotes the convolution operation. The term y^j signifies the j-th output neuron, x^i represents the i-th input feature, K^{ij} denotes the convolution kernel involved in the convolution operation of this specific layer, and b^j indicates the bias associated with the j-th feature.

The pooling layer, analogous to the complex cells in the primary visual cortex, exhibits robustness and insensitivity to positional shifts in input data. This layer's function compresses the data dimensions, enhancing the computational efficiency of the network model. There are two primary forms of pooling operations:

1. Average Pooling: This method calculates the mean value within the sampling window, thus determining the feature's eigenvalue.
2. Max Pooling: This approach identifies the most salient feature by selecting the maximum value within the sampling window.

In this research, the max pooling technique was employed, expressed as:

$$f(X_k) = \max_{a \in X_k}\{a_1, \dots, a_s\} \tag{2}$$

Here, eq. (2) utilizes max pooling for signal dimensionality reduction. It maps the feature Y, derived from the convolutional layer, into multiple non-overlapping regions X, with k representing indices 1 through K. Each region size is defined as $1 \times s$.

Within Convolutional Neural Networks (CNNs), after undergoing a series of convolutional and pooling layer operations, feature maps are transformed into a one-dimensional vector. This vector subsequently passes through a fully connected layer, where it is assigned weights corresponding to different categories. The output from this layer is then matched against the true labels to assess the effectiveness of the model. The level of deviation between the predicted output and actual labels is calculated using the cross-entropy loss function, which is mathematically expressed in the following way [26]:

$$E = \frac{1}{N}\sum_i L_i = -\frac{1}{N}\sum_i \sum_{i=1}^{M} y_{ic} \log\left(p_{ic}\right) \tag{3}$$

In this formulation, y_{ic} is a binary indicator used to denote whether the true class of the i-th sample corresponds to class c, taking the value of 1 if it does and 0 otherwise. The term p_{ic} signifies the model's predicted probability that the i-th sample observed falls into class c. Here, M represents the total count of categories within the dataset, while N denotes the total number of samples under consideration. E is used to represent the error value, with a lower E value indicative of a more successful learning outcome achieved by the model.

The model's weights and biases are updated using gradient descent after calculating the output category's error. The Adaptive Moment Estimation (Adam) algorithm optimizes this process, influencing the training speed and results [27].

The Internet addiction prediction model, which leverages deep learning techniques, underwent a comprehensive process of feature extraction and optimization of its weights before it was fully trained. This model is structured with an input layer, followed by five convolutional layers, three pooling layers, two layers for normalization, and culminates in a fully connected layer. The convolution kernels are designed with dimensions of 1×2 in the initial two layers, changing to 1×3 in the subsequent three layers. To address the potential problems of vanishing and exploding gradients, the Rectified Linear Unit (ReLU) function was chosen as the activation function. Additionally, batch normalization was introduced post the first and second convolutional layers to enhance the rate at which the model converges. Important parameters of the model include a learning rate set at 0.001, a batch size of 64, the employment of the Adam optimizer, and a training regimen spanning 500 epochs.

2.3 Experiment platform and evaluation indexes

The network model was trained and evaluated on a local laptop equipped with a 12th generation Intel® Core™ i5-12500H processor, 16 GB of RAM, and a 1 TB SSD. It featured an NVIDIA GeForce RTX 3060 Laptop GPU with 4 GB of dedicated memory. The operating system used was Ubuntu 18.04, and the model was developed using Python 3.10 in conjunction with the PyTorch 1.10.0 deep learning framework.

The model's performance was assessed using four key metrics: accuracy, precision, recall, and the F1 score. Accuracy represents the ratio of correctly predicted instances to the overall number of instances. Precision is the measure of the proportion of true positive predictions in all positive predictions made. Recall reflects the proportion of actual positives that have been correctly identified. The F1 score, which is the harmonic mean of precision and recall, is scaled between 0 and 1, where higher values denote better performance of the model.

3 Results

3.1 Internet addiction prediction results

The efficacy of the proposed model for predicting Internet addiction is demonstrated in Fig. 2. During the training phase,

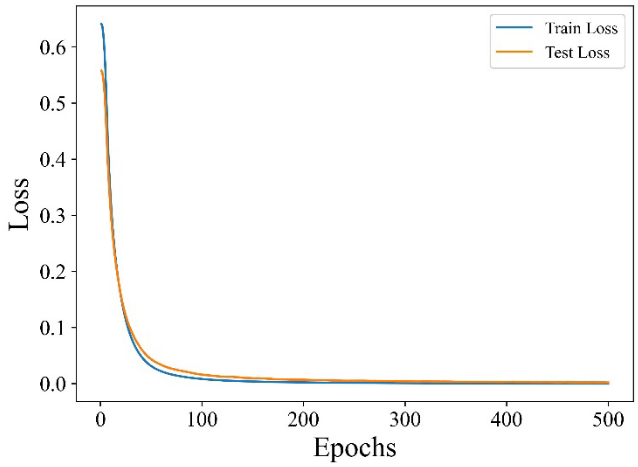

Figure 2. Training results of Internet addiction prediction model.
Source: The authors.

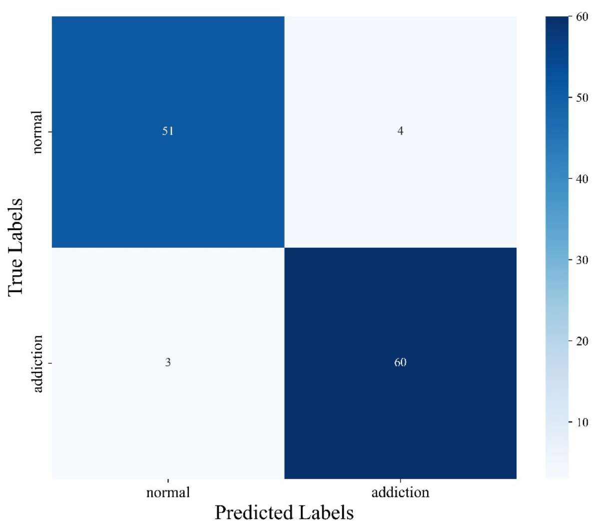

Figure 3. Prediction results of Internet addiction model.
Source: The authors.

the model's loss function initially decreased rapidly, signaling swift learning of the primary data features and robust data-fitting capability. As training progressed, the model reached a state of convergence. In the testing phase, the loss value steadily aligned with the training outcomes, eventually approaching zero, which underscores the model's enhanced predictive proficiency.

For practical evaluation, 118 samples were randomly selected as test data, with results detailed in the confusion matrix in Fig. 3. Of the normal Internet users, 51 were accurately identified, with 3 misclassifications. For users exhibiting Internet addiction, 60 were correctly predicted, with 4 misclassified. This resulted in a prediction accuracy of 94.07%, reflecting the model's substantial generalization capacity.

3.2 Prediction results of different models

The effectiveness of the 1D-CNN model was compared with that of Support Vector Machine (SVM) and Multilayer Perceptron (MLP) algorithms in a detailed assessment. This comparison was based on identical training data and model parameters, focusing on predicting normal and Internet-addicted users. As per the data in Table 2, the 1D-CNN model achieved a 92.77% accuracy, outperforming the SVM and MLP models by 5.30% and 3.21%, respectively.

In terms of precision, recall, and F1 scores for normal Internet users, the 1D-CNN model scored 93.59%, 93.75%, and 0.937, respectively. For Internet-addicted users, the scores were 99.74%, 98.73%, and 0.992, respectively. The F1 scores for normal users predicted by the 1D-CNN were 0.05 and 0.036 higher than those of the SVM and MLP models, while for addicted users, the improvements were 0.07 and 0.06, respectively. These results establish the 1D-CNN model's superior predictive capabilities compared to SVM and MLP models, making it a viable tool for early detection and prevention of Internet addiction among college students.

3.3 Internet addiction significantly affects the results of factors

In Table 1, influences with statistically significant differences were selected as input characteristics for the prediction model, namely, first grade (C1), second grade (C2), single-parent family (C3), non-single-parent family (C4), poor family (C5), rich family (C6), negative behavior (C7), withdrawal behavior (C8), neuroticism (C9), extraversion (C10), no depression (C11), depressed (C12), overly intrusive (C13), understanding and caring (C14), anxiety (C15), no anxiety (C16), poor social support (C17), and very good social support (C18).

Table 2.
Prediction results of different models.

Model	A	Forecast result			Internet addiction		
		P	*R*	*F*1	*P*	*R*	*F*1
SVM	87.47	83.78	93.67	0.884	91.67	93.45	0.925
MLP	89.56	87.46	93.85	0.901	95.85	92.72	0.932
1D-CNN	92.77	93.59	93.75	0.937	99.74	98.73	0.992

Note: A indicates Accuracy, *P* indicates Precision rate, *R* indicates Recall rate, *F*1 indicates F1 score.
Source: The authors.

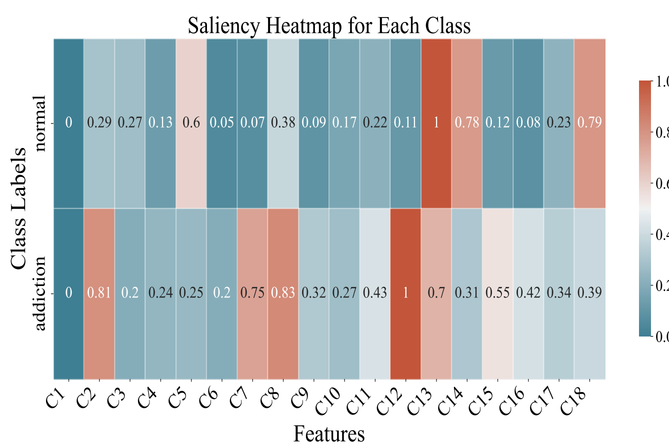

Figure 4. Significant factors affecting Internet addiction.
Source: The authors.

As can be seen in Fig. 4, first grade (C1): there is no difference between Internet addiction and normal Internet users with an impact factor of 0; second grade (C2): the impact factor of 0.81 for Internet addiction is significantly higher than that of 0.27 for normal internet users; single-parent family (C3) and non-single-parent family (C4): the impact factor was less than 0.5, with no significant effect on Internet addiction; poor family (C5): the impact factor for normal Internet users is 0.6; rich family (C6): no difference between Internet addiction and normal Internet users; negative behavior (C7) and withdrawal behavior (C8): significant influence on Internet addiction with an impact factor of 0.75 and 0.83 respectively among Internet addiction, and a smaller weighting among normal Internet users; neuroticism (C9) and extraversion (C10): there is less influence between Internet addiction and normal Internet users; no depression (C11) and depressed (C12): those depressed are more likely to be Internet addicted with an impact factor of 1.0; overly intrusive (C13) and understanding and caring (C14): overly intrusive families have an effect on both Internet addiction and normal Internet users, and understanding and caring families have a positive effect on normal Internet users; anxiety (C15) and no anxiety (C16): those with anxiety have a greater impact on Internet addiction, with an impact factor of 0.55; poor social support (C17) and very good social support (C18): individuals with very good social support have a greater weight among normal Internet users.

Based on the above analyses, the priority of the influence factors of Internet addiction is: C12 (1.0) > C8 (0.83) > C2 (0.81) > C7 (0.75) > C13 (0.7) > C15 (0.55) > C11 (0.43) > C16 (0.42) > C18 (0.39) > C17 (0.34) > C9 (0.32) > C14 (0.31) > C10 (0.27) > C5 (0.25) > C4 (0.24) > C3 (0.2) > C6 (0.2) > C1 (0), and the priority of the influence factors of normal Internet users is: C13 (1.0) > C18 (0.79) > C14 (0.78) > C5 (0.6) > C8 (0.38) > C2 (0.29) > C3 (0.27) > C17 (0.23) > C11 (0.22) > C10 (0.17) > C4 (0.13) > C15 (0.12) > C12 (0.11) > C9 (0.09) > C16 (0.08) > C7 (0.07) > C6 (0.05) > C1 (0). This ranking can help to understand the different effects of the factors on Internet addiction and normal Internet users, and provide direction for future interventions and research.

4 Conclusions

The statistical analysis results of the questionnaire show that the Internet addiction rate among college students is 20.71%. There are no statistical differences in Internet addiction among age, gender, and being an only child, $P > 0.05$. There are statistically significant differences in Internet addiction among grade, whether a single parent is a single parent, family economic conditions, depression, anxiety, and mobile phone addiction, $P < 0.05$. There are significant statistical differences in Big Five personality status, parenting styles, and social support, $P < 0.001$.

Using statistical results as original data, we built an Internet addiction prediction model based on one-dimensional convolutional neural network addiction. The results show that the accuracy of the model is 92.77%; the accuracy of predicting normal Internet users is 93.59%, and the recall rate is 93.59%. is 93.75, and the F1 score is 0.937; the precision rate of predicting Internet addiction is 99.74%, the recall rate is 98.73%, and the F1 score is 0.992. Compared with the Internet addiction prediction model constructed by SVM and MLP algorithms, its accuracy is 5.30 and 3.21 percentage points higher respectively. The F1 scores of normal Internet users are 0.05 and 0.036 higher respectively, and the F1 scores of Internet addiction are 0.07 and 0.06 higher respectively.

On this basis, the influencing factors with statistically significant differences were selected as the input features of the prediction model, and the gradient of the input features was calculated using the method of gradient back propagation to get the degree of influence of the input features on the category, and the results showed that second-grade students, negative behaviors and withdrawal behaviors, having depression, over-interfering, and having anxiety had a greater influence on Internet addiction, and the influence factors were all greater than 0.5. Family poverty, excessive interference, understanding and caring family and good social support have a greater impact on normal Internet use and can help students avoid Internet addiction.

After the above analysis, more attention should be paid to second-year students, students with negative and withdrawal behaviors, students with depression, students with over-interfering families and students with anxiety when preventing internet addiction in college to provide a research basis for preventing internet addiction in college students.

5 Discussion

The Internet's integral role in daily life has led to concerns regarding Internet addiction, negatively impacting academic, physical, and mental health. Di et al. [12] developed a predictive model using a support vector machine, identifying neuroticism and poor planning as key contributors to this addiction. Huang et al. [28] found a positive correlation between narrative disorders and cell phone addiction. Chemnad et al. [8] examined environmental influences on Internet addiction, using demographic data, diagnostic questionnaires, and scales related to family relationships and health behaviors. They discovered that family and school environments significantly predict Internet addiction, with a 29.64% prevalence rate. Xie et al. [29] demonstrated a positive correlation between anxiety, depression, and Internet addiction. Kuss et al. [30] focused on adolescents, noting that traits like low emotional stability and introversion increase susceptibility to Internet addiction. Chemnad et al. [8] also studied the link between social media usage and problematic Internet behaviors, suggesting the need for monitoring individual application usage for effective intervention strategies. Ioannidis et al. [31] used Logistic Regression, Random Forest, and Bayesian Machine Learning to create predictive models for problematic Internet use, validating the use of impulsivity and compulsivity as key factors in these models.

In this study, a prediction model for Internet addiction was constructed based on a one-dimensional convolutional neural network. One-dimensional convolutional neural networks (1D-CNN) demonstrate superior efficiency in feature acquisition and robustness in model construction with traditional data. Current Internet addiction questionnaires often rely on self-reporting, which may not accurately reflect addiction behaviors due to students avoiding direct questions. This approach fails to identify and intervene in early stages of addiction considering psychological, social, and familial factors. Our method evaluates Internet addiction based on abnormalities in these factors, employing machine learning and deep learning techniques for precise predictions. Specifically, 1D-CNN effectively identifies the risk of Internet addiction among college students in a university in Shandong Province, aiming to provide targeted early intervention and reduce addiction rates. Future research could implement predictive software in real-time applications to monitor Internet addiction symptoms among college students, thus contributing to early prevention strategies in university settings.

References

[1] Duong, X.L., Liaw, S.Y., and Augustin, J.L.P.M., How has Internet Addiction Tracked Over the Last Decade? A Literature Review and 3C Paradigm for Future Research. International journal of preventive medicine, 11, pp. 175, 2020. DOI: https://10.4103/ijpvm.IJPVM_212_20

[2] Bickham, D.S., Current Research and Viewpoints on Internet Addiction in Adolescents. Current Pediatrics Reports, 9(1), pp. 1-10, 2021. DOI: https://10.1007/s40124-020-00236-3

[3] Chou, C., Condron, L., and Belland, J.C., A Review of the Research on Internet Addiction. Educational Psychology Review, 17(4), pp. 363-388, 2005. DOI: https://10.1007/s10648-005-8138-1

[4] King, D.L., Chamberlain, S.R., Carragher, N., Billieux, J., Stein, D., Mueller, K., Potenza, M.N., Rumpf, H.J., Saunders, J., Starcevic, V., Demetrovics, Z., Brand, M., Lee, H.K., Spada, M., Lindenberg, K., Wu, A.M.S., Lemenager, T., Pallesen, S., Achab, S., Kyrios, M., Higuchi, S., Fineberg, N.A., and Delfabbro, P.H., Screening and assessment tools for gaming disorder: A comprehensive systematic review. Clinical Psychology Review, 77, pp. 101831, 2020. DOI: https://10.1016/j.cpr.2020.101831

[5] Saunders, J.B., Hao, W., Long, J., King, D.L., Mann, K., Fauth-Bühler, M., Rumpf, H.J., Bowden-Jones, H., Rahimi-Movaghar, A., Chung, T., Chan, E., Bahar, N., Achab, S., Lee, H.K., Potenza, M., Petry, N., Spritzer, D., Ambekar, A., Derevensky, J., Griffiths, M.D., Pontes, H.M., Kuss, D., Higuchi, S., Mihara, S., Assangangkornchai, S., Sharma, M., Kashef, A.E., Ip, P., Farrell, M., Scafato, E., Carragher, N., and Poznyak, V., Gaming disorder: Its delineation as an important condition for diagnosis, management, and prevention. Journal of Behavioral Addictions, 6(3), pp. 271-279, 2017. DOI: https://10.1556/2006.6.2017.039

[6] Kardefelt-Winther, D., A conceptual and methodological critique of internet addiction research: Towards a model of compensatory internet use. Computers in Human Behavior, 31, pp. 351-354, 2014. DOI: https://10.1016/j.chb.2013.10.059

[7] Yang, K., Chen, J., and Liu, J., Relationship between personality traits, depression, anxiety and Internet overuse in junior middle school students. Journal of Psychiatry, 32(6), pp. 429-432, 2019. DOI: https://10.3969/j.issn.2095-9346.2019.06.007

[8] Chemnad, K., Aziz, M., Belhaouari, S.B., and Ali, R., The interplay between social media use and problematic internet usage: Four behavioral patterns. Heliyon, 9(5), pp. e15745, 2023. DOI: https://10.1016/j.heliyon.2023.e15745

[9] Lai, C.M., Mak, K.K., Watanabe, H., Ang, R.P., Pang, J.S., and Ho, R.C.M., Psychometric Properties of the Internet Addiction Test in Chinese Adolescents. Journal of Pediatric Psychology, 38(7), pp. 794-807, 2013. DOI: https://10.1093/jpepsy/jst022

[10] Davies, A., Veličković, P., Buesing, L., Blackwell, S., Zheng, D., Tomašev, N., Tanburn, R., Battaglia, P., Blundell, C., Juhász, A., Lackenby, M., Williamson, G., Hassabis, D., and Kohli, P., Advancing mathematics by guiding human intuition with AI. Nature, 600(7887), pp. 70-74, 2021. DOI: https://10.1038/s41586-021-04086-x

[11] Han, Y., Liao, Y., Ma, X., Guo, X., Li, C., and Liu, X., Analysis and prediction of the penetration of renewable energy in power systems using artificial neural network. Renewable Energy, 215, pp. 118914, 2023. DOI: https://10.1016/j.renene.2023.118914

[12] Di, Z., Gong, X., Shi, J., Ahmed, H.O.A., and Nandi, A.K., Internet addiction disorder detection of Chinese college students using several personality questionnaire data and support vector machine. Addictive Behaviors Reports, 10(1), pp. 100200, 2019. DOI: https://10.1016/j.abrep.2019.100200

[13] Chaudhury, P., and Kumar, T.H., A Study on impact of smartphone addiction on academic performance. International Journal of Engineering & Technology, 7(2.6), pp. 50, 2018. DOI: https://10.14419/ijet.v7i2.6.10066

[14] Shae, Z.Y., and Tsai, J.J.P., In Deep Learning Mechanism for Pervasive Internet Addiction Prediction, 2020 IEEE Second International Conference on Cognitive Machine Intelligence (CogMI), Atlanta, GA, USA, IEEE: Atlanta, GA, USA, 2020.

[15] Mak, K.K., Lai, C.M., Ko, C.H., Chou, C., Kim, D.I., Watanabe, H., and Ho, R.C.M., Psychometric Properties of the Revised Chen Internet Addiction Scale (CIAS-R) in Chinese Adolescents. Journal of Abnormal Child Psychology, 42(7), pp. 1237-1245, 2014. DOI: https://10.1007/s10802-014-9851-3

[16] Zhang, X., Wang, M.C., He, L., Jie, L., and Deng, J., The development and psychometric evaluation of the Chinese Big Five Personality Inventory-15. PLOS ONE, 14(8), pp. e0221621, 2019. DOI: https://10.1371/journal.pone.0221621

[17] Huang, Q., Li, Y., Huang, S., Qi, J., Shao, T., Chen, X., Liao, Z., Lin, S., Zhang, X., Cai, Y., and Chen, H., Smartphone Use and Sleep Quality in Chinese College Students: A Preliminary Study. Frontiers in Psychiatry, 11, pp. 352, 2020. DOI: https://10.3389/fpsyt.2020.00352

[18] Dunstan, D.A., and Scott, N., Norms for Zung's Self-rating Anxiety Scale. BMC Psychiatry, 20(1), pp. 90, 2020. DOI: https://10.1186/s12888-019-2427-6

[19] Khalid, A., Zhang, Q., Wang, W., Ghaffari, A.S., and Pan, F., The relationship between procrastination, perceived stress, saliva alpha-amylase level and parenting styles in Chinese first year medical students. Psychology Research and Behavior Management, 12(11), pp. 489-498, 2019. DOI: https://10.2147/prbm.S207430

[20] Zhang, Q., Ran, G., and Ren, J., Parental Psychological Control and Addiction Behaviors in Smartphone and Internet: The Mediating Role of Shyness among Adolescents. International Journal of Environmental Research and Public Health, 19(24), pp. 16702, 2022. DOI: https://10.3390/ijerph192416702

[21] Kerres Malecki, C. and Kilpatrick Demary, M., Measuring perceived social support: Development of the child and adolescent social support scale (CASSS). Psychology in the Schools, 39(1), pp. 1-18, 2001. DOI: https://10.1002/pits.10004

[22] Imamverdiyev, Y., and Sukhostat, L., Lithological facies classification using deep convolutional neural network. Journal of Petroleum Science and Engineering, 174, pp. 216-228, 2019. DOI: https://10.1016/j.petrol.2018.11.023

[23] Liu, J.J., and Liu, J.C., Integrating deep learning and logging data analytics for lithofacies classification and 3D modeling of tight sandstone reservoirs. Geoscience Frontiers, 13(1), pp. 101311, 2022. DOI: https://10.1016/j.gsf.2021.101311

[24] Guo, W., Shen, W., Zhou, S., Xue, H., Liu, D., and Wang, N., Shale favorable area optimization in coal-bearing series: A case study from the Shanxi Formation in Northern Ordos Basin, China. Energy Exploration & Exploitation, 36(5), pp. 1295-1309, 2017. DOI: https://10.1177/0144598717748951

[25] Najafabadi, M.M., Villanustre, F., Khoshgoftaar, T.M., Seliya, N., Wald, R., and Muharemagic, E., Deep learning applications and challenges in big data analytics. Journal of big data, 2(1), pp. 1-21, 2015. DOI: https://10.1186/s40537-014-0007-7

[26] Kulkarni, S.R., and Harman, G., Statistical learning theory: a tutorial. WIREs Computational Statistics, 3(6), pp. 543-556, 2011. DOI: https://10.1002/wics.179

[27] Kingma, D.P., and Ba, J., Adam: A Method for Stochastic Optimization. arXiv.org, pp. 2014. DOI: https://10.48550/arXiv.1412.6980

[28] Huang, H., Wan, X., Lu, G., Ding, Y., and Chen, C., The Relationship Between Alexithymia and Mobile Phone Addiction Among Mainland Chinese Students: A Meta-Analysis. Frontiers in Psychiatry, 13(11), pp. 2022. DOI: https://10.3389/fpsyt.2022.754542

[29] Xie, X., Cheng, H., and Chen, Z., Anxiety predicts internet addiction, which predicts depression among male college students: A cross-lagged comparison by sex. Frontiers in Psychology, 13, pp. 1102066, 2023. DOI: https://10.3389/fpsyg.2022.1102066
[30] Kuss, D.J., van Rooij, A.J., Shorter, G.W., Griffiths, M.D., and van de Mheen, D., Internet addiction in adolescents: Prevalence and risk factors. Computers in Human Behavior, 29(5), pp. 1987-1996, 2013. DOI: https://10.1016/j.chb.2013.04.002
[31] Ioannidis, K., Chamberlain, S.R., Treder, M.S., Kiraly, F., Leppink, E.W., Redden, S.A., Stein, D.J., Lochner, C., and Grant, J.E., Problematic internet use (PIU): Associations with the impulsive-compulsive spectrum. An application of machine learning in psychiatry. Journal of Psychiatric Research, 83, pp. 94-102, 2016. DOI: https://10.1016/j.jpsychires.2016.08.010

Xi Wang, received her Master's degree in Social and Organizational Psychology from the University of Exeter, UK. She is currently pursuing her PhD at the School of Educational Sciences, Ludong University. Her areas of interest include positive psychology, corrective psychology, and statistical analysis.
ORCID: 0009-0008-7887-0539

Enyou Zhang, received his Master's degree in Education from Henan University, China. He majors in criminal psychology, correctional psychology and police officers' mental health.
ORCID: 0009-0003-0124-4664

Yingjun Cui, received her Master's degree in Agricultural Engineering from Shandong University of Technology, China. She is currently pursuing her PhD at the School of Agricultural Engineering and Food Science, Shandong University of Technology. Her areas of interest include statistical analysis.
ORCID: 0009-0002-8290-2368

Jie Huang, received his Master's degree in Agricultural Engineering from Shandong University of Technology, China. He is currently pursuing her PhD at the School of Agricultural Engineering and Food Science, Shandong University of Technology. His areas of interest include statistical analysis.
ORCID: 0009-0007-4137-7508

Meng Cheng, received her PhD. in Agricultural Engineering from Shandong University of Technology, China. She is the corresponding author and currently a teacher at Shandong University of Technology. Her research work is aimed at mental health education of college students and big data analysis.
ORCID: 0000-0003-3316-3204