Exploring Factors Influencing Adolescent Short Video Addiction: A Latent Profile Analysis and Machine Learning Approach

Huiying Chen [1], Chengyi Zhao [2], Ting Jiang [2], Yueyang Wang [2], Yizhen Ren [2, *]

[1, 2] Faculty of Psychology, Beijing Normal University, Beijing 100875, China

[2] Department of Psychology, Guangzhou University, Guangzhou 510006, China

* Corresponding Author Email: 874096692@qq.com

Abstract. The proliferation of short-form video platforms has precipitated mounting concerns regarding addiction among adolescents. This investigation examines Short-Form Video Addiction (SFVA) through the integration of Latent Profile Analysis (LPA) and machine learning techniques. Analysis of data from 11,687 adolescents revealed four distinct user profiles: low-risk self-controlled (18.03%), high-impact low-awareness (24.60%), high-risk dependent (25.12%), and moderate-risk potential (32.25%) groups. Subsequent machine learning analysis, utilizing random forest algorithm with 89% accuracy, identified peer attachment, parental overprotection, and parenting styles as paramount predictors of SFVA development. These findings illuminate the multifaceted nature of adolescent SFVA and provide empirically-grounded recommendations for targeted intervention strategies.

Keywords: short video addiction; machine learning; latent profile analysis; random forest algorithm; nonlinear relationships.

1. Introduction

The rise of short video platforms has led to an increase in user engagement, particularly among young people. These videos, often produced by independent creators, captivate audiences with their brief and engaging content. However, excessive viewing can negatively impact adolescents' values and overall development.

Adolescence is a crucial period for growth in cognitive abilities, emotional management, and social skills. Overindulgence in digital media can distract from important developmental tasks, potentially leading to Short-Form Video Addiction (SFVA).

This study aims to explore the factors contributing to SFVA using Latent Profile Analysis and machine learning methods. While previous research has centered on internet and gaming addictions, there has been less focus on short video addiction specifically. By comparing various machine learning algorithms, this research seeks to identify the most effective model for analyzing SFVA data and the key predictors across personal, educational, familial, and psychological dimensions.

Understanding SFVA is crucial for protecting adolescent well-being and supporting healthier development. Effective interventions can guide young people in critically engaging with online content, fostering self-identity, and building psychological resilience.

The insights gained from this research will be valuable for families, educators, and policymakers, emphasizing the need for a balanced approach to technology use. By addressing the complexities of SFVA, we can help ensure that adolescent’s benefit from their digital experiences while maintaining their personal and social growth.

2. Research ideas and methods

2.1. Research Question

This study seeks to explore the heterogeneity of adolescent Short-Form Video Addiction (SFVA) behaviors and the factors that predict these behaviors across different domains. Specifically, it aims to address the following research questions:

1. Is there significant variability in SFVA behaviors among adolescents, and can they be classified into distinct risk groups based on their SFVA scale scores?

2. Which individual, family, school, and peer factors serve as strong predictors of adolescent SFVA, and what are the nonlinear relationships among these factors?

To achieve these objectives, the study will employ Latent Profile Analysis (LPA) on collected survey data to uncover diverse SFVA behaviors and identify distinct risk groups, thereby informing targeted intervention strategies. Additionally, machine learning models and SHAP (SHapley Additive exPlanations) analysis will be utilized to pinpoint key nonlinear predictors across multiple dimensions, including personal, family, school, and peer factors.

2.2. Measurement and Latent Profile Analysis

A six-item scale developed by Choi and Lim (2016) was used to measure Short-Form Video Addiction (SFVA) in adolescents. Adapted from the Social Media Disorder Scale by Van et al. (2016), this scale includes unique aspects of short videos. Participants rated items on a 7-point Likert scale (1 = not at all true, 7 = completely true), with higher scores indicating more severe addiction symptoms like preoccupation and social dysfunction.

The scale's reliability and validity are well-established. Choi and Lim (2016) confirmed its unidimensionality through factor analyses, reporting strong internal consistency (Cronbach's α = 0.89).

Table 1. Profile Weighted Mean and Standard Deviation Values

Latent Class		X1	X2	X3	X4	X5	X6
Class 1	Mean	2.226	1.395	1.415	1.564	1.261	1.327
	SD	1.34	0.969	1.03	1.264	0.902	0.907
Class 2	Mean	5.612	5.713	5.37	3.381	1.445	1.517
	SD	1.34	0.969	1.03	1.264	0.902	0.907
Class 3	Mean	4.16	3.771	3.603	3.875	3.442	3.632
	SD	1.34	0.969	1.03	1.264	0.902	0.907
Class 4	Mean	5.768	5.784	5.61	5.643	5.287	5.443
	SD	1.34	0.969	1.03	1.264	0.902	0.907

Latent Profile Analysis (LPA) was conducted using Mplus 8.4 to identify behavioral patterns of Short-Form Video Addiction (SFVA) among adolescents and delineate risk groups. This method effectively uncovers complex patterns related to social network use and addiction. Various models were compared using fit indices (AIC, BIC, Entropy) to determine the optimal number of profiles, revealing four distinct patterns of short video usage. The model fit indices indicated that the four-category model was optimal. Table 1 presents the mean and variance estimates for each category across six dimensions of short video addiction, along with 95% confidence intervals. Additionally, Figure 1 illustrates the latent profile plots for these four categories.

Descriptive statistics show differences in short video usage among teenagers, with average scores ranging from 2.738 to 3.954. Key findings across six dimensions of short video addiction include: difficulty focusing on academics (mean = 3.954, SD = 1.897), insomnia resulting from excessive use (mean = 3.545, SD = 1.941), and interference with social activities (mean = 3.413, SD = 1.882). Analysis of skewness and kurtosis for all variables suggests a relatively normal distribution of the data.

The classifications are as follows:

Low-Risk Self-Control Group (Class 1, 18.03%): Scored low on across all dimensions, particularly X5 (“I feel anxious if I cannot access SNS”) and X6 (“I have attempted to spend less time on SNS but have not succeeded”), reflecting strong self-control and a low risk of addiction.

High Impact, Low Awareness Group (Class 2, 24.60%): Scored high on academic, sleep, and social dimensions but low on anxiety and control, suggesting a significant influence from short videos without a full understanding of this impact.

Medium-Risk Potential Group (Class 3, 32.25%): This group scored moderately high, especially on X4 (“My family or friends think that I spend too much time on SNS”) and X6. This indicates a frequent use of SNS with potential risks for future development.

High-Risk Dependence Group (Class 4, 25.12%): This group scored high across all dimensions, particularly on X5and X6, indicating a severe dependence that disrupts daily functioning.

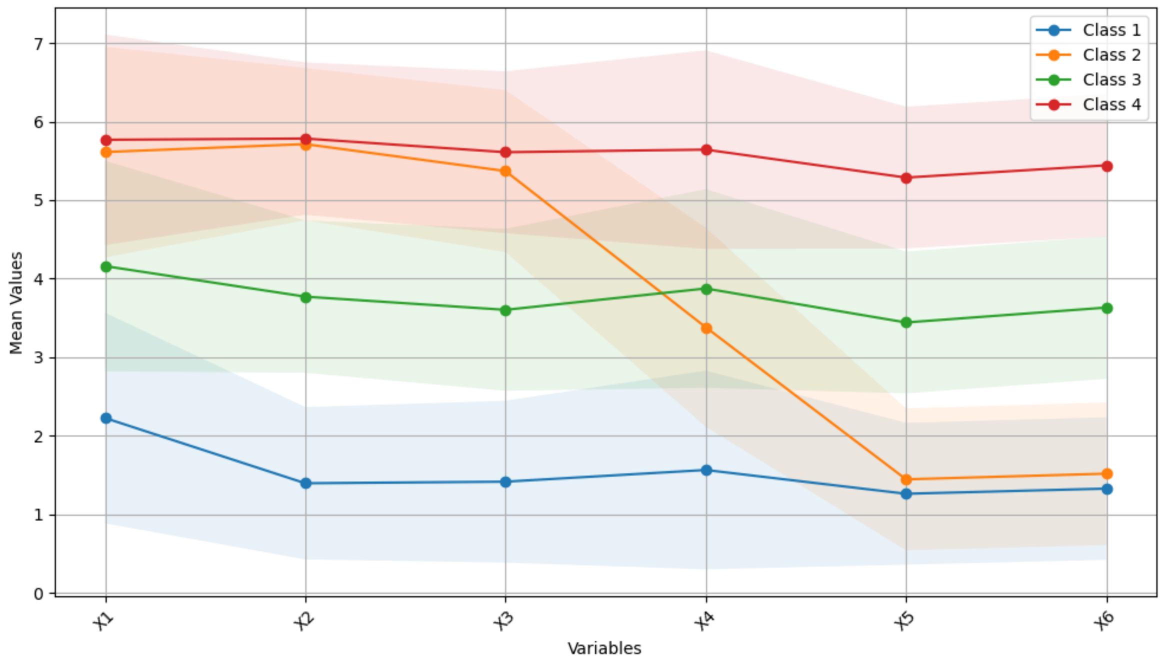

Fig. 1 Latent Profile Plot

2.3. Machine Learning

Study 1 divides adolescent addiction groups into four categories through latent profile analysis. On this basis, Study 2 investigated the factors that affect short video addiction, including family upbringing, parent-child relationships, peer interaction, and school environment, including self-factors such as depression (PHQ-9), generalized anxiety disorder (GAD-7), childhood negative experiences, and complex post-traumatic stress disorder, as well as family and peer factors such as parenting styles, strict discipline, and peer attachment. The data was collected through "Wenjuanxing" in 14 urban areas of Jiangsu and Henan in June 2023 to ensure no missing values and obtained approval from the ethics committee.

To evaluate the performance of different machine learning methods in predicting the risk of short video addiction, six classic algorithms were selected, including random forest, support vector machine, neural network, logistic regression, decision tree, and naive Bayes. To assess model performance comprehensively, four key evaluation metrics were employed:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}. \quad (1)$$

$$Precision = \frac{TP}{TP+FP}. \quad (2)$$

$$Recall = \frac{TP}{TP+FN}. \quad (3)$$

$$F1 = \frac{2*Precision*Recall}{Precision+Recall}. \quad (4)$$

where accuracy measures the overall correct prediction rate, precision indicates the proportion of true positive predictions among all positive predictions, recall represents the proportion of actual positive cases correctly identified, and F1 score provides a balanced measure between precision and recall. The selection of these algorithms takes into account their theoretical maturity and interpretability, in order to compare the performance of different algorithm families and ensure comprehensive and reliable research results. SMOTE oversampling was implemented to handle imbalanced data, ensuring fairness and practicality of the results. Meanwhile, statistical testing confirmed the significance of the results, enhancing the persuasiveness of the conclusion.

Table 2. Performance Metrics of Different Machine Learning Models

Model	accuracy	precision	recall	F1 score
Decision Tree	65.42%	65.42%	65.64%	65.51%
Random Forest	89.08%	89.19%	89.22%	89.39%
LightGBM	67.96%	68.02%	68.04%	68.16%
Support Vector Machine (SVM)	56.75%	57.18%	56.43%	56.67%
Naive Bayes	39.83%	42.67%	40.66%	37.81%
Logistic Regression	43.49%	43.51%	42.76%	42.87%

After testing and comparing six machine learning algorithms: random forest, support vector machine (SVM), neural network, logistic regression, decision tree, and naive Bayes—the random forest model demonstrated superior performance across all evaluation metrics, achieving accuracy, precision, recall, and F1 score around 89%, significantly outpacing other models. The decision tree and LightGBM models exhibited relatively similar performance, with accuracy ranging from 65% to 68%, and other metrics also being fairly balanced, but they were still notably inferior to the random forest model. This discrepancy may be attributed to the differing methods of data partitioning and feature selection employed by the decision tree and LightGBM, affecting their performance on this dataset. In contrast, the support vector machine, naive Bayes, and logistic regression models performed poorly, with accuracy below 60%.

3. Results

3.1. Heterogeneous Patterns of Short-Form Video Addiction: A Latent Profile Analysis

The Latent Profile Analysis (LPA) revealed heterogeneous patterns in adolescent short-form video addiction (SFVA), identifying four distinct profiles: low-risk self-controlled, high-impact low-awareness, high-risk dependent, and moderate-risk potential groups. This classification extends beyond the initial three-group hypothesis, particularly with the emergence of the high-impact low-awareness group, characterized by significant academic, sleep, and social impairments despite limited risk perception.

The identification of moderate-risk potential and high-risk dependent groups highlights populations requiring immediate intervention. These findings enhance our understanding of SFVA's complexity and provide empirical evidence for targeted intervention strategies. While LPA offers valuable insights into behavioral pattern diversity, it primarily focuses on observable response patterns rather than underlying causal mechanisms.

Recognizing SFVA as a complex psychosocial phenomenon influenced by individual, familial, peer, and academic factors, subsequent research will employ machine learning methods and SHAP (Shapley Additive Explanations) analysis. This approach aims to quantify feature importance and reveal non-linear relationships between various factors and SFVA, thereby establishing a foundation for more effective, targeted interventions by educators, mental health professionals, and parents.

3.2. Effectiveness of Machine Learning Approach in SFVA Prediction

The application of machine learning techniques, particularly the random forest algorithm, demonstrated robust capability in mining complex patterns within the SFVA dataset. The superior performance of the random forest model, achieving approximately 89% accuracy across all evaluation metrics, validates its effectiveness in capturing intricate relationships between multiple risk factors. This methodological approach not only ensures reliable prediction but also offers interpretable insights through feature importance analysis. Notably, the random forest's ability to quantify feature importance provides valuable insights into the relative contribution of various factors, facilitating the construction of a comprehensive adolescent SFVA model. The consistency of results across different evaluation metrics (accuracy, precision, recall, and F1 score) further substantiates the credibility of our findings, as illustrated in the comparative analysis of confusion matrices (see Fig. 2), indicating that machine learning serves as a reliable tool for understanding and modeling the multifaceted nature of adolescent SFVA.

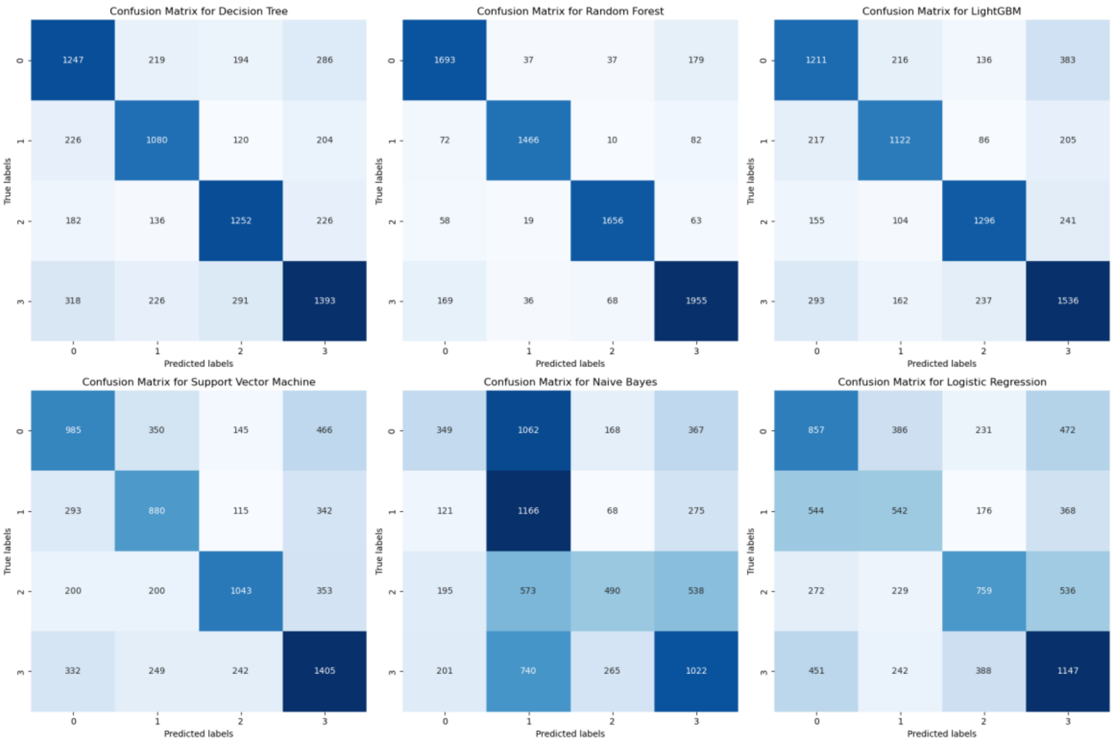

Fig. 2 Comparative Analysis of Confusion Matrices

3.3. Identification of Key Risk Factors Through Feature Importance Analysis

The random forest model identified and quantified the relative contributions of various risk factors associated with adolescent SFVA. The top-ranking predictors, in descending order of importance, are as follows:

Peer attachment (0.026)
Maternal overprotection (0.026)
Authoritarian parenting style (0.025)
Permissive parenting style (0.024)
Paternal overprotection (0.024)
Authoritative parenting style (0.023)

These findings underscore the predominant influence of peer relationships and parenting patterns in SFVA development. Notably, both attachment-related factors (peer attachment) and parenting styles (authoritarian, permissive, and authoritative) emerge as critical determinants, suggesting that intervention strategies should prioritize addressing these psychosocial domains.

The identified risk factors align with established psychological theories and developmental frameworks. The prominence of peer attachment as the leading predictor (0.026) corresponds with attachment theory and adolescent development literature, suggesting that teenagers with insecure peer attachments might turn to short-form videos as a substitute for social connection and emotional regulation. This finding resonates with the social compensation hypothesis in digital media use (Valkenburg & Peter, 2009).

Maternal and paternal overprotection (0.026 and 0.024 respectively) emerged as significant predictors, consistent with the self-determination theory (Ryan & Deci, 2020). Overprotective parenting may restrict adolescents' autonomy and competence development, potentially driving them towards excessive online engagement as a means of exercising personal control and independence. The high ranking of authoritarian parenting (0.025) further supports this interpretation, as rigid parental control might prompt adolescents to seek psychological escape through digital platforms.

The significant influence of permissive parenting (0.024) presents an interesting contrast, suggesting that insufficient behavioral boundaries might also contribute to SFVA. This aligns with the behavioral addiction framework (Brand et al., 2019), where lack of external regulation may impair the development of self-regulatory capabilities necessary for healthy media consumption (Koning et al., 2018).

4. Conclusions

This investigation enriches our understanding of adolescent short-form video addiction (SFVA) through a dual-methodological framework. Employing Latent Profile Analysis (LPA), we delineated four distinct SFVA patterns: low-risk self-controlled, high-impact low-awareness, high-risk dependent, and moderate-risk potential groups, illuminating the heterogeneous manifestations of adolescent SFVA and establishing a foundation for targeted interventions.

Through advanced machine learning techniques, notably the random forest algorithm which demonstrated exceptional performance with 89% accuracy, we systematically quantified the relative contributions of diverse risk factors. The feature importance analysis unveiled peer attachment, parental overprotection, and parenting styles as the paramount predictors of SFVA. These empirical findings corroborate established psychological frameworks, including attachment theory (Allen & Tan, 2016) and self-determination theory (Ryan & Deci, 2020), suggesting that both compromised social connections and suboptimal parenting practices may catalyze SFVA development.

The synthesis of LPA and machine learning methodologies not only advances SFVA research but also yields substantial practical implications. Our findings indicate that prevention and intervention initiatives should prioritize enhancing peer relationships, optimizing parenting practices, and fostering healthy parent-child dynamics. These outcomes provide evidence-based directives for educators, mental health professionals, and parents in addressing adolescent SFVA. Subsequent research endeavors could expand upon these findings by examining the temporal stability of SFVA profiles and investigating potential protective mechanisms that might attenuate addiction risks, thereby establishing a robust framework for future investigations in behavioral addiction research.

References

[1] Allen, J. P., & Tan, J. (2016). The multiple facets of attachment in adolescence. *In J. Cassidy & P. Shaver (Eds.), Handbook of attachment* (3rd ed., pp. 399–415). New York NY: Guilford.

[2] Brand M, Wegmann E, Stark R, Müller A, Wölfling K, Robbins TW, Potenza MN. The Interaction of Person-Affect-Cognition-Execution (I-PACE) model for addictive behaviors: Update, generalization to addictive behaviors beyond internet-use disorders, and specification of the process character of addictive behaviors. *Neurosci Biobehav Rev*. 2019 Sep; 104:1-10.

[3] Choi, S. B., & Lim, M. S. (2016). Effects of social and technology overload on psychological well-being in young South Korean adults: The mediatory role of social network service addiction. *Computers in Human Behavior*, 61, 245–254.

[4] Koning IM, Peeters M, Finkenauer C, van den Eijnden RJJM. Bidirectional effects of Internet-specific parenting practices and compulsive social media and Internet game use. *J Behav Addict*. 2018 Sep 1;7(3):624-632.

[5] Ryan, R. M., & Deci, E. L. (2020). Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions. *Contemporary educational psychology*, 61, 101860.

[6] Valkenburg, P. M., & Peter, J. (2009). Social consequences of the Internet for adolescents: *A decade of research. Current Directions in Psychological Science,* 18(1), 1–5.

[7] Van den Eijnden, R. J., Lemmens, J. S., & Valkenburg, P. M. (2016). The social media disorder scale. *Computers in human behavior*, 61, 478-487.