computation

MDPI

Article

Predicting Pornography Use Among Adolescents in Spain: Findings from Ensemble Tree Models and Explainable Machine Learning

Jorge de Andrés-Sánchez [1,*], Ángel Belzunegui-Eraso [2], Inma Pastor-Gosálbez [2] and Anna Sánchez-Aragón [2]

[1] Social and Business Research Laboratory, Department of Business Administration, Universitat Rovira i Virgili, Campus de Bellissens, 43204 Reus, Spain
[2] Social and Business Research Laboratory, Department of History and History of Art, Universitat Rovira i Virgili, Campus Catalunya, 43002 Tarragona, Spain; angel.belzunegui@urv.cat (Á.B.-E.); inma.pastor@urv.cat (I.P.-G.); annamaria.sanchez@urv.cat (A.S.-A.)
* Correspondence: jorge.deandres@urv.cat

Abstract

Pornography use during adolescence is a relevant issue from social, educational, and public health perspectives. As with other behaviours, its correlates may involve complex and nonlinear relationships. This study uses data from the 2023 Spanish Survey on Drug Use in Secondary Education (ESTUDES), a major source for analysing potentially addictive behaviours among adolescents in Spain because of its large sample size (original sample: N = 42,208; complete-case analytical sample: N = 33,543). Pornography use was modelled as a binary outcome using logistic regression, XGBoost, LightGBM, and CatBoost. The models included sociodemographic characteristics, family-related factors, parental control, substance use, selected sexual behaviours, indicators of mental well-being, and addictive Internet use. The four methods showed similar predictive performance. CatBoost achieved the highest AUC and the lowest Brier score and log loss, whereas LightGBM obtained the highest accuracy, specificity, and precision. Logistic regression yielded the highest sensitivity, negative predictive value, balanced accuracy, and F1-score. SHAP analysis identified sex as the most influential predictor, followed by addictive Internet use, cannabis and alcohol use, family conflict, and selected risky sexual behaviours. The findings suggest that pornography use among adolescents forms part of a broader behavioural and psychosocial profile. Whereas sex operated mainly as a strong direct predictor, problematic Internet use showed substantial interactions with age and cannabis and alcohol use. This study demonstrates how explainable machine learning can complement conventional regression by identifying the main correlates, nonlinear patterns, and interactions associated with adolescent pornography use.

Keywords: adolescent pornography use; addictive use of Internet; risk behaviours; adolescent health; explainable machine learning; XGBoost; LightGBM; CatBoost; Shapley Additive Explanations

Check for updates

Academic Editor: Francesco Cauteruccio

Received: 28 June 2026
Revised: 24 July 2026
Accepted: 2 August 2026
Published: 4 August 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Pornography use among adolescents is a relevant research topic from educational, social, and public health perspectives because of its multifaceted nature [1]. The expansion of Internet access and mobile devices has made sexually explicit content readily accessible, privately consumable, and often subject to limited supervision [2]. Understanding which

factors are associated with adolescent pornography use is important because the literature has linked this behaviour to several dimensions of adolescent development, including risky sexual behaviours, psychosomatic symptoms, sexual socialisation, and inaccurate sexual beliefs [3–5]. Longitudinal research also suggests that the onset of pornography use may be associated with sex, age, previous sexual experience, gender-role identity, and exposure to sexualised content on social media [6]. Identifying the factors associated with pornography use is therefore relevant not only for describing its prevalence but also for understanding which adolescent profiles show a greater propensity to engage in this behaviour, some of whom may also be at greater risk of developing more frequent or problematic patterns of use [1].

Pornography use becomes particularly concerning when it is difficult to regulate or begins to interfere with emotional, relational, educational, or sexual functioning [7]. Evidence also suggests that high-frequency or problematic use may be associated with cognitive-affective distress, psychosomatic symptoms, maladaptive emotion regulation, sexual dissatisfaction, sexually aggressive attitudes, and beliefs that reinforce sex inequality [4,8,9]. An additional factor complicating the study of pornography use is that this behaviour may occur in a context characterised by privacy, anonymity, and limited adult oversight, often accompanied by moral disapproval and concealment of the behaviour [10].

Existing studies point to a wide range of factors potentially associated with pornography use, including sociodemographic characteristics, family functioning, parental supervision, Internet use, leisure patterns, substance use, risk behaviours, and psychological distress [11]. Pereda and Águila-Otero [12] report a high prevalence of intentional access to pornography among adolescents, with significant differences according to sex and age. Similarly, systematic reviews indicate that adolescents' use of sexually explicit online material is consistently associated with being male, older age, pubertal development, sexual attitudes, peer influence, relational independence, and problematic Internet use [11,13].

As shown in Figure 1, this study considers several groups of variables that previous research has identified as relevant to characterising pornography use among adolescents: sociodemographic factors, family relationships and conflict, parental control inside and outside the home, substance use, risky sexual behaviours, indicators of psychological distress, and addictive Internet use [11,13].

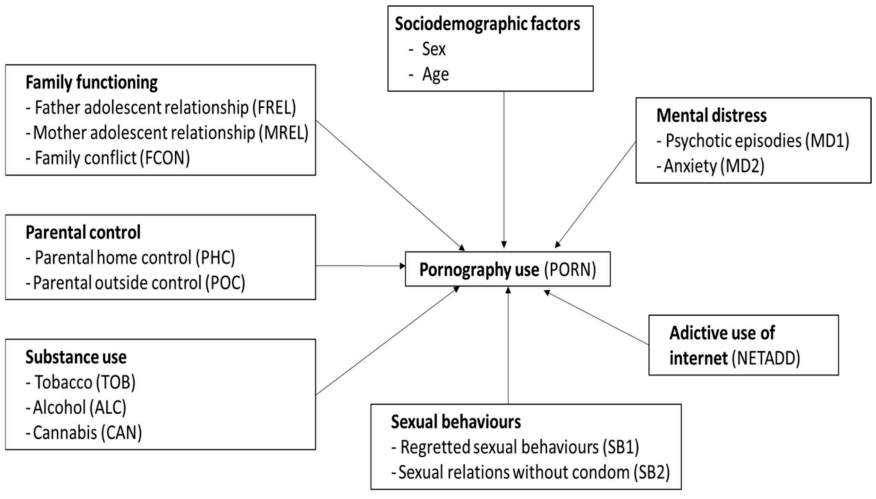


Figure 1. Conceptual groundwork used in this paper.

Large population-based datasets make it possible to analyse these factors robustly and compare their relative importance. Accordingly, this study examines pornography use among adolescents using data from the 2023 Spanish Survey on Drug Use in Sec-

ondary Education (ESTUDES 2023), conducted by the Spanish Observatory on Drugs and Addictions [14]. Its objective is to identify the main personal, family-related, digital, and behavioural factors associated with intentional pornography use among Spanish adolescents. ESTUDES is part of a biennial survey programme conducted by the OEDA to monitor addictive behaviours among Spanish adolescents and constitutes a key reference source for their study in Spain [1].

Research on potentially problematic behaviours during adolescence has traditionally relied on conventional statistical methods, particularly linear and logistic regression. This pattern is also evident in research on adolescent pornography use, both in studies applying linear regression [4,15,16] and in those using logistic regression [1,17–19]. The present study combines logistic regression with tree-based ensemble learning methods, which generally have a greater capacity to capture nonlinear relationships and complex interactions among variables. Specifically, three gradient boosting methods are employed: Extreme Gradient Boosting (XGBoost) [20], Light Gradient Boosting Machine (LightGBM) [21] and Categorical Boosting (CatBoost) [22].

Gradient boosting models are particularly useful for predicting and explaining human behaviour [23] and have recently demonstrated their value in analysing potentially problematic behaviours during adolescence. Applications include the prediction of addictive Internet use using LightGBM and CatBoost [24], smartphone addiction using XGBoost and LightGBM [25,26], and substance use or susceptibility to e-cigarette use using XGBoost [27,28]. However, to the best of our knowledge, these techniques have not yet been applied to the analysis of pornography use among adolescents. Nor has the potential predictive advantage of machine-learning methods been sufficiently assessed using major Spanish datasets on adolescent behaviour, such as ESTUDES 2023 [14], whose findings provide an important source of evidence for the design and evaluation of public policies in Spain. Consequently, the first research question is formulated as follows:

RQ1: How does the predictive performance of gradient boosting models compare with that of logistic regression when applied to ESTUDES 2023 data?

In behavioural research, analytical techniques should provide not only adequate predictive performance but also interpretability. Beyond achieving strong predictive performance, an analysis should identify which variables are relevant and which are secondary or even negligible in predicting the behaviour under examination. For this purpose, the present study uses Shapley Additive Explanations (SHAP) [29], which have already demonstrated their usefulness in improving the understanding of potentially problematic behaviours during adolescence [24,25,28]. Accordingly, a second research question is proposed:

RQ2. Which predictors make the greatest contribution to predicting pornography use among adolescents, and how do their predictive contributions interact within the model?

Unlike logistic regression odds ratios, which summarise the overall association between each predictor and the outcome, SHAP enables predictions to be interpreted both globally, as required to address RQ2, and at more specific levels, including individual observations. This property makes it possible to examine not only the overall importance of a correlate but also whether its predictive importance varies according to factors traditionally considered potential moderators of adolescent behaviour, such as sex [30,31] and age [32,33]. This issue is particularly relevant because, although the literature identifies both characteristics as two of the most consistent predictors of pornography use during adolescence [1], their role may not be limited to direct effects but may also involve interactions with other predictors. The relative predictive importance of other correlates may differ across sex and age groups [34]. Consequently, a third research question is formulated:

RQ3. To what extent does the predictive importance of the proposed correlates vary according to adolescents' sex and age?

Although the present study does not propose a new machine-learning algorithm, it makes both a methodological contribution to computational social science and a substantive contribution to research on adolescent risk behaviours. Studies in this field have traditionally relied mainly on conventional regression models, whereas the use of explainable machine-learning approaches remains comparatively limited. From a computational social science perspective, the methodological contribution lies in implementing a rigorous and reproducible analytical framework that integrates logistic regression, multiple ensemble tree-based models, formal statistical comparisons of predictive performance, calibration assessment, robustness analyses, missing-data sensitivity analysis, and global, interaction-based, and local SHAP explanations. This study applies this framework to predict adolescent pornography use using ESTUDES 2023, a large and widely used Spanish survey that has been scarcely exploited through machine-learning techniques. The comparison provides evidence on both the potential and the limits of these methods: XGBoost, LightGBM, and CatBoost produced statistically significant improvements in discrimination and probabilistic accuracy, but the absolute gains over logistic regression were modest. Thus, the findings show that ensemble methods can provide additional predictive value while also indicating that logistic regression remains a competitive and interpretable approach for this particular problem. More broadly, the proposed analytical framework illustrates how explainable machine learning can be used to analyse large-scale social survey data while preserving predictive rigour and substantive interpretability, and it may be extended to other problematic adolescent behaviours, including substance use, risky sexual behaviour, and problematic digital-media use.

To address these questions, the remainder of the article is structured as follows. First, the conceptual framework underpinning the selection of variables is presented. The sample, variables, and analytical methodology are then described, followed by the empirical results. Finally, the article concludes with a discussion of the findings, their implications, the study's limitations, and directions for future research.

2. Materials and Methods

2.1. Conceptual Groundwork

As shown in Figure 1, adolescents' self-reported pornography use (PORN) is associated with seven groups of variables. The first block of correlates comprises adolescents' sex and age. The literature identifies both factors as particularly consistent predictors of pornography use during adolescence. Systematic reviews show that the use of sexually explicit online material is consistently associated with being male and with older age or more advanced pubertal development [11,13]. In the Spanish context, Pereda and Águila Otero [12] also report significant sex- and age-related differences in intentional access to pornography. Moreover, longitudinal studies indicate that boys are more likely to initiate pornography use and that developmental factors, such as age and previous sexual experience, are involved in this initiation process [6]. Sex and age are therefore considered not merely basic sociodemographic variables, but central dimensions for understanding the heterogeneity of adolescent pornography use [1].

The second block concerns family functioning. The family constitutes one of the primary contexts of socialisation during adolescence and may be associated with lower or higher levels of pornography use. Greater family cohesion, closer parent–adolescent relationships, and open communication are associated with lower exposure to pornography, whereas family conflict, poorly structured family environments, and low perceived support are related to higher levels of use [35]. Similarly, it has been reported that a positive

intrafamilial environment is associated with lower pornography use, whereas negative family environments and relational independence are associated with higher use [16].

The third block includes parental control. These variables are incorporated because adolescent pornography use commonly occurs in private digital environments, where family regulation may be more difficult, although by no means irrelevant. From the perspective of the Differential Susceptibility to Media Effects Model, parental monitoring has been shown to be negatively associated with problematic outcomes related to the use of sexually explicit material, while higher levels of parental supervision are associated with less frequent use and lower acceptance of sexual permissiveness [36]. Various reports have also highlighted the importance of parental supervision of Internet exposure and of family attitudes towards pornography and sexuality [37].

The fourth block comprises substance use. The inclusion of these variables is justified because pornography use is frequently associated with adolescent lifestyles characterised by greater experimentation, sensation seeking, and exposure to risk [13]. Adolescents who intentionally access pornography have been reported to engage more frequently in risk behaviours, including alcohol use [12]. Farré et al. [34] find that pornography use is associated with higher levels of substance use.

The fifth block comprises variables related to sexual experience and risk-taking in sexual practices. The literature suggests that pornography use may be connected to adolescent sexual socialisation, both as a source of sexual information and through its association with sexual attitudes and behaviours. Exposure to sexually explicit websites is associated with greater sexual permissiveness, a higher number of sexual partners, and certain sexual practices [3]. Drawing on the sexual script acquisition, activation, and application model, research has shown that pornography use and dependence are associated with inaccurate sexual beliefs among adolescents [5]. Variables concerning specific sexual practices, such as engaging in sexual relations that the adolescent subsequently regrets or failing to use a condom, may therefore be relevant for capturing the relationship between pornography use and sexual behaviour.

The sixth block incorporates indicators of psychological distress. Their inclusion is justified because part of the literature links high-frequency or problematic pornography use to emotional difficulties, psychosomatic symptoms, and maladaptive emotion-regulation strategies. Longitudinal studies have reported that more frequent pornography use is associated with psychosomatic symptoms, although its relationship with depressive symptoms appears to be more complex [4]. Models of problematic use describe pornography use as a behaviour that, among some adolescents, may be characterised by impaired control, salience, use for mood modification, and persistence despite negative consequences [38]. Other studies indicate that relief from psychological stress and coping with negative emotions may serve as motives for pornography use [8]. The inclusion of indicators of anxiety, emotional distress, and psychological difficulties therefore makes it possible to examine whether pornography use is associated with profiles of greater emotional vulnerability among adolescents, without assuming that all pornography use is necessarily problematic.

The seventh block concerns addictive Internet use. Its inclusion is motivated by the fact that pornography is predominantly accessed in digital environments and may form part of broader patterns of intensive, dysregulated, or difficult-to-control Internet use [13]. The systematic review conducted by Chaise et al. [11] identifies Internet addiction as one of the factors associated with adolescents' use of sexually explicit material. A preference for viewing pornography online has also been found to be associated with addictive Internet use [39,40]. The relationship between these behaviours may reflect an overlap in difficulties with behavioural self-control, as dysfunctional Internet use is associated with greater

exposure to online pornography, while impulsivity and sensation seeking have been linked to both pornography use and problematic pornography use [41].

2.2. Sampling and Demographic Characteristics of the Sample

The study uses secondary data from the 2023 Spanish Survey on Drug Use in Secondary Education (ESTUDES 2023) [14], a cross-sectional, descriptive, nationwide survey. The target population comprised students aged 14–18 enrolled in compulsory secondary education, upper secondary education, intermediate-level vocational training, or basic vocational education. Exceptionally, 19-year-old students who remained enrolled in these programmes were also included.

A two-stage, stratified cluster sampling design was employed. Stratification was based on the autonomous community or autonomous city, school ownership, and type of educational programme. In the first stage, schools were randomly selected; in the second stage, classrooms within those schools were selected. The questionnaire was administered to all students present in the selected classrooms on the day of data collection. The national sample comprised 42,208 students from approximately 2000 classrooms and more than 800 schools. Using Cochran's formula and assuming $p = q = 0.5$, the sampling error for Spain as a whole was estimated at $\pm 0.5\%$ at the 95% confidence level.

Data were collected through a self-administered questionnaire completed collectively and in person in the classroom. Depending on the autonomous community, the questionnaire was available in Spanish, Catalan, Valencian, Basque, and Galician. Participation was fully anonymous, and no personally identifying information was collected. This procedure yielded a large and representative sample of students enrolled in secondary education in Spain, while the anonymous survey format was intended to facilitate candid responses to sensitive questions concerning pornography and substance use. Table 1 presents the sociodemographic composition of the sample.

Table 1. Sample characteristics.

Variable	Category	n	%
Sex	Male	21,247	50.34
	Female	20,961	49.66
Year of birth	2004	1450	3.44
	2005	9454	22.4
	2006	10,665	25.27
	2007	10,928	25.89
	2008	9684	22.94
	2009	27	0.06
Country of birth	Spain	38,570	91.38
	Foreign country	3463	8.2
	Missing	175	0.41
Language	Spanish	34,317	81.3
	Catalan	3817	9.04
	Valencian	1402	3.32
	Basque	1303	3.09
	Galician	1369	3.24
Municipality size	Fewer than 10,000 inhabitants	5757	13.64
	10,000–100,000 inhabitants	19,953	47.27
	100,001–400,000 inhabitants	10,728	25.42
	More than 400,000 inhabitants	5770	13.67
Education level	Compulsory secondary education	23,157	54.86
	Upper secondary education	15,123	35.83
	Intermediate vocational training	1177	2.79
	Advanced vocational training	2751	6.52

2.3. Items Used, Data Preprocessing, and Variable Measurement

Table 2 presents the items used to measure each group of variables, the wording of the corresponding questions, their coding, and the observed absolute and relative frequencies. Regarding the outcome variable, 16,731 adolescents, representing 39.60% of the original sample, reported having viewed or used pornography during the previous 30 days, whereas 22,025 (52.20%) reported that they had not, and 3452 (8.18%) had missing information. The sample was almost evenly distributed by sex, with 50.34% male and 49.66% female participants. Most participants were aged between 15 and 18 years, an age range that accounted for 96.50% of the original sample.

The first step in the analysis was to determine the final analytical sample. We first excluded 3452 participants who had not provided information on pornography use, leaving 38,756 adolescents with observed values for the outcome variable. Of these, 33,543 (86.55%) had complete information for all predictors and were included in the main analytical sample. Complete information was required for all variables except the items used to construct the Internet-use index. For this composite measure, participants were required to have answered at least 7 of the 14 items. Overall, 8665 of the 42,208 adolescents in the original sample were excluded: 3452 because pornography-use information was missing and a further 5213 because at least one predictor was unavailable.

The final analytical sample therefore comprised 33,543 complete observations. Despite this reduction, the nominal margin of error remained small under the assumption of simple random sampling. Applying Cochran's formula at the 95% confidence level and assuming maximum variability ($p = q = 0.5$), the estimated margin of error was approximately $\pm 0.54\%$, under the assumption of simple random sampling.

As shown in Table A1 in Appendix A, differences between included and excluded participants were generally limited. All absolute standardised mean differences were below 0.25. The largest difference concerned cannabis use, which was higher among excluded participants (mean = 1.522) than among included participants (mean = 0.994; SMD = −0.230). Excluded participants also reported slightly higher tobacco and alcohol use, regretted sexual intercourse, intercourse without a condom, and self-reported psychotic episodes. They were also somewhat more likely to be male and reported slightly less favourable relationships with their parents. Pornography use was reported by 46.88% of excluded participants and 42.59% of included participants, although the corresponding standardised difference was negligible (SMD = −0.087). Overall, these results suggest that the complete-case analysis may have slightly underrepresented adolescents with higher-risk behavioural profiles, although most differences were small.

The variables measuring Internet use require specific consideration. A composite NETADD score was calculated as the mean of 14 items (NETADD1–NETADD14), using all available responses provided that at least 7 items had been answered. Valid NETADD scores were obtained for 40,744 participants (96.53%), whereas 1464 observations (3.47%) had missing values. The resulting index had a mean of 1.29 and a standard deviation of 0.81. The composite score, which could range from 0 to 4, was subsequently normalised to the [0, 1] interval. The internal consistency of the 14-item scale was assessed using Cronbach's alpha, which yielded a value of 0.965, indicating excellent reliability. NETADD should be interpreted as a self-reported index of problematic Internet use rather than as a clinical diagnosis of Internet addiction.

Table 2. Variables and items used in the analysis: definitions, coding, and descriptive statistics.

Domain	Variable and Description	Survey Item	Coding and Descriptive Statistics, n (%)
Outcome	Pornography use (PORN)	Have you used or viewed pornography during the last 30 days?	0 = No: 22,025 (52.20%); 1 = Yes: 16,731 (39.60%); Missing: 3452 (8.18%)
Sociodemographic factors	Adolescent's Sex (SEX)	Sex of the adolescent	0 = Male: 21,247 (50.34%); 1 = Female: 20,961 (49.66%)
	Adolescent's Age (AGE)	Adolescent's age, calculated from the reported year of birth: AGE = 2023 − year of birth	14 years: 27 (0.06%); 15 years: 9684 (22.94%); 16 years: 10,928 (25.89%); 17 years: 10,665 (25.27%); 18 years: 9454 (22.40%); 19 years: 1450 (3.44%)
Family functioning	Father–Adolescent Relationship (FREL)	How would you describe your relationship with your father?	0 = No father: 1495 (3.54%); 1 = Low or very low closeness: 1709 (4.05%); 2 = Moderate closeness: 6866 (16.27%); 3 = High or very high closeness: 30,600 (72.50%); Missing: 1538 (3.64%).
	Mother–Adolescent Relationship (MREL)	How would you describe your relationship with your mother?	0 = No mother: 448 (1.06%); 1 = Low or very low closeness: 885 (2.10%); 2 = Moderate closeness: 5170 (12.25%); 3 = High or very high closeness: 34,408 (81.52%); Missing: 1297 (3.07%).
	Family Conflict (FCON)	During the last 12 months, have you had a major conflict or argument with your parents or siblings?	0 = No: 24,942 (59.09%); 1 = Yes: 16,234 (38.46%); Missing: 1032 (2.45%)
Parental control	PHC—Parental Home Control	My parents establish clear rules about what I am allowed to do at home	0 = Sometimes, rarely, or never: 25,320 (59.99%); 1 = Always or almost always: 15,586 (36.93%); Missing: 1302 (3.08%).
	Parental Outside Control (POC)	My parents establish clear rules about what I am allowed to do outside the home	0 = Sometimes, rarely, or never: 21,119 (50.04%); 1 = Always or almost always: 19,698 (46.67%); Missing: 1391 (3.30%).
Substance use	Tobacco Use (TOB)	During the last 30 days, how frequently have you smoked cigarettes?	0 = Never: 33,786 (80.05%); 1 = Less than one day per week: 2464 (5.84%); 2 = Some days per week, but not daily: 2389 (5.66%); 3 = Daily: 2880 (6.82%); Missing: 689 (1.63%)
	Alcohol Use (ALC)	Number of weekends on which the adolescent consumed alcoholic beverages during the last 30 days	0 = No weekends: 21,537 (51.03%); 1 = 1 weekend: 7809 (18.50%); 2 = 2 weekends: 6110 (14.48%); 3 = 3 weekends: 3151 (7.47%); 4 = Every weekend: 2974 (7.05%); Missing: 627 (1.49%)
	Cannabis Use (CAN)	On how many days in your lifetime have you used hashish or marijuana?	0 = No days: 31,268 (74.08%); 1 = 1 day: 1581 (3.75%); 2 = 2 days: 834 (1.98%); 3 = 3 days: 738 (1.75%); 4 = 4–5 days: 966 (2.29%); 5 = 6–9 days: 746 (1.77%); 6 = 10–19 days: 953 (2.26%); 7 = 20–39 days: 774 (1.83%); 8 = 40 days or more: 2080 (4.93%); Missing: 2268 (5.37%)
Sexual behaviours	Sexual behaviour 1 (SB1)	I have had sexual intercourse that I regretted the following day	0 = No: 38,009 (90.05%); 1 = Yes: 3049 (7.22%); Missing: 1150 (2.72%)
	Sexual behaviour 2 (SB2)	I have had sexual intercourse without using a condom	0 = No: 34,745 (82.32%); 1 = Yes: 6316 (14.96%); Missing: 1147 (2.72%)

Table 2. *Cont.*

Domain	Variable and Description	Survey Item	Coding and Descriptive Statistics, n (%)
Mental distress	Mental distress 1 (MD1)	I have experienced a psychotic episode, such as hallucinations or delusions	0 = No: 39,555 (93.72%); 1 = Yes: 1425 (3.38%); Missing: 1228 (2.91%)
	Mental distress 2 (MD2)	I have experienced an anxiety attack	0 = No: 27,419 (64.96%); 1 = Yes: 13,664 (32.37%); Missing: 1125 (2.67%)
Addictive use of Internet	Internet Use Index (NETADD)	Composite index calculated as the mean of NETADD1–NETADD14. The mean was calculated when at least seven valid responses were available	Item range: 0–4. Valid observations: 40,744 (96.53%); Missing: 1464 (3.47%); Mean = 1.29; SD = 0.810; Cronbach's alpha = 0.96.
	Difficulty Stopping Internet Use (NETADD1)	How often have you found it difficult to stop using the Internet once you were online?	0 = Never: 7062 (16.73%); 1 = Rarely: 9012 (21.35%); 2 = Sometimes: 13,257 (31.41%); 3 = Often: 7777 (18.43%); 4 = Very frequently: 3645 (8.64%); Missing: 1455 (3.45%)
	Continued Internet Use Despite Wanting to Stop (NETADD2)	How often have you remained online despite wanting to stop?	0 = Never: 11,624 (27.54%); 1 = Rarely: 9406 (22.28%); 2 = Sometimes: 10,025 (23.75%); 3 = Often: 6488 (15.37%); 4 = Very frequently: 3080 (7.30%); Missing: 1585 (3.76%)
	Advice to Reduce Internet Use (NETADD3)	How often have your parents or friends told you that you should spend less time online?	0 = Never: 9095 (21.55%); 1 = Rarely: 10,193 (24.15%); 2 = Sometimes: 10,822 (25.64%); 3 = Often: 6707 (15.89%); 4 = Very frequently: 3430 (8.13%); Missing: 1961 (4.65%)
	Preference for Internet over Social Interaction (NETADD4)	How often do you prefer going online rather than spending time with other people, such as parents or friends?	0 = Never: 14,400 (34.12%); 1 = Rarely: 13,411 (31.78%); 2 = Sometimes: 8107 (19.21%); 3 = Often: 3052 (7.23%); 4 = Very frequently: 1345 (3.19%); Missing: 1893 (4.48%)
	Sleep Loss Due to Internet Use (NETADD5)	How often do you sleep less because you are online?	0 = Never: 13,660 (32.36%); 1 = Rarely: 10,870 (25.75%); 2 = Sometimes: 8717 (20.65%); 3 = Often: 4811 (11.40%); 4 = Very frequently: 2382 (5.64%); Missing: 1768 (4.19%)
	Internet Preoccupation (NETADD6)	How often do you find yourself thinking about the Internet even when you are not online?	0 = Never: 16,375 (38.79%); 1 = Rarely: 12,894 (30.55%); 2 = Sometimes: 7131 (16.90%); 3 = Often: 2872 (6.80%); 4 = Very frequently: 1161 (2.75%); Missing: 1775 (4.21%)
	Desire to Go Online (NETADD7)	How often do you look forward to going online?	0 = Never: 11,784 (27.92%); 1 = Rarely: 12,888 (30.53%); 2 = Sometimes: 10,023 (23.75%); 3 = Often: 4308 (10.21%); 4 = Very frequently: 1445 (3.42%); Missing: 1760 (4.17%)
	Perceived Need to Reduce Internet Use (NETADD8)	How often do you think that you should use the Internet less?	0 = Never: 8484 (20.10%); 1 = Rarely: 10,042 (23.79%); 2 = Sometimes: 12,258 (29.04%); 3 = Often: 6575 (15.58%); 4 = Very frequently: 3039 (7.20%); Missing: 1810 (4.29%)
	Failed Attempts to Reduce Internet Use (NETADD9)	How often have you tried to spend less time online but failed to do so?	0 = Never: 15,201 (36.01%); 1 = Rarely: 10,715 (25.39%); 2 = Sometimes: 8883 (21.05%); 3 = Often: 3842 (9.10%); 4 = Very frequently: 1676 (3.97%); Missing: 1891 (4.48%)
	Rushing Tasks to Go Online (NETADD10)	How often do you rush to finish your work so that you can go online?	0 = Never: 14,485 (34.32%); 1 = Rarely: 11,364 (26.93%); 2 = Sometimes: 8805 (20.86%); 3 = Often: 3924 (9.30%); 4 = Very frequently: 1716 (4.07%); Missing: 1914 (4.53%)

Table 2. *Cont.*

Domain	Variable and Description	Survey Item	Coding and Descriptive Statistics, n (%)
	Neglect of Obligations Due to Internet Use (NETADD11)	How often do you neglect responsibilities, such as schoolwork or spending time with your family, because you prefer going online?	0 = Never: 15,551 (36.84%); 1 = Rarely: 11,676 (27.66%); 2 = Sometimes: 8207 (19.44%); 3 = Often: 3418 (8.10%); 4 = Very frequently: 1573 (3.73%); Missing: 1783 (4.22%)
	Internet Use When Feeling Down (NETADD12)	How often do you go online when you are feeling down?	0 = Never: 9465 (22.42%); 1 = Rarely: 7708 (18.26%); 2 = Sometimes: 10,198 (24.16%); 3 = Often: 8136 (19.28%); 4 = Very frequently: 4909 (11.63%); Missing: 1792 (4.25%)
	Internet Use to Escape Negative Feelings (NETADD13)	How often do you go online to forget your worries or negative feelings?	0 = Never: 11,876 (28.14%); 1 = Rarely: 8240 (19.52%); 2 = Sometimes: 8671 (20.54%); 3 = Often: 6695 (15.86%); 4 = Very frequently: 4609 (10.92%); Missing: 2117 (5.02%)
	Distress When Unable to Use the Internet (NETADD14)	How often do you feel restless, frustrated, or irritated when you cannot use the Internet?	0 = Never: 19,933 (47.23%); 1 = Rarely: 11,126 (26.36%); 2 = Sometimes: 5579 (13.22%); 3 = Often: 2259 (5.35%); 4 = Very frequently: 1394 (3.30%); Missing: 1917 (4.54%)

2.4. Data Analysis

The data analysis was conducted in five stages. First, the 33,543 complete observations comprising the main analytical sample were divided into training and independent test samples using joint stratification by pornography use and sex. This procedure preserved both the sex distribution and the proportion of positive and negative cases. A total of 26,833 observations (80%) were allocated to the training sample, while the remaining 6710 observations (20%) were reserved for the final assessment of predictive performance. The test sample was not used for model estimation, hyperparameter selection, or early stopping.

In the final analytical sample, 19,256 adolescents (57.41%) had not used pornography during the previous 30 days, whereas 14,287 (42.59%) had done so. Given the relatively balanced distribution of the outcome, no oversampling, undersampling, synthetic data generation, or differential class weighting was applied. Pornography use was treated as the positive class throughout the analysis.

In the second stage, four predictive models were developed using the training sample: logistic regression, XGBoost [20], LightGBM [21], and CatBoost [22]. Logistic regression was included as an interpretable statistical benchmark, allowing its coefficients to be reported as odds ratios with 95% confidence intervals. For logistic regression, XGBoost, and LightGBM, non-binary predictors were rescaled to the [0, 1] interval using their predefined ranges, whereas binary predictors remained unchanged. Consequently, the odds ratios of non-binary predictors represent the change associated with moving from the minimum to the maximum of the corresponding scale. CatBoost was estimated using the original coding of the variables, with binary predictors explicitly identified as categorical features.

Hyperparameter selection for the three boosting algorithms was performed exclusively within the training sample using five-fold cross-validation jointly stratified by pornography use and sex. The search ranges were selected to cover commonly used values from relatively conservative to moderately complex model specifications, while keeping the computational burden manageable and applying the same number of candidate configurations to each boosting algorithm. The same folds were used for all candidate configurations to ensure comparability. For XGBoost, 48 configurations were evaluated by combining learning rates of 0.03, 0.05, and 0.10; maximum tree depths of 3, 4, 5, and 6; minimum child weights of

1 and 5; and subsample proportions of 0.70 and 0.90. The proportion of predictors used in each tree was fixed at 0.80, gamma at 0, L2 regularisation at 1, and L1 regularisation at 0.

For LightGBM, 48 configurations were formed by combining learning rates of 0.03, 0.05, and 0.10; maximum numbers of leaves of 7, 15, 31, and 63; minimum numbers of observations per leaf of 20 and 50; and feature fractions of 0.70 and 0.90. The bagging fraction was fixed at 0.80 and the bagging frequency at one.

For CatBoost, 48 configurations were evaluated by combining learning rates of 0.03, 0.05, and 0.10; tree depths of 4, 5, 6, and 7; L2 leaf regularisation values of 3 and 10; and bagging temperatures of 0 and 1. Bayesian bootstrapping, symmetric trees, a random-strength parameter of 1, and a feature-sampling proportion of 1 were used. For all three boosting algorithms, a maximum of 1000 iterations was allowed, and early stopping was applied after 30 rounds without improvement.

The primary model-selection criterion was the mean cross-validated area under the receiver operating characteristic curve (AUC). When configurations produced similar AUC values, lower log loss and Brier scores were used as secondary criteria, followed by lower model complexity. Once the optimal configuration had been identified, each boosting algorithm was refitted using the complete training sample and the selected number of boosting iterations.

In the third stage, the four models were evaluated once on the independent test sample. Class predictions were obtained using a probability threshold of 0.50. Discrimination was evaluated using AUC, while classification performance was assessed using accuracy, sensitivity, specificity, precision, negative predictive value, balanced accuracy, and the F1-score. The Brier score and log loss were used to assess probabilistic accuracy. Calibration was examined through the calibration intercept, calibration slope, expected calibration error based on ten equal-width probability intervals, and calibration plots constructed from deciles of predicted risk.

Pairwise differences in AUC were tested using DeLong's procedure. Differences in classification accuracy at the 0.50 threshold were examined using paired McNemar tests. Holm correction was applied separately to the pairwise DeLong and McNemar comparisons to control for multiple testing. These analyses addressed RQ1.

In the fourth stage, several robustness and sensitivity analyses were conducted. First, five-fold cross-validation was repeated within the original training sample using ten different random seeds, while retaining the selected hyperparameters and number of boosting iterations. The independent test sample remained untouched during this robustness analysis. This analysis evaluated the stability of performance with respect to changes in fold composition and stochastic model training. Out-of-fold AUC, Brier score, log loss, and threshold-dependent classification metrics were calculated for each seed.

The potential influence of missing predictor data was evaluated. Adolescents included in the complete-case analysis were compared with those who had observed pornography-use data but at least one missing predictor using standardised mean differences. Multiple imputation by chained equations was then conducted as a sensitivity analysis among the 38,756 adolescents with observed values for pornography use. The dependent variable was not imputed. Ten imputed datasets were generated using predictive mean matching for ordinal or numerical predictors and logistic regression imputation for binary predictors. The same logistic regression model was estimated in every imputed dataset, and the coefficients were combined using Rubin's rules. The pooled estimates were compared with those obtained from the 33,543 complete cases.

Variance inflation factors were obtained from the complete-case logistic regression model to assess potential multicollinearity among the predictors.

In the fifth stage, the boosting model showing the best overall combination of test-set performance metrics was selected for the interpretability analysis using SHAP [29], with primary consideration given to AUC, the Brier score, and log loss. The selected model was the version fitted exclusively on the training sample and was not refitted using either the test observations or the full analytical sample. SHAP values were subsequently calculated for all 6710 observations in the independent test sample. Each SHAP value represents the contribution of a predictor to the model output for a specific observation. For the binary classification model, these contributions were obtained on the raw model-output scale. Positive SHAP values increase the model output and, consequently, the predicted probability of pornography use, whereas negative values decrease it. Global predictive importance was quantified as the mean absolute SHAP value of each predictor, and the distribution, magnitude, and direction of individual contributions were represented using a beeswarm plot.

SHAP dependence plots were produced for the most influential predictors to examine potential nonlinearities and conditional patterns. Global feature-interaction strengths were estimated using the native SHAP interaction procedure of the selected gradient boosting model. Dependence plots were coloured according to the predictor showing the strongest interaction with the focal variable. Finally, local explanations were generated for representative true-positive, true-negative, false-positive, and false-negative classifications to illustrate how predictor contributions combined in individual predictions. These analyses addressed RQ2.

To evaluate heterogeneity in predictive contributions, mean absolute SHAP values were compared between male and female adolescents, excluding the SHAP contribution of sex itself. Differences across age groups were examined after excluding the contribution of age. Welch's t-tests and Cohen's d were used for the comparisons by sex, whereas analysis of variance and eta squared were used for the comparisons across age groups. Given the large sample size, the substantive interpretation focused primarily on effect sizes and the stability of predictor rankings rather than on statistical significance alone. These analyses addressed RQ3.

All analyses were conducted in R 4.4.2 on a Windows 10 64-bit computer equipped with an 11th-generation Intel Core i5-1135G7 processor, four physical cores, eight logical cores, and 15.6 GB of RAM. Four processing threads were used. The main packages included xgboost 3.1.3.1, lightgbm 4.6.0, catboost 1.2.10, pROC 1.19.0.1, mice 3.19.0, car 3.1.3, dplyr 1.1.4, and ggplot2 4.0.1. Model-fitting times were measured over five repeated fits after an initial warm-up, using the selected hyperparameters and excluding preprocessing and tuning. The time required to calculate CatBoost SHAP values was measured over three repetitions after an initial warm-up. The random seed used for the principal sample partition and cross-validation procedure was 123.

3. Results

3.1. Logistic Regression Results

After dividing the sample into training and test subsamples (Step 1), LR was estimated using the same training sample employed for the gradient boosting models (N = 26,833). The results are reported in Table 3. It provided a statistically significant improvement in fit over the null model (χ^2 = 11,018, $p < 0.001$). Nagelkerke's pseudo-R^2 of 0.376 indicates a moderate level of explanatory performance.

SEX showed the strongest association with pornography use. Because it was coded as 0 = male and 1 = female, its odds ratio (OR) of 0.072 ($p < 0.001$) indicates that the odds of pornography use among female adolescents were 92.8% lower than among male adolescents. NETADD also showed a particularly strong positive association: an increase

from its minimum to its maximum value multiplied the odds of pornography use by 5.307 ($p < 0.001$).

Table 3. Logistic regression results for adolescent pornography use.

Variable	b	SE	z-Ratio	*p*-Value	OR	95% CI for OR
Intercept	0.011	0.100	0.11	0.915	1.011	0.831–1.229
SEX	−2.631	0.032	−82.87	<0.001	0.072	0.068–0.077
AGE	0.345	0.061	5.68	<0.001	1.412	1.253–1.590
FREL	−0.136	0.086	−1.59	0.111	0.873	0.738–1.032
MREL	−0.122	0.060	−2.03	0.042	0.885	0.787–0.996
FCON	0.314	0.031	10.28	<0.001	1.369	1.289–1.454
PHC	−0.032	0.030	−1.07	0.286	0.969	0.914–1.027
POC	−0.048	0.029	−1.65	0.098	0.953	0.901–1.009
TOB	−0.070	0.065	−1.08	0.280	0.932	0.820–1.059
ALC	0.695	0.050	14.00	<0.001	2.004	1.818–2.209
CAN	0.670	0.067	10.00	<0.001	1.955	1.714–2.229
SB1	0.247	0.055	4.47	<0.001	1.280	1.148–1.426
SB2	0.218	0.042	5.18	<0.001	1.244	1.145–1.351
MD1	−0.036	0.078	−0.47	0.639	0.964	0.828–1.123
MD2	0.059	0.034	1.74	0.081	1.061	0.993–1.133
NETADD	1.669	0.072	23.23	<0.001	5.307	4.610–6.109

Note. N = 26,833, b = coefficient, SE = standard error, OR = odds ratio and CI = interval coefficient. The likelihood-ratio test indicated that the model was statistically significant, likelihood ratio test $\chi^2(15) = 11{,}018$, $p < 0.001$. The log-likelihood values were −17,371.97 for the estimated model and −22,880.83 for the null model. Model fit statistics were AIC = 34,775.94 and BIC = 34,910.67. The pseudo-R^2 values were 0.241 (McFadden), 0.280 (Cox–Snell), and 0.376 (Nagelkerke).

Among the other statistically significant variables, alcohol use (OR = 2.004, $p < 0.001$) and cannabis use (OR = 1.955, $p < 0.001$) showed the largest associations. Older age was also associated with higher odds of pornography use (OR = 1.412, $p < 0.001$), as was family conflict (OR = 1.369, $p < 0.001$). Certain sexual behaviours were also relevant, including having had sexual intercourse that the adolescent regretted the following day (OR = 1.280, $p < 0.001$) and having had sexual intercourse without using a condom (OR = 1.244, $p < 0.001$).

The relationship with the mother showed a smaller negative association. A better mother–adolescent relationship was associated with lower odds of pornography use (OR = 0.885, $p = 0.042$). None of the remaining variables reached statistical significance at the 5% level.

3.2. Optimisation of the Gradient Boosting Models and Predictive Comparison with Logistic Regression (Research Question 1)

Table 4 presents the optimal hyperparameters obtained after tuning the three gradient boosting models. The cross-validation results were very similar, with mean AUC values of 0.8222 for XGBoost, 0.8223 for LightGBM, and 0.8225 for CatBoost. Their standard deviations were also small and comparable, ranging from 0.0024 to 0.0025, indicating stable predictive performance across the validation folds. Thus, none of the three models showed a clear advantage during the tuning stage.

Figures A1–A3 in Appendix B show the training and validation learning curves for the selected XGBoost, LightGBM, and CatBoost configurations, respectively. For all three algorithms, AUC increased, and log loss decreased rapidly during the initial boosting iterations, after which validation performance stabilised. Although training performance continued to improve gradually, the validation curves showed only minor additional gains. Early stopping selected 75 iterations for XGBoost, 77 for LightGBM, and 249 for CatBoost. The limited divergence between the training and validation curves indicates

that the selected configurations achieved stable predictive performance with only modest evidence of overfitting.

Table 4. Optimal Hyperparameters Identified through Tuning of the Gradient Boosting Models.

Model	Optimal Hyperparameters	CV AUC, Mean (SD)
XGBoost	eta = 0.10; max_depth = 3; min_child_weight = 1; subsample = 0.70; colsample_bytree = 0.80; gamma = 0; lambda = 1; alpha = 0; nrounds = 75	0.8222 (0.0025)
LightGBM	learning_rate = 0.10; num_leaves = 7; min_data_in_leaf = 20; feature_fraction = 0.90; bagging_fraction = 0.80; bagging_freq = 1; max_depth = −1; lambda_L1 = 0; lambda_L2 = 0; nrounds = 77	0.8223 (0.0025)
CatBoost	learning_rate = 0.05; depth = 4; l2_leaf_reg = 3; bagging_temperature = 0; bootstrap_type = Bayesian; boosting_type = Plain; grow_policy = SymmetricTree; random_strength = 1; rsm = 1; border_count = 128; iterations = 249	0.8225 (0.0024)

Note. CV AUC = cross-validated area under the ROC curve. Values are reported as the mean, with the standard deviation in parentheses.

The confusion matrices reported in Table 5 compare the results obtained from logistic regression and the three gradient boosting models. Logistic regression correctly identified a larger number of adolescents who used pornography, as reflected in its lower number of false negatives and higher number of true positives. By contrast, XGBoost, LightGBM, and CatBoost classified non-users more accurately, yielding fewer false positives and more true negatives.

Table 5. Confusion Matrices for the Methods Evaluated on the Test Sample.

	LR		XGBoost		LightGBM		CatBoost	
Actual | Predicted	**0**	**1**	**0**	**1**	**0**	**1**	**0**	**1**
0	2858	994	2992	860	2997	855	2979	873
1	677	2181	796	2062	798	2060	790	2068

Note. 0 stands for non-users of pornography and 1 for users.

Table 6 provides a broader evaluation of discrimination, classification performance, probabilistic accuracy, and calibration. Overall, the differences among the four models were small, and no method clearly dominated across all criteria. LR achieved the highest sensitivity, negative predictive value, balanced accuracy, and F1-score, indicating a greater ability to identify adolescents who had used pornography. By contrast, the three gradient boosting models achieved higher specificity, precision, and AUC values, reflecting better identification of non-users and slightly greater discriminative ability. LightGBM obtained the highest overall accuracy, specificity, and precision, whereas CatBoost achieved the highest AUC and the lowest Brier score and log loss. Calibration was satisfactory for all models, as shown by intercepts close to zero and slopes close to one. LightGBM produced the lowest expected calibration error, although the values were low for all three boosting models. The calibration curves (Figure A4 of Appendix B) closely followed the line of perfect calibration for all four models, consistent with the near-zero calibration intercepts, slopes close to one, and low ECE values reported in Table 6.

Table 7 shows that the DeLong tests indicated that XGBoost, LightGBM, and CatBoost each achieved a significantly higher AUC than LR. Nevertheless, the absolute improvements were small, ranging from approximately 0.0048 to 0.0055 AUC points. No significant differences in AUC were found among the three gradient boosting models. Likewise, the McNemar tests showed no significant pairwise differences in classification accuracy at the 0.50 probability threshold. These results therefore indicate that the boosting models pro-

vided a small but statistically detectable improvement in discrimination over LR, although none showed a clear advantage in overall classification accuracy.

Table 6. Predictive performance of LR, XGBoost, LightGBM and CatBoost in the test sample.

Model	AUC	Accuracy	Sensitivity	Specificity	Precision	NPV	Balanced Accuracy
LR	0.816	0.751	0.7631	0.742	0.6869	0.8085	0.7525
XGBoost	0.8208	0.7532	0.7215	0.7767	0.7057	0.7899	0.7491
LightGBM	0.8212	0.7537	0.7208	0.778	0.7067	0.7897	0.7494
CatBoost	0.8215	0.7522	0.7236	0.7734	0.7032	0.7904	0.7485

Model	F1	Brier Score	Log Loss	Calibration Intercept	Calibration Slope	ECE
LR	0.723	0.1721	0.5190	−0.0031	0.9848	0.0203
XGBoost	0.7135	0.1699	0.5134	−0.0070	1.0053	0.0101
LightGBM	0.7137	0.1697	0.5128	−0.0048	1.003	0.0092
CatBoost	0.7132	0.1695	0.5123	−0.0050	1.0034	0.0107

Note. Pornography use was treated as the positive class, and classifications were obtained using a probability threshold of 0.50. AUC = area under the receiver operating characteristic curve; NPV = negative predictive value; ECE = expected calibration error calculated using ten equal-width probability bins. Higher values indicate better performance for AUC, accuracy, sensitivity, specificity, precision, NPV, balanced accuracy, and F1. Lower values indicate better performance for the Brier score, log loss, and ECE. A calibration intercept close to 0 and a calibration slope close to 1 indicate better calibration.

Table 7. Pairwise comparisons of AUC and classification accuracy in the independent test sample.

Model Comparison	AUC Difference	95% CI	DeLong (*p*-Value)	Accuracy Difference	McNemar χ^2	McNemar (*p*-Value)
LR − XGBoost	−0.00476	[−0.007, −0.002]	<0.001	−0.00224	0.693	1
LR − LightGBM	−0.00516	[−0.008, −0.003]	<0.001	−0.00268	1.047	1
LR − CatBoost	−0.00552	[−0.008, −0.003]	<0.001	−0.00119	0.19	1
XGBoost − LightGBM	−0.00039	[−0.001, 0.001]	0.597	−0.00045	0.089	1
XGBoost − CatBoost	−0.00076	[−0.002, 0.002]	0.188	0.00104	0.632	1
LightGBM − CatBoost	−0.00036	[−0.001, 0.001]	0.597	0.00149	1.761	1

Note. Differences were calculated as the performance of the first model minus that of the second model. DeLong tests compared paired AUC values, whereas McNemar tests compared classification accuracy using a probability cutoff of 0.50. Reported *p*-values were adjusted for six pairwise comparisons using Holm's procedure. CI = confidence interval.

CatBoost was selected for the subsequent interpretability analysis because it achieved the best nominal combination of AUC, Brier score, and log loss in the independent test sample. However, the formal comparisons did not indicate that CatBoost was statistically superior to XGBoost or LightGBM, whose predictive performance was virtually identical. Therefore, CatBoost was selected as a representative gradient boosting model rather than as a clearly dominant method. Logistic regression remained competitive overall and may be preferable when maximising sensitivity is the primary objective.

The median fitting times over five repeated runs were 0.11 s for logistic regression, 0.12 s for XGBoost, 0.22 s for LightGBM, and 1.28 s for CatBoost. These times excluded data preprocessing and hyperparameter tuning and correspond to the hardware configuration described in Section 2.4.

The robustness analyses reported in Appendix A showed the following patterns:

- The results were stable across alternative cross-validation partitions and treatments of missing data. Across ten repetitions of five-fold cross-validation, the mean AUC values were 0.8170 for logistic regression, 0.8218 for XGBoost, 0.8217 for LightGBM, and 0.8221 for CatBoost, with very small standard deviations. CatBoost achieved the highest AUC in nine of the ten repetitions, although the differences among the three boosting models remained minimal (Table A2).

- The multiple-imputation sensitivity analysis also produced estimates very similar to those obtained from the complete-case sample. Changes in the odds ratios were below 5% for all predictors, and the substantive conclusions remained unchanged, except that the weak association with the mother–adolescent relationship was no longer statistically significant after imputation (Table A3).
- All variance inflation factors were below 2.1, providing no evidence of problematic multicollinearity (Table A3).

3.3. Shapley Additive Explanations Analysis and Variation in the Predictive Importance of Pornography-Use Correlates by Sex and Age

3.3.1. Global Analysis of SHAP Values (Research Question 2)

Figure 2 presents the beeswarm plot of the SHAP values obtained from the CatBoost model. This plot shows the direction in which each variable contributes to the prediction of pornography use. Figure 3 reports the mean absolute SHAP values, thereby providing a clearer representation of the relative predictive importance of the variables. The calculation of CatBoost SHAP values for all 6710 observations in the independent test sample required a median of 0.38 s across three repetitions.

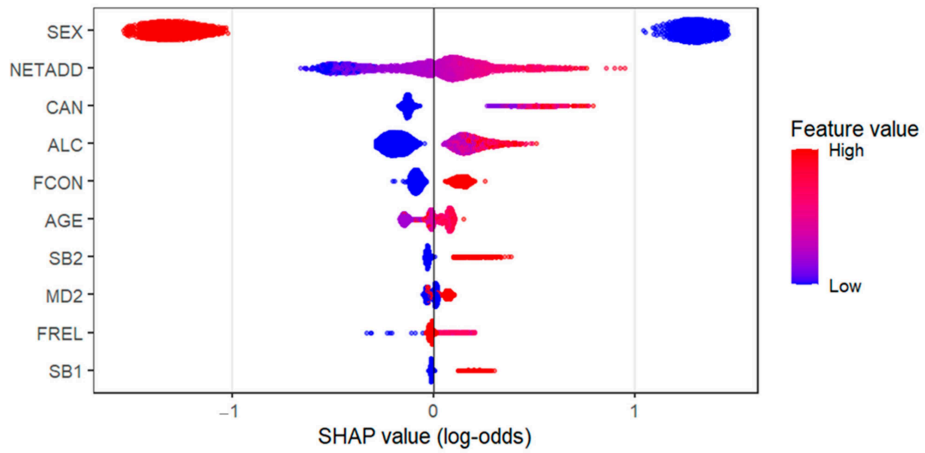


Figure 2. Beeswarm of SHAP values estimated with CatBoost.

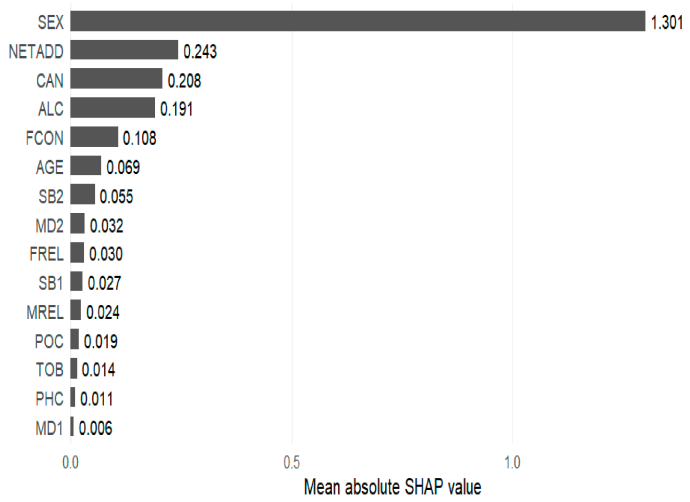


Figure 3. Global predictor importance based on mean absolute SHAP values for the CatBoost model of adolescent pornography use.

SEX was by far the most influential variable in the model's predictions. Low values, corresponding to male adolescents, were concentrated on the positive side of the plot,

whereas high values, corresponding to female adolescents, were clearly located on the negative side. NETADD was the second most important variable. High levels of addictive Internet use were mainly associated with positive SHAP contributions, whereas low levels were concentrated on the negative side. A similar pattern was observed for CAN and ALC: higher levels of cannabis and alcohol use shifted predictions towards pornography use, whereas lower levels contributed towards predicted non-use. FCON also showed a clear pattern, with the presence of family conflict associated with positive contributions to the prediction of pornography use and its absence being linked with negative SHAP values.

Having had sexual intercourse without using a condom (SB2) and having had sexual intercourse that the adolescent subsequently regretted (SB1) are also linked to the likelihood of pornography use. Older age tended to shift the prediction moderately towards use. However, the contributions of these variables were smaller than those of NETADD, CAN, ALC, and FCON.

High values of MD2, corresponding to having experienced an anxiety attack, generally showed positive contributions. By contrast, more favourable relationships with the father and mother (FREL and MREL, respectively) tended to contribute negatively to the prediction of pornography use. Higher levels of tobacco use also showed a limited positive contribution. Nevertheless, the predictive contributions of all these variables were substantially smaller.

Finally, MD1 and the variables related to parental control produced SHAP values close to zero for most observations, indicating that they made only a limited contribution to the model's predictions.

Figure 4 shows that NETADD was not only one of the most influential individual predictors but also the main interaction hub in the CatBoost model. It appeared in six of the ten strongest interactions, particularly in combination with age, cannabis use, alcohol use, tobacco use, and the quality of the father–adolescent relationship. This suggests that the predictive contribution of problematic Internet use varied according to adolescents' developmental stage, substance-use profile, and family context. By contrast, SEX did not appear among the ten strongest interactions, despite being by far the most important predictor in the global SHAP analysis. This pattern suggests that sex operated mainly as a strong direct predictive component, whereas NETADD contributed more through its joint relationships with other characteristics. These interaction patterns are predictive rather than causal and should therefore be interpreted as indications of modelled heterogeneity.

Figure 5 displays the SHAP dependence plots corresponding to the four strongest interactions identified by CatBoost. The upper-left panel shows a clear nonlinear and predominantly increasing relationship between NETADD and the model output. Very low NETADD values contributed negatively to the prediction of pornography use, whereas SHAP values became positive at approximately 1 on the original 0–4 NETADD scale and continued to increase thereafter. Colouring observations by AGE suggests some heterogeneity in the magnitude of this contribution, although the overall increasing pattern was present across age groups.

The upper-right panel shows the CAN $\times$ NETADD interaction. The largest change occurred between no lifetime CAN use and any reported use, after which the contribution remained positive and increased more gradually across higher CAN categories. NETADD introduced additional dispersion within each CAN level, indicating that the predictive contribution of CAN was partly conditional on NETADD.

A similar pattern was observed for the ALC $\times$ NETADD interaction. Adolescents reporting no ALC use generally showed negative SHAP values, whereas any ALC use shifted the model output towards pornography use. The largest increase occurred between no consumption and one weekend of ALC use, followed by a more gradual rise at higher

levels. Variation in NETADD again contributed to differences in SHAP values within the same ALC category.

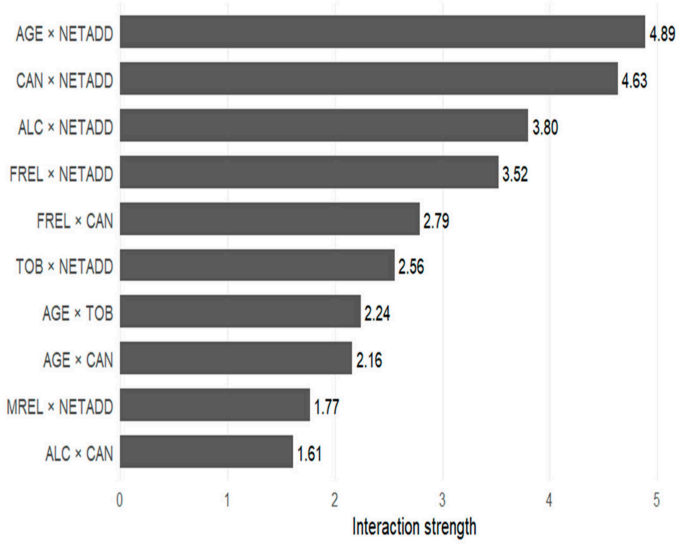


Figure 4. Ten strongest feature interactions identified by the CatBoost model.

Finally, the FREL × NETADD interaction displayed a non-monotonic pattern. The absence of a father was associated with negative SHAP contributions, low-to-moderate FREL scores showed positive contributions, and the highest FREL category produced values close to zero or slightly negative. The colouring by NETADD indicates that the contribution of FREL also varied according to NETADD. Overall, these plots show that NETADD was relevant not only as an individual predictor but also as a contextual variable associated with variation in the predictive contribution of AGE, CAN, ALC, and FREL. These patterns are predictive and should not be interpreted causally.

Local SHAP explanations for representative true positives, true negatives, false positives, and false-negative cases are reported in Figure A5 of Appendix B. In the illustrated misclassified cases, prediction errors occurred when the dominant contribution of SEX opposed the combined contributions of NETADD, CAN, ALC, and family-related variables.

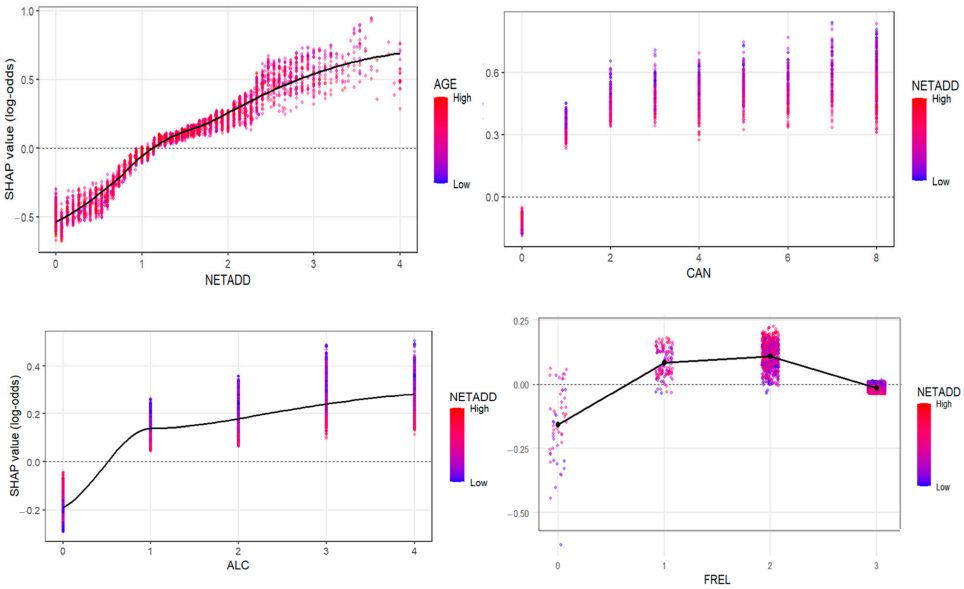


Figure 5. SHAP dependence plots for the four strongest feature interactions in the CatBoost model.

3.3.2. Analysis of Sex- and Age-Related Differences in the Predictive Contributions of the Variables to Pornography Use (Research Question 3)

Table 8 presents the mean absolute SHAP values, the ranking of the explanatory variables, and the differences between male and female adolescents in their contributions to the prediction of pornography use. The variables are presented according to the overall order of importance shown in Figure 2. In this regard, differences involving the most influential predictors, such as NETADD, ALC, CAN, FCON and AGE, are of greater substantive interest than those observed for variables with very limited overall predictive contributions, such as those related to parental control. Cohen's d quantifies the standardised difference between male and female adolescents, but it does not reflect the overall importance of a predictor within the model. Consequently, a large effect size for a variable with very small mean absolute SHAP values indicates a marked relative difference between the two groups, but not necessarily a substantively important contribution to the prediction of pornography use.

Differences were observed in the highest-ranked variable: ALC was the most important predictor among male adolescents, whereas NETADD ranked first among female adolescents. In addition, ALC ranked third among female adolescents and second among male adolescents, while CAN ranked second among female adolescents and third among male adolescents.

Sex differences in mean absolute SHAP values were statistically significant for eleven factors ($p < 0.001$). The exceptions were CAN, AGE and SB1. However, given the large sample size, statistical significance may be reached even for differences of little practical relevance. Interpretation should therefore focus primarily on the magnitude of the effect sizes [42]. Among the five most important variables overall, the differences in NETADD, CAN and AGE were negligible, even though the difference in NETADD was statistically significant. By contrast, ALC and FCON showed large effect sizes. These findings indicate that, although the general structure of predictor importance was relatively similar across sexes, substantive differences emerged in the predictive weight of several key variables.

Table 8. SHAP Values and Ranking of the Explanatory Variables of Pornography Use by Sex.

Variable	Mean |SHAP| (Females)	Mean |SHAP| (Males)	Difference	t-Ratio	*p*-Value	|Cohen's d|	Interpretation
NETADD	0.257 (1)	0.229 (2)	−0.028	−6.37	<0.001	−0.1554	Negligible
CAN	0.209 (2)	0.207 (3)	−0.002	−0.41	0.685	−0.0099	Negligible
ALC	0.148 (3)	0.234 (1)	0.086	64.59	<0.001	1.5803	Large
FCON	0.124 (4)	0.091 (4)	−0.033	−46.90	<0.001	−1.1441	Large
AGE	0.069 (5)	0.069 (5)	0	−0.08	0.939	−0.0019	Negligible
SB2	0.063 (6)	0.047 (6)	−0.017	−11.15	<0.001	−0.2718	Small
MD2	0.051 (7)	0.012 (11)	−0.040	−105.57	<0.001	−2.5654	Large
FREL	0.043 (8)	0.017 (9)	−0.025	−27.05	<0.001	−0.6587	Moderate
MREL	0.040 (9)	0.008 (13)	−0.033	−55.63	<0.001	−1.3516	Large
SB1	0.027 (10)	0.027 (7)	−0.001	−0.51	0.612	−0.0124	Negligible
POC	0.017 (11)	0.020 (8)	0.003	8.89	<0.001	0.2173	Small
TOB	0.016 (12)	0.013 (10)	−0.003	−6.78	<0.001	−0.1656	Negligible
PHC	0.011 (13)	0.010 (12)	−0.001	−4.90	<0.001	−0.1195	Negligible
MD1	0.008 (14)	0.005 (14)	−0.002	−5.52	<0.001	−0.1345	Negligible

Note. The correlates are presented in descending order of their overall relative importance in Figure 3. The position occupied by each variable in the sex-specific importance ranking is shown in parentheses. The Spearman rank correlation between the variable rankings for male and female adolescents was 0.82. The magnitude of the differences between the two groups was interpreted using the absolute value of Cohen's d (|Cohen's d|), following the criteria: $|d| < 0.20$ indicates a negligible effect; $0.20 \leq |d| < 0.50$, a small effect; $0.50 \leq |d| < 0.80$, a moderate effect; and $|d| \geq 0.80$, a large effect [43].

Table 9 presents the mean absolute SHAP values for each variable across age groups, together with the results of the ANOVA used to assess differences between these groups. It shows that only SEX, CAN, FCON, SB2, MD2, POC and TOB exhibited differences of some substantive relevance, although their effect sizes remained small. The predictive contribution of SEX was greater in the intermediate age groups. The importance of CAN and SB2 increased progressively with age, whereas that of FCON and MD2 decreased. Taken together, these findings indicate that age introduces some variation in the predictive strength of specific correlates but does not substantially alter their overall ranking.

Table 9. Mean |SHAP| Values by Age Group and ANOVA Test of Between-Group Differences.

	Mean |SHAP|								
Variable	**≤15 Years**	**16 Years**	**17 Years**	**18 Years**	**19 Years**	**F**	***p*-Value**	**η^2**	**Interpretation**
SEX	1.156 (1)	1.187 (1)	1.188 (1)	1.184 (1)	1.143 (1)	183.95	<0.001	0.027	Small
NETADD	0.260 (2)	0.251 (2)	0.240 (2)	0.235 (2)	0.247 (3)	22.92	<0.001	0.003	Negligible
CAN	0.155 (4)	0.174 (4)	0.198 (3)	0.220 (3)	0.258 (2)	286.73	<0.001	0.041	Small
ALC	0.182 (3)	0.176 (3)	0.180 (4)	0.175 (4)	0.188 (4)	14.19	<0.001	0.002	Negligible
FCON	0.125 (5)	0.112 (5)	0.107 (5)	0.106 (5)	0.098 (5)	267.55	<0.001	0.038	Small
SB2	0.039 (6)	0.047 (6)	0.049 (6)	0.056 (6)	0.070 (6)	104.48	<0.001	0.015	Small
MD2	0.036 (7)	0.035 (7)	0.029 (8)	0.026 (8)	0.024 (11)	169.52	<0.001	0.025	Small
FREL	0.023 (10)	0.026 (8)	0.030 (7)	0.031 (7)	0.036 (7)	34.56	<0.001	0.005	Negligible
SB1	0.023 (9)	0.022 (11)	0.025 (9)	0.025 (10)	0.025 (9)	37.79	<0.001	0.006	Negligible
MREL	0.027 (8)	0.025 (9)	0.025 (10)	0.024 (11)	0.025 (10)	13.90	<0.001	0.002	Negligible
POC	0.019 (12)	0.022 (10)	0.025 (11)	0.026 (9)	0.030 (8)	91.26	<0.001	0.013	Small
TOB	0.021 (11)	0.019 (12)	0.017 (12)	0.015 (12)	0.014 (12)	194.73	<0.001	0.028	Small
PHC	0.003 (13)	0.003 (13)	0.003 (13)	0.004 (13)	0.004 (13)	4.39	<0.001	0.001	Negligible
MD1	0.003 (14)	0.003 (14)	0.003 (14)	0.003 (14)	0.003 (14)	23.24	<0.001	0.003	Negligible

Note. The correlates are presented in descending order of their overall relative importance in Figure 3. η^2 denotes the effect size of the mean differences between age groups. Following Cohen [43], $\eta^2 < 0.01$ indicates a negligible effect; $0.01 \leq \eta^2 < 0.06$, a small effect; $0.06 \leq \eta^2 < 0.14$, a moderate effect; and $\eta^2 \geq 0.14$, a large effect.

4. Discussion

4.1. General Questions

This study examined the ability of logistic regression and three ensemble tree methods, XGBoost, LightGBM, and CatBoost, to predict pornography use among Spanish adolescents. The analysis used data from ESTUDES 2023 [14], one of Spain's main nationwide surveys on substance use and other potentially problematic behaviours among secondary school students.

Regarding RQ1, the results indicated a slight nominal advantage for CatBoost, although this advantage cannot be considered conclusive. CatBoost achieved the highest AUC and the lowest Brier score and log loss, whereas LightGBM obtained the highest accuracy, specificity, and precision. Logistic regression achieved the highest sensitivity, negative predictive value, balanced accuracy, and F1-score. Thus, the identification of the best-performing model depended partly on the metric prioritised. The boosting models were somewhat more effective at identifying adolescents who did not use pornography, whereas logistic regression identified a higher proportion of users.

The paired DeLong tests showed that all three boosting models achieved significantly higher AUC values than logistic regression. However, the absolute differences were small, ranging from approximately 0.005 to 0.006. No significant AUC differences were found among XGBoost, LightGBM, and CatBoost, and the McNemar tests detected no significant differences in classification accuracy between any pair of models. The repeated cross-validation analysis also indicated highly stable performance, with CatBoost obtaining the highest mean AUC in nine of the ten repetitions. Overall, these findings suggest a small but consistent improvement in discrimination from boosting methods, rather than a decisive superiority over logistic regression.

The competitiveness of logistic regression relative to more complex models is not surprising. Some systematic reviews have found no general advantage of tree-based ensemble methods over logistic regression in clinical prediction models [44]. More recently, it has been argued that no approach is universally superior, as predictive performance depends on the characteristics of the dataset, including sample size, number of predictors, linearity of the relationships, class imbalance, and, particularly, data quality [45].

Because CatBoost provided the best nominal combination of AUC, Brier score, and log loss, together with stable repeated cross-validation performance, it was selected as the reference model for addressing RQ2 and RQ3 through SHAP values. This choice does not imply that CatBoost was unequivocally superior to XGBoost, LightGBM, or logistic regression. Rather, it provided a representative and well-calibrated boosting model for examining the importance, direction, interactions, and heterogeneity of the predictive correlates. Logistic regression nevertheless remained highly competitive and may be preferable when sensitivity and the identification of pornography users are prioritised.

Regarding RQ2, sex clearly emerged as the variable with the greatest predictive importance. The model assigned substantially higher predicted probabilities of pornography use to boys than to girls. By contrast, although age showed a statistically significant association in the logistic regression, its SHAP importance was considerably lower and far below that of the main predictors. The findings therefore strongly support the relevance of sex but provide much more limited support for age as a central factor in pornography use, thereby qualifying the conclusions of some previous studies and reports [11–13].

After sex, addictive Internet use (NETADD) was the most important predictive correlate. This result is in agreement with the systematic review by Chaise et al. [11] and with studies associating pornography use with intensive or problematic patterns of Internet use [39–41]. Internet use provides a context characterised by accessibility, availability, and privacy and may form part of a broader behavioural pattern characterised by self-regulation difficulties, persistent use, and high exposure to digital content [39].

The next group of relevant variables comprised cannabis and alcohol use. Tobacco use, by contrast, made only a very limited predictive contribution. The association of alcohol and cannabis with the outcome is consistent with profiles characterised by greater propensity for experimentation, sensation seeking, and exposure to risk behaviours [12,13,34]. Nevertheless, the limited importance of tobacco indicates that not all forms of substance use are related to the outcome in the same way.

Variables related to family functioning, sexual behaviour, and psychological distress ranked below those discussed in the preceding paragraphs. Moreover, indicators belonging to the same domain did not exhibit uniform predictive importance. Within the family-functioning domain, the presence of family conflict was the most relevant predictor, whereas the variables measuring relationships with the father and mother made much smaller contributions. This finding is therefore partly consistent with previous research linking pornography use to problematic family functioning [16,37], although it suggests that overt family conflict may be more informative than general assessments of parent–adolescent relationships or parental control.

Regarding psychological distress, only anxiety attacks showed some predictive relevance, whereas the indicator concerning psychotic episodes had an almost negligible contribution. This result is partially in accordance with Kostopoulou [8], who suggests that relief from psychological stress and coping with negative emotions may serve as motives for pornography use. However, the limited relative importance of psychological distress within the model qualifies this interpretation.

Both sexual behaviours examined showed statistically significant associations in the logistic regression. Nevertheless, the SHAP values assigned considerably greater impor-

tance to sexual intercourse without using a condom. The findings are broadly consistent with previous research linking pornography use to certain risky sexual behaviours [3,5,8]. In any case, both sexual and psychological distress indicators showed substantially less predictive importance than family conflict.

Finally, the parental-control variables made only residual contributions. Although some previous studies have suggested that parental supervision may be linked with lower exposure to pornography [36,37], the present findings do not support considering parental control one of the main predictive correlates once the remaining individual, behavioural, and family-related characteristics are considered simultaneously.

The interaction analysis further showed that NETADD played a central role in the joint predictive structure of the model. The strongest interaction was observed between AGE and NETADD, followed by the interactions of NETADD with CAN, ALC, and FREL. As shown in Figure 5, the contribution of NETADD increased nonlinearly across its range, while CAN and ALC produced their largest changes when moving from no use to any reported use. The FREL pattern was non-monotonic, with intermediate categories showing greater positive contributions than the highest category. These results indicate that the predictive contributions of behavioural and family-related variables varied according to NETADD and, to a lesser extent, AGE. These interactions describe predictive patterns and should not be interpreted as causal or moderating effects.

Regarding RQ3, in addition to being the strongest overall predictor, sex defined subgroups in which several other correlates showed markedly different predictive contributions. The most notable result concerned alcohol use. Among boys, alcohol became the most important correlate, ranking slightly above addictive Internet use, which occupied second place in the overall sample. Relevant sex-related differences were also observed in the importance of family conflict. These findings suggest that the predictive importance of these correlates differed between boys and girls.

By contrast, age-related differences were considerably more limited. The effect sizes obtained from comparisons across age groups were predominantly negligible or small. Thus, although age showed a statistical association with pornography use, it did not appear to substantially alter the importance of the remaining predictors.

These results are congruent with studies identifying sex or gender as a potential moderator of factors associated with risk behaviours during adolescence [30,31]. By contrast, they provide less support for studies assigning age a prominent moderating role in risk-taking and in the influence of social or contextual factors [32,33]. In line with Farré et al. [34], the results suggest that sex is associated with meaningful differences in the profile of predictive correlates of pornography use, whereas age was linked with substantially less predictive heterogeneity than sex.

4.2. Theoretical and Practical Implications

This study offers several theoretical implications for research on pornography use during adolescence and for comparisons between traditional statistical methods and machine-learning techniques. First, the findings challenge the assumption that machine-learning models should automatically replace logistic regression. Although the boosting models achieved slightly better predictive performance in terms of AUC, specificity, and the quality of the estimated probabilities, logistic regression showed higher sensitivity and, therefore, a greater ability to identify adolescents who reported using pornography. This finding is compatible with research showing that complex algorithms do not always substantially outperform logistic regression and that method selection should depend on the analytical objective, data quality, and the type of error to be minimised [44–46].

Model comparison should therefore not be reduced to determining which algorithm achieves the highest value on a single metric. Logistic regression remains relevant because of its parsimony, transparency, and ability to estimate associations that can be readily interpreted through coefficients and odds ratios. Machine-learning models, in turn, provide greater flexibility for representing nonlinear relationships, interactions, and complex patterns. The two approaches should be regarded as complementary: logistic regression allows global associations to be analysed and theoretically grounded relationships to be tested, whereas boosting models may improve classification and identify predictive structures that are not adequately represented by additive linear relationships.

Second, the application of SHAP extends the theoretical contribution of the study by showing that the importance of a variable need not be uniform across all adolescents. SHAP explanations estimate how each characteristic contributes to each individual prediction and facilitate the analysis of heterogeneity across subgroups. In the present study, this approach was initially applied to sex and age, but it could also be extended to other relevant criteria, such as school type (public or private), urban or rural setting, geographical area, and specific family or socioeconomic characteristics. Nevertheless, differences in SHAP values should be interpreted as predictive heterogeneity rather than as causal effects or formal evidence of moderation unless specific statistical analyses are conducted to test such differences.

Finally, the findings support a multidimensional conceptualisation of adolescent pornography use. Its predictive relationships with Internet use, alcohol and cannabis consumption, family conflict, and certain sexual behaviours suggest that it should not be studied as a completely isolated behaviour. Rather, it may form part of a broader network of interconnected digital, family-related, sexual, and substance-use practices. This perspective supports integrative theoretical models combining individual, family, digital, and behavioural factors, rather than explanations focused exclusively on a single dimension.

From an applied perspective, the results indicate that model selection should be guided by the intended analytical or screening objective, at least within the dataset analysed in this study. When the priority is to identify as many adolescents who use pornography as possible (for example, in exploratory studies or broad preventive programmes), logistic regression may be particularly useful because of its higher sensitivity. However, this advantage also entails a larger number of false positives, that is, adolescents classified as users when they are actually non-users. In settings where each positive classification triggers interviews, assessments, counselling, or other specialised resources, this type of error may generate unnecessary costs and inefficient use of the available resources.

By contrast, XGBoost and LightGBM showed higher specificity and produced fewer false positives, which may make them more appropriate when avoiding unnecessary interventions is the main priority. However, this choice entails accepting a greater number of false negatives. Consequently, no model is universally superior: the decision should take into account the balance between sensitivity and specificity, the costs of each type of error, and the intended purpose of the tool.

SHAP explanations may also contribute to the design of more tailored interventions. Rather than applying a single preventive programme to the entire adolescent population, educational and health authorities could use aggregate explanatory patterns to identify which factors are most relevant across different adolescent profiles. For example, the observed subgroup patterns may help inform the design of preventive programmes by indicating whether greater emphasis should be placed on digital literacy and Internet-use regulation, substance use, family conflict, sexual education, or emotional well-being. Similar analyses could be conducted by sex, age, school type, geographical area, or socioeconomic context, provided that algorithmic predictions are not converted into stigmatising labels or automated decisions concerning individual adolescents.

The findings also suggest that preventive policies should not address pornography use in isolation or rely exclusively on punitive approaches. The observed relationships with alcohol and cannabis use, risky sexual behaviours, family conflict, and intensive or addictive Internet use support a more comprehensive form of intervention. School- and community-based programmes should combine affective and sexual education, digital literacy, critical reflection on pornographic content, substance-use prevention, the promotion of safer sexual practices, and the strengthening of family communication.

Within this framework, families and educators need resources to address the issue without relying solely on prohibition or surveillance. Regulators and digital platforms, for their part, should strengthen age-assurance systems, protection against unwanted exposure, and the transparency of recommendation mechanisms. These measures should be complemented by counselling and psychological support services for adolescents who experience loss of control, emotional distress, or negative consequences related to pornography use.

5. Conclusions

5.1. Main Takeaways

This study examined the ability of logistic regression, XGBoost, LightGBM and CatBoost to predict pornography use during the previous 30 days among Spanish adolescents, using a large sample drawn from ESTUDES 2023 [14]. The four methods showed good and very similar predictive performance. CatBoost achieved the highest AUC and the lowest Brier score and log loss, whereas LightGBM obtained the highest accuracy, specificity, and precision. Logistic regression achieved the highest sensitivity, negative predictive value, balanced accuracy, and F1-score. The findings do not support replacing logistic regression automatically with machine-learning models. Method selection depends on the applied objective: logistic regression is particularly competitive when the aim is to maximise the detection of adolescents who use pornography, whereas gradient boosting models produce fewer false positives and classify non-users more accurately.

The SHAP analysis showed that SEX was by far the most influential predictor, followed by NETADD, CAN, ALC, FCON, and certain risky sexual behaviours. Whereas SEX operated mainly as a strong direct predictor, NETADD showed substantial interactions with AGE, CAN, ALC, and FREL. These findings suggest that pornography use is not an isolated behaviour but forms part of a broader digital, family-related, and behavioural profile. In addition, the importance of some correlates varied by sex, particularly ALC and FCON, whereas age-related differences were generally small.

Overall, the study contributes to the literature by combining predictive model comparison, global and local interpretation, interaction analysis, and the assessment of predictive heterogeneity using SHAP. The findings suggest that comprehensive preventive approaches could integrate affective and sexual education, digital literacy, substance-use prevention, the promotion of safer sexual behaviours, and family support. Nevertheless, these models should be used to inform policies and allocate resources at the population level rather than to diagnose, sanction, or stigmatise individual adolescents.

In this regard, it should be noted that sampling error increases when analyses focus on smaller subgroups, such as specific autonomous communities, school years, or age groups. ESTUDES therefore provides highly reliable estimates at the national level, but the findings should be interpreted more cautiously when analyses are conducted on smaller territorial or population subgroups.

5.2. Limitations and Future Research Directions

This study has several limitations that should be considered when interpreting its findings. First, the cross-sectional design of ESTUDES 2023 prevents temporal or causal relation-

ships from being established between the predictors and pornography use. The observed associations may reflect bidirectional influences or the presence of unmeasured factors. Problematic Internet use and pornography use may be bidirectionally correlated, but the cross-sectional design does not permit their temporal ordering to be established. The findings should therefore be interpreted as predictive relationships rather than causal effects.

The dependent variable distinguishes only between adolescents who did and did not use pornography during the previous 30 days. This measure does not capture frequency, duration, type of content, age of onset, context of access, motivations, or potential negative consequences. Nor does it distinguish between occasional and problematic use. Future research should employ ordinal or continuous outcomes and incorporate specific measures of problematic pornography use.

The data were based on self-reported responses to sensitive questions. Although anonymity may encourage honest reporting, the responses may still be affected by recall bias, social desirability, concealment, or differences in how participants interpret the questions. In addition, NETADD is a self-reported index of problematic Internet-use symptoms and should not be interpreted as a clinical diagnosis of Internet addiction. Furthermore, the sample included only adolescents enrolled in formal education.

The complete-case approach reduced the original sample and may have introduced bias if missingness was not completely random. In addition, the analysis did not explicitly incorporate sampling weights or the complex survey design, which calls for caution when generalising the estimates to the broader Spanish adolescent population. Generalisability to other countries is also limited because of cultural, regulatory, and digital-context differences.

Future research could examine whether more complex deep-learning architectures provide additional predictive value for the analysis of problematic adolescent behaviours. In particular, hybrid neural models combining Transformer-based components with Explainable AI techniques may be useful when textual, longitudinal, social-media, or multimodal information is available. Chechkin et al. [47], for example, illustrate the potential of hybrid neural architectures for detecting and classifying destructive digital content. Applying comparable approaches to adolescent pornography use and related risk behaviours could help capture complex contextual and temporal patterns that cannot be fully represented by conventional cross-sectional survey data.

Differences in model performance were small. The use of a 0.50 classification threshold also influenced the resulting sensitivity and specificity. Future research should conduct temporal or external validation, compare alternative classification thresholds, and assess model calibration across subgroups. SHAP values likewise do not represent causal effects or definitive evidence of moderation. These analyses should be complemented by explicit interaction terms, longitudinal models, and comparisons across sex, school type, geographical area, socioeconomic status, and family context.

Author Contributions: Conceptualization, Á.B.-E. and A.S.-A.; methodology, I.P.-G.; software, J.d.A.-S.; validation, Á.B.-E., I.P.-G. and A.S.-A.; formal analysis, A.S.-A.; investigation, Á.B.-E.; resources, I.P.-G.; data curation, Á.B.-E.; writing—original draft preparation, Á.B.-E. and A.S.-A.; writing—review and editing, J.d.A.-S. and I.P.-G.; visualization, J.d.A.-S. and A.S.-A.; supervision, I.P.-G.; project administration, Á.B.-E.; funding acquisition, I.P.-G. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: This paper relies on secondary data available at https://pnsd.sanidad.gob.es/profesionales/sistemasInformacion/sistemaInformacion/encuestas_ESTUDES.htm (accessed on 1 March 2026).

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AGE	Adolescent age
AIC	Akaike information criterion
ALC	Alcohol use
ANOVA	Analysis of variance
AUC	Area under the receiver operating characteristic curve
BIC	Bayesian information criterion
CAN	Cannabis use
CI	Confidence interval
CV	Cross-validation
ESTUDES	Spanish Survey on Drug Use in Secondary Education
F1	F1-score
FCON	Family conflict
FREL	Father–adolescent relationship
LightGBM	Light Gradient Boosting Machine
LR	Likelihood ratio
MD1	Mental distress indicator 1 (psychotic episode)
MD2	Mental distress indicator 2 (anxiety attack)
MREL	Mother–adolescent relationship
MSC	Mathematics Subject Classification
NETADD	Addictive Internet use index
NETADD1–NETADD14	Items measuring addictive Internet use
NPV	Negative predictive value
OEDA	Spanish Observatory on Drugs and Addictions
OR	Odds ratio
PHC	Parental home control
POC	Parental outside-home control
PORN	Pornography use
ROC	Receiver operating characteristic
RQ	Research question
SB1	Sexual behaviour 1 (regretted sexual intercourse)
SB2	Sexual behaviour 2 (sexual intercourse without a condom)
SD	Standard deviation
SE	Standard error
SEX	Adolescent sex
SHAP	Shapley Additive Explanations
TOB	Tobacco use
XGBoost	Extreme Gradient Boosting

Appendix A

Table A1. Comparison between adolescents included in the complete-case analytical sample and those excluded because of missing predictor values.

Variable	Excluded Valid	Included Mean/Proportion	Excluded Mean/Proportion	SMD	*p*-Value
PORN	5213	0.426	0.469	−0.087	<0.001
SEX	5213	0.503	0.423	0.16	<0.001
AGE	5213	16.588	16.554	0.029	0.056
FREL	4299	2.81	2.713	0.187	<0.001
MREL	4083	2.65	2.532	0.162	<0.001
FCON	4543	0.398	0.397	0.001	0.948
PHC	4354	0.379	0.391	−0.024	0.147

Table A1. *Cont.*

Variable	Excluded Valid	Included Mean/Proportion	Excluded Mean/Proportion	SMD	*p*-Value
POC	4280	0.483	0.491	−0.017	0.305
TOB	4603	0.376	0.485	−0.123	<0.001
ALC	4788	0.999	1.123	−0.098	<0.001
CAN	3240	0.994	1.522	−0.230	<0.001
SB1	4460	0.072	0.103	−0.118	<0.001
SB2	4469	0.153	0.192	−0.106	<0.001
MD1	4391	0.033	0.056	−0.126	<0.001
MD2	4481	0.336	0.326	0.022	0.159
NETADD	4484	1.297	1.263	0.041	0.019

Note. The comparison was restricted to adolescents with observed information on pornography use. The complete-case analytical sample contained 33,543 observations, whereas 5213 observations were excluded because at least one predictor was missing. For binary variables, means represent proportions. SEX was coded as 0 = male and 1 = female. SMD denotes the standardised mean difference and was calculated as the mean among included participants minus the mean among excluded participants, divided by the pooled standard deviation. Because of the large sample size, substantive interpretation should focus primarily on the magnitude of the SMD rather than on statistical significance.

Table A2. Repeated five-fold cross-validation results across ten random seeds.

Model	AUC, Mean (SD)	Accuracy	Sensitivity	Specificity	Balanced Accuracy	F1-Score	Brier Score	Log Loss
LR	0.8170 (0.00012)	0.7513	0.7588	0.7457	0.7523	0.7221	0.1716	0.5183
XGBoost	0.8218 (0.00023)	0.7559	0.7189	0.7834	0.7512	0.7151	0.1692	0.5118
LightGBM	0.8217 (0.00030)	0.7561	0.717	0.7852	0.7511	0.7147	0.1691	0.5117
CatBoost	0.8221 (0.00028)	0.7564	0.7246	0.7799	0.7523	0.717	0.169	0.5113

Note. Values other than the AUC standard deviation are means across ten repetitions of five-fold cross-validation conducted within the original training sample. The selected hyperparameters and numbers of boosting iterations were held constant, while the random seed governing fold allocation and stochastic model fitting was varied. Higher values indicate better performance for AUC, accuracy, sensitivity, specificity, balanced accuracy, and F1-score; lower values indicate better performance for the Brier score and log loss. Logistic regression and CatBoost obtained the same balanced accuracy after rounding.

Table A3. Sensitivity of logistic regression estimates to multiple imputation and multicollinearity assessment.

Predictor	Complete-Case OR	Multiple-Imputation OR	Absolute Change (%)	VIF
SEX	0.072	0.0753	4.58	1.393
AGE	1.0714	1.0662	0.49	1.137
FREL	0.9556	0.9778	2.32	1.146
MREL	0.9603	0.9729	1.31	1.15
FCON	1.369	1.3854	1.2	1.283
PHC	1.0322	1.0323	0.01	1.188
POC	1.0491	1.0406	0.81	1.208
TOB	0.9768	0.9848	0.82	2.055
ALC	1.1898	1.1937	0.33	1.425
CAN	1.0874	1.0807	0.62	2.019
SB1	1.2796	1.2447	2.73	1.19
SB2	1.2438	1.2527	0.72	1.337
MD1	0.9643	0.9636	0.07	1.066
MD2	1.0606	1.0525	0.76	1.412
NETADD	1.5178	1.5083	0.63	1.204

Note. Complete-case estimates were based on 33,543 observations. Multiple-imputation estimates were pooled across ten imputed datasets comprising 38,756 adolescents with observed pornography-use information. Absolute change was calculated relative to the complete-case odds ratio. The direction and substantive interpretation of the estimates remained stable, and all changes in the odds ratios were below 5%. MREL was weakly significant in the complete-case analysis ($p = 0.042$) but not after multiple imputation ($p = 0.137$). VIFs were calculated from the complete-case logistic regression model; all values were below 2.1, indicating no problematic multicollinearity. OR = odds ratio; VIF = variance inflation factor. Unlike Table 3, the complete-case and multiple-imputation models reported here used the predictors in their original measurement scales to ensure compatibility across the imputed datasets. Therefore, the odds ratios of non-binary predictors are not directly comparable with those reported in Table 3.

Appendix B

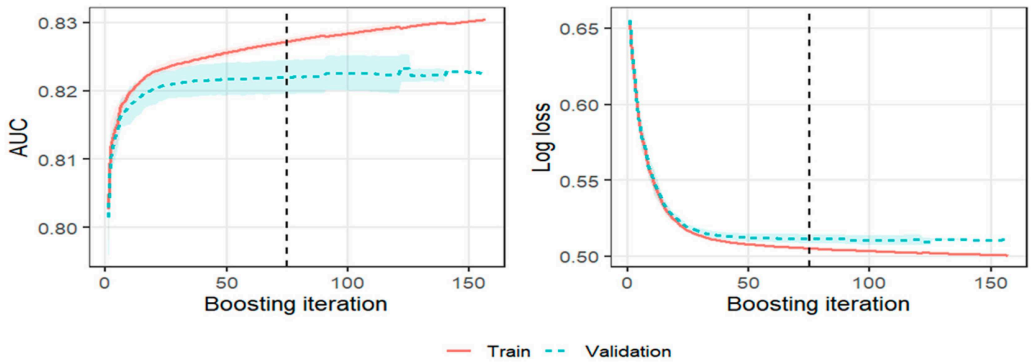


Figure A1. Training and validation learning curves for the selected XGBoost model. The vertical black dashed line indicates the number of boosting iterations selected by early stopping. The shaded area represents variability across the validation folds.

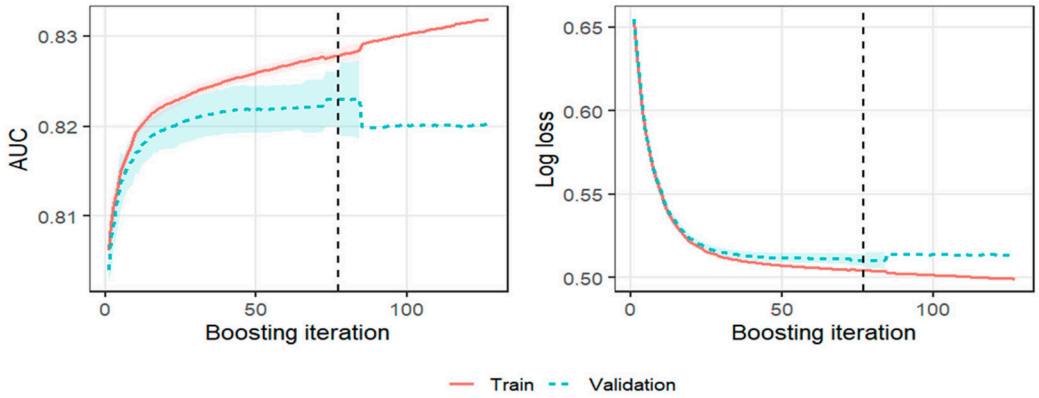


Figure A2. Training and validation learning curves for the selected LightGBM model. The vertical black dashed line indicates the number of boosting iterations selected by early stopping. The shaded area represents variability across the validation folds.

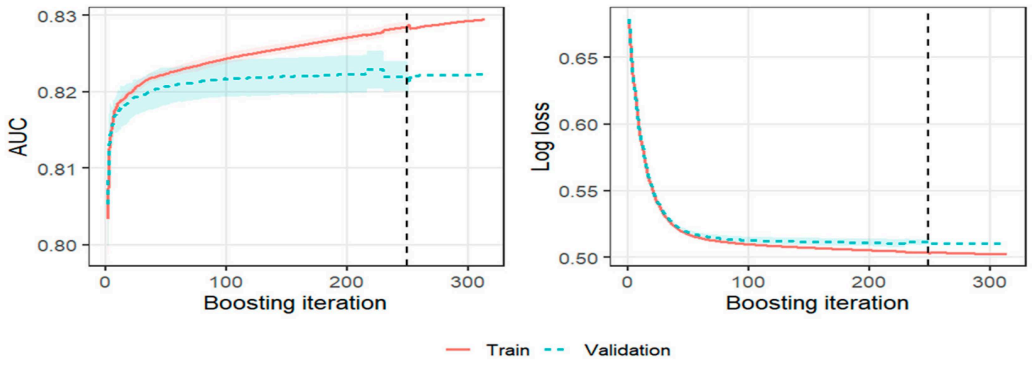


Figure A3. Training and validation learning curves for the selected CatBoost model. The vertical black dashed line indicates the number of boosting iterations selected by early stopping. The shaded area represents variability across the validation folds.

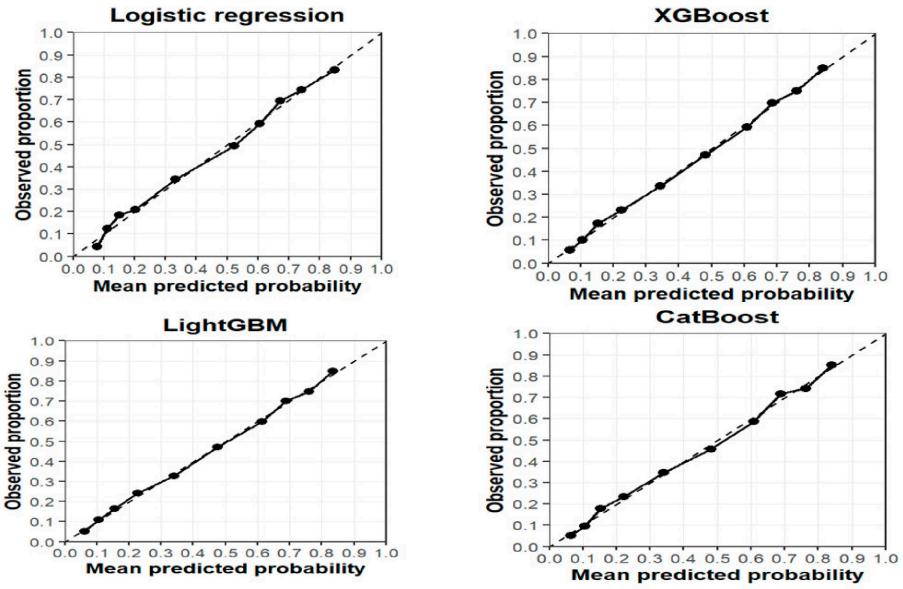


Figure A4. Calibration curves for the four predictive models in the independent test sample.



Figure A5. Local SHAP explanations for representative true-positive, true-negative, false-positive, and false-negative predictions.

References

1. Belzunegui Eraso, Á.G.; Pastor-Gosálbez, I.; Sánchez-Aragón, A. Pornografía, sustancias y pantallas: Un análisis de los predictores del consumo de pornografía digital en la adolescencia española. In Proceedings of the XVI Congreso Español de Sociología, Bilbao, Spain, 15–17 July 2026.
2. Pham, T.; Goto, D.; Kakinaka, M. Adolescents' Risky Online Behaviors and Social Desirability Bias in Vietnam: Privacy Disclosure and Pornography Use. *Cyberpsychol. Behav. Soc. Netw.* **2025**, *28*, 266–274. [CrossRef] [PubMed]
3. Braun-Courville, D.K.; Rojas, M. Exposure to sexually explicit web sites and adolescent sexual attitudes and behaviors. *J. Adolesc. Health* **2009**, *45*, 156–162. [CrossRef] [PubMed]
4. Mattebo, M.; Tydén, T.; Häggström-Nordin, E.; Nilsson, K.W.; Larsson, M. Pornography consumption and psychosomatic and depressive symptoms among Swedish adolescents: A longitudinal study. *Ups. J. Med. Sci.* **2018**, *123*, 237–246. [CrossRef] [PubMed]
5. Wright, P.J.; Tokunaga, R.S.; Herbenick, D.; Paul, B. Pornography vs. sexual science: The role of pornography use and dependency in U.S. teenagers' sexual illiteracy. *Commun. Monogr.* **2022**, *89*, 332–353. [CrossRef]
6. Vandenbosch, L.; Peter, J. Antecedents of the Initiation of Watching Sexually Explicit Internet Material: A Longitudinal Study Among Adolescents. *Mass Commun. Soc.* **2016**, *19*, 499–521. [CrossRef]

7. Bőthe, B.; Vaillancourt-Morel, M.-P.; Dion, J.; Štulhofer, A.; Bergeron, S. Validity and reliability of the short version of the Problematic Pornography Consumption Scale for Adolescents (PPCS-6-A). *Psychol. Addctv. Behav.* **2021**, *35*, 486–500. [CrossRef] [PubMed]
8. Kostopoulou, E. Impact of COVID-19 on adolescent sexual life and attitudes: Have we considered all the possible secondary effects of the pandemic? *Eur. J. Pediatr.* **2023**, *182*, 2459–2469. [CrossRef] [PubMed]
9. Privara, M.; Bob, P. Pornography consumption and cognitive-affective distress. *J. Nerv. Ment. Dis.* **2023**, *211*, 641–646. [CrossRef] [PubMed]
10. Kabaha, M.; Efrati, Y. Profiles of Alcohol Use Disorders, Problematic Pornography Use, and Compulsive Sexual Behavior: The Role of Mental Health, Moral Disapproval, Thought Suppression, and Self-Concealment Among Israeli–Arab Adolescents. *Arch. Sex. Behav.* **2025**, *54*, 2269–2286. [CrossRef] [PubMed]
11. Chaise, R.F.; Cardoso, C.P.; Wingert, F.F.; Piltcher, T.T.; Rocha, K.B.; Martín López, M.J. Factors associated with sexually explicit Internet material use among adolescents: A systematic review. *Arch. Sex. Behav.* **2024**, *53*, 3993–4029. [CrossRef] [PubMed]
12. Pereda, N.; Águila-Otero, A. Prevalence of and factors associated with intentional pornography access among Spanish adolescents. *Sex. Res. Soc. Policy* **2026**. [CrossRef]
13. Paulus, F.W.; Nouri, F.; Ohmann, S.; Möhler, E.; Popow, C. The impact of Internet pornography on children and adolescents: A systematic review. *Encephale* **2024**, *50*, 649–662. [CrossRef] [PubMed]
14. Observatorio Español de las Drogas y las Adicciones. *Informe Sobre Adicciones Comportamentales y Otros Trastornos Adictivos 2023: Indicador de Admitidos a Tratamiento por Adicciones Comportamentales. Juego Con Dinero, Uso de VIDEOJUEGOS, Uso Problemático de Internet y Otros Trastornos Adictivos en las Encuestas de Drogas en España EDADES y ESTUDES 2023*; Observatorio Español de las Drogas y las Adicciones: Madrid, Spain, 2023.
15. Hald, G.M.; Kuyper, L.; Adam, P.C.G.; de Wit, J.B.F. Does viewing explain doing? Assessing the association between sexually explicit materials use and sexual behaviors in a large sample of Dutch adolescents and young adults. *J. Sex. Med.* **2013**, *10*, 2986–2995. [CrossRef] [PubMed]
16. Rivera, R.; Santos-Velasco, D.; Cabrera-García, V.; Docal-Millán, M.d.C. Online and offline pornography consumption in Colombian adolescents. *Comunicar* **2016**, *24*, 37–45. [CrossRef]
17. Donevan, M.; Jonsson, L.; Bladh, M.; Priebe, G.; Fredlund, C.; Svedin, C.G. Adolescents' Use of Pornography: Trends over a Ten-Year Period in Sweden. *Arch. Sex. Behav.* **2022**, *51*, 1125–1140. [CrossRef] [PubMed]
18. Calatrava, M.; de Magalhaes, P.A.; Vidaurreta, M.; Rivas, S.; López-del Burgo, C.; Belintxon, M. Parental competence and pornography use among Hispanic adolescents. *Behav. Sci.* **2024**, *14*, 926. [CrossRef] [PubMed]
19. Feijóo, S.; Portela, V.; Rial, A. Online Pornography Consumption, Risky Behaviors, and Sexist Attitudes in Adolescence: A Cross-Sectional Survey Study. *Arch. Sex. Behav.* **2025**, *54*, 3223–3233. [CrossRef] [PubMed]
20. Chen, T.; Guestrin, C. XGBoost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Association for Computing Machinery, San Francisco, CA, USA, 13–17 August 2016; pp. 785–794. [CrossRef]
21. Ke, G.; Meng, Q.; Finley, T.; Wang, T.; Chen, W.; Ma, W.; Ye, Q.; Liy, T.-Y. LightGBM: A highly efficient gradient boosting decision tree. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 3146–3154.
22. Prokhorenkova, L.; Gusev, G.; Vorobev, A.; Dorogush, A.V.; Gulin, A. CatBoost: Unbiased boosting with categorical features. In Proceedings of the 32nd International Conference on Neural Information Processing Systems, Red Hook, NY, USA, 3–8 December 2018; pp. 6639–6649.
23. Andrés-Sánchez, J.; Torres-Coronas, T. Workers' perceptions of the economic impact of generative AI on sustainable economic development: Evidence from ensemble tree models and explainable machine learning. *Qual. Quant.* **2026**. [CrossRef]
24. Huang, K.; Yang, Y.; Wang, L.; Li, J.; Qu, D.; Chen, R.; Chi, X. The Dual Effects of Individual and Contextual Factors on Adolescent Problematic Internet Use: Machine Learning Approaches and SHAP Explanations. *J. Behav. Addict* **2026**. [CrossRef] [PubMed]
25. Li, Z.; Long, Q.; Hao, M.; Liu, L.; Weng, Y.; Yang, C.; Wu, J.; Huang, B.; Qiu, J.; Huang, Y.; et al. Discovery of Factors Associated with Smartphone Addiction Among High School Adolescents: Using Machine Learning and Network Analysis. *J. Affect Disord.* **2026**, *392*, 120096. [CrossRef] [PubMed]
26. Lee, K.; Seo, W.; Jung, S.Y. Prediction of Smartphone Addiction Among Korean Adolescents Based on Physical Activity and Mental Health: A Machine Learning Analysis Using LASSO and SHAP From the Korea Youth Risk Behavior Survey. *Alpha Psychiatry* **2026**, *27*, 46201. [CrossRef] [PubMed]
27. Kim, S.; Kim, H.; Kim, S.; Lee, H.; Hammoodi, A.; Choi, Y.; Kim, H.J.; Smith, L.; Kim, M.S.; Fond, G.; et al. Machine Learning-Based Prediction of Substance Use in Adolescents in Three Independent Worldwide Cohorts: Algorithm Development and Validation Study. *J. Med. Internet Res.* **2025**, *27*, e62805. [CrossRef] [PubMed]
28. Seelpipat, N.; Sutivong, D. Analyzing adolescents' e-cigarette initiation tendency using explainable machine learning. *J. Comput Soc. Sci.* **2025**, *8*, 95. [CrossRef]
29. Lundberg, S.M.; Lee, S.-I. A unified approach to interpreting model predictions. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 4765–4774.

30. Kincaid, C.; Jones, D.J.; Sterrett, E.; McKee, L. A Review of Parenting and Adolescent Sexual Behavior: The Moderating Role of Gender. *Clin. Psychol. Rev.* **2012**, *32*, 177–188. [CrossRef] [PubMed]
31. Pinquart, M. Associations of Parenting Dimensions and Styles with Externalizing Problems of Children and Adolescents: An Updated Meta-Analysis. *Dev. Psychol.* **2017**, *53*, 873–932. [CrossRef] [PubMed]
32. Gardner, M.; Steinberg, L. Peer Influence on Risk Taking, Risk Preference, and Risky Decision Making in Adolescence and Adulthood: An Experimental Study. *Dev. Psychol.* **2005**, *41*, 625–635. [CrossRef] [PubMed]
33. Giletta, M.; Scholte, R.H.J.; Engels, R.C.M.E.; Ciairano, S.; Prinstein, M.J. Adolescent Non-Suicidal Self-Injury: A Cross-National Study of Community Samples from Italy, the Netherlands and the United States. *Psychiatry Res.* **2012**, *197*, 66–72. [CrossRef] [PubMed]
34. Farré, J.M.; Montejo, A.L.; Agulló, M.; Granero, R.; Chiclana Actis, C.; Villena, A.; Maideu, E.; Sánchez, M.; Fernández-Aranda, F.; Jiménez-Murcia, S.; et al. Pornography use in adolescents and its clinical implications. *J. Clin. Med.* **2020**, *9*, 3625. [CrossRef] [PubMed]
35. Villena-Moya, A.; Chiclana-Actis, C.; Granero, R.; Fernández-Aranda, F.; Potenza, M.N.; Jiménez-Murcia, S.; Mestre-Bach, G. Pornography use and associated factors in adolescents: A cross-jurisdictional approach (Spain vs. Mexico). *Arch. Sex. Behav.* **2025**, *54*, 733–748. [CrossRef] [PubMed]
36. Tomić, I.; Burić, J.; Štulhofer, A. Associations between Croatian adolescents' use of sexually explicit material and sexual behavior: Does parental monitoring play a role? *Arch. Sex. Behav.* **2018**, *47*, 1881–1893. [CrossRef] [PubMed]
37. Moya, A.V.; Efrati, Y.; Chiclana Actis, C.; Testa, G. Family and Parental Factors Related to Pornography Use and Problematic Pornography Use Among Adolescents and Young Adults. *Curr. Addict Rep.* **2025**, *12*, 38. [CrossRef]
38. Jiang, X.; Wu, Y.; Zhang, K.; Bőthe, B.; Hong, Y.; Chen, L. Symptoms of problematic pornography use among help-seeking male adolescents: Latent profile and network analysis. *J. Behav. Addict* **2022**, *11*, 912–927. [CrossRef] [PubMed]
39. Alexandraki, K.; Stavropoulos, V.; Burleigh, T.L.; King, D.L.; Griffiths, M.D. Internet pornography viewing preference as a risk factor for adolescent Internet addiction: The moderating role of classroom personality factors. *J. Behav. Addict* **2018**, *7*, 423–432. [CrossRef] [PubMed]
40. Andrie, E.K.; Sakou, I.I.; Tzavela, E.C.; Richardson, C.; Tsitsika, A.K. Adolescents' online pornography exposure and its relationship to sociodemographic and psychopathological correlates: A cross-sectional study in six European countries. *Children* **2021**, *8*, 925. [CrossRef] [PubMed]
41. Bocci Benucci, S.; Di Gesto, C.; Ghinassi, S.; Casale, S.; Fioravanti, G. Pornography Use, Problematic Pornography Use, Impulsivity, and Sensation Seeking: A Meta-Analysis. *J. Sex. Med.* **2024**, *21*, 922–939. [CrossRef] [PubMed]
42. Lin, M.; Lucas, H.C.; Shmueli, G. Too big to fail: Large samples and the p-value problem. *Inf. Syst. Res.* **2013**, *24*, 906–917. [CrossRef]
43. Cohen, J. *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed.; Lawrence Erlbaum Associates: Hillsdale, NJ, USA, 1988.
44. Christodoulou, E.; Ma, J.; Collins, G.S.; Steyerberg, E.W.; Verbakel, J.Y.; Van Calster, B. A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. *J. Clin. Epidemiol.* **2019**, *110*, 12–22. [CrossRef] [PubMed]
45. Hu, Y.; Zhang, X.; Slavin, V.; Belsti, Y.; Tiruneh, S.A.; Callander, E.; Enticott, J. Beyond comparing machine learning and logistic regression in clinical prediction modelling: Shifting from model debate to data quality. *J. Med. Internet Res.* **2025**, *27*, e77721. [CrossRef] [PubMed]
46. Wu, T.; Wei, Y.; Wu, J.; Yi, B.; Li, H. Logistic regression technique is comparable to complex machine learning algorithms in predicting cognitive impairment related to post-intensive care syndrome. *Sci. Rep.* **2023**, *13*, 2485. [CrossRef] [PubMed]
47. Chechkin, A.; Pleshakova, E.; Gataullin, S. A Hybrid Neural Network Transformer for Detecting and Classifying Destructive Content in Digital Space. *Algorithms* **2025**, *18*, 735. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.