International Scientific Conference 20 UNITECH 26

INTERNATIONAL SCIENTIFIC CONFERENCE
19-21 November 2026
GABROVO

TRADITIONAL STATISTICS OR MACHINE LEARNING? RATIONALES FOR USING MACHINE LEARNING IN CYBERPSYCHOLOGY RESEARCH: A NARRATIVE REVIEW

Berfin Kılınç[1]*, Eylem Seyhan[1], Deniz Mertkan Gezgin[2]

[1]Trakya University, Department of Computer Engineering, Edirne, Türkiye
[2]Trakya University, Department of Instructional Technologies, Edirne, Türkiye
**Corresponding author: berfinkilinc10@trakya.edu.tr*

Abstract
Traditional statistics typically test theoretically specified relationships, whereas machine learning (ML) emphasizes prediction. This narrative review examined rationales for using ML instead of, or alongside, traditional statistics in cyberpsychology. Four complementary searches were conducted in Scopus (312 records) and the Web of Science Core Collection (269 records); 23 empirical studies formed the core corpus. Studies were compared by topic, data source, analytic aims, methods, and explainable artificial intelligence (XAI) use; validation practices were noted but not systematically coded. Most studies used cross-sectional self-reports; only two drew on non-self-report data (fMRI and Facebook digital traces). Seven studies used SHapley Additive exPlanations (SHAP), seven combined ML with inferential approaches, and 13 compared complex models with logistic, linear, or penalized regression. ML was chiefly motivated by prediction and classification, multiple predictors, high-dimensional data, and interpretability. However, complex ML did not consistently outperform regression, which sometimes performed similarly or better. Traditional statistics and ML thus serve complementary purposes. This synthesis proposes selecting methods according to the research question, data quality, and validation needs rather than algorithmic novelty.

Keywords: Cyberpsychology, machine learning, traditional statistical methods, XAI, narrative review.

1. INTRODUCTION

Digital technologies have transformed the ways in which individuals communicate, learn, work, and spend their leisure time, giving rise to cyberpsychology, which examines technology-related human behavior and brings together psychology, human–computer interaction, computer science, and engineering [1]. Its principal areas of inquiry are summarized in Figure 1 [2].

Among the most frequently examined topics in the field are internet addiction, problematic social media use, and gaming disorder [3], together with technology-mediated problematic behaviors such as cyberchondria, cyberbullying, and cyberloafing [4]. In recent years, smartphone addiction, nomophobia, and fear of missing out (FoMO) have also been studied extensively [5].

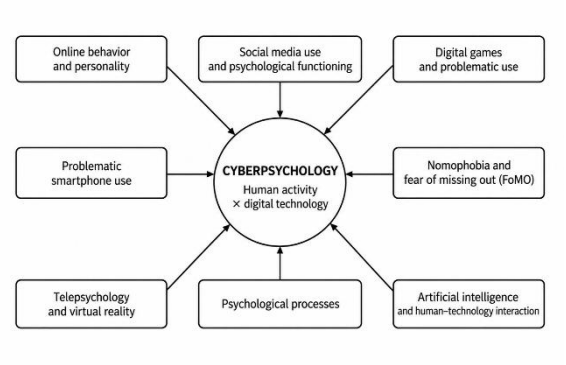


Figure 1. Core research areas of cyberpsychology

Note. Adapted from Norman [6], Ancis [2], and Attrill-Smith et al. [1].

This is an open access article licensed under Creative Commons Attribution 4.0 International
Print ISSN: 1313-230X, Online ISSN: 3033-1404

For a long time, these phenomena were examined predominantly through data collected with self-report scales and analyzed with traditional statistical methods [7]. In recent years, however, machine learning (ML) methods have become increasingly visible in the field. This shift also concerns the nature of research questions: in addition to asking "Is there a significant relationship among the variables?", researchers are increasingly asking, "How accurately can the behavior of a new individual be predicted?"

1.1. Methods Used in Traditional Research

Traditional statistical methods estimate the direction and magnitude of relationships among variables and test them against a theoretical model, serving explanation, inference, and hypothesis testing. Hierarchical regression assesses the stepwise contribution of variable blocks to explained variance (e.g., [8]). Mediation and moderation analyses examine the mechanisms and boundary conditions of effects [9]. Confirmatory factor analysis (CFA) verifies measurement structures, and structural equation modeling (SEM) and partial least squares SEM (PLS-SEM) evaluate direct and indirect relationships within a single model [10].

These approaches emphasize prespecified models, but their explanatory power does not necessarily generalize to unseen cases [11]. Selecting among numerous predictors and managing complex models can also become challenging.

1.2. Machine Learning

In this study, traditional statistics and ML were distinguished by the purpose and evaluation of the analysis rather than by the algorithm [12], [13]. Analyses that explain relationships through significance testing were classified as traditional methods, whereas models whose predictive performance was evaluated through cross-validation or a test set were classified as ML. Accordingly, logistic regression, LASSO, ridge, and elastic net were also considered ML when employed for predictive purposes [11], [14]. Thus, "ML–regression comparison" refers to comparing complex algorithms (e.g., random forest [RF], support vector machines [SVM], gradient boosting methods such as XGBoost, and artificial neural networks [ANN]) with regression-based models.

ML encompasses supervised, unsupervised, and semi-supervised learning [14]. Predictive pipelines typically handle missing data, encoding, scaling, and feature selection before fitting models on training data and evaluating performance on held-out data or through cross-validation [11], [15]. Accuracy, F1, and AUC capture different aspects of performance, especially when classes are imbalanced.

Because high performance does not imply that a model is psychologically interpretable, performance and interpretability should be considered together [14], [16]. Explainable artificial intelligence (XAI) methods make predictions and feature contributions comprehensible: permutation importance and partial dependence plots depict global relationships, SHapley Additive exPlanations (SHAP) quantifies feature contributions to individual (and aggregated) predictions, and local interpretable model-agnostic explanations (LIME) explains single observations locally [17]. However, XAI outputs do not demonstrate causality and may be distorted by measurement errors and intercorrelated predictors. They should therefore be evaluated in conjunction with psychological theory, research design, and traditional statistics [16], [17].

1.3. The Present Study

ML should be selected according to the research question and data structure, not novelty. In cyberpsychology, it expands—rather than replaces—traditional explanation by enabling prediction, classification, and pattern discovery. Reviews of clinical psychology and psychiatry have emphasized diagnosis and prognosis [18], while the rationales for adopting ML in cyberpsychology remain less clear. This review compares when and why the approaches are used and how they may complement each other. Four questions guided the synthesis:

RQ1. Which research requirements are invoked to justify the use of ML in cyberpsychology research?

RQ2. Under which conditions do ML models outperform, or fail to outperform, traditional models?

RQ3. How are traditional statistical methods and ML employed together within the same studies?

RQ4. How should the two approaches be positioned in relation to the nature of the research question?

2. METHOD

2.1. Research Design

This narrative review interpretively compared studies with different designs [19], [20]. Unlike a systematic review, it did not seek exhaustive, reproducible selection. The SANRA criteria informed planning and reporting [21].

2.2. Search Strategy

Scopus and Web of Science (WoS) Core Collection were searched on September 21, 2026, using four groups: (1) explanation versus prediction in psychology (Scopus 103; WoS 86); (2) regression versus ML in problematic internet, smartphone, gaming, and social media use (17; 14); (3) ML combined with SEM, mediation/moderation, hierarchical regression, or network analysis (49; 43); and (4) high dimensionality, nonlinearity, many predictors, and big data (143; 126). Full strings are available upon request.

The searches yielded 312 Scopus and 269 WoS records, with possible overlap across groups. No publication-year restriction was imposed. Broader behavioral-science sources also informed the methodological framing.

2.3. Study Selection Process

After merging the records and removing duplicates based on DOI, title, and bibliographic information, 403 unique records remained; these were screened by title and abstract and then at full text for topical and methodological relevance. The principal criteria for inclusion in the core literature set were that a study addressed a cyberpsychological phenomenon, explained the rationale for adopting ML, and reported its methodology with sufficient clarity. Studies contributing to at least one of the rationale themes outlined below were prioritized. The mere use of an ML algorithm was insufficient; the study had to justify ML explicitly in its introduction or method with reference to the research question, data structure, or a predictive requirement, and "sufficient clarity" meant reporting the algorithms, target variable, and performance metrics. The rationales in Table 1 are the reviewers' thematic interpretations of these justifications rather than verbatim statements. Ultimately, 23 studies were included in the core literature set. Unlike in systematic reviews, the reasons for exclusion were not reported at each stage. Methodological sources on the prediction–explanation distinction, measurement, method complementarity, and XAI, drawn mainly from the first search group, were used to support interpretation.

2.4. Data Extraction and Analysis

For each study, the author(s), year, topic or data source, methods, aim, and rationale for adopting ML were extracted, together with prediction and classification tasks, model comparisons, feature selection, validation, and XAI use (Table 1). The coding framework was derived from the research questions and the methodological literature [11], [14], [22]. The question "Why was ML adopted?" was examined under the following headings: prediction and classification, many predictors, complex relationships, novel or high-dimensional data, explainability, the complementarity of methods, and situations in which ML did not prove superior. Studies could contribute to more than one theme. Coding was conducted by two researchers, and disagreements were resolved through consultation with a third researcher. Study-level coded characteristics are presented in the final column of Table 1.

The synthesis compared research aims, data structures, methods, and evaluation practices rather than pooling incompatible accuracy estimates. Self-report, digital-trace, diary, and neuroimaging evidence were considered in relation to each study's design and aims.

Table 1. *Methodological Characteristics of the 23 Studies Included in the Review and Rationales for the Shift Toward Machine Learning*

Study	Topic / data source	Method / algorithm	Methodological rationale identified in the review	Codes
McKee et al. [23]	FoMO and maladaptive behavior	Hierarchical regression with interactions + ML (LR, DT, RF, SVM)	Predicting behavioral classes of new individuals beyond explaining relationships	HYB; REG
Gan et al. [24]	Internet behavioral addiction in adolescents	LR, MLP, RF, KNN, SVM, XGBoost	Selecting strong predictors among many candidates; individual-level prediction	PUA; REG
Zhou et al. [25]	Smartphone addiction in adolescents with depression	RF, SVM, LR, XGBoost, SHAP, mediation/moderation	ML selects predictors; traditional analyses test mechanisms	PUA; SHAP; HYB; REG
Abu-Taieh, AlHadid, Kaabneh et al. [26]	Smartphone addiction and social isolation	SEM + ML (linear regression, SMO, bagging REPTree, ANN, RF)	Predictive power alongside SEM-tested theoretical relationships	PUA; HYB; REG
Abu-Taieh, AlHadid, Masa'deh et al. [27]	Social network use, anxiety, and depression	SEM + extended TAM + ML (as above; 10-fold CV)	Prediction and variable selection beyond theoretical modeling	HYB; REG
Ioannidis et al. [28]	Problematic internet use	LR, RF, Naive Bayes	Out-of-sample prediction and cross-center transferability	PUA; REG
Li et al. [29]	Smartphone addiction in high school students	ML + network analysis	Data-driven selection among many candidate predictors	PUA; NET
Lee et al. [30]	Smartphone addiction in Korean adolescents	LASSO, LR, RF, XGBoost, LightGBM, SHAP	Variable selection and prediction from many physical and psychological variables (large sample)	PUA; SHAP; REG
Gui et al. [31]	Internet addiction and cognitive emotion regulation	LASSO, LR, RF, SVM, GBM, MLP, SHAP	Predicting co-occurring psychological conditions from multiple risk factors; XAI explanation	PUA; SHAP; REG
Lee and Kim [32]	Self-reported usage patterns and smartphone addiction (Korea, N = 29,712)	DT, RF, XGBoost + hypothesis tests	Prediction from many usage characteristics (large sample); ML complements hypothesis testing	PUA
Namlı et al. [33]	Smartphone and gaming addiction in adolescents	Regression vs. LASSO, elastic net, RF, GBM, AdaBoost, SVR, KNN, MLP, LightGBM, GBM–RF hybrid	Predictive accuracy of classical regression vs. ML	PUA; REG
Çiftci and Yıldız [34]	Social media addiction, happiness, and life satisfaction	SEM, mediation + ML (incl. elastic net) + SHAP	Prediction and variable importance beyond relational analyses	PUA; SHAP; HYB; REG
Xie and Li [35]	Cyberbullying victimization	Multiple ML + network analysis + SHAP	Joint evaluation of many predictors across ecological levels	SHAP; NET
Zhou et al. [36]	Daily problematic mobile phone use	ML + network analysis	Next-day prediction from dynamic and static diary predictors	PUA; NET
Liu et al. [37]	Problematic smartphone use in adolescents	Multiple ML models	Heterogeneous risk factors and prediction (nationally representative sample)	PUA

Study	Topic / data source	Method / algorithm	Methodological rationale identified in the review	Codes
Zhan et al. [38]	Internet gaming disorder in adolescents	ML + SHAP + network analysis	Selecting key variables among many risk factors; explaining results	PUA; SHAP; NET
Liang et al. [39]	Internet gaming disorder in adolescents	ML + path analysis	ML screens variables; path analysis tests relationships	PUA; HYB
Tie et al. [40]	fMRI, stress, self-control, and problematic smartphone use	CPM + CV + mediation	Prediction from high-dimensional neuroimaging patterns	PUA; HYB
Vera Cruz et al. [41]	Smartphone applications, mental health, and behavioral predictors	ML + LR	ML selects predictors; LR examines relationships	_
Güzel et al. [42]	Smartphone addiction and temporomandibular disorders	Correlation; LR, linear, ridge, RF, XGBoost, SVM/SVR, KNN, GB, ensemble subspace discriminant; SHAP	Traditional analyses for relationships; ML for prediction and classification	PUA; SHAP; REG
Xiao et al. [43]	Problematic smartphone use and FoMO in adolescents (N = 2,100)	LASSO, ridge, elastic net, other ML	Relative importance of many predictors (penalized regression best)	PUA; REG
Marengo et al. [44]	Facebook digital footprints and social network use disorder tendency (N = 1,094)	ML ensemble of linear and nonlinear algorithms	Predicting individual differences from objective, high-volume digital footprints	PUA
Giraldo-Jiménez et al. [45]	Smartphone addiction in university students (Colombia, N = 1,228)	RF, LR, SVM, other classifiers	Data-driven alternative to self-report/expert diagnosis (similar accuracy, 76%–77%)	PUA; REG

Note. Rationales are the reviewers' interpretations, not verbatim author statements. Güzel et al. [42] was included for its focus on smartphone addiction and combined analyses. Codes are nonexclusive: PUA = problematic use/addiction as an outcome or key variable (n = 19); SHAP = SHAP used (n = 7); HYB = ML with SEM, mediation/moderation, path analysis, or hierarchical regression (n = 7); NET = network analysis (n = 4); REG = regression compared with complex ML (n = 13). Vera Cruz et al. [41] was not coded REG because its logistic regression examined ML-selected predictors. Abbreviations: ANN = artificial neural network; CPM = connectome-based predictive modeling; CV = cross-validation; DT = decision tree; GB/GBM = gradient boosting/machine; KNN = k-nearest neighbours; LASSO = least absolute shrinkage and selection operator; LR = logistic regression; MLP = multilayer perceptron; RF = random forest; SEM = structural equation modeling; SMO = sequential minimal optimization; SVM = support vector machine; SVR = support vector regression; TAM = technology acceptance model; XGBoost = extreme gradient boosting.

3. RESULTS

The 23 studies were published in 2016 (n = 1), 2021 (n = 1), 2022 (n = 5), 2023 (n = 3), 2025 (n = 4), and 2026 (n = 9), a distribution that cannot establish temporal prevalence in this narrative sample. Nineteen addressed problematic internet, smartphone, gaming, or social media use as a core variable. Four concerned FoMO/maladaptive behavior [23], social networking and psychological outcomes [27], cyberbullying [35], or mental health applications [41]. All employed supervised ML; conclusions primarily concern problematic technology use.

Cross-sectional self-report data predominated. Only Tie et al. [40] (fMRI) and Marengo et al. [44] (Facebook traces) used non-self-report sources; Zhou et al. [36] used repeated diary measures. Seven studies used SHAP, seven combined ML with inferential models, four used network analysis alongside ML, and 13 compared regression with complex ML. Codes overlap (Table 1).

3.1. The Need for Prediction and Classification

The most salient common feature is the expansion of research questions from relationships among variables to predicting

and classifying individuals or behaviors. McKee et al. [23] separated the two aims within one study, using traditional statistics for relationships and ML to classify new individuals. Lee and Kim [32] and Giraldo-Jiménez et al. [45] predicted risk and behavior, and Ioannidis et al. [28] tested out-of-sample prediction of problematic internet use across centers in two countries. Gan et al. [24] likewise compared ML with logistic regression (Section 3.6).

3.2. Assessing Multiple Predictors and Complex Relationships

The second rationale is that outcomes are associated with many variables whose associations may be nonlinear. Li et al. [29], Liu et al. [37], and Xie and Li [35] examined smartphone addiction, heterogeneous predictors of problematic smartphone use, and cyberbullying victimization, respectively, through multiple predictors and network patterns. Xiao et al. [43] and Namlı et al. [33] focused on the relative contribution of numerous predictors and on regression–ML comparison (Section 3.6). Here, ML served both prediction and the examination of variable importance and complex patterns.

Evidence of nonlinear gains was limited because interactions and nonlinearity were not consistently isolated. Penalized regression performed best in Xiao et al. [43], illustrating that complexity alone does not guarantee improvement. Small samples, measurement error, non-independence, and missingness remain important challenges [22], [46].

3.3. Use of Novel and High-Dimensional Data Sources

The diversification of data sources also supports the shift toward ML. Tie et al. [40] predicted problematic use from fMRI connectivity data, Marengo et al. [44] from Facebook digital footprints, and Zhou et al. [36] from 14-day diary data. Here, the nature and structure of the data, not only their volume, changed. By contrast, Lee and Kim [32] jointly evaluated numerous usage characteristics in a self-report dataset (N = 29,712).

3.4. The Need for Explainability

Interpretability was a further rationale: researchers sought to understand which features drove prediction. SHAP was used in addiction and cyberbullying studies; Zhou et al. [25] tested SHAP-selected factors with mediation/moderation, while Gui et al. [31] combined ML and SHAP. Such attributions answer which inputs affect a fitted prediction, not whether they cause the outcome [17].

3.5. Complementary Use of Traditional Statistics and Machine Learning

The approaches were often complementary. Zhou et al. [25], Vera Cruz et al. [41], and Liang et al. [39] used ML to select predictors, followed by mediation/moderation, logistic regression, or path analysis. Both Abu-Taieh teams [26], [27] combined SEM with ML. Traditional models examined theoretical mechanisms; ML addressed prediction and variable selection.

3.6. Machine Learning Is Not Superior in All Cases

Complex ML models were not consistently superior. Of 13 regression–ML comparisons, four illustrated this issue: Giraldo-Jiménez et al. [45] reported similar accuracies for RF, logistic regression, and SVM (76%–77%); Xiao et al. [43] favored penalized regression; and Gan et al. [24] reported AUCs of 0.847 (logistic regression), 0.846 (SVM), and 0.836 (XGBoost). Ioannidis et al. [28] also found comparable results, whereas Namlı et al. [33] favored ML. Model comparisons depend on data and validation. Because validation designs were not coded systematically, these patterns remain preliminary.

4. DISCUSSION

The reviewed shift toward ML reflects changing questions more than consistently better algorithms, although requiring an explicit ML rationale at selection may have shaped this finding. As cyberpsychology moves from explaining associations toward predicting individual risk, held-out

performance becomes central [11]. Multiple psychological and behavioral predictors, together with emerging non-questionnaire data, broaden methodological needs. Digital traces can also predict personality-related differences [47].

Hybrid studies illustrate the complementary roles of inference and prediction. XAI can interpret model contributions but cannot establish mechanisms [17]. Testing SHAP-identified variables with mediation, moderation, or path analysis can address distinct explanatory questions.

Complex ML is not invariably better: a systematic review found no general advantage over logistic regression for clinical prediction [48]. External validation remains crucial [18].

Here, logistic and penalized regression counted as ML when evaluated predictively. Therefore, the relevant contrast is often simple versus complex predictive models, not statistics versus ML. The same logistic regression can support theoretical inference when coefficients are interpreted, or prediction when judged on held-out performance [12], [13].

The methodological appraisal of the reviewed literature highlights three points: (1) because most studies relied on cross-sectional self-report data, predictive performance may have been influenced by shared method variance; (2) internal cross-validation is insufficient for establishing generalizability, and external validation in an independent sample provides stronger evidence [11], [28]; and (3) data leakage may occur if scaling, feature selection, and hyperparameter tuning are performed before the training–test split [49]. In ML studies, accuracy or AUC values should not be interpreted independently of the data type and validation design. As validation strategies were not coded, the extent of external validation and leakage risk in the reviewed studies could not be determined. The TRIPOD+AI statement [50] provides a reporting checklist for these issues.

5. LIMITATIONS

Because this is a narrative review, study selection is not reproducible, and the findings are not based on a representative sample of the literature. Restricting the search to Scopus and the WoS Core Collection may have omitted relevant studies indexed elsewhere (e.g., Google Scholar, PsycINFO, PubMed, IEEE Xplore), a risk heightened by the interdisciplinary nature of the field. Because ML-related characteristics (prediction, high-dimensional data, explainability, and hybrid methods) were considered during selection, the themes may partly reflect the selection criteria. Consequently, the findings do not establish the reasons for adopting ML; rather, they describe how these rationales appear in the reviewed literature. Furthermore, the corpus consists largely of studies on problematic technology use and addiction (19 of the 23 studies); the rationales for using ML may differ in other areas of cyberpsychology, such as virtual reality, online identity, or human–artificial intelligence interaction.

Heterogeneous populations, outcomes, algorithms, and metrics prevented direct performance comparison or tests on a common dataset. Rationales were interpretive codes and may reflect reviewer judgment. Validation strategies, sample sizes, and metrics were not systematically coded, limiting performance-related conclusions. Rapid developments in digital phenotyping, generative AI, multimodal data, and XAI may also change these patterns.

6. CONCLUSION

This narrative review of 23 cyberpsychology studies found five recurring rationales for ML: prediction/classification, many predictors, complex patterns, high-dimensional data, and explainability. Complex ML did not consistently outperform regression in the available comparisons. Hybrid studies used ML for prediction or selection and traditional statistics to examine theoretical relationships and mechanisms.

Method selection should thus follow the question: traditional inference for explanation and theory testing, ML for validated prediction, and both when both aims matter. Complexity is not a proxy for

accuracy. Data structure, sample size, measurement, and validation should guide modeling, with transparent reporting of out-of-sample performance and interpretation.

A practical checklist follows: (1) Is the aim explanation, prediction, or both? (2) Do sample size and measurement quality support the model? (3) Are training, validation, testing, and leakage prevention sound, ideally with independent validation? (4) Are XAI outputs interpreted as model-specific contributions rather than causal effects?

REFERENCES

[1] A. Attrill-Smith, C. Fullwood, M. Keep, and D. J. Kuss, Eds., *The Oxford Handbook of Cyberpsychology*. Oxford, U.K.: Oxford Univ. Press, 2019.

[2] J. R. Ancis, "The age of cyberpsychology: An overview," *Technology, Mind, and Behavior*, vol. 1, no. 1, pp. 9–23, 2020, doi: 10.1037/tmb0000009.

[3] N. Männikkö, H. Ruotsalainen, A. Tolvanen, and M. Kääriäinen, "Problematic gaming is associated with some health-related behaviors among Finnish vocational school students," *International Journal of Mental Health and Addiction*, vol. 18, pp. 993–1007, 2020, doi: 10.1007/s11469-019-00100-6.

[4] V. Starcevic and E. Aboujaoude, "Cyberchondria, cyberbullying, cybersuicide, cybersex: 'New' psychopathologies for the 21st century?" *World Psychiatry*, vol. 14, no. 1, pp. 97–100, 2015, doi: 10.1002/wps.20195.

[5] J. D. Elhai, H. Yang, and C. Montag, "Fear of missing out (FOMO): Overview, theoretical underpinnings, and literature review on relations with severity of negative affectivity and problematic technology use," *Brazilian Journal of Psychiatry*, vol. 43, no. 2, pp. 203–209, 2021, doi: 10.1590/1516-4446-2020-0870.

[6] K. L. Norman, *Cyberpsychology: An Introduction to Human-Computer Interaction*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 2017.

[7] P. Elosua, F. J. Abad, and A. Hernández, "Support of the dimensionality and internal consistency of the Generalized Problematic Internet Use Scale-2: Systematic review and meta-analysis," *Cyberpsychology, Behavior, and Social Networking*, vol. 28, no. 1, pp. 7–17, 2025, doi: 10.1089/cyber.2024.0162.

[8] M. S. Islam et al., "Problematic smartphone and social media use among Bangladeshi college and university students amid COVID-19: The role of psychological well-being and pandemic related factors," *Frontiers in Psychiatry*, vol. 12, Art. no. 647386, 2021, doi: 10.3389/fpsyt.2021.647386.

[9] Y. Zhang, "Direct and indirect effects of neuroticism on Internet addiction in college students: A structure equation modeling analysis," *Psychological Reports*, vol. 124, no. 2, pp. 611–626, 2021, doi: 10.1177/0033294120918806.

[10] M. Sarstedt, J. F. Hair, Jr., C. Nitzl, C. M. Ringle, and M. C. Howard, "Beyond a tandem analysis of SEM and PROCESS: Use of PLS-SEM for mediation analyses!" *International Journal of Market Research*, vol. 62, no. 3, pp. 288–299, 2020, doi: 10.1177/1470785320915686.

[11] T. Yarkoni and J. Westfall, "Choosing prediction over explanation in psychology: Lessons from machine learning," *Perspectives on Psychological Science*, vol. 12, no. 6, pp. 1100–1122, 2017, doi: 10.1177/1745691617693393.

[12] L. Breiman, "Statistical modeling: The two cultures," *Statistical Science*, vol. 16, no. 3, pp. 199–231, 2001, doi: 10.1214/ss/1009213726.

[13] G. Shmueli, "To explain or to predict?" *Statistical Science*, vol. 25, no. 3, pp. 289–310, 2010, doi: 10.1214/10-STS330.

[14] G. Orrù, M. Monaro, C. Conversano, A. Gemignani, and G. Sartori, "Machine learning in psychometrics and psychological research," *Frontiers in Psychology*, vol. 10, Art. no. 2970, 2020, doi: 10.3389/fpsyg.2019.02970.

[15] M. A. Al-Garadi et al., "Predicting cyberbullying on social media in the Big Data era using machine learning algorithms: Review of literature and open challenges," *IEEE Access*, vol. 7, pp. 70701–70718, 2019, doi: 10.1109/ACCESS.2019.2918354.

[16] D. W. Joyce, A. Kormilitzin, K. A. Smith, and A. Cipriani, "Explainable artificial intelligence for mental health through transparency and interpretability for understandability," *npj Digital Medicine*, vol. 6, Art. no. 6, 2023, doi: 10.1038/s41746-023-00751-9.

[17] R. Lavelle-Hill, G. Smith, H. Deininger, and K. Murayama, "An explainable artificial intelligence handbook for psychologists: Methods, opportunities, and

challenges," *Psychological Methods*, early access, 2025, doi: 10.1037/met0000772.
[18] D. B. Dwyer, P. Falkai, and N. Koutsouleris, "Machine learning approaches for clinical psychology and psychiatry," *Annual Review of Clinical Psychology*, vol. 14, pp. 91–118, 2018, doi: 10.1146/annurev-clinpsy-032816-045037.
[19] J. Sukhera, "Narrative reviews: Flexible, rigorous, and practical," *Journal of Graduate Medical Education*, vol. 14, no. 4, pp. 414–417, 2022, doi: 10.4300/JGME-D-22-00480.1.
[20] T. Greenhalgh, S. Thorne, and K. Malterud, "Time to challenge the spurious hierarchy of systematic over narrative reviews?" *European Journal of Clinical Investigation*, vol. 48, no. 6, Art. no. e12931, 2018, doi: 10.1111/eci.12931.
[21] C. Baethge, S. Goldbeck-Wood, and S. Mertens, "SANRA—a scale for the quality assessment of narrative review articles," *Research Integrity and Peer Review*, vol. 4, Art. no. 5, 2019, doi: 10.1186/s41073-019-0064-8.
[22] R. Lavelle-Hill, G. Smith, and K. Murayama, "Bridging traditional-statistics and machine-learning approaches in psychology: Navigating small samples, measurement error, nonindependent observations, and missing data," *Advances in Methods and Practices in Psychological Science*, vol. 8, no. 3, pp. 1–31, 2025, doi: 10.1177/25152459251345696.
[23] P. C. McKee, C. J. Budnick, K. S. Walters, and I. Antonios, "College student Fear of Missing Out (FoMO) and maladaptive behavior: Traditional statistical modeling and predictive analysis using machine learning," *PLOS ONE*, vol. 17, no. 10, Art. no. e0274698, 2022, doi: 10.1371/journal.pone.0274698.
[24] Y. Gan et al., "Application of machine learning in predicting adolescent Internet behavioral addiction," *Frontiers in Psychiatry*, vol. 15, Art. no. 1521051, 2025, doi: 10.3389/fpsyt.2024.1521051.
[25] Y. Zhou et al., "Predictors of smartphone addiction in adolescents with depression: Combing the machine learning and moderated mediation model approach," *Behaviour Research and Therapy*, vol. 189, Art. no. 104749, 2025, doi: 10.1016/j.brat.2025.104749.
[26] E. M. Abu-Taieh et al., "Predictors of smartphone addiction and social isolation among Jordanian children and adolescents using SEM and ML," *Big Data and Cognitive Computing*, vol. 6, no. 3, Art. no. 92, 2022, doi: 10.3390/bdcc6030092.
[27] E. M. Abu-Taieh, I. AlHadid, R. Masa'deh, R. S. Alkhawaldeh, S. Khwaldeh, and A. Alrowwad, "Factors affecting the use of social networks and its effect on anxiety and depression among parents and their children: Predictors using ML, SEM and extended TAM," *International Journal of Environmental Research and Public Health*, vol. 19, no. 21, Art. no. 13764, 2022, doi: 10.3390/ijerph192113764.
[28] K. Ioannidis et al., "Problematic internet use (PIU): Associations with the impulsive-compulsive spectrum. An application of machine learning in psychiatry," *Journal of Psychiatric Research*, vol. 83, pp. 94–102, 2016, doi: 10.1016/j.jpsychires.2016.08.010.
[29] Z. Li et al., "Discovery of factors associated with smartphone addiction among high school adolescents: Using machine learning and network analysis," *Journal of Affective Disorders*, vol. 392, Art. no. 120096, 2026, doi: 10.1016/j.jad.2025.120096.
[30] K. Lee, W. Seo, and S. Y. Jung, "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*, vol. 27, no. 1, Art. no. 46201, 2026, doi: 10.31083/AP46201.
[31] J. Gui et al., "Interpretable machine learning for identifying co-occurring internet addiction and maladaptive cognitive emotion regulation in older adults: A cross-sectional study," *Frontiers in Public Health*, vol. 14, Art. no. 1901765, 2026, doi: 10.3389/fpubh.2026.1901765.
[32] J. Lee and W. Kim, "Prediction of problematic smartphone use: A machine learning approach," *International Journal of Environmental Research and Public Health*, vol. 18, no. 12, Art. no. 6458, 2021, doi: 10.3390/ijerph18126458.
[33] S. Namlı et al., "The relationship between smartphone and game addiction, leisure time management, and the enjoyment of physical activity: A comparison of regression analysis and machine learning models," *Healthcare*, vol. 13, no. 15, Art. no. 1805, 2025, doi: 10.3390/healthcare13151805.
[34] N. Çiftci and M. Yıldız, "The relationship between social media addiction, happiness, and life satisfaction in adults: Analysis with machine learning approach," *International Journal of Mental Health and Addiction*, vol. 21, no. 5, pp. 3500–3516, 2023, doi: 10.1007/s11469-023-01118-7.

[35] K. Xie and C. Li, "Ecological predictors and network patterns of cyberbullying victimisation among Chinese adolescents: Integrating machine learning and network analysis," *Child Abuse Review*, vol. 35, no. 4, Art. no. e70161, 2026, doi: 10.1002/car.70161.

[36] H. Zhou, Q. Zheng, H. Jiang, and J. Lu, "Examining stable and dynamic factors associated with daily problematic mobile phone use: Evidence from a 14-day diary study using machine learning and network analysis," *Addictive Behaviors*, vol. 182, Art. no. 108774, 2026, doi: 10.1016/j.addbeh.2026.108774.

[37] S. Liu, Y. Lin, W. Hu, Y. Meng, and X. Liu, "Mapping heterogeneous predictors of problematic smartphone use in Chinese adolescents: A machine learning approach with a nationally representative sample," *Journal of Affective Disorders*, vol. 410, Art. no. 122007, 2026, doi: 10.1016/j.jad.2026.122007.

[38] L. Zhan, J. Lu, T. Fan, Y. Wang, Y. Liu, and Y. Shen, "Identifying key risk factors of adolescent Internet gaming disorder using explainable machine learning and network analysis," *Journal of Affective Disorders*, vol. 415, Art. no. 122495, 2026, doi: 10.1016/j.jad.2026.122495.

[39] L. Liang, S. Liu, C. Zheng, K. Liu, J. Fei, and S. Mei, "An investigation into factors associated with adolescent internet gaming disorder: Utilizing machine learning and path analysis model," *Journal of Affective Disorders*, vol. 408, Art. no. 121862, 2026, doi: 10.1016/j.jad.2026.121862.

[40] B. Tie et al., "Smartphone and the brain: Stress and self-control mediate the association between the connectome-based predictive modeling of fMRI brain network and problematic smartphone use," *Computers in Human Behavior*, vol. 165, Art. no. 108531, 2025, doi: 10.1016/j.chb.2024.108531.

[41] G. Vera Cruz et al., "Smartphone apps for mental health and wellbeing: A usage survey and machine learning analysis of psychological and behavioral predictors," *Digital Health*, vol. 9, Art. no. 20552076231152164, 2023, doi: 10.1177/20552076231152164.

[42] H. Ç. Güzel et al., "Smartphone addiction and temporomandibular disorders among university students: A machine learning based multiple regression analysis study," *Medicine*, vol. 105, no. 10, Art. no. e47931, 2026, doi: 10.1097/MD.0000000000047931.

[43] B. Xiao, N. Parent, L. Rahal, and J. Shapka, "Using machine learning to explore the risk factors of problematic smartphone use among Canadian adolescents during COVID-19: The important role of fear of missing out (FoMO)," *Applied Sciences*, vol. 13, no. 8, Art. no. 4970, 2023, doi: 10.3390/app13084970.

[44] D. Marengo, C. Montag, A. Mignogna, and M. Settanni, "Mining digital traces of Facebook activity for the prediction of individual differences in tendencies toward social networks use disorder: A machine learning approach," *Frontiers in Psychology*, vol. 13, Art. no. 830120, 2022, doi: 10.3389/fpsyg.2022.830120.

[45] C. F. Giraldo-Jiménez, J. Gaviria-Chavarro, M. Sarria-Paja, L. A. Bermeo Varón, J. J. Villarejo-Mayor, and A. L. F. Rodacki, "Smartphones dependency risk analysis using machine-learning predictive models," *Scientific Reports*, vol. 12, Art. no. 22649, 2022, doi: 10.1038/s41598-022-26336-2.

[46] R. Jacobucci and K. J. Grimm, "Machine learning and psychological research: The unexplored effect of measurement," *Perspectives on Psychological Science*, vol. 15, no. 3, pp. 809–816, 2020, doi: 10.1177/1745691620902467.

[47] J. Hinds and A. N. Joinson, "Digital data and personality: A systematic review and meta-analysis of human perception and computer prediction," *Psychological Bulletin*, vol. 150, no. 6, pp. 727–766, 2024, doi: 10.1037/bul0000430.

[48] E. Christodoulou, J. Ma, G. S. Collins, E. W. Steyerberg, J. Y. Verbakel, and B. Van Calster, "A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models," *Journal of Clinical Epidemiology*, vol. 110, pp. 12–22, 2019, doi: 10.1016/j.jclinepi.2019.02.004.

[49] S. Kapoor and A. Narayanan, "Leakage and the reproducibility crisis in machine-learning-based science," *Patterns*, vol. 4, no. 9, Art. no. 100804, 2023, doi: 10.1016/j.patter.2023.100804.

[50] G. S. Collins et al., "TRIPOD+AI statement: Updated guidance for reporting clinical prediction models that use regression or machine learning methods," *BMJ*, vol. 385, Art. no. e078378, 2024, doi: 10.1136/bmj-2023-078378.