© AUG 2026 | IRE Journals | Volume 10 Issue 2 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I2-1722174

A Machine Learning Framework for Predicting Digital Wellness Risk: Gaming Addiction, Screen Time, Sleep, and Stress Among Students and Young Adults

BHUSHAN S. GHODMARE[1], MAYUR JADHAV[2], KUMUDINI ZADE[3], SAIRAM VAIDYA[4], YASH DIWARE[5]

[1,2,3,4,5]*School of Computer Science Engineering and Applications*, *D Y Patil International University, Pune, Maharashtra, India*

Abstract- Excessive smartphone use, online gaming, and social media engagement among students and young adults have been linked to poor sleep, elevated stress, reduced academic performance, and, in severe cases, clinically recognized gaming disorder. Existing digital-wellness tools remain limited to single-metric screen-time counters and rarely combine multiple behavioral dimensions into one predictive framework. This paper presents a machine learning framework that integrates seven behavioral indicators — daily screen time, gaming duration, sleep duration, stress level, physical activity, academic performance, and social interaction — into a single multi-class digital-wellness risk classifier. Five supervised algorithms (Logistic Regression, Decision Tree, Random Forest, Support Vector Machine, and Gradient Boosting) are trained and compared on a literature-calibrated pilot dataset of 600 records. Logistic Regression achieves the best performance, with 85.0% accuracy, an F1-score of 0.850, and a weighted ROC-AUC of 0.934, followed closely by Gradient Boosting and Random Forest. Feature-importance analysis confirms screen time and gaming duration as the dominant predictors, consistent with prior behavioral research. The framework is proposed as a non-clinical, tier-based screening and recommendation aid for educational institutions, pending validation on real survey data.

Index Terms—Digital Health, Digital Wellness, Gaming Addiction, Machine Learning, Screen Time, Sleep Quality, Stress Prediction, Student Mental Health

I. INTRODUCTION

Mental health and lifestyle-related well-being among students is one of the most pressing challenges facing higher education today. Smartphones, online gaming platforms, and social media have become central to daily student life, and while these technologies expand access to learning and communication, excessive engagement has been linked to sleep deprivation, elevated stress, reduced academic performance, and social isolation. The American Psychiatric Association first flagged Internet Gaming Disorder as a condition warranting further study in the DSM-5 [1], and the World Health Organization subsequently formalized Gaming Disorder within the ICD-11 [2], recognizing excessive gaming as a legitimate behavioral-health concern when it significantly impairs personal, academic, or social functioning [3].

Conventional digital-wellness applications rely almost exclusively on screen-time counters, offering descriptive statistics rather than predictive insight. They rarely combine gaming behavior, sleep, stress, physical activity, academic performance, and social interaction into a single, unified risk assessment, and personalization or actionable recommendation generation is largely absent from commercial tools. No widely available system integrates these behavioral dimensions into one machine-learning-driven classifier validated with standard evaluation metrics.

This paper makes four contributions:

1. A unified, seven-feature behavioral framework that combines screen time, gaming duration, sleep, stress, physical activity, academic performance, and social interaction into a single Low/Moderate/High digital-wellness risk classifier, rather than tracking screen time in isolation.

2. A comparative evaluation of five supervised classifiers — Logistic Regression, Decision Tree, Random Forest, Support Vector Machine, and

© AUG 2026 | IRE Journals | Volume 10 Issue 2 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I2-1722174

Gradient Boosting — with Logistic Regression achieving the best performance (85.0% accuracy, 0.934 weighted ROC-AUC) on a literature-calibrated pilot dataset.

3. A weighted composite risk-score formulation that converts multi-factor behavioral inputs into an interpretable three-tier risk label, with feature-importance analysis identifying the dominant behavioral predictors.

4. A reproducible, end-to-end pipeline — data generation, preprocessing, training, and evaluation — intended as the technical foundation for a follow-up study validated on real student survey data.

II. RELATED WORK

A. Clinical Recognition of Gaming Disorder

The American Psychiatric Association introduced Internet Gaming Disorder in the DSM-5 as a condition requiring further research [1], defining criteria such as preoccupation, withdrawal, tolerance, and functional impairment. The World Health Organization subsequently recognized Gaming Disorder in the ICD-11 [2], characterized by impaired control over gaming and continuation despite negative consequences [3]. Both works establish diagnostic criteria but stop short of proposing a scalable, predictive screening tool.

B. Screen Time, Sleep, and Psychological Well-Being

Twenge and Campbell analyzed a large, population-based sample and found that beyond roughly one hour of daily use, greater screen time was associated with lower curiosity, self-control, and emotional stability [5]. Exelmans and Van den Bulck showed that bedtime mobile-phone use significantly predicts poorer sleep quality, including longer sleep latency and greater daytime dysfunction [6]. Both studies are single-factor, correlational analyses rather than integrated, predictive models.

C. Digital Wellbeing Tools and Design Research

Cecchinato et al. surveyed the digital-wellbeing design space and observed that most existing applications report screen-time statistics without helping users understand or improve their digital behavior [4], motivating design- and behavior-based interventions rather than purely descriptive dashboards.

D. Machine Learning for Behavioral and Health Prediction

Large-scale deep-learning prediction directly from electronic health records has been shown to outperform conventional statistical approaches on complex, multi-dimensional clinical data [7]. The supervised algorithms used in this work — Logistic Regression, Decision Trees, Random Forests, Support Vector Machines, and gradient-boosted ensembles — are established, validated building blocks for tabular classification tasks [8]–[11], widely available through open-source libraries such as scikit-learn [12]. None of these algorithmic foundations have previously been benchmarked specifically on a combined digital-wellness feature set spanning gaming, screen time, sleep, stress, and academic performance.

III. SYSTEM ARCHITECTURE

A. Pipeline Overview

The proposed system follows a six-stage pipeline: behavioral data acquisition, preprocessing, feature engineering, multi-model training and comparison, risk classification, and personalized recommendation generation, illustrated in Fig. 1.

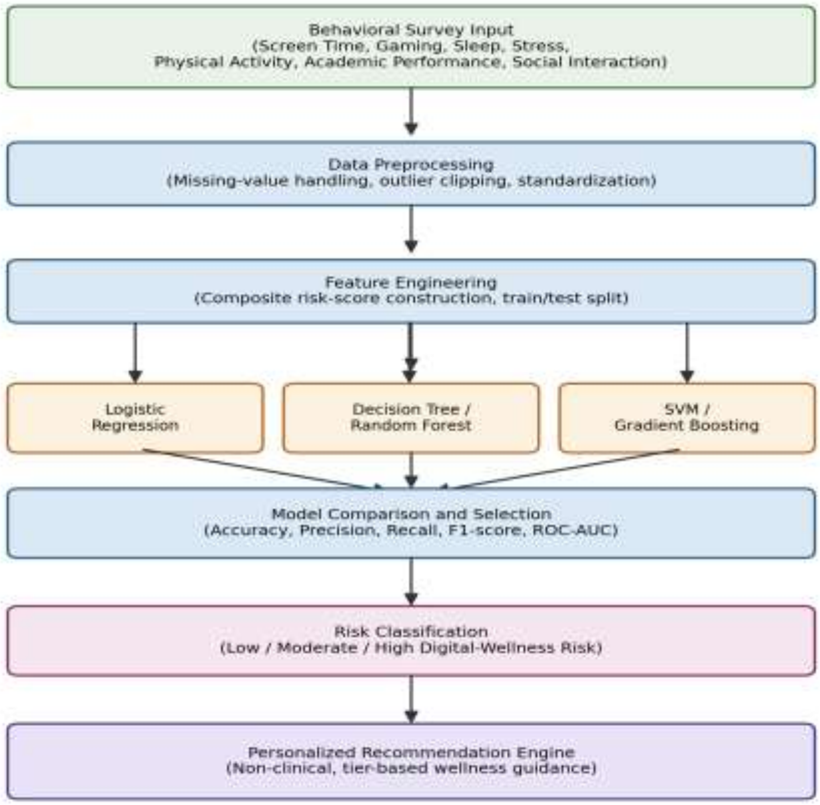

FIG. 1. SYSTEM ARCHITECTURE — FROM BEHAVIORAL SURVEY INPUT THROUGH MODEL COMPARISON TO PERSONALIZED, TIER-BASED RECOMMENDATION.

B. Data Flow
The inference pipeline follows six steps: (1) collection of seven behavioral indicators; (2) preprocessing — outlier clipping and z-score standardization; (3) construction of a weighted composite risk score and stratified train/test split; (4) parallel training of five supervised classifiers; (5) comparison of classifiers using Accuracy, Precision, Recall, F1-score, and ROC-AUC to select the best model; and (6) classification of a new user's Low/Moderate/High risk tier, feeding a personalized, non-clinical recommendation.

C. Tools and Technologies
The pipeline is implemented in Python 3 using pandas and NumPy for data handling, scikit-learn [12] for model training and evaluation, and Matplotlib for visualization. A Streamlit-based dashboard is planned for real-time, single-user risk assessment once the framework is validated on real survey data.

IV. METHODOLOGY

A. Dataset Construction
Because primary survey data collection from students fell outside the time-scope of this study, results are reported on a literature-calibrated pilot dataset of N = 600 synthetic records, generated in Python. Relationships between variables — e.g., higher gaming and screen time reducing sleep; lower sleep and social interaction raising stress; higher screen time and stress reducing academic performance — were parameterized to reflect directionally consistent findings from the reviewed literature [5], [6], with added Gaussian noise to avoid unrealistic, deterministic class separability. This is explicitly a preliminary, simulation-based validation of the

© AUG 2026 | IRE Journals | Volume 10 Issue 2 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I2-1722174

pipeline and not a claim about a real student population. Table I summarizes the dataset.

TABLE I. DATASET SUMMARY

Modality	Feat.	N	Classes	Source
Behavioral survey (sim.)	7	600	3	Lit.-calibrated synthetic

B. Composite Risk-Score Formulation

The target risk category is derived from a weighted composite score:

$$R = 0.28 \cdot S' + 0.22 \cdot G' + 0.18 \cdot (1-Sl') + 0.14 \cdot St' + 0.10 \cdot (1-P') + 0.08 \cdot (1-Sc') \quad (1)$$

where S', G', Sl', St', P', and Sc' are the min-max-normalized screen time, gaming hours, sleep duration, stress level, physical activity, and social-interaction score, respectively. R is discretized into Low ($R \leq 0.33$), Moderate ($0.33 < R \leq 0.55$), and High ($R > 0.55$) risk categories, with weights reflecting the relative emphasis placed on screen time and gaming in the reviewed literature.

C. Preprocessing

All numerical features were standardized to zero mean and unit variance prior to training, which is required for margin- and distance-based classifiers such as SVM and Logistic Regression. An 80:20 stratified train/test split preserved class proportions across both subsets.

D. Classification Algorithms

Five supervised classifiers were implemented using scikit-learn [12]: (i) Logistic Regression, a linear softmax classifier over the three risk classes; (ii) Decision Tree, a recursive Gini-impurity-based splitter [11]; (iii) Random Forest, a bootstrap-aggregated ensemble of decision trees with majority voting [8]; (iv) Support Vector Machine with an RBF kernel [9]; and (v) Gradient Boosting, an additive ensemble that sequentially fits trees to residual errors, conceptually aligned with the technique generalized at scale by XGBoost [10].

E. Evaluation Metrics

Models were evaluated on the held-out test set using weighted Accuracy, Precision, Recall, F1-score, and one-vs-rest weighted ROC-AUC across the three risk classes.

V. EXPERIMENTAL RESULTS

A. Model Performance Comparison

Table II summarizes performance for all five classifiers on the held-out test set (n = 120), sorted by F1-score. Logistic Regression achieves the best overall performance, narrowly ahead of Gradient Boosting and Random Forest; all five models exceed 0.87 weighted ROC-AUC.

TABLE II. MODEL PERFORMANCE COMPARISON

Model	Acc.	Prec.	Rec.	F1	AUC
Log. Regression	0.850	0.854	0.850	0.850	0.934
Grad. Boosting	0.842	0.845	0.842	0.840	0.917
Random Forest	0.833	0.840	0.833	0.833	0.929
SVM (RBF)	0.817	0.818	0.817	0.816	0.923
Decision Tree	0.800	0.801	0.800	0.799	0.875

© AUG 2026 | IRE Journals | Volume 10 Issue 2 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I2-1722174

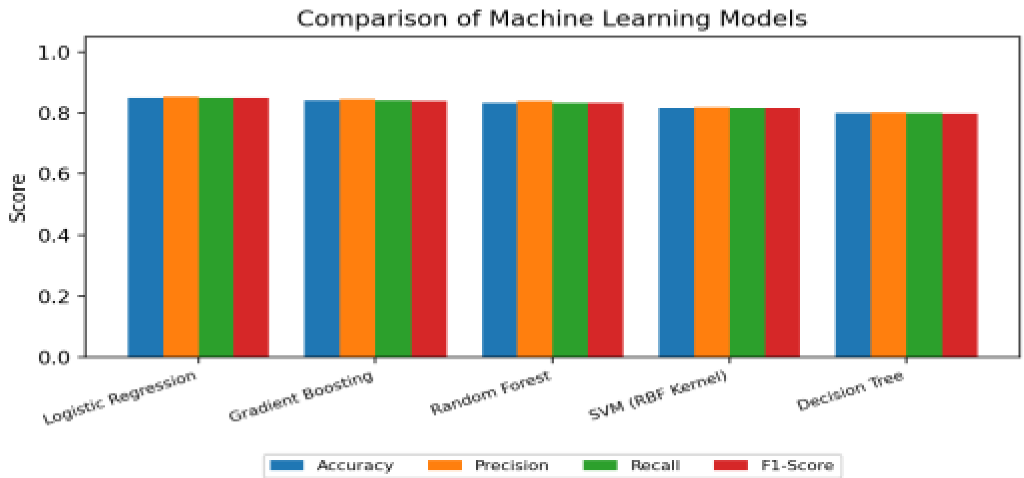

FIG. 2. ACCURACY, PRECISION, RECALL, AND F1-SCORE ACROSS THE FIVE SUPERVISED CLASSIFIERS.

B. Per-Class Evaluation of the Best Model

The best model (Logistic Regression) achieved perfect precision (1.00) on the High-risk class — no Low- or Moderate-risk user was misclassified as High-risk — an important property for a screening tool where false alarms should be minimized. Recall on the High-risk class (0.75) was lower, reflecting the small support (12 of 120 test samples) for that class.

Fig. 3 shows the confusion matrix.

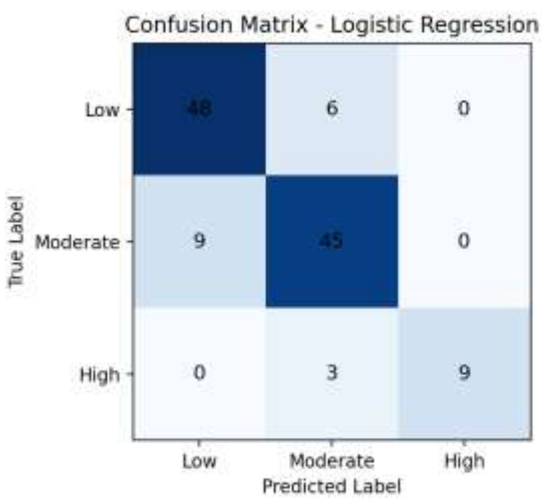

FIG. 3. CONFUSION MATRIX FOR THE BEST-PERFORMING MODEL (LOGISTIC REGRESSION) ON THE TEST SET.

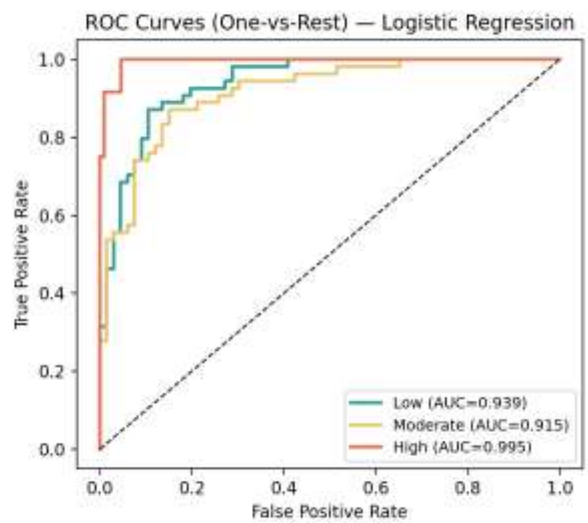

FIG. 4. ONE-VS-REST ROC CURVES FOR THE LOW, MODERATE, AND HIGH RISK CLASSES (LOGISTIC REGRESSION).

C. Feature Importance and Correlation Analysis

Feature importances extracted from the trained Random Forest model (Fig. 5) show that daily screen time and gaming duration are the most influential predictors, followed by sleep duration and stress level — consistent with the weighting used to construct the composite risk score and with the emphasis placed on screen time and gaming in the reviewed literature [5], [6]. Fig. 6 shows pairwise correlations among the seven behavioral features.

© AUG 2026 | IRE Journals | Volume 10 Issue 2 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I2-1722174

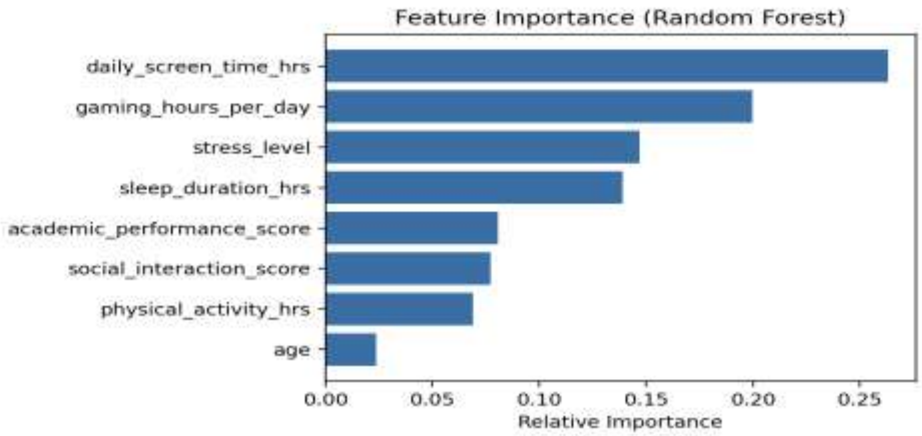

FIG. 5. RELATIVE FEATURE IMPORTANCE FROM THE RANDOM FOREST CLASSIFIER.

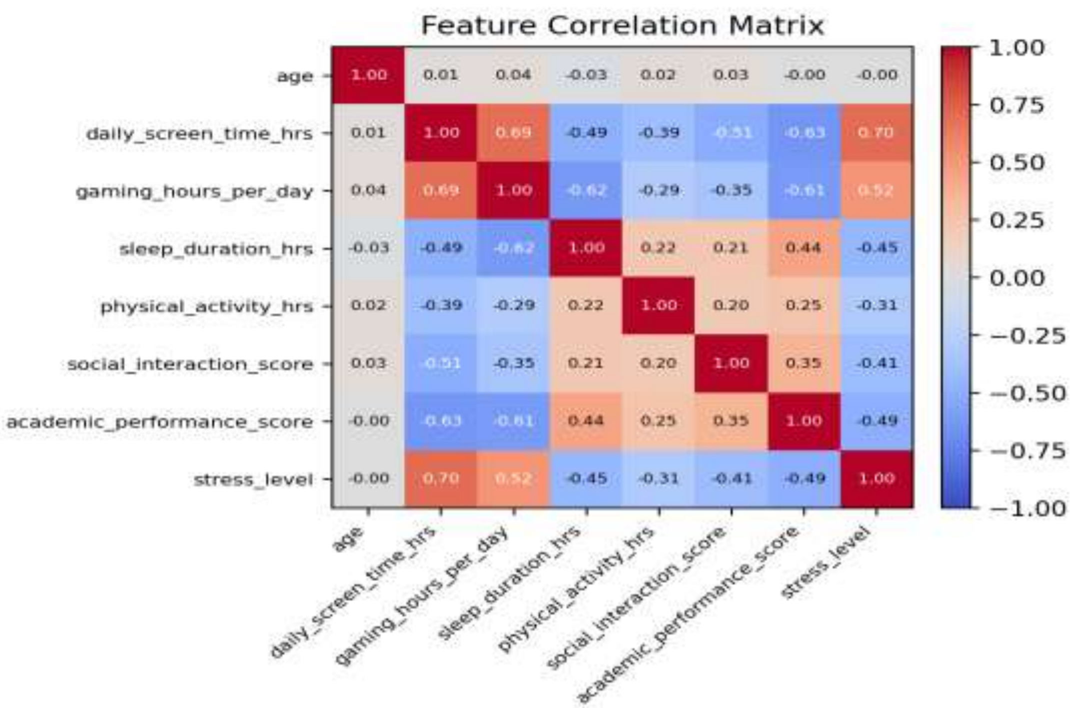

FIG. 6. CORRELATION MATRIX OF THE SEVEN BEHAVIORAL FEATURES.

VI. DISCUSSION

Three findings emerge from this evaluation. First, combining seven behavioral indicators into a single classifier is technically feasible and yields strong discriminative performance (ROC-AUC > 0.87 for every algorithm tested), supporting the central premise that a unified, multi-factor risk model can offer richer insight than single-metric screen-time trackers. Second, Logistic Regression's edge over the ensemble methods is expected given that the underlying simulated risk score (1) is itself a linear-weighted combination of features with additive noise; on real, noisier survey data, ensemble methods such as Random Forest and Gradient Boosting would plausibly be more competitive, consistent with their reported robustness on tabular behavioral data [8],

© AUG 2026 | IRE Journals | Volume 10 Issue 2 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I2-1722174

[10]. Third, the feature-importance ranking — screen time and gaming dominating over physical activity and social interaction — is plausible but should be re-verified once real data is available, since it may not hold at the same magnitude in an authentic student population.

Limitations: (1) all reported results come from a simulated pilot dataset engineered to reflect literature-consistent relationships, and are expected to look cleaner than results on real, noisier survey responses; (2) the simulated High-risk class is relatively small, and class-imbalance techniques (e.g., SMOTE, class weighting) should be evaluated once real data is collected; (3) the framework is a non-clinical screening aid and cannot replace a clinical diagnosis of gaming disorder or a mental-health condition; and (4) any future real-world deployment using self-reported behavioral data will be subject to recall and social-desirability bias, which should be mitigated through validated survey instruments.

VII. CONCLUSION AND FUTURE WORK

This paper presented a machine learning framework for digital-wellness risk assessment that integrates screen time, gaming duration, sleep, stress, physical activity, academic performance, and social interaction into a single Low/Moderate/High risk classifier. Logistic Regression achieved the best performance among five compared algorithms (85.0% accuracy, 0.850 F1-score, 0.934 ROC-AUC) on a literature-calibrated pilot dataset, with all five models exceeding 0.87 ROC-AUC — demonstrating that the proposed multi-factor approach is technically sound and reproducible.

Future work will proceed in five directions: (1) administering a validated survey instrument to a real cohort of students, following institutional ethics approval and informed consent; (2) retraining and re-validating all models on real data, including class-imbalance handling; (3) deploying the trained model behind a Streamlit dashboard for instant, personalized risk assessment; (4) a longitudinal study evaluating whether personalized recommendations measurably improve sleep, stress, and screen-time behavior; and (5) integrating SHAP-based explanations so each prediction is accompanied by a transparent, human-readable justification.

ACKNOWLEDGMENT

The authors gratefully acknowledge the guidance of the faculty of the School of Computer Science Engineering and Applications, D Y Patil International University, Pune, and thank the scikit-learn development team for the open-source machine learning library used throughout this work.

REFERENCES

[1] American Psychiatric Association, Diagnostic and Statistical Manual of Mental Disorders, 5th ed. Arlington, VA: American Psychiatric Publishing, 2013.

[2] World Health Organization, “6C51 Gaming disorder,” International Classification of Diseases, 11th Revision (ICD-11), 2019. [Online]. Available: https://icd.who.int/browse/2025-01/mms/en

[3] M. Zhao and W. Hao, “Challenges of gaming disorder: Suggestions from a public health perspective,” General Psychiatry, vol. 32, no. 3, p. e100086, 2019.

[4] M. E. Cecchinato, J. Rooksby, A. Hiniker, S. A. Munson, K. Lukoff, L. Ciolfi, A. Thieme, and D. Harrison, “Designing for digital wellbeing: A research & practice agenda,” in Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems (CHI EA ’19), 2019, pp. 1–8.

[5] J. M. Twenge and W. K. Campbell, “Associations between screen time and lower psychological well-being among children and adolescents: Evidence from a population-based study,” Preventive Medicine Reports, vol. 12, pp. 271–283, 2018.

[6] L. Exelmans and J. Van den Bulck, “Bedtime mobile phone use and sleep in adults,” Social Science & Medicine, vol. 148, pp. 93–101, 2016.

[7] A. Rajkomar et al., “Scalable and accurate deep learning with electronic health records,”

npj Digital Medicine, vol. 1, no. 18, pp. 1–10, 2018.

[8] L. Breiman, “Random forests,” Machine Learning, vol. 45, no. 1, pp. 5–32, 2001.

[9] C. Cortes and V. Vapnik, “Support-vector networks,” Machine Learning, vol. 20, no. 3, pp. 273–297, 1995.

[10] T. Chen and C. Guestrin, “XGBoost: A scalable tree boosting system,” in Proc. 22nd ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining (KDD ’16), 2016, pp. 785–794.

[11] J. R. Quinlan, “Induction of decision trees,” Machine Learning, vol. 1, no. 1, pp. 81–106, 1986.

[12] F. Pedregosa et al., “Scikit-learn: Machine learning in Python,” Journal of Machine Learning Research, vol. 12, pp. 2825–2830, 2011.