Jurnal Minfo Polgan (JMP)
Volume 15, Nomor 1, Maret 2026
DOI : https://doi.org/10.33395/jmp.v15i1.16178

e-ISSN : 2797-3298
p-ISSN : 2089-9424

# Predictive Modeling of Smartphone Addiction: Performance Evaluation of KNN, XGBoost, and Naive Bayes on Behavioral Patterns

**[1*]M. Rhifky Wayahdi, [2]Fahmi Ruziq, [3]Auliana Nasution, [4]Ellanda Purwawijaya**
[1*,2]Program Studi Sistem Informasi, Fakultas Teknologi, Universitas Battuta, Medan, Indonesia
[3,4]Program Studi Informatika, Fakultas Teknologi, Universitas Battuta, Medan, Indonesia
**Korespondensi*: muhammadrhifkywayahdi@gmail.com

**Submitted:**Feb 7, 2026 | **Accepted:** Mar 9, 2026 | **Published:** Mar 22, 2026

***ABSTRACT***

*Excessive smartphone use has triggered a global crisis in the form of smartphone addiction, which negatively impacts mental health and productivity. Most current detection methods still rely on subjective questionnaires that are prone to bias. Therefore, this study aims to evaluate and compare the performance of machine learning-based predictive models—namely K-Nearest Neighbors (KNN), Naive Bayes, and Extreme Gradient Boosting (XGBoost)—in objectively classifying addiction levels based on user behavioral patterns. The research methodology adopts a standard machine learning workflow encompassing data preprocessing, model training, and performance evaluation using a dataset of 3,300 user activity log entries. Empirical results demonstrate that XGBoost yields superior classification performance, achieving an accuracy of 97.27% and an F1-Score of 96.70%, significantly outperforming the KNN (94.54%) and Naive Bayes (89.09%) algorithms. Further feature importance analysis confirms that App Usage Time is the most crucial predictor in detecting addiction. This study concludes that the XGBoost architecture is highly robust in handling non-linear behavioral feature interactions, enabling high-precision predictions. The implications of these findings provide a solid technical foundation for the development of automated early detection systems. Future research should consider expanding the dataset to include application categorization and integrating XGBoost modeling into real-time digital wellbeing application prototypes.*

***Keywords****: Behavior Patterns*; *K-Nearest Neighbors*; *Machine Learning*; *Naive Bayes*; *Smartphone Addiction*; *XGBoost.*

**ABSTRAK**

Penggunaan *smartphone* yang berlebihan telah memicu krisis global berupa kecanduan (*smartphone addiction*) yang berdampak negatif pada kesehatan mental dan produktivitas. Mayoritas metode deteksi saat ini masih bergantung pada kuesioner subjektif yang rentan terhadap bias. Oleh karena itu, penelitian ini bertujuan untuk mengevaluasi dan membandingkan performa model prediktif berbasis *machine learning*, yaitu *K-Nearest Neighbors* (KNN), *Naive Bayes*, dan *Extreme Gradient Boosting* (XGBoost), dalam mengklasifikasikan tingkat kecanduan secara objektif berdasarkan pola perilaku pengguna (*behavioral patterns*). Metode penelitian mengadopsi alur kerja *machine learning* standar yang mencakup pra-pemrosesan data, pelatihan model, dan evaluasi performa menggunakan dataset berukuran 3.300 entri log aktivitas pengguna. Hasil pengujian secara empiris membuktikan bahwa XGBoost menghasilkan performa klasifikasi paling superior dengan tingkat akurasi mencapai 97,27% dan *F1-Score* 96,70%, secara signifikan mengungguli algoritma KNN (94,54%) dan Naive Bayes (89,09%). Analisis *feature importance* lebih lanjut mengonfirmasi bahwa intensitas waktu penggunaan aplikasi (*App Usage Time*) merupakan prediktor absolut yang paling krusial dalam mendeteksi kecanduan. Kesimpulan penelitian ini menegaskan bahwa arsitektur XGBoost sangat tangguh dalam menangani interaksi fitur perilaku non-linear, sehingga mampu menghasilkan prediksi yang sangat presisi. Implikasi dari temuan ini memberikan landasan teknis yang kuat bagi pengembangan sistem deteksi dini terotomatisasi. Untuk penelitian selanjutnya, direkomendasikan perluasan dataset dengan mempertimbangkan kategorisasi jenis aplikasi serta integrasi pemodelan XGBoost ke dalam purwarupa aplikasi *digital wellbeing* secara *real-time*.

**Kata Kunci**: *K-Nearest Neighbors*; Kecanduan *Smartphone*; *Machine Learning*; *Naive Bayes*; Pola Perilaku; *XGBoost*.

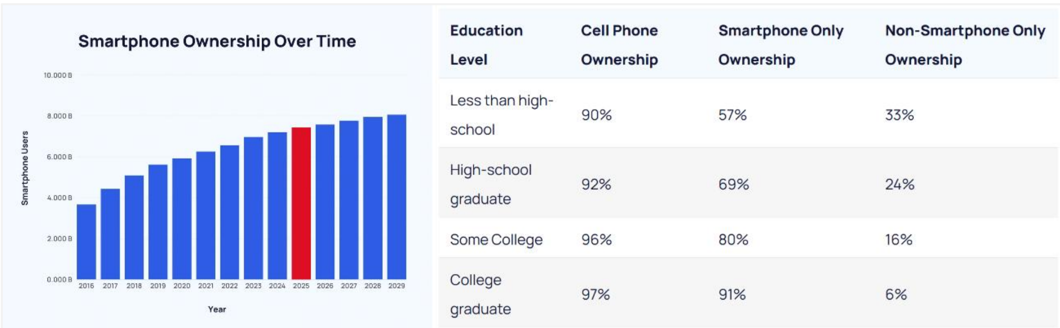
## INTRODUCTION

The presence of smartphones has fundamentally altered patterns of human interaction (Osorio-Molina et al., 2021; Wayahdi & Ruziq, 2025). Conversely, this trend has triggered a global crisis in the form of nomophobia or smartphone addiction (Osorio-Molina et al., 2021; Ratan et al., 2021; Sunday et al., 2021). Numerous studies have demonstrated that excessive device usage exerts destructive impacts across various age demographics, ranging from a deterioration in quality of life (Okur et al., 2025; Sela et al., 2022) and the onset of psychological issues such as anxiety and depression (Asare et al., 2021; Nikolic et al., 2023; Osorio-Molina et al., 2021; Ratan et al., 2021; Sarhan, 2024) to diminished academic and professional productivity (Abbasi et al., 2021; Nikolic et al., 2023; Osorio-Molina et al., 2021; Rathakrishnan et al., 2021; Sunday et al., 2021) and disrupted sleep patterns (Nikolic et al., 2023; Ratan et al., 2021; Rathakrishnan et al., 2021; Sohn et al., 2021). Furthermore, this condition is correlated with neurological and musculoskeletal disorders (Ratan et al., 2021; Schmitgen et al., 2022; Singh et al., 2025), a decline in social skills (Ahmed et al., 2025; Osorio-Molina et al., 2021; Sharma et al., 2023; Yogesh et al., 2024), and increased feelings of loneliness and aggressive behavior (Doo & Kim, 2022; Hu et al., 2024; Yilmaz et al., 2023). The use of gaming and social media as maladaptive coping mechanisms further exacerbates this phenomenon within the broader community (Abbasi et al., 2021; Wayahdi & Ruziq, 2025).


| Rank | Country | Smartphone Users | Total Population | Smartphone Penetration |
|---|---|---|---|---|
| 1 | China | 974.69 million | 1.43 billion | 68.4% |
| 2 | India | 659 million | 1.42 billion | 46.5% |
| 3 | United States | 276.14 million | 338.29 million | 81.6% |
| 4 | Indonesia | 187.7 million | 275.5 million | 68.1% |
| 5 | Brazil | 143.43 million | 215.31 million | 66.6% |
| 6 | Russia | 106.44 million | 144.71 million | 73.6% |
| 7 | Japan | 97.44 million | 123.95 million | 78.6% |
| 8 | Nigeria | 83.34 million | 218.54 million | 38.1% |
| 9 | Mexico | 78.37 million | 127.5 million | 61.5% |
| 10 | Pakistan | 72.99 million | 235.82 million | 31% |

Figure 1. Smartphone Users in the Top 10 Countries

Digital penetration in Indonesia is characterized as massive, ranking fourth globally with a total of 187.7 million users. Increasing accessibility has rendered these devices central, indispensable instruments in the daily activities of the population.




| Education Level | Cell Phone Ownership | Smartphone Only Ownership | Non-Smartphone Only Ownership |
|---|---|---|---|
| Less than high-school | 90% | 57% | 33% |
| High-school graduate | 92% | 69% | 24% |
| Some College | 96% | 80% | 16% |
| College graduate | 97% | 91% | 6% |

Figure 2. Smartphone Ownership Trends

Global trends indicate consistent escalation, with projections exceeding 7 billion smartphone users by 2025. The exponential increase in screen time intensity across all demographics underscores the research urgency regarding the identification of risk factors to prevent the escalation of smartphone addiction.

Despite extensive research on the impacts of smartphone addiction, the methodologies used to measure addiction levels still face several limitations. The majority of studies rely on psychometric questionnaire instruments, which are susceptible to respondent subjectivity and recall bias (James et al., 2023). Meanwhile, studies based on application usage logs often encounter generalizability issues (Li et al., 2022) due to a focus on highly specific samples or evaluation within a narrow scale. Reported addiction prevalence rates in the general public vary significantly and continue to escalate globally (Freitas et al., 2024; Mokhtarinia et al., 2022; Olson et al., 2022), necessitating a more objective and precise early detection system.

As a solution to the limitations of self-reporting observational methods, the utilization of machine learning based on behavioral pattern analysis offers a far more measurable approach (Joseph & Maheswari, 2025; Kim et al., 2024). Through predictive modeling, hidden patterns from multidimensional behavioral data—such as total daily usage duration, screen unlock frequency, specific duration on entertainment/social applications, and device activity during sleep hours—can be extracted automatically (Duan et al., 2021; Wayahdi & Ruziq, 2025). Transforming actual behavioral data into computational features enables systems to accurately classify addiction risk levels without depending on user honesty or perception.

The accuracy of such predictive modeling depends heavily on the selection of classification algorithms capable of handling the heterogeneity and complexity of behavioral log data. This research comprehensively evaluates and compares three algorithms with distinct computational characteristics: K-Nearest Neighbors (KNN), XGBoost, and Naive Bayes. KNN represents instance-based learning, which operates optimally in detecting the proximity of subjects with similar daily behavioral patterns (Ramadhani & Wayahdi, 2024; Ruziq & Wayahdi, 2025; Wayahdi & Ruziq, 2025). Meanwhile, Naive Bayes offers computational time efficiency through a probabilistic approach based on Bayes' Theorem, proving reliable in processing the independent probabilities of addiction-triggering features (Kumar & Michael, 2026; Putri et al., 2026). On the other hand, Extreme Gradient Boosting (XGBoost) represents an advanced ensemble learning architecture that has been proven robust in minimizing overfitting and handling non-linear behavioral data patterns and class imbalances within datasets (Duan et al., 2021; Joseph & Maheswari, 2025; Wayahdi & Ruziq, 2022).

To date, the performance comparison of these three algorithms, specifically utilizing behavioral pattern feature extraction in general user populations, requires extensive exploration to identify the optimal model. Therefore, this study aims to evaluate the performance of KNN, XGBoost, and Naive Bayes predictive models in classifying smartphone addiction. The results of this comparison are expected to yield a classification model with superior performance metrics, which can subsequently be implemented as an Artificial Intelligence foundation for automated early detection systems to support public mental health in the digital era.

## RESEARCH METHOD

This study adopts a standard machine learning workflow to evaluate and compare the performance of smartphone addiction level classification models. The research framework consists of five primary stages: Data Acquisition, Data Preprocessing, Modeling Scenario, Performance Evaluation, and Result Interpretation.

### Data Acquisition

The initial phase focuses on collecting a representative dataset containing user-device interaction history. The analyzed dataset comprises 3,300 user behavior entries. This dataset includes 11 multidimensional features, covering both demographic attributes and technical operational metrics: User ID, Device Model, Operating System, App Usage Time (min/day), Screen on Time (hours/day), Battery Drain (mAh/day), Number of Apps Installed, Data Usage (MB/day), Age, Gender, and the target variable, User Behavior Class, which is categorized on a scale of 1 to 5 (representing addiction severity).

### Data Preprocessing

Raw data is processed systematically to ensure quality and compatibility with the three algorithms under evaluation. The preprocessing sequence includes:

a. Cleaning and Transformation: Handling missing values and transforming categorical variables (such as Device Model and Operating System) into numerical representations using one-hot encoding.

b. Feature Selection: To measure linear relationships between variables and eliminate redundant features, Pearson Correlation Coefficient ($r$) analysis is applied. The Pearson correlation formula is defined in Equation (1).

$$r_{xy} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}} \tag{1}$$

c. Data Normalization and Standardization: Algorithms sensitive to scale, such as KNN, require feature range alignment. Metric features are normalized using Min-Max Normalization (range 0 to 1) as shown in Equation (2), or Z-Score Standardization (distribution with a mean of 0 and standard deviation of 1) as shown in Equation (3).

$$x'_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{2}$$

$$z = \frac{x - \mu}{\sigma} \tag{3}$$

Once the data is standardized and cleaned, the dataset is partitioned into two independent sets: 70–80% is allocated as the training set to build the models, and 20–30% serves as the testing set to evaluate the models' generalization capability.

**Modeling Scenario**

This stage forms the core of the comparison, where three classification algorithms with distinct computational approaches are trained using the same training set:

a. K-Nearest Neighbors (KNN): This model is configured using an instance-based learning technique. Hyperparameter optimization focuses on determining the optimal value for the nearest neighbors ($k$). The distance calculation between behavioral pattern data points is executed using the Euclidean distance metric, as formulated in Equation (4).

$$d(p\ q) = \sqrt{\sum_{i=1}^{n}(p_i - q_i)^2} \tag{4}$$

b. Naive Bayes: This algorithm models classification based on the probabilistic principles of Bayes' Theorem. It predicts the probability of class $c$ (addiction level) based on input feature $x$ (behavioral patterns) under the "naive" assumption that each feature is independent of the others. The posterior probability formula for Naive Bayes classification is shown in Equation (5).

$$P(c|x) = \frac{P(x|c)P(c)}{P(x)} \tag{5}$$

Where $P(c|x)$ is the posterior probability of the predicted class, $P(c)$ is the class prior probability, and $P(x|c)$ is the likelihood of the input features given that class.

c. Extreme Gradient Boosting (XGBoost): Representing a decision-tree-based ensemble learning algorithm, XGBoost is trained additively to handle complex non-linear interactions. XGBoost minimizes an objective function ($\mathcal{L}$) that combines a prediction loss function with a regularization penalty term to prevent overfitting, as shown in Equation (6).

$$\mathcal{L}^{(t)} = \sum_{i=1}^{n} l\left(y_i\ \widehat{y_i^{(t)}}\right) + \sum_{k=1}^{t} \Omega(f_k) \tag{6}$$

Where $l$ represents the classification loss function (the difference between prediction $\widehat{y_i}$ and actual target $y_i$), while $\Omega(f_k)$ is the regularization penalty that controls tree complexity.

**Performance Evaluation**

To measure the accuracy and robustness of each model's predictions, testing is validated using cross-validation techniques. Prediction error analysis is performed using a Confusion Matrix to map True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN). Quantitative evaluation is calculated using standard performance metrics:

a. Accuracy: Formulated in Equation (7) to measure the overall ratio of correct predictions.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{7}$$

b. Precision: Formulated in Equation (8) to measure the proportion of accurate positive predictions.

$$Precision = \frac{TP}{TP + FP} \tag{8}$$

c. Recall: Formulated in Equation (9) to measure the proportion of actual positive cases successfully detected.

$$Recall = \frac{TP}{TP + FN} \tag{9}$$

d. F1-Score: Formulated in Equation (10) as the harmonic mean of Precision and Recall.

$$F1\text{-}Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{10}$$

**Result Interpretation**

The final stage involves an extensive comparative analysis of the three algorithms. The model with the highest combination of performance metrics—specifically the average F1-Score to account for potential imbalanced data—is designated as the superior algorithm. Beyond measuring technical accuracy, this process extracts feature importance (e.g., using Permutation Feature Importance for KNN or the Gain metric for XGBoost) to identify the specific behavioral patterns most significant in triggering a user's high addiction tendency.

## RESULTS AND DISCUSSION

**Research Results**

a. Exploratory Data Analysis (EDA)

Based on the initial extraction of 3,300 user behavior data points, demographic exploration reveals a relatively uniform distribution across various age groups. As visualized in Figure 3, the age range of users spans from teenagers to those over 60 years old, with a mean age of 40.4 years.

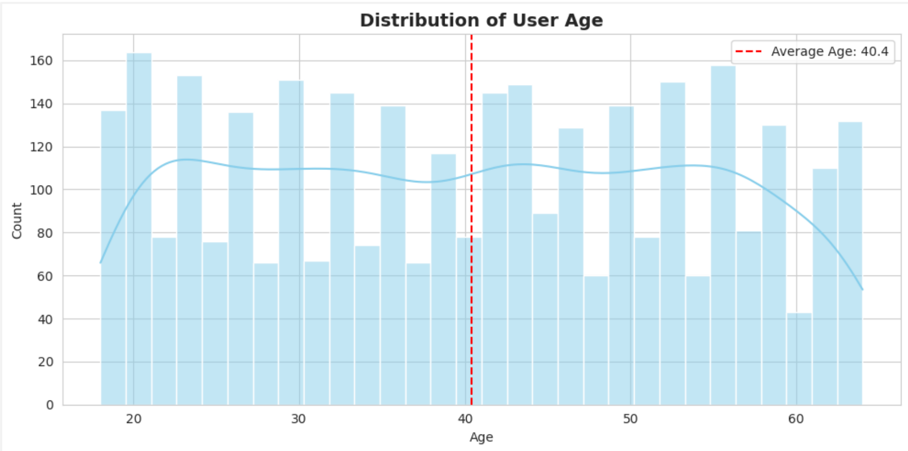



Figure 3. Age Distribution of Smartphone Users

Furthermore, a Pearson correlation analysis between operational and demographic features was conducted to map the dependencies of variables to be trained in the models (Figure 4).

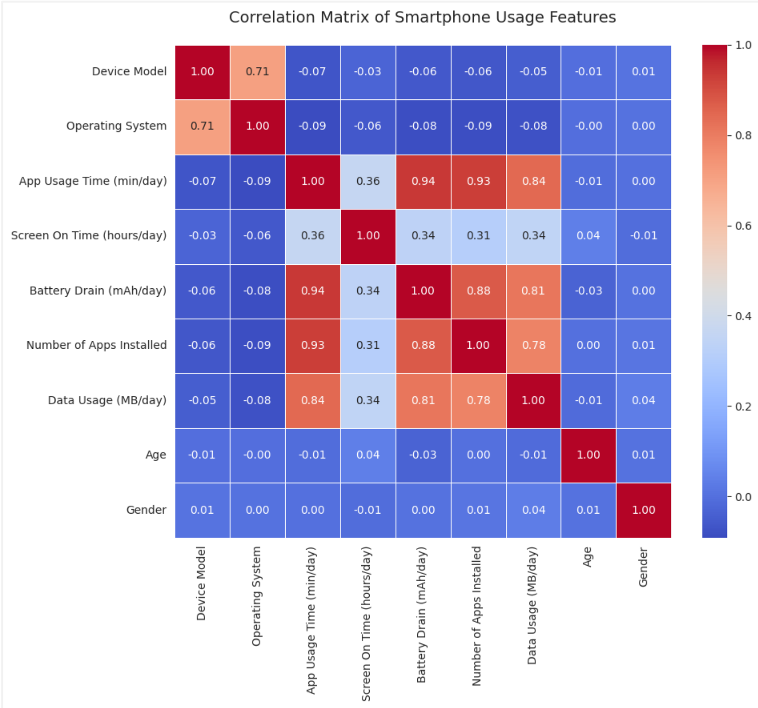



Figure 4. Correlation Matrix of User Behavior Features

The correlation matrix indicates a very strong positive linear relationship between App Usage Time and Battery Drain, with a correlation value of 0.94, as well as with the Number of Apps Installed at 0.93. A high correlation was also found between battery consumption and internet data usage (0.81). Conversely, demographic factors such as Age and Gender show correlations nearly touching zero across all usage features, proving that device intensity is not dictated by inherent demographic profiles. This high correlation between operational features serves as a critical foundation for evaluating how each classification algorithm handles behavioral variable dependencies.

b. Model Performance Evaluation and Comparison

The three algorithms were trained and tested using a data partition ratio of 70:30. Performance evaluation was conducted using macro-average metrics to accommodate multi-class classification (addiction scales 1 to 5). The performance metric comparisons for the three models are presented in Table 1.

Table 1. Comparison of Classification Algorithm Performance Metrics

| Algorithm | Accuracy | Precision | Recall | F1-Score |
|---|---|---|---|---|
| **K-Nearest Neighbors (KNN)** | 0,9454 | 0,9371 | 0,9407 | 0,9387 |
| **Naive Bayes** | 0,8909 | 0,8727 | 0,9147 | 0,8899 |
| **XGBoost** | 0,9727 | 0,9683 | 0,9663 | 0,9670 |

Based on Table 1 and the F1-Score comparison visualized in Figure 3, the XGBoost algorithm demonstrates highly superior classification performance, achieving the highest accuracy of 97.27% and an F1-Score of 96.70%. In second place, the KNN model shows robust performance with an accuracy of 94.54%. Conversely, the Naive Bayes algorithm yielded the lowest performance among the three, with an accuracy of 89.09%.

Table 2. Detailed Performance Evaluation Metrics per Class (Classification Report)

| Algorithm | Class (Addiction Level) | Precision | Recall | F1-Score | Support |
|---|---|---|---|---|---|
| **K-Nearest** | 1 | 0.92 | 0.96 | 0.94 | 106 |
| **Neighbors** | 2 | 0.93 | 0.90 | 0.92 | 175 |

| Algorithm | Class (Addiction Level) | Precision | Recall | F1-Score | Support |
|---|---|---|---|---|---|
| | 3 | 0.96 | 0.97 | 0.96 | 463 |
| | 4 | 0.92 | 0.90 | 0.91 | 143 |
| | 5 | 0.95 | 0.97 | 0.96 | 103 |
| | **Overall Accuracy** | | | **0.95** | **990** |
| **Naive Bayes** | 1 | 0.90 | 0.94 | 0.92 | 106 |
| | 2 | 0.82 | 0.93 | 0.87 | 175 |
| | 3 | 0.98 | 0.85 | 0.91 | 463 |
| | 4 | 0.74 | 0.92 | 0.82 | 143 |
| | 5 | 0.92 | 0.94 | 0.93 | 103 |
| | **Overall Accuracy** | | | **0.89** | **990** |
| **XGBoost** | 1 | 0.95 | 0.98 | 0.97 | 106 |
| | 2 | 0.98 | 0.97 | 0.97 | 175 |
| | 3 | 0.98 | 0.98 | 0.98 | 463 |
| | 4 | 0.93 | 0.96 | 0.94 | 143 |
| | 5 | 1.00 | 0.94 | 0.97 | 103 |
| | **Overall Accuracy** | | | **0.97** | **990** |

The detailed metrics in Table 2 illustrate the classification capability distribution of each algorithm across all addiction levels. XGBoost exhibits exceptional consistency, evidenced by a perfect precision score (1.00) in predicting Class 5 (the highest addiction level). Conversely, the primary weakness of the Naive Bayes model is evident in its ability to identify Class 4, where precision dropped significantly to 0.74, indicating a high rate of False Positives for that category.

Further analysis using the Confusion Matrix (Figure 6) confirms the robustness of XGBoost. The XGBoost model was able to map nearly all classes perfectly, as proven by the minimal False Positive and False Negative ratios. In contrast, the Naive Bayes Confusion Matrix displays a wider dispersion of prediction errors, particularly in distinguishing between Class 3 and Class 4 addiction levels.

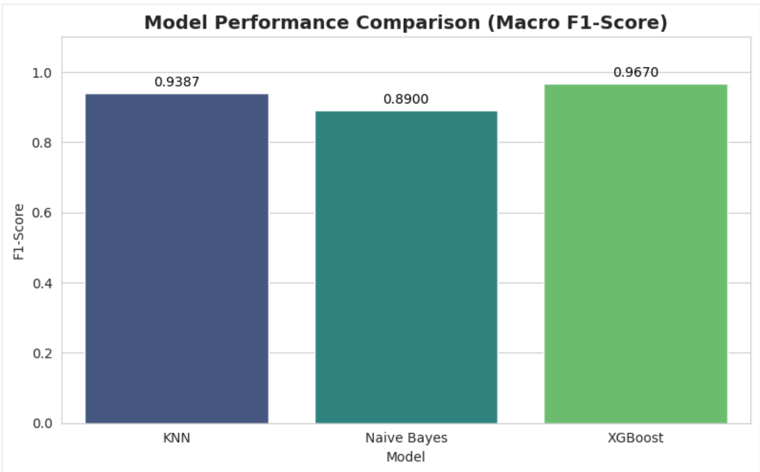



Figure 5. Model Performance Comparison based on Macro F1-Score

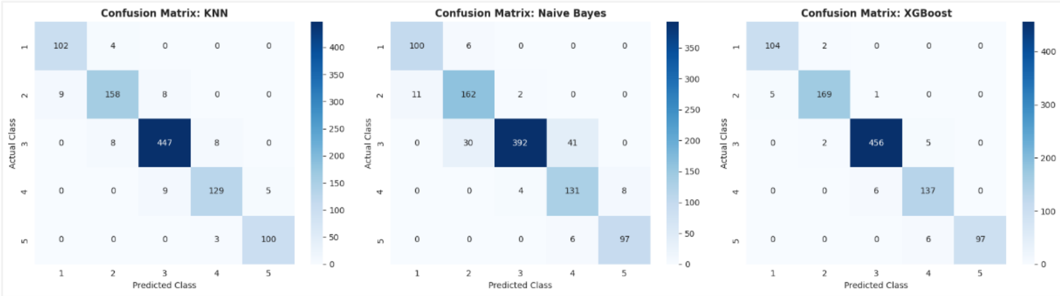



Figure 6. Confusion Matrix for KNN, Naive Bayes, and XGBoost Models

c. Feature Significance Analysis (Feature Importance)
Feature importance extraction using the XGBoost model based on the Gain metric (Figure 7) reveals the specific behavioral patterns that serve as primary predictors for determining addiction levels.

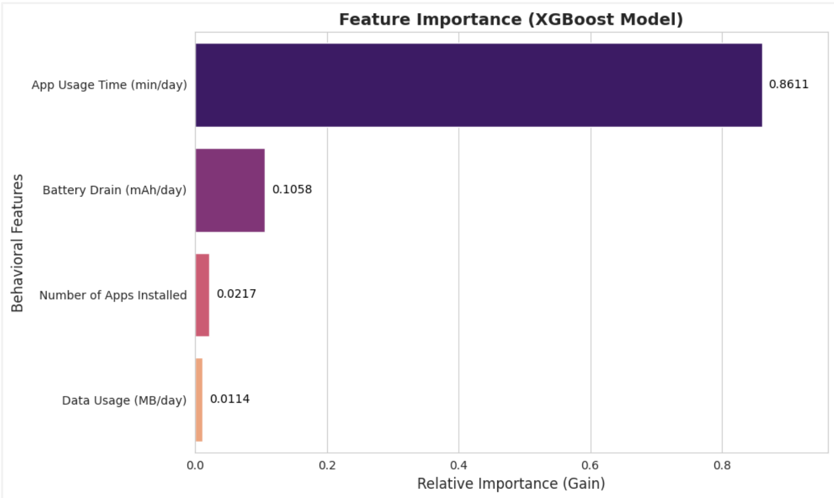



Figure 7. Feature Importance (Gain Metric) in the XGBoost Model

There is a sharp disparity in significance among operational features. App Usage Time dominates overwhelmingly with an importance value of 0.8611. This feature is followed by Battery Drain (0.1058). In contrast, the Number of Apps Installed (0.0217) and Data Usage (0.0114) proved to offer very minor predictive contributions toward determining behavior classes.

**Discussion**

a. Comparison of Algorithm Strengths and Weaknesses
Empirical testing proves that the choice of algorithmic architecture significantly impacts the accuracy of smartphone addiction detection.
1) XGBoost Advantages: The superiority of XGBoost in this study validates the robustness of decision-tree-based ensemble learning in handling complex behavioral data. XGBoost effectively minimizes overfitting and is highly adaptive in recognizing non-linear decision boundaries, making it the most recommended model for this dataset.
2) KNN Performance: The KNN algorithm provided highly competitive results (94.54%). As an instance-based learning approach, KNN successfully grouped similar user characteristics based on Euclidean distance. This high performance suggests that users with similar addiction levels tend to have clusters of closely related usage metrics (duration and battery).
3) Naive Bayes Weaknesses: The lower performance of Naive Bayes (89.09%) can be explained theoretically through EDA (Figure 4) and the algorithm’s fundamental assumptions. Naive Bayes assumes feature independence; however, in smartphone addiction, App Usage Time is strongly correlated (0.94) with Battery Drain. Violating this independence assumption led to decreased classification probability accuracy,

resulting in frequent prediction errors between adjacent classes.

b. Interpretation of Behavioral Patterns
The Feature Importance findings (Figure 5) provide objective insight that smartphone addiction is directly determined by the actual duration of interaction (App Usage Time), rather than the variety of applications installed or demographic factors. The dominance of the usage duration feature (0.8611) confirms that the most valid indicator for detecting addiction is active screen intensity. This demonstrates that machine learning modeling using raw behavior logs can produce detection systems that are far more objective and precise than traditional psychometric questionnaires, which are prone to recall bias or respondent dishonesty.

c. Research Limitations and Recommendations
While this predictive modeling achieved a reliability level above 97%, several operational limitations remain. The current dataset does not categorize the types of applications used (e.g., differentiating duration between social media, gaming, or educational productivity apps). Specific information regarding "content type" could potentially refine user addiction profiles. Future research is encouraged to expand the dataset with detailed application categories and integrate this XGBoost model into digital wellbeing application prototypes capable of real-time monitoring to provide an early warning system for user mental health.

## CONCLUSION

This study concludes that machine learning-based predictive modeling, utilizing behavioral pattern metrics, serves as a highly effective and objective approach for detecting smartphone addiction levels. Empirical performance comparisons demonstrate that Extreme Gradient Boosting (XGBoost) is the superior classification model, achieving an accuracy of 97.27% and an F1-Score of 96.70%. It significantly outperformed K-Nearest Neighbors (94.54%) and Naive Bayes (89.09%) due to its robustness in handling complex non-linear feature interactions. Furthermore, feature importance analysis confirms that screen time intensity (App Usage Time) is the primary predictor of addiction, effectively neutralizing assumptions regarding the influence of user demographic factors. These findings offer a significant practical contribution as an artificial intelligence foundation for automated early detection systems. Future research is recommended to expand the dataset through application categorization and to integrate the XGBoost architecture into real-time digital wellbeing application prototypes capable of mitigating addiction risks as they occur.

## REFERENCES

Abbasi, G. A., Jagaveeran, M., Goh, Y. N., & Tariq, B. (2021). The impact of type of content use on smartphone addiction and academic performance: Physical activity as moderator. *Technology in Society*, *64*, 101521. https://doi.org/10.1016/j.techsoc.2020.101521

Ahmead, M., Maqboul, E., Alshawish, E., & Dweib, M. (2025). The prevalence of smartphone addiction and its related risk factors among Palestinian high school students: a cross-sectional study. *Frontiers in Psychiatry*, *16*, 1636080. https://doi.org/10.3389/fpsyt.2025.1636080

Asare, K. O., Terhorst, Y., Vega, J., Peltonen, E., Lagerspetz, E., & Ferreira, D. (2021). Predicting depression from smartphone behavioral markers using machine learning methods, hyperparameter optimization, and feature importance analysis: exploratory study. *JMIR mHealth and uHealth*, *9*(7), e26540. https://doi.org/10.2196/26540

Doo, E. Y., & Kim, J. H. (2022). Parental smartphone addiction and adolescent smartphone addiction by negative parenting attitude and adolescent aggression: A cross-sectional study. *Frontiers in public health*, *10*, 981245. https://doi.org/10.3389/fpubh.2022.981245

Duan, L., He, J., Li, M., Dai, J., Zhou, Y., Lai, F., & Zhu, G. (2021). Based on a decision tree model for exploring the risk factors of smartphone addiction among children and adolescents in China during the COVID-19 pandemic. *Frontiers in Psychiatry*, *12*, 652356. https://doi.org/10.3389/fpsyt.2021.652356

Freitas, B. H. B. M., Gaíva, M. A. M., Diogo, P. M. J., & Bortolini, J. (2024). Self-reported smartphone addiction among Brazilian adolescents in the COVID-19 pandemic context: a mixed-method study. *Trends in Psychology*, *32*(3), 1007-1026. https://doi.org/10.1007/s43076-022-00208-0

Hu, Z., Li, X., & Xiang, Y. (2024). Developmental trajectories of childhood emotional maltreatment and smartphone addiction in primary school children: Based on a three-year longitudinal study. *Journal of Aggression, Maltreatment & Trauma*, *33*(6), 758-775. https://doi.org/10.1080/10926771.2023.2290719

James, R. J., Dixon, G., Dragomir, M. G., Thirlwell, E., & Hitcham, L. (2023). Understanding the construction of 'behavior'in smartphone addiction: A scoping review. *Addictive behaviors*, *137*, 107503. https://doi.org/10.1016/j.addbeh.2022.107503

Joseph, C., & Maheswari, P. U. (2025). Facial emotion based smartphone addiction detection and prevention using deep learning and video based learning. *Scientific Reports*, *15*(1), 18025. https://doi.org/10.1038/s41598-025-99681-7

Kim, K., Yoon, Y., & Shin, S. (2024). Explainable prediction of problematic smartphone use among South Korea's children and adolescents using a Machine learning approach. *International journal of medical informatics*, *186*, 105441. https://doi.org/10.1016/j.ijmedinf.2024.105441

Kumar, N. V., & Michael, G. (2026). Predicting the accuracy of software defects using grid search CV in comparison with the Naive Bayes. In *AIP Conference Proceedings* (Vol. 3341, No. 1, p. 020049). AIP Publishing LLC. https://doi.org/10.1063/5.0317148

Li, T., Xia, T., Wang, H., Tu, Z., Tarkoma, S., Han, Z., & Hui, P. (2022). Smartphone app usage analysis: datasets, methods, and applications. *IEEE communications surveys & tutorials*, *24*(2), 937-966. https://doi.org/10.1109/COMST.2022.3163176

Mokhtarinia, H. R., Torkamani, M. H., Farmani, O., Biglarian, A., & Gabel, C. P. (2022). Smartphone addiction in children: patterns of use and musculoskeletal discomfort during the COVID-19 pandemic in Iran. *BMC pediatrics*, *22*(1), 681. https://doi.org/10.1186/s12887-022-03748-7

Nikolic, A., Bukurov, B., Kocic, I., Vukovic, M., Ladjevic, N., Vrhovac, M., ... & Sipetic, S. (2023). Smartphone addiction, sleep quality, depression, anxiety, and stress among medical students. *Frontiers in public health*, *11*, 1252371. https://doi.org/10.3389/fpubh.2023.1252371

Okur, S., Engin, M. Ç., Kütük, H., & Satıcı, S. A. (2025). Cross-lagged relations between smartphone addiction and flourishing in adolescents. *Personality and Individual Differences*, *236*, 113008. https://doi.org/10.1016/j.paid.2024.113008

Olson, J. A., Sandra, D. A., Colucci, É. S., Al Bikaii, A., Chmoulevitch, D., Nahas, J., ... & Veissière, S. P. (2022). Smartphone addiction is increasing across the world: A meta-analysis of 24 countries. *Computers in Human Behavior*, *129*, 107138. https://doi.org/10.1016/j.chb.2021.107138

Osorio-Molina, C., Martos-Cabrera, M. B., Membrive-Jiménez, M. J., Vargas-Roman, K., Suleiman-Martos, N., Ortega-Campos, E., & Gómez-Urquiza, J. L. (2021). Smartphone addiction, risk factors and its adverse effects in nursing students: A systematic review and meta-analysis. *Nurse education today*, *98*, 104741. https://doi.org/10.1016/j.nedt.2020.104741

Putri, S. R., Hasibuan, A. K., Sinaga, C. A., Manullang, E. N., Turnip, A., Dharma, A., & Turnip, M. (2026). Analysis of brain activity to methamphetamine stimulus using electroencephalography technology with Naive Bayes algorithm. *Bulletin of Electrical Engineering and Informatics*, *15*(2), 1463-1472. https://doi.org/10.11591/eei.v15i2.10385

Ramadhani, S., & Wayahdi, M. R. (2024). K-nearest neighbor and random forest algorithms in loan approval prediction. *Jurnal Minfo Polgan*, *13*(1), 1307-1313. https://doi.org/10.33395/jmp.v13i1.14345

Ratan, Z. A., Parrish, A. M., Zaman, S. B., Alotaibi, M. S., & Hosseinzadeh, H. (2021). Smartphone addiction and associated health outcomes in adult populations: a systematic review. *International journal of environmental research and public health*, *18*(22), 12257. https://doi.org/10.3390/ijerph182212257

Rathakrishnan, B., Bikar Singh, S. S., Kamaluddin, M. R., Yahaya, A., Mohd Nasir, M. A., Ibrahim, F., & Ab Rahman, Z. (2021). Smartphone addiction and sleep quality on academic performance of university students: An exploratory research. *International journal of* *environmental research and public health*, *18*(16), 8291. https://doi.org/10.3390/ijerph18168291

Ruziq, F., & Wayahdi, M. R. (2025). Web-Based Diabetes Risk Prediction System Using K-NN on Kaggle Early Stage Diabetes Dataset. *Jurnal Teknik Informatika (Jutif)*, *6*(5), 3217-3229. https://doi.org/10.52436/1.jutif.2025.6.5.5277

Sarhan, A. L. (2024). The relationship of smartphone addiction with depression, anxiety, and stress among medical students. *SAGE Open Medicine*, *12*, 20503121241227367. https://doi.org/10.1177/20503121241227367

Schmitgen, M. M., Wolf, N. D., Sambataro, F., Hirjak, D., Kubera, K. M., Koenig, J., & Wolf, R. C. (2022). Aberrant intrinsic neural network strength in individuals with "smartphone addiction": an MRI data fusion study. *Brain and behavior*, *12*(9), e2739. https://doi.org/10.1002/brb3.2739

Sela, A., Rozenboim, N., & Ben-Gal, H. C. (2022). Smartphone use behavior and quality of life: What is the role of awareness?. *PloS one*, *17*(3), e0260637. https://doi.org/10.1371/journal.pone.0260637

Sharma, D., Goel, N. K., Sidana, A., Kaura, S., & Sehgal, M. (2023). Prevalence of smartphone addiction and its relation with depression among school-going adolescents. *Indian Journal of Community Health*, *35*(1), 27-31. https://doi.org/10.47203/IJCH.2023.v35i01.006

Singh, A., Ali, A., & Paul, F. A. (2025). Neurobiological contributions to addiction: a narrative review of adolescent and adult vulnerabilities. *Journal of Addictive Diseases*, 1-11. https://doi.org/10.1080/10550887.2025.2513142

Sohn, S. Y., Krasnoff, L., Rees, P., Kalk, N. J., & Carter, B. (2021). The association between smartphone addiction and sleep: a UK cross-sectional study of young adults. *Frontiers in psychiatry*, *12*, 629407. https://doi.org/10.3389/fpsyt.2021.629407

Sunday, O. J., Adesope, O. O., & Maarhuis, P. L. (2021). The effects of smartphone addiction on learning: A meta-analysis. *Computers in Human Behavior Reports*, *4*, 100114. https://doi.org/10.1016/j.chbr.2021.100114

Wayahdi, M. R., & Ruziq, F. (2022). KNN and XGBoost Algorithms for Lung Cancer Prediction. *Journal of Science Technology (JoSTec)*, *4*(1). https://doi.org/10.55299/jostec.v4i1.251

Wayahdi, M. R., & Ruziq, F. (2025). Predicting Smartphone Addiction Levels with K-Nearest Neighbors Using User Behavior Patterns. *Jurnal Teknik Informatika (Jutif)*, *6*(5), 3379-3391. https://doi.org/10.52436/1.jutif.2025.6.5.4905

Yilmaz, F. G. K., Avci, U., & Yilmaz, R. (2023). The role of loneliness and aggression on smartphone addiction among university students. *Current psychology*, *42*(21), 17909-17917. https://doi.org/10.1007/s12144-022-03018-w

Yogesh, M., Ladani, H., & Parmar, D. (2024). Associations between smartphone addiction, parenting styles, and mental well-being among adolescents aged 15–19 years in Gujarat, India. *BMC Public Health*, *24*(1), 2462. https://doi.org/10.1186/s12889-024-19991-9