International Journal of Management Science and Information Technology
IJMSIT

E-ISSN: 2774-5694 P-ISSN: 2776-7388

Volume 6 (1), January-June 2026, 586-594
DOI: https://doi.org/10.35870/ijmsit.v6i1.7098

Continuous Regression Models for Mapping the Smartphone Addiction Spectrum Using Random Forest Regressor

I Putu Eka Aditya Saputra [1*], I Made Gede Sunarya [2], Putu Hendra Suputra [3]
[1*,2,3] Computer Science Study Program, Graduate Faculty, Universitas Pendidikan Ganesha, Buleleng Regency, Bali Province, Indonesia

Email: eka.aditya.saputra@student.undiksha.ac.id [1*], sunarya@undiksha.ac.id [2], hendra.suputra@undiksha.ac.id [3]

Article history:
Received May 4, 2026
Revised May 18, 2026
Accepted May 20, 2026

Abstract

This study proposes a predictive modeling approach to measure the level of smartphone addiction in adolescents by transforming a conventional binary classification model into a continuous regression model. The use of categorical labels often fails to capture the complex spectrum of addictive behaviors, so this study implemented the Random Forest Regressor algorithm to predict addiction scores on a scale of 1.0 to 10.0. The experimental results show that the regression model is able to provide high prediction accuracy, as evidenced by the coefficient of determination obtained R^2 of 0.8607 and a Mean Absolute Error (MAE) of 0.2854. These findings confirm that the regression approach offers better data resolution in mapping the degree of digital dependency than classification methods. In practice, this model produces a continuous score that provides a dynamic tool for mental health professionals. This approach allows for objective monitoring of patient's behavioral progress during recovery. Furthermore, this model can facilitate multilevel psychological interventions and tailored care, from early prevention to therapy for high-risk addicts.

Keywords:
Smartphone Addiction; Random Forest Regressor; Predictive Modeling; Continuous Score; CRISP-DM.

1. INTRODUCTION

Mental health issues in adolescents have becomes a worldwide concern. Over the past several decades, research on mental health issues in adolescents has been conducted (Jiwa Permana et al., 2025; Windarwati et al., 2022). Smartphone addiction is academically represented as an individual's chronic inability to regulate the intensity, duration, and frequency of their mobile device use, which at a certain equilibrium point leads to negative consequences that impair psychosocial well-being and daily functional performance (Li et al., 2020; Marciano & Camerini, 2022; Ndayambaje & Okereke, 2025). Some of the causes of smartphone addiction include low self-control, loneliness and sensation seeking behavior (Ratuaki et al., 2023). This deep dependency often manifests clinically in the form of nomophobia, a syndrome of anxiety, agitation, or irrational fear that strikes individuals when they are disconnected from their smartphone, lose battery power, or are out of range of internet connectivity (Agung et al., 2024; Srivastava et al., 2025). Its impact extends beyond individual well-being and significantly impacts the daily functioning and effectiveness of healthcare practitioners. Healthcare workers frequently encounter situations that require empathy, compassion, and emotional support for patients and their families, even amidst personal stress, grief, and difficult circumstances (Jim et al., 2024).

The widespread adoption of digital technology has raised significant concerns regarding problematic screen device use. In Indonesia, the prevalence of smartphone usage expanded to 63.53% of the population in 2019, corresponding with a substantial rise in internet accessibility (Nurmala et al., 2022; Statistik-Telekomunikasi-Indonesia-2019, n.d.). When left unmanaged, this high level of digital engagement presents severe developmental and psychological risks. Research indicates that unsupervised device exposure among children, particularly during routine activities such as mealtimes, contributes to the development of attention-deficit disorders (Park & Park, 2021) Beyond early developmental impacts, excessive smartphone use consistently diminishes individual well-being across broader demographics (Hanafi et al., 2019; Yosep Fredi et al., 2026) At the behavioral level, smartphone addiction is frequently associated with poor psychological adjustment, occupational difficulties, heightened interpersonal conflict, and increased vulnerability to cyberbullying (Stone, 2020). Furthermore, the persistent cognitive load generated by constant digital notifications impairs the brain's algorithmic information processing capabilities. This neurological disruption serves as the underlying mechanism for the linear degradation of academic performance, computational focus, and overall quality of life (Upshaw et al., 2022)

Research subjects, especially demographic groups that are on the high end of the addiction spectrum, often experience the phenomenon of denial or social desirability bias (Crowhurst & Hosseinzadeh, 2024; Zhou & Feng, 2025). This leads to under-reporting, where users report application usage duration, unlock frequency, or anxiety levels that are much lower than the actual data (Shiferaw et al., 2025). This is where the discipline of Computer Science, particularly through the advanced paradigms of the sub-fields of Machine Learning, Pattern Recognition, and Data Mining, offers high-level computational resolution (Likhith et al., 2024; Joseph & Maheswari, 2025). Modern computing eliminates the subjective biases of human neurocognitive behaviour through an objective approach to behavioural data extraction.

The rapid advancement of Artificial Intelligence (AI) has facilitated the automated inference of mental states through multi-modal data using machine learning (Aditra Pradnyana et al., 2025). Early computational frameworks attempted to quantify Problematic Smartphone Use (PSU) on a continuous spectrum; for example, (Elhai et al., 2020) applied shrinkage-based linear algorithms—such as Lasso, Ridge, and Elastic Net Regression identifying Fear of Missing Out (FOMO) as the dominant predictor of addiction severity. While treating the output as a continuous score was a critical methodological step, this linear approach inherently struggles to capture the complex, non-linear interactions characterizing psychological variables. To accommodate these non-linear dynamics, subsequent studies transitioned toward advanced ensemble methods, though they predominantly reverted to binary classification paradigms (Lee & Kim, 2021) analyzed empirical smartphone usage logs from 29,712 respondents in South Korea, demonstrating that Random Forest and XGBoost significantly outperformed baseline Decision Trees in categorizing individuals into dichotomous "problematic" or "normal" cohorts.

Subsequent research developments increasingly emphasize model scalability and interpretation capabilities (Hong et al., 2024) using a large sample of approximately 3,000 Chinese college students to build a machine learning-based prediction model for the risk of smartphone addiction. The developed model achieved an accuracy of 76.68% and revealed that perfectionism is a key preSdictor often overlooked in previous studies. In the same year, a study (Kim et al., 2024) conducted a comprehensive test of eight algorithms simultaneously: Logistic Regression, Random Forest, GBM, XGBoost, LightGBM, CatBoost, MLP, and CNN using South Korean smartphone addiction survey data from 2017 to 2024. The results showed that XGBoost achieved the highest precision, 87.60%, in identifying high-risk groups. This study also successfully identified new, previously undiscovered risk factors, namely the consumption of game content, webtoons, and web novels, as significant predictors.

Although previous studies have made significant contributions to the development of machine learning-based smartphone addiction prediction models, a research gap consistently emerges. Nearly all previous studies have used a binary classification approach that can only output categorical labels, namely "addicted" or "not addicted." This approach, while useful for initial screening, has a fundamental limitation because it cannot quantify the gradation or degree of addiction numerically. Smartphone addiction is not a dichotomous condition, but rather a spectrum that moves continuously from normal use to highly problematic use. The only study that comes close to a continuous score-prediction approach is (Elhai et al., 2020) However, this study only utilised a shrinkage-based linear regression algorithm, which has limitations in modelling complex nonlinear relationships between behavioural and psychological variables. Thus, there is a significant research gap in the application of decision tree-based ensemble algorithms, particularly Random Forest Regressors, for continuous prediction of addiction scores.

2. RESEARCH METHOD

The Cross-Industry Standard Process for Data Mining (CRISP-DM) is a structured approach to data mining project development and predictive model creation (Blasi & Alsuwaiket, 2020). CRISP-DM was chosen because it is able to provide a systematic, structured, and replicable research process, starting from understanding the problem and ending with model evaluation (Ardyaputra et al., 2026). Athis method consists of six main phases as shown in Figure 1. The phases are (1) Business understanding, (2) Data understanding, (3) Data preparation, (4) Modelling, (5) Evaluation, and (6) Deployment. In this research, five phases of the Deployment phase have been completed.

2.1. Business Understanding

With the aim of identifying people with smartphone addiction characteristics, a social literature review was conducted on the social context to analyze the characteristics of people with smartphone addiction.

2.2. Data Understanding

The dataset used in this study is a secondary dataset sourced from the Kaggle repository, synthetically designed to simulate real-world smartphone usage behavior among adolescents. This dataset includes 3000 individual data entries of adolescents aged 13 to 19. This data provides a comprehensive overview of the interaction between digital habits and indicators of mental health, academic achievement, and daily lifestyle patterns, which is highly relevant for measuring the level of technology addiction in vulnerable age groups.

Although the dataset utilized is synthetically designed for computational benchmarking, its construct validity is strongly anchored in established empirical literature. The variable architecture within the datasetencompassing psychosocial indicators such as anxiety, depression, and specific digital consumption metrics accurately mirrors the validated real-world parameters of problematic smartphone use identified in prior clinical. By utilizing a dataset structurally mapped to theoretically sound variables, this study ensures high construct validity for model training, while simultaneously circumventing the ethical and privacy constraints associated with collecting large-scale, sensitive psychological data from real adolescents

Structurally, this dataset consists of 25 features that can be categorized into four main domains. First is the demographic profile, which includes variables such as age, gender, location, and education level. Second are technology usage metrics, which include daily usage duration, phone check frequency, screen time before bed, and specific usage durations within the categories of social media, gaming, and education. Third are psychosocial indicators such as anxiety levels, depression levels, self-esteem, and quality of family communication. Finally, there are performance and lifestyle variables such as academic grades, sleep duration, and exercise hours.

The target variable, Addiction_Level, is represented on a numerical scale from 1.0 to 10.0, indicating the intensity of the subject's addiction. The data show the distribution of daily mobile phone usage, with an average of 5 hours and some extreme cases exceeding 11 hours. Furthermore, there is a categorical feature, Phone_Usage_Purpose, that summarizes the primary purpose of device use, ranging from simple browsing to educational activities, providing a qualitative dimension to the quantitative analysis.

Pearson correlation analysis in Figure 2 was conducted to identify the linear relationship between the numerical features and the target variable Addiction_Level, where it was found that the duration of daily use (Daily_Usage_Hours) had the strongest positive correlation of 0.60, followed by the number of applications used (0.31) and time on social media 0.30. In contrast, a significant negative correlation was found in the duration of sleep (Sleep_Hours) of -0.21, which indicates that higher levels of addiction tend to be followed by a decrease in the quality of adolescents' rest.

2.3. Data Preparation

Data preprocessing is a crucial step because raw data often contains errors, inconsistencies, missing values, or irrelevant information. This stage processes the raw data through several steps, which can then be used as input for data visualization, machine learning, deep learning, and other applications (Indrawan et al., 2022; Ketut et al., 2026). The data preparation process begins with the elimination of features or columns deemed statistically irrelevant to the prediction target. Data preprocessing was conducted systematically to ensure model objectivity and computational efficiency. First, three non-predictive identifiers ID, Name, and Location were removed, reducing the dataset's dimensionality from 25 to 22 features. Next, a programmatic data integrity check identified 210 records, 7.00% of the total 3,000 instances containing missing or Null (NaN) values. To maintain data fidelity and prevent potential bias from artificial imputation techniques, a strict listwise deletion strategy was applied. This cleaning step successfully removed all incomplete rows, resulting in a final data shape of 2,790 valid observations across 22 variables. This rigorous procedure ensured that the subsequent modeling was trained exclusively on complete and robust behavioral records. This aims to optimize model parameters, especially if the missing data is random, and to avoid potential bias caused by inappropriate imputation techniques (Little & Rubin, 2019; Emmanuel et al., 2021).

After the dataset has been cleaned, the next step is to define the research variables. The dataset is separated into feature variables X, which encompass various behavioural and psychological indicators, and the target variable Y, Addiction_Level. Next, the data is divided (data splitting) using a ratio of 80% as the training set and 20% as the testing set. This splitting is performed before further transformation to ensure that information from the testing data does not leak into the model training process (data leakage).

Stage of data preparation involves transforming features to enable optimal processing by machine learning algorithms (Ahsan et al., 2021). Categorical variables were converted into numeric representations using the Label Encoding technique. This process was applied to the Gender, School_Grade, and Phone_Usage_Purpose features. Given that the variables in the dataset have a wide range of values, feature scaling was performed using the StandardScaler. This technique transforms the data to have a mean of zero and a standard deviation of one, which is crucial for aligning the contribution of each feature in the predictive modelling process.

To minimize noise from irrelevant variables and improve predictive performance, feature selection was conducted based on the Pearson correlation matrix in Figure 2. Specifically, only features demonstrating a meaningful linear relationship with the target variable defined by an absolute correlation coefficient of 0.19 or higher were retained for the second experiment. This process resulted in a subset of six primary features

Daily_Usage_Hours, Apps_Used_Daily, Time_on_Social_Media, Time_on_Gaming, Phone_Checks_Per_Day, and Sleep_Hours. The selection of these variables is theoretically sound, as they capture key behavioral indicators of digital dependency. For example, extended time spent on social media and gaming reflects intensive screen engagement, while reduced sleep duration often serves as a recognized physiological consequence of addiction severity.

The integration of these establishes a multidimensional framework essential for mapping the addiction spectrum. Within this architecture, technology usage metrics function as direct behavioral predictors. Concurrently, psychosocial and lifestyle variables contextualize these behaviors by accounting for underlying emotional vulnerabilities and physiological impacts. Incorporating this diverse feature space enables the Random Forest algorithm to effectively capture the complex, non-linear interactions between digital consumption patterns and adolescent psychosocial functioning during the initial model training.

2.4. Modeling

At this stage, a predictive model was built using the Random Forest Regressor algorithm to estimate the smartphone addiction score (Addiction_Level). Random Forest was chosen for its superiority as an ensemble learning method that combines many decision trees to improve prediction accuracy and stability. This algorithm is very effective in handling datasets with complex feature interactions, such as the relationship between application usage duration, sleep patterns, and mental health indicators (Yu et al., 2025) and has good resistance to overfitting through the bagging mechanism (Namlı et al., 2025), (Lee & Kim, 2021).

The modelling process was implemented using training data, representing 80% of the total dataset, that had undergone preprocessing and standardization. During training, the Random Forest Regressor builds multiple decision trees and averages their outputs to produce the final prediction. The use of default parameters and certain optimizations was carried out to ensure the model could capture non-linear patterns in features such as Daily_Usage_Hours and Social_Interactions. The trained model was then ready to be tested on 20% of the testing data to measure the extent of its generalization ability in predicting addiction levels in new data.

2.5. Evaluation

The evaluation phase aims to measure the predictive performance of the Random Forest Regressor model in objectively estimating smartphone addiction levels. Model effectiveness is assessed using R-Squared (R^2), Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE). R^2 plays a role in measuring the extent to which the variance in the target variable Addiction_Level is explained by the model's features. Meanwhile, MAE provides an overview of the average absolute error on the same scale as the target variable, facilitating interpretation of prediction deviations, while RMSE imposes a greater penalty on large prediction errors, highlighting the model's sensitivity to outliers. Through the integration of these three metrics, the accuracy, stability, and reliability of the model in predicting smartphone addiction behavior can be comprehensively validated.

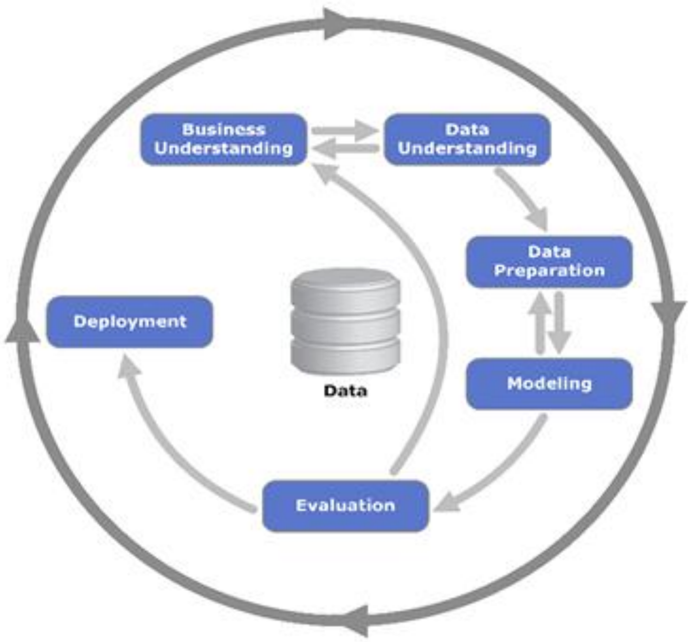

Figure 1. CRISP-DM

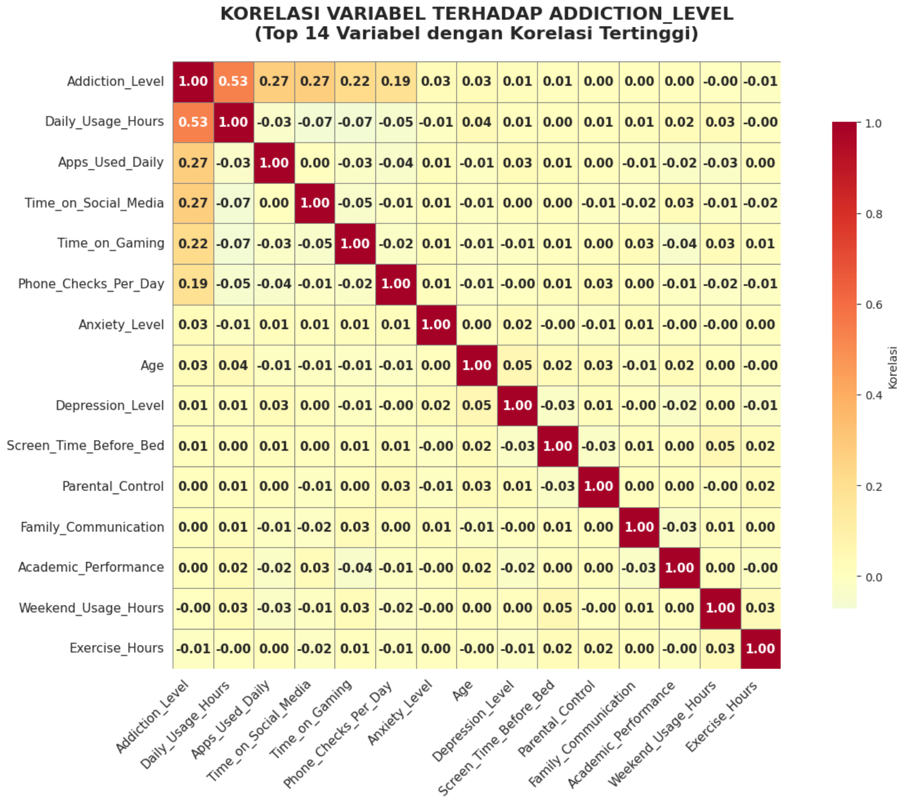

Figure 2. Heatmap Correlation of Top 14 Variables to Addiction Level Variable

3. RESULTS AND DISCUSSION

After the training, the evaluation results, including the algorithm's RMSE and MAE, are shown in Table 1. Based on the testing conducted on the testing set, the Random Forest Regressor model shows solid generalization ability in estimating the level of smartphone addiction. The coefficient of determination value R^2 A value of 0.8067 indicates that the model is able to explain 80.67% of the variability in the target variable through the features used. This figure indicates a high level of goodness-of-fit, considering that human behavior and addiction often involve dynamic psychological variables. Analysis of the prediction error magnitude shows an MAE of 0.3532 and an RMSE of 0.5090. The low MAE indicates that the average deviation of the model's predictions from the actual scores is only about 0.35 points on the addiction scale. Meanwhile, the RMSE value is slightly higher than the MAE, indicating the presence of some error variance with a larger magnitude, but remains within tolerable limits without disrupting the overall stability of the model.

Table 1. Prediction Results

Model Evaluation	Score
R^2	0.8067
RMSE	0.5090
MAE	0.3532

The plot in Figure 3 shows that on the training data, the model achieves very good performance, with a coefficient of determination R2 of 0.969, indicating that 96.9% of the variation in the training data is explained by the model. The training data plot shows a distribution of points that closely follow the ideal regression line, indicating high prediction accuracy and minimal prediction error. In contrast, on the test data, the model obtained an R2 value of 0.8067, indicating the model's generalization ability is still good but has decreased in performance compared to the training data. The test data plot shows a more dispersed distribution of points around the regression line, especially in the actual value range of 8.5–10.0, indicating higher prediction variability. A slight overfitting phenomenon is evident from the performance difference between the training and test data, but the model still shows strong overall predictive ability with an R^2 value above 0.80 on previously unseen data.

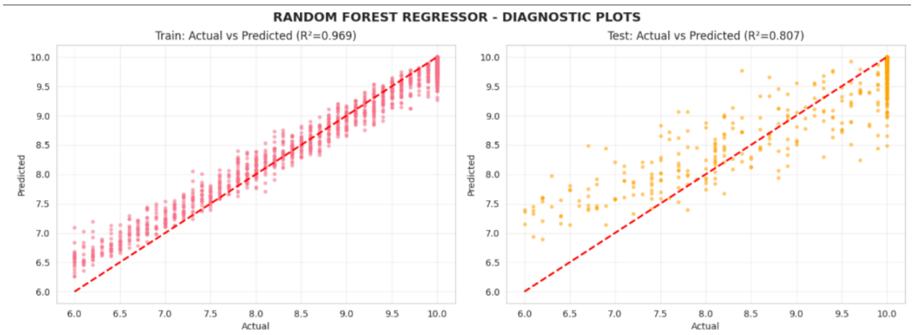

Figure 3. Plot Comparison of Train vs Test

In the second experiment, which used six features with the highest correlation: Daily_Usage_Hours, Time_on_Social_Media, Time_on_Gaming, Apps_Used_Daily, Phone_Checks_Per_Day, and Sleep_Hours. The results of the second experiment showed that implementing the feature selection technique provided a significant performance improvement for the Random Forest Regression model. By focusing the input variables on the main features, the model was able to produce a coefficient of determination R^2 of 0.8607 on the test data. These results show an improvement compared to the initial experiment which only obtained an R^2 of 0.8067. In addition to the increase in the R^2 value, the prediction error rate also decreased markedly: the RMSE decreased from 0.5090 to 0.4321, and the MAE decreased from 0.3532 to 0.2854. Although there is a difference between the performance on the training and test data, these results show that selecting features that specifically represent daily gadget usage behaviour can increase the model's precision in predicting the level of smartphone addiction. The comparison results are shown in Table 2.

Table 2. Model Performance Comparison for Before and After Feature Selection

Experiment Scenario	R^2	RMSE	MAE
Random Forest Regresion (all feature)	0.8067	0,5090	0,3532
Random Forest Regresion (Feature Selection)	0,8607	0,4321	0,2854

Visualization of model performance in Figure 4 through a scatter plot between actual and predicted values shows a significant difference between the two experiments. In the plot of the model's results after feature selection, the distribution of data points shows a denser, more consistent pattern along the diagonal. The density of the data distribution in the blue plot indicates that the second model has a lower error variance than the first model. This visually demonstrates that using dominant features improves prediction linearity and reduces noise introduced by less relevant features, enabling the model to map the level of smartphone addiction more accurately.

From a practical standpoint across both clinical and educational settings, the high predictive accuracy R^2 = 0.8607 and minimized error rates MAE = 0.2854 achieved post-feature selection carry significant operational value. In educational environments, this continuous regression model provides school counselors with an objective instrument to identify early cognitive and behavioral degradation, allowing for intervention before a student fall into severe addiction. The superior performance of the model in the second experiment is directly attributed to the elimination of demographic variables such as age and gender which demonstrated negligible correlation with addiction levels. In the initial model, these demographic attributes essentially acted as computational noise, obscuring the underlying patterns. Conversely, the accuracy improved significantly when the algorithm was focused exclusively on active behavioral features. As empirically validated by the correlation matrix, daily usage duration (Daily_Usage_Hours) exhibited the strongest mathematical relationship with the target variable r = 0.53, serving as the most dominant contributor representing the sheer volume of screen exposure. This primary predictor was further strengthened by other highly correlated variables, specifically the number of apps used daily r = 0.27 and time spent on social media r = 0.27 and gaming r = 0.22, which act as specific psychological triggers for compulsive screen attachment. Furthermore, the inclusion of sleep duration (Sleep_Hours), which exhibited a significant negative correlation r = -0.21, provided the algorithm with a crucial physiological parameter to measure the real-world impact of digital dependency. With a margin of error of merely 0.28 points on a 1.0 to 10.0 scale, clinicians and educational practitioners can effectively mitigate the risk of false positives or misdiagnoses commonly caused by the social desirability biases of self-reported questionnaires. Ultimately, the resulting continuous scores enable both schools and clinics to dynamically track an adolescent's recovery progress over time. This facilitates the design of proportionately tailored, tiered interventions—ranging from preventive counseling for moderate dependency to intensive clinical therapy for high-risk individuals.

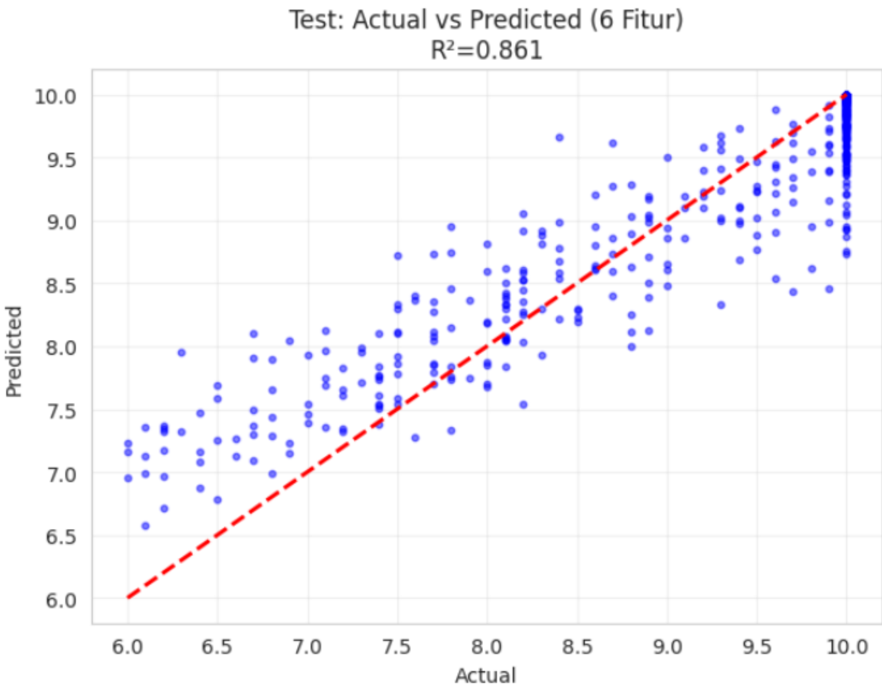

Figure 4. Plot Comparison of Train vs Test model with Selection Feature

4. CONCLUSION

This study successfully implemented the Random Forest Regressor algorithm to predict the level of smartphone addiction in adolescents using a continuous score approach, which has been shown to provide more granular computational resolution than conventional binary classification models. Based on the experimental results, this research establishes three specific key findings. First, the integration of feature selection techniques—based on the heatmap correlation matrix significantly improved model performance. By isolating six main digital behavior variables Daily_Usage_Hours, Time_on_Social_Media, Time_on_Gaming, Apps_Used_Daily, Phone_Checks_Per_Day, and Sleep_Hours), the model achieved a high determination coefficient of 0.8607 and a decreased Mean Absolute Error (MAE) of 0.2854. Thus, reducing the dimensionality of the data not only improved computational efficiency but also strengthened the model's validity. Second, these findings confirm that daily smartphone usage patterns are much more deterministic predictors in mapping the addiction spectrum than demographic factors. Demographic attributes essentially acted as computational noise, whereas active behavioral metrics provided direct predictive signals. Third, with its minimized error margin, the model demonstrates strong and stable generalization capabilities. This makes it an effective objective diagnostic tool for clinicians and educational practitioners to conduct early, targeted interventions on technology use behavior. As a future development step, this study recommends further exploration of ensemble boosting algorithms, such as XGBoost and LightGBM, to compare their computational efficiency with Random Forest Reggresor, especially on high-dimensional datasets with higher noise levels. Furthermore, the application of automated hyperparameter optimization techniques, such as Bayesian Optimization, is recommended to mitigate potential overfitting more precisely and to improve the model's generalizability.

REFERENCES

Aditra Pradnyana, G., Anggraeni, W., Mulyanto Yuniarno, E., & Purnomo, M. H. (2025). Revealing Depression Through Social Media via Adaptive Gated Cross-Modal Fusion Augmented with Insights from Personality Traits. IEEE Access, 13, 133465–133482. https://doi.org/10.1109/ACCESS.2025.3593273

Agung, A. A. G., Gading, I. K., Agetania, N. L. P., Prawira, A. A. G. O. A., Deng, J.-B., & Werang, B. R. (2024). Exploring Work-Related Stress among Indonesian Primary School Teachers: A Study in the Post-Covid-19 Era. Journal of Ecohumanism, 3(4), 805–815. https://doi.org/10.62754/joe.v3i4.3577

Ahsan, M., Mahmud, M., Saha, P., Gupta, K., & Siddique, Z. (2021). Effect of Data Scaling Methods on Machine Learning Algorithms and Model Performance. Technologies, 9(3), 52. https://doi.org/10.3390/technologies9030052

Ardyaputra, G. Y., Bagus, I., Pascima, N., & Suputra, P. H. (2026). Analisis Sentimen Penggunaan Cekat.AI dalam Menggantikan Customer Service Menggunakan Logistic Regression dan TF-IDF. Journal of Computer System and Informatics (JoSYC), 7(2), 47–56. https://doi.org/10.47065/josyc.v7i2.9292

Blasi, A. H., & Alsuwaiket, M. (2020). Analysis of Students' Misconducts in Higher Education using Decision Tree and ANN Algorithms. Engineering, Technology & Applied Science Research, 10(6), 6510–6514. https://doi.org/10.48084/etasr.3927

Crowhurst, S., & Hosseinzadeh, H. (2024). Risk Factors of Smartphone Addiction: A Systematic Review of Longitudinal Studies. Public Health Challenges, 3(2). https://doi.org/10.1002/puh2.202

Elhai, J. D., Yang, H., Rozgonjuk, D., & Montag, C. (2020). Using machine learning to model problematic smartphone use severity: The significant role of fear of missing out. Addictive Behaviors, 103. https://doi.org/10.1016/j.addbeh.2019.106261

Emmanuel, T., Maupong, T., Mpoeleng, D., Semong, T., Mphago, B., & Tabona, O. (2021). A survey on missing data in machine learning. Journal of Big Data, 8(1), 140. https://doi.org/10.1186/s40537-021-00516-9

Hanafi, E., Siste, K., Wiguna, T., Kusumadewi, I., & Nasrun, M. W. (2019). Temperament profile and its association with the vulnerability to smartphone addiction of medical students in Indonesia. PLOS ONE, 14(7), e0212244. https://doi.org/10.1371/journal.pone.0212244

Hong, Y., Rong, X., & Liu, W. (2024). Construction of influencing factor segmentation and intelligent prediction model of college students' cell phone addiction model based on machine learning algorithm. Heliyon, 10(8), e29245. https://doi.org/10.1016/j.heliyon.2024.e29245

Indrawan, G., Setiawan, H., & Gunadi, A. (2022). Multi-class SVM Classification Comparison for Health Service Satisfaction Survey Data in Bahasa. HighTech and Innovation Journal, 3(4), 425–442. https://doi.org/10.28991/HIJ-2022-03-04-05

Jim, E. L., Pio, R. J., Asaloei, S. I., Leba, S. M. R., Angelianawati, D., & Werang, B. R. (2024). Work-Related Stress, Emotional Exhaustion, Job Satisfaction, and Organizational Commitment of Indonesian Healthcare Workers. International Journal of Religion, 5(5), 308–316. https://doi.org/10.61707/6fzykj38

Jiwa Permana, A. A., Sudarma, M., Sukarsa, I. M., & Hartati, R. S. (2025). Development of a Life Story-Based Digital Counseling Model to Detect Student Depression Using LSTM. JOIV: International Journal on Informatics Visualization, 9(2), 550. https://doi.org/10.62527/joiv.9.2.2642

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

Ketut, N., Artalia, A., Made, I., Maysanjaya, D., & Suputra, P. H. (2026). SENTIMENT ANALYSIS OF BALI'S RAPID TOURISM GROWTH USING INDOBERT WITH INSET LEXICON AUTOMATIC LABELLING. 15(1).

Kim, K., Yoon, Y., & Shin, S. (2024). Explainable prediction of problematic smartphone uses 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

Lee, J., & Kim, W. (2021). Prediction of problematic smartphone use: A machine learning approach. International Journal of Environmental Research and Public Health, 18(12). https://doi.org/10.3390/ijerph18126458

Li, L., Griffiths, M. D., Mei, S., & Niu, Z. (2020). Fear of Missing Out and Smartphone Addiction Mediates the Relationship Between Positive and Negative Affect and Sleep Quality Among Chinese University Students. Frontiers in Psychiatry, 11. https://doi.org/10.3389/fpsyt.2020.00877

Likhith, S., Chitteti, C., Dharani, M., Nivedhitha, V., Geethika, N. G., & Godwin, V. (2024). Machine Learning Model for Prediction of Smartphone Addiction. 2024 International Conference on Expert Clouds and Applications (ICOECA), 924–929. https://doi.org/10.1109/ICOECA62351.2024.00163

Little, R., & Rubin, D. (2019). Statistical Analysis with Missing Data, Third Edition. Wiley. https://doi.org/10.1002/9781119482260

Namlı, S., Çar, B., Kurtoğlu, A., Yılmaz, E., Tekkurşun Demir, G., Güvendi, B., Batu, B., & Aldhahi, M. I. (2025). 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, 13(15), 1805. https://doi.org/10.3390/healthcare13151805

Ndayambaje, E., & Okereke, P. U. (2025). The Psychopathology of Problematic Smartphone Use (PSU): A Narrative Review of Burden, Mediating Factors, and Prevention. Health Science Reports, 8(5). https://doi.org/10.1002/hsr2.70843

Nurmala, I., Nadhiroh, S. R., Pramukti, I., Tyas, L. W., Zari, A. P., Griffiths, M. D., & Lin, C.-Y. (2022). Reliability and validity study of the Indonesian Smartphone Application-Based Addiction Scale (SABAS) among college students. Heliyon, 8(8), e10403. https://doi.org/10.1016/j.heliyon.2022.e10403

Park, J. H., & Park, M. (2021). Smartphone use patterns and problematic smartphone use among preschool children. PLOS ONE, 16(3), e0244276. https://doi.org/10.1371/journal.pone.0244276

Ratuaki, F., Gading, I. K., & Suarni, N. K. (2023). Pengaruh Konseling Realita untuk Menurunkan kecanduan terhadap smartphone dan meningkatkan Prestasi belajar. Jurnal EDUCATIO: Jurnal Pendidikan Indonesia, 9(1), 37. https://doi.org/10.29210/1202322638

Shiferaw, B. D., Tang, J., Wang, Y., Wang, Y., Wang, Y., Mackay, L. E., Luo, Y., Yan, N., Shen, X., Zhou, T., Zhu, Y., Cai, J., Wang, Q., Yan, W., Gao, X., Pan, H., & Wang, W. (2025). Impact of digital addiction on youth health: A systematic review and meta-analysis. Journal of Behavioral Addictions, 14(3), 1129–1158. https://doi.org/10.1556/2006.2025.00081

Srivastava, S., Verma, N., Kumar, D., Singh, N., & Kumar, K. (2025). Nomophobia as an Emerging Psychopathology Psychophysiological Mechanisms and Clinical Implications. Annals of Neurosciences. https://doi.org/10.1177/09727531251351082

statistik-telekomunikasi-indonesia-2019. (n.d.).

Stone, R. (2020). The Effects of Cyberbullying as it Relates to social media: A California High School Assistant Principal and High School Counselor Perspective. https://digitalcommons.umassglobal.edu/edd_dissertations

Upshaw, J. D., Stevens, C. E., Ganis, G., & Zabelina, D. L. (2022). The hidden cost of a smartphone: The effects of smartphone notifications on cognitive control from a behavioral and electrophysiological perspective. PLOS ONE, 17(11), e0277220. https://doi.org/10.1371/journal.pone.0277220

Windarwati, H. D., Lestari, R., Wicaksono, S. A., Kusumawati, M. W., Ati, N. A. L., Ilmy, S. K., Sulaksono, A. D., & Susanti, D. (2022). Relationship between stress, anxiety, and depression with suicidal ideation in adolescents. Jurnal Ners, 17(1). https://doi.org/10.20473/jn.v17i1.31216

Yosep Fredi, I Made Tegeh, & Ketut Agustini. (2026). Smartphone Use, Learner Autonomy, and Social Context in Student Creative Thinking. Indonesian Journal of Innovation Studies. https://ijins.umsida.ac.id/index.php/ijins/article/view/1758/2216

Yu, C., Kong, X., Yu, W., Ni, X., Chen, J., & Liao, X. (2025). Machine learning models for predicting the risk of depressive symptoms in Chinese college students. Frontiers in Psychiatry, 16. https://doi.org/10.3389/fpsyt.2025.1648585

Zhou, X., & Feng, B. (2025). Social anxiety and smartphone addiction among college students: the mediating role of loneliness. Frontiers in Psychiatry, 16. https://doi.org/10.3389/fpsyt.2025.1621900