



“A Comparative Analysis of K-Nearest Neighbors and Logistic Regression for Classifying the Effectiveness of the Do Not Disturb Feature in Supporting Student Focus and Productivity”

Nadhifa Syakira Amanada¹, Aurelia Citra Dewi Permadani², Maulidya Miftahul Magfiroh³,
Ramya Kinanthi⁴, Aisyah Aszahra Alfani⁵, and Murad Ali^{6,*}

^{1,2,3}Department of Data Science, UPN “Veteran” Jawa Timur, Indonesia, ⁴Department of Medical Technology, Institut Teknologi Sepuluh Nopember, Indonesia, ⁵Department of Physiotherapy, Universitas Airlangga, Indonesia, ⁶Department of English Literature, UIN Sunan Ampel Surabaya, Indonesia

¹25083010044@student.upnjatim.ac.id, ²25083010046@student.upnjatim.ac.id,
³25083010056@student.upnjatim.ac.id, ⁴5049251057@student.its.ac.id, ⁵aisyah.aszahra.alfani-2025@vokasi.unair.ac.id, ⁶drmuradjalalbacha@gmail.com

ABSTRACT

Smartphone notifications and device-use habits can disrupt students' concentration during academic activities. This study compares the K-Nearest Neighbors (KNN) and Logistic Regression algorithms for classifying students' perceived effectiveness of the Do Not Disturb (DND) feature. Data were collected through an online questionnaire involving 150 students at Universitas Pembangunan Nasional “Veteran” Jawa Timur in September 2026. Eleven predictor variables were used, covering study-focus duration, digital-device use, DND activation frequency, DND usage duration, frequency of activating and using DND when concentration is needed, focus duration without checking notifications, perceived assistance of DND in maintaining focus, task completion speed, task completion timeliness, and perceived productivity improvement. The target variable, perceived DND effectiveness, was categorized as ineffective (scores 1–3) or effective (scores 4–5). The dataset was divided into training and testing sets using a stratified 70:30 split, and stratified five-fold cross-validation was applied during model development. KNN achieved a mean cross-validation accuracy of 89.52% with an optimal (k) value of 17, whereas Logistic Regression achieved 82.86%. On the 45-observation test set, Logistic Regression achieved an accuracy of 84.44%, precision of 88.24%, recall of 90.91%, and F1-score of 89.55% for the effective class, with an ROC-AUC of 0.8990. KNN achieved an accuracy of 80.00%, precision of 81.58%, recall of 93.94%, and F1-score of 87.32% for the effective class, with an ROC-AUC of 0.9318. Logistic Regression demonstrated higher accuracy, precision, and F1-score, while KNN achieved higher recall for the effective class and a higher ROC-AUC. These findings suggest that Logistic Regression provided more balanced classification performance on the test set, although neither algorithm outperformed the other across all evaluation metrics. The results reflect students' perceived DND effectiveness and should be interpreted cautiously because some predictors conceptually overlap with the target variable and evaluation was based on a single train-test split.

Keywords: Do Not Disturb, K-Nearest Neighbors, Logistic Regression, machine learning, student productivity

* Corresponding author.

E-mail address: drmuradjalalbacha@gmail.com
<http://doi.org/10.33005/jasid.v2i2.76>

I. INTRODUCTION

The development of digital technology has made smartphones an important part of students' academic activities. However, notifications and the urge to check devices can disrupt concentration and hinder task completion. These distractions may arise from external stimuli, such as notifications, or internal urges to use devices [1]. In academic contexts, the use of notification-management features while completing programming tasks may increase engagement and productivity for some students [2]. These conditions highlight the importance of managing smartphone use to support students' focus and productivity.

One feature that can be used to reduce these disruptions is Do Not Disturb (DND), which limits notifications according to the user's settings. However, its effectiveness has not been consistent. DND use may increase productivity for some students [2], but hiding notifications alone does not necessarily improve learning performance because internal distractions and device-use patterns remain influential [1]. Another study reported a positive relationship between DND use and academic performance, with a correlation coefficient of 0.06 [3]. Therefore, DND effectiveness needs to be examined in relation to study-focus characteristics, device use, DND-use habits, and perceptions of focus and productivity.

These characteristics can be analyzed using machine-learning approaches, including the K-Nearest Neighbors (KNN) and Logistic Regression algorithms. Previous research found that Logistic Regression performed better than KNN when predicting computer science students' performance under the tested configuration [4]. Both algorithms have also been applied to classify student performance based on online learning activities [5]. Conversely, KNN performed better than Ordinal Logistic Regression in school-accreditation classification [6]. The application of KNN to student-performance and user classification has also been examined in other studies [7], [8]. These differing results indicate the need to compare the two algorithms across different data characteristics and classification problems.

No study directly comparing KNN and Logistic Regression to classify DND effectiveness based on device-use characteristics, study focus, and student productivity was identified among the eight references reviewed [9]. Therefore, this study aims to compare the performance of the two algorithms in classifying DND effectiveness using eleven predictor variables covering study-focus duration, digital-device use, DND-use behavior, focus without notifications, and perceptions of focus and productivity. The findings are expected to identify the algorithm with the best classification performance for predicting DND-effectiveness categories among students.

II. RESEARCH METHODS

A. Research Design

This study uses a quantitative approach with machine-learning classification to compare the performance of K-Nearest Neighbors (KNN) and Logistic Regression in predicting the effectiveness category of the Do Not Disturb (DND) feature based on eleven predictor variables. The research stages include data inspection and preprocessing, identification of predictor and target variables, division of the data into training and testing sets, model development using stratified five-fold cross-validation, and evaluation and comparison of both algorithms based on the selected classification metrics. The research workflow draws on the application of classification methods and model evaluation in previous studies [4], [5].

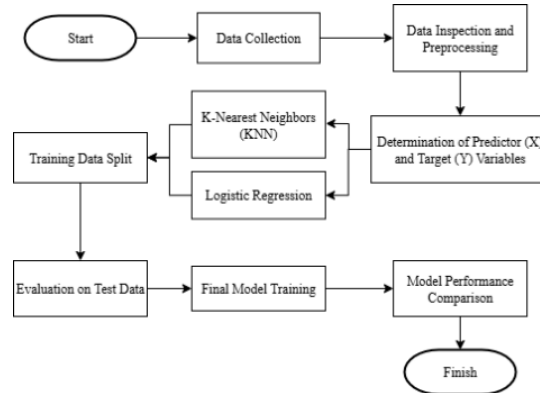


Figure 1. Research Flowchart

B. Data Collection

The research data were obtained through an online questionnaire using Google Forms, distributed to students at Universitas Pembangunan Nasional “Veteran” Jawa Timur in September 2026. The questionnaire included an informed-consent statement and yielded 150 responses with no missing values in the variables analyzed. The instrument covered digital-device use, DND-use habits, study focus, and perceptions of task completion and productivity. Perception items were measured using a five-point Likert scale, while device use was recorded numerically. Respondents’ identities were not used as predictors, and the data were checked for format, completeness, and consistency before analysis. Participation in this study was voluntary, and the collected data were kept confidential and used solely for research purposes.

C. Data Preprocessing

Data preprocessing involved standardizing column names according to the variable mapping, checking data format and completeness, and separating eleven predictor variables (X) and one target variable (Y), namely DND_Effectiveness. All 150 observations were used because there were no missing values in the variables analyzed. The data were then divided into 105 training observations and 45 testing observations. Predictor variables were standardized within the training and cross-validation pipeline to align measurement scales while preventing information leakage from validation or test data. The processed data were then used to develop the KNN and Logistic Regression models.

D. Classification

This study uses the K-Nearest Neighbors (KNN) and Logistic Regression algorithms to classify the effectiveness of the Do Not Disturb (DND) feature based on eleven predictor variables. The two algorithms use different approaches—distance between observations and class-probability estimation—and have been used in studies classifying student performance [4], [5].

1) K-Nearest Neighbors (KNN)

K-Nearest Neighbors (KNN) is a distance-based classification algorithm that determines an observation’s class based on the majority class among its nearest neighbors in the training data. Distances between observations are calculated using Euclidean distance based on standardized predictor variables. The Euclidean distance equation is expressed as follows [4], [5].

$$d(a, b) = \sqrt{\sum_{j=1}^p (a_j - b_j)^2} \quad (1)$$

Where:

$d(a, b)$ =Distance between observations a and b

a_j, b_j =Value of predictor j for an observation

p =Number of predictors

After calculating the distances, the model selects the k observations with the smallest distances and determines the predicted class through majority voting. In this study, odd values of k from 1 to 25 were tested to reduce the possibility of ties. The best k value was selected based on mean accuracy using stratified five-fold cross-validation on the training data. In each iteration, four folds were used for training and one fold for validation, until each fold had served as the validation set. The KNN model used uniform weights, so each neighbor contributed equally to class determination. After selecting the best k value, the model was retrained using the entire training set and then evaluated on the test set. The use of KNN with Euclidean distance and the selection of k through model evaluation have also been applied in previous classification studies [4], [5].

2) Logistic Regression

Binary Logistic Regression is used to estimate the probability that an observation belongs to the effective class based on eleven predictor variables. The probability of the effective class is calculated using the following logistic function [4].

$$p_i = \frac{1}{1 + \exp[-(\beta_0 + \sum_{j=1}^p \beta_j x_{ij})]} \quad (2)$$

Where:

p_i = Probability that a respondent belongs to the effective class

β_0 = Intercept

β_j = Coefficient of predictor j

x_{ij} = Value of predictor j for an observation

p = Number of predictors

The class is determined using a threshold of 0.5: effective if and ineffective if . The coefficients can be converted into odds ratios (ORs) using the following equation. $p_i \geq 0.5$ $p_i < 0.5$

$$OR_j = e^{\beta_j} \quad (3)$$

A value of indicates an increase in the odds of the effective class, whereas indicates a decrease in the odds, with the other predictors held constant. Because the predictors are standardized, this interpretation applies to a one-standard-deviation increase in the relevant predictor and does not directly establish a causal relationship. The model was implemented using scikit-learn with L2 regularization, $C=1$, and a maximum of 1,000 iterations. Standardization was performed within the pipeline during stratified five-fold cross-validation on the training data to prevent information leakage. $OR_j > 1$ $OR_j < 1$

E. Model Validation and Evaluation

Evaluation was conducted to compare KNN and Logistic Regression on the 45 test observations that were not used during model development. The evaluation used accuracy, precision, recall, F1-score, and the Area Under the Receiver Operating Characteristic Curve (ROC-AUC). These metrics have been used in previous studies to evaluate and compare classification algorithms in educational problems [4]–[6].

1) Accuracy

Accuracy measures the proportion of correct predictions among all observations and is calculated using Equation (4).

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (4)$$

2) Precision

Precision measures the proportion of predicted effective-class observations that are correct, as shown in Equation (5).

$$\text{Precision} = \frac{TP}{TP + FP} \quad (5)$$

3) Recall

Recall measures the model's ability to identify observations that actually belong to the effective class, as shown in Equation (6).

$$\text{Recall} = \frac{TP}{TP + FN} \quad (6)$$

4) F1-Score

The F1-score is the harmonic mean of precision and recall, as expressed in Equation (7).

$$\text{F1 - Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (7)$$

Where:

TP = Number of correctly predicted effective-class observations

TN = Number of correctly predicted ineffective-class observations

FP = Number of ineffective-class observations predicted as effective

FN = Number of effective-class observations predicted as ineffective

5) ROC-AUC

ROC-AUC measures the model's ability to distinguish between the effective and ineffective classes based on probability scores across different classification thresholds. ROC-AUC is calculated as the area under the ROC curve, as shown in Equation (8). In addition, baseline accuracy is used for comparison by predicting every test observation as the majority class determined from the training data. Baseline accuracy is calculated using Equation (9). The estimated proportion of the effective class, 74%, should be verified against the actual data distribution. The performance of KNN and Logistic Regression is compared using all evaluation metrics so that the assessment does not rely solely on accuracy but also considers the model's ability to distinguish between the two classes [4]–[6].

ROC-AUC equation:

$$\text{AUC} = \text{TPR}(\text{FPR})d(\text{FPR}) \quad (8)$$

Baseline accuracy equation:

$$\text{Baseline Accuracy} = \frac{n_{c^*}}{N_{\text{test}}} \quad (9)$$

III. RESULTS AND DISCUSSION

A. Research Data Characteristics

This study uses 150 observations with eleven predictor variables and one target variable, DND_Effectiveness. Data inspection showed that there were no missing values in the twelve variables analyzed. Descriptive statistics are presented in Table I to illustrate the distribution and variation of each variable.

TABLE I. DESCRIPTIVE STATISTICS OF RESEARCH VARIABLES

Variable	Mean	Std	Min	Median	Max
Study_Focus_Duration_Hours	3.13	2.88	1	2	23
Digital_Device_Hours	8.25	3.88	1	8	23
DND_Activation_Frequency	4.21	4.74	0	3	24
DND_Duration_Hours	6.71	5.69	0	6	23
Frequently_Activates_DND	3.57	1.32	1	4	5

Variable	Mean	Std	Min	Median	Max
Frequently_Uses_DND	3.77	1.24	1	4	5
Focus_Without_Notifications	5.21	4.74	0	4	23
DND_Helps_Maintain_Focus	3.99	1.02	1	4	5
Tasks_Completed_Faster	3.79	1.03	1	4	5
Tasks_Completed_On_Time	3.75	1.02	1	4	5
Productivity_Improvement	3.81	1.03	1	4	5
DND_Effectiveness	4.02	0.96	1	4	5

The target variable was then converted into two classes: ineffective (scores 1–3), labeled 0, and effective (scores 4–5), labeled 1. This grouping was adopted as an operational definition for the binary classification task in this study. Responses with scores of 1–3 were categorized as ineffective, whereas responses with scores of 4–5 were categorized as effective. This classification was based on the five-point response scale used to measure perceived DND effectiveness. However, converting the original five-point scale into two classes may reduce information about differences between individual response levels. Therefore, the classification results should be interpreted in accordance with this operational definition. The grouping showed that 111 respondents (74%) belonged to the effective class, while 39 respondents (26%) belonged to the ineffective class. This distribution is shown in Figure 2. The dominance of the effective class should be considered during evaluation so that model performance is not assessed solely by overall accuracy.

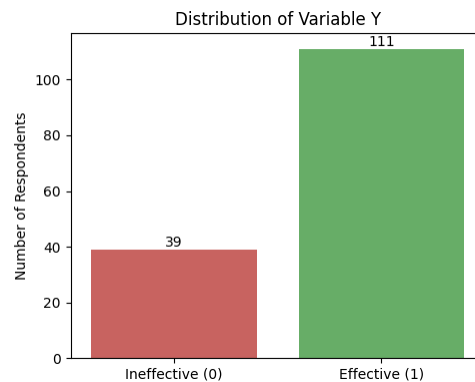


Figure 2. Distribution of DND Effectiveness Classes

The data were then divided at a 70:30 ratio using a stratified split, yielding 105 training observations and 45 test observations. Standardization was applied within each model's pipeline to align predictor scales and prevent information from the test data from entering the training process.

B. K-Nearest Neighbors Classification Results

The KNN model was developed using StandardScaler and GridSearchCV with stratified five-fold cross-validation. The candidate values of (k) were odd integers ranging from 1 to 25. The parameter search identified (k = 17) as the optimal value, achieving a mean cross-validation accuracy of 89.52%.

Figure 3 illustrates the cross-validation accuracy obtained for each candidate value of (k).

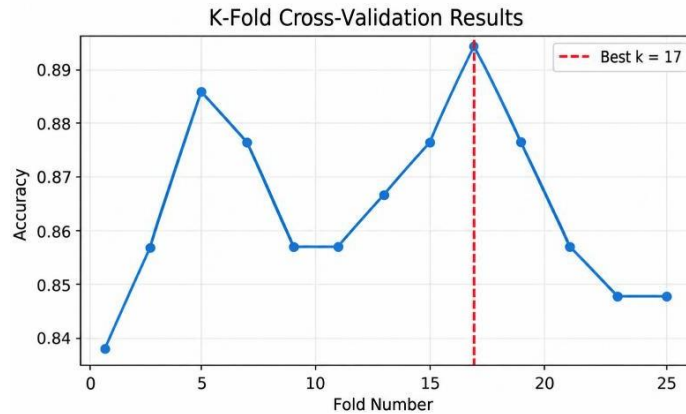


Figure 3. KNN Hyperparameter Selection

Evaluation on the 45 test observations yielded an accuracy of 80.00% and an ROC-AUC of 0.9318. Class-specific evaluation results are presented in Table II.

TABLE II. KNN CLASSIFICATION REPORT

Class	Precision	Recall	F1-Score	Support
Ineffective	0.71	0.42	0.53	12
Effective	0.82	0.94	0.87	33
Accuracy	-	-	0.80	45
Macro average	0.77	0.68	0.70	45
Weighted average	0.79	0.80	0.78	45

Table II presents the class-specific classification performance of KNN on the 45 test observations. The support values indicate that the test set contained 12 observations in the ineffective class and 33 observations in the effective class. KNN achieved a recall of 0.94 for the effective class, correctly identifying 31 of the 33 observations, but its recall for the ineffective class was only 0.42, with 5 of the 12 observations correctly identified. This difference indicates that KNN was more effective at recognizing the effective class than the ineffective class. Therefore, the overall accuracy of 0.80 should be interpreted alongside the class-specific metrics, particularly because the test set contained more effective than ineffective observations.

C. Logistic Regression Classification Results

The Logistic Regression model was developed using eleven predictor variables, with standardization within the pipeline and stratified five-fold cross-validation. Cross-validation yielded a mean accuracy of 0.8286, or 82.86%. On the test set of 45 observations, the model achieved an accuracy of 84.44% and an ROC-AUC of 0.8990. Class-specific evaluation results are presented in Table III.

TABLE III. LOGISTIC REGRESSION CLASSIFICATION REPORT

Class	Precision	Recall	F1-Score	Support
Ineffective	0.73	0.67	0.70	12
Effective	0.88	0.91	0.90	33
Accuracy	-	-	0.84	45
Macro average	0.80	0.79	0.80	45
Weighted average	0.84	0.84	0.84	45

Table III presents the class-specific classification performance of Logistic Regression on the 45 test observations. The support values show that 12 observations belonged to the ineffective class and 33

belonged to the effective class. Logistic Regression correctly identified 8 of the 12 ineffective observations, resulting in a recall of 0.67, while it correctly identified 30 of the 33 effective observations, resulting in a recall of 0.91. Compared with KNN, Logistic Regression achieved higher recall for the ineffective class, indicating a greater ability to recognize respondents in this category. However, its recall for the effective class was slightly lower than that of KNN. These results demonstrate why class-specific metrics are important when comparing classification models, rather than relying on accuracy alone.

In addition to class predictions, the Logistic Regression model produced coefficients and odds ratios for each predictor variable. The results are presented in Table IV.

TABLE IV. LOGISTIC REGRESSION COEFFICIENTS AND ODDS RATIOS

Variable	Coefficients	Odds Ratios
DND_Helps_Maintain_Focus	1.0974	2.9965
Tasks_Completed_On_Time	0.6479	1.9116
Tasks_Completed_Faster	0.5944	1.8120
Frequently_Uses_DND	0.4218	1.5247
Focus_Without_Notifications	0.3639	1.4390
Digital_Device_Hours	0.2942	1.3421
Productivity_Improvement	0.2426	1.2745
DND_Activation_Frequency	0.0739	1.0767
Frequently_Activates_DND	-0.1780	0.8370
DND_Duration_Hours	-0.4662	0.6274
Study_Focus_Duration_Hours	-0.7125	0.4904

Among the predictors, DND_Helps_Maintain_Focus had the largest positive coefficient ($\beta = 1.0974$), corresponding to an odds ratio of 2.9965. Within the fitted model, a one-standard-deviation increase in this standardized predictor was associated with approximately 3.00 times the odds of being classified in the effective category, holding the other predictors constant. Conversely, Study_Focus_Duration_Hours had a negative coefficient ($\beta = -0.7125$) and an odds ratio of 0.4904. These values describe associations within the fitted model and should not be interpreted as evidence of causal effects or statistically significant relationships, because p-values and confidence intervals for the coefficients were not reported. The results should also be considered in light of the conceptual overlap between some predictors and the target variable.

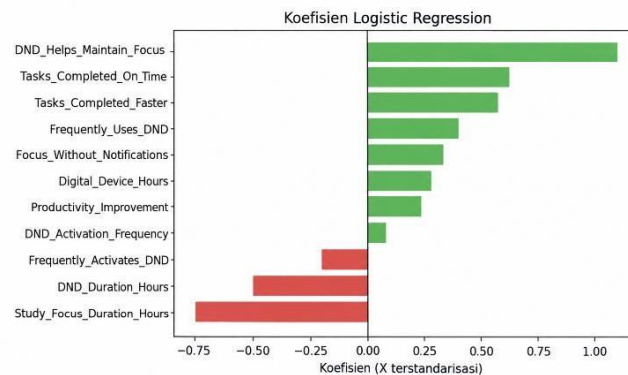


Figure 4. Logistic Regression Coefficients

D. Model Performance Comparison

The two algorithms were compared using evaluation metrics on the 45 test observations. A summary of the results is presented in Table V.

TABLE V. COMPARISON OF CLASSIFICATION PERFORMANCE

Metric	KNN (k=17)	Logistic Regression
Cross-validation accuracy	89.52%	82.86%
Accuracy	80.00%	84.44%
Precision	81.58%	88.24%
Recall	93.94%	90.91%
F1-score	87.32%	89.55%
ROC-AUC	0.9318	0.8990

The results show that Logistic Regression achieved higher accuracy, precision, and F1-score than KNN on the test data. Conversely, KNN achieved higher recall for the effective class and a higher ROC-AUC. The baseline accuracy was calculated by assigning every observation in the test set to the majority class identified in the training data. Because the effective category was the majority class, this baseline achieved an accuracy of 73.33% (33 correct predictions out of 45 test observations). Both KNN and Logistic Regression exceeded this baseline, achieving test accuracies of 80.00% and 84.44%, respectively. However, exceeding the baseline accuracy alone does not establish that a model performs well across both classes; precision, recall, F1-score, and ROC-AUC must also be considered. The difference in accuracy between the two models was 4.44 percentage points, equivalent to two additional correct predictions among the 45 test observations.

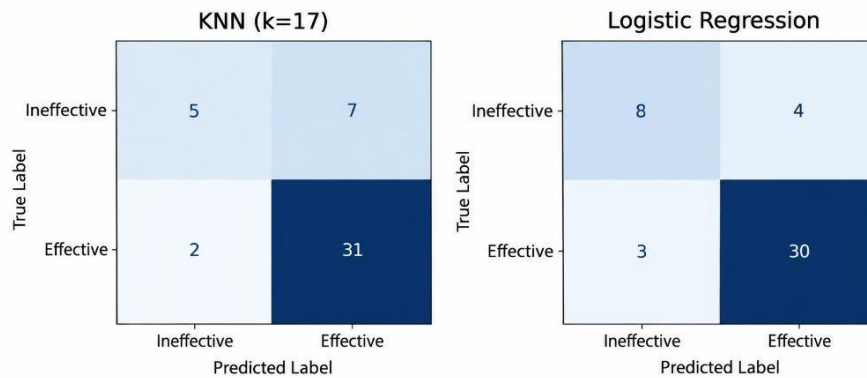


Figure 5. Confusion Matrices of KNN and Logistic Regression

The difference in the models' performance is also visible in the ROC curves in Figure 6. KNN achieved an AUC of 0.9318, whereas Logistic Regression achieved 0.8990. This indicates that KNN had greater discriminatory ability based on probability scores on the test data, despite its lower classification accuracy.

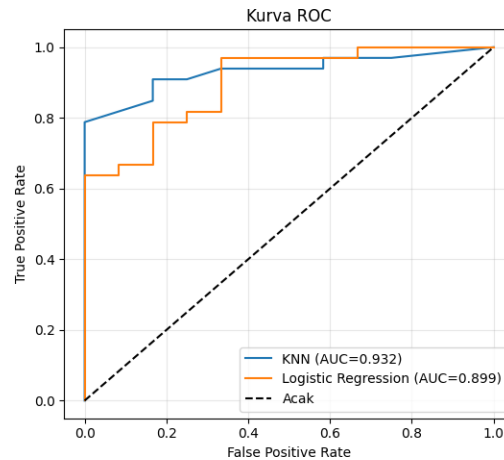


Figure 6. ROC Curves of KNN and Logistic Regression

E. Discussion

The results show that neither algorithm outperformed the other across all evaluation metrics. Logistic Regression achieved higher accuracy and F1-score and was better at identifying the ineffective class. Meanwhile, KNN achieved higher recall for the effective class and a higher ROC-AUC. This finding is consistent with Ibrahim et al., who reported that Logistic Regression outperformed KNN in classifying student performance under the configuration tested [4]. However, it differs from the study by Syukri et al., which reported that KNN performed better than Ordinal Logistic Regression in school-accreditation classification [6]. These differences confirm that algorithm performance depends on the characteristics of the dataset and the classification problem.

The Logistic Regression coefficients indicate that the perception that DND helps maintain focus had the strongest positive association with the predicted effectiveness category, followed by perceptions of completing tasks on time and more quickly. However, these findings should be interpreted cautiously because several predictors are conceptually related to the target variable, DND Effectiveness. Specifically, DND_Helps_Maintain_Focus, Tasks_Completed_On_Time, Tasks_Completed_Faster, and Productivity_Improvement capture aspects of focus and productivity that may overlap with respondents' perceptions of DND effectiveness. Consequently, the models' classification performance may partly reflect this conceptual similarity rather than their ability to predict effectiveness using clearly distinct information. This overlap does not necessarily indicate data leakage, but it limits the interpretation of the results. Furthermore, the accuracy difference between the models represents only two additional correct predictions on the 45-observation test set. Because the evaluation relied on a single train-test split, the findings do not establish that either algorithm is universally superior. Future research should use larger samples, evaluate model stability through repeated data splits, and consider predictors that are more clearly distinguished from the target construct.

CONCLUSION

This study compared the K-Nearest Neighbors (KNN) and Logistic Regression algorithms to classify the effectiveness of the Do Not Disturb (DND) feature using eleven predictor variables related to digital-device use, study focus, DND-use behavior, and students' perceived productivity. Testing on 45 observations showed that Logistic Regression achieved an accuracy of 84.44%, precision of 88.24%, and an F1-score of 89.55%, higher than KNN, which achieved 80.00%, 81.58%, and 87.32%, respectively. Conversely, KNN achieved a recall of 93.94% for the effective class and an ROC-AUC of 0.9318, higher than Logistic Regression, which achieved a recall of 90.91% and an ROC-AUC of 0.8990. Logistic Regression was also better at identifying the ineffective class, with a recall of 66.67%, compared with 41.67% for KNN. Thus, Logistic Regression showed more balanced classification performance on the test data, while KNN was better at identifying the effective class and distinguishing classes based on probability scores.

The Logistic Regression coefficients showed that the perception that DND helps maintain focus had

the largest positive coefficient, 1.0974, with an odds ratio of 2.9965. This finding indicates a positive association between this variable and the model's prediction of the effective class, but it does not establish causality. In addition, the performance difference between the algorithms should be interpreted cautiously because evaluation was conducted on a single test split containing 45 observations, and some predictors are conceptually similar to the target variable. Future research should use a larger number of respondents, evaluate model stability through repeated testing, and re-examine relationships among variables. Overall, Logistic Regression may be considered when the research priority is more balanced classification performance, whereas KNN is relevant when the main focus is identifying the effective class and ROC-AUC.

REFERENCES (10PT, IEEE STYLE)

- [1] H.-H. Liao, T.-C. Liu, and Y.-C. Lin, "External and internal smartphone distractions: Effects on phone use and learning performance in college classrooms," *Instructional Science*, vol. 54, no. 1, Art. no. 41, 2026, doi: 10.1007/s11251-026-09794-3.
- [2] K. Hart, C. M. Warren, S. Poulsen, and J. Edwards, "Phone use while programming," in *Proc. 17th Int. Conf. Educational Data Mining (EDM)*, Atlanta, GA, USA, 2024, pp. 270–278, doi: 10.5281/zenodo.12729814.
- [3] E. A. Amadi, P. C. Ogwunte, and E. C. Njoku, "Digital detoxification and academic performance of Business Education undergraduate students in Rivers State universities," *International Journal of Innovative Social Sciences & Humanities Research*, vol. 13, no. 4, pp. 281–288, Oct.–Dec. 2025, doi: 10.5281/zenodo.17601200.
- [4] S. A. Ibrahim, U. T. AbdulRauf, I. O. Mustapha, and A. A. Oni, "Comparing the efficiency of Logistic Regression Classifier and K-Nearest Neighbours Classifiers for predicting students' performance in Computer Science programme," *The Nexus (Science Edition)*, vol. 3, no. 1, pp. 64–67, Jun. 2024. [Online]. Available: https://www.researchgate.net/publication/381977891_COMPARING_THE_EFFICIENCY_OF_LOGISTIC_REGRESSION_CLASSIFIER_AND_K-NEAREST_NEIGHBOURS_CLASSIFIERS_FOR_PREDICTING_STUDENTS%27_PERFORMANCE_IN_COMPUTER_SCIENCE_PROGRAMME. [Accessed: Oct. 15, 2026]
- [5] M. A. H. Alias, M. A. A. Aziz, N. Hambali, and M. N. Taib, "Development of machine learning algorithms in student performance classification based on online learning activities," *International Journal of Electrical and Computer Engineering*, vol. 14, no. 6, pp. 7126–7136, Dec. 2024, doi: 10.11591/ijece.v14i6.pp7126-7136.
- [6] N. Syukri, Y. P. Hiola, M. R. Putri, and B. Susetyo, "School accreditation prediction based on literacy and numeracy: Ordinal Logistic Regression vs KNN," *Bulletin of Computer Science Research*, vol. 6, no. 1, pp. 491–501, Dec. 2025, doi: 10.47065/bulletincsr.v6i1.861.
- [7] E. Ahmed, "Student performance prediction using machine learning algorithms," *Applied Computational Intelligence and Soft Computing*, vol. 2024, Art. no. 4067721, 2024, doi: 10.1155/2024/4067721.
- [8] A. Huang, R. Xu, Y. Chen, and M. Guo, "Research on multi-label user classification of social media based on ML-KNN algorithm," *Technological Forecasting and Social Change*, vol. 188, Art. no. 122271, 2023, doi: 10.1016/j.techfore.2022.122271.
- [9] Z. Deng, Z. Cheng, P. Ferreira, and P. A. Pavlou, "From smartphones to smart students: Learning vs. distraction using smartphones in the classroom," *Information Systems Research*, vol. 37, no. 1, pp. 276–293, 2026, doi: 10.1287/isre.2022.0078.
- [10] S. Isak-Zatega, A. Lipovac, and V. Lipovac, "Logistic regression based in-service assessment of mobile web browsing service quality acceptability," *EURASIP Journal on Wireless Communications and Networking*, vol. 2020, Art. no. 96, May 2020, doi: 10.1186/s13638-020-01708-2.
- [11] N. Srimaneekarn, A. Hayter, W. Liu, and C. Tantipoj, "Binary response analysis using logistic regression in dentistry," *International Journal of Dentistry*, vol. 2022, Art. no. 5358602, Mar. 2022, doi: 10.1155/2022/5358602.