Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Issue 1, Januari 2026
DOI : https://doi.org/10.33395/sinkron.v10i1.15583
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Classification of Instagram and TikTok Addiction Levels among University Students Using the Naive Bayes Classifier

Indri Monica Cristiani Silalahi[1)], Ummi Athiyah[2)], Diandra Chika Fransisca[3)*]
[1,3)] Informatics Engineering Study Program, Telkom University, Purwokerto Indonesia,
[2)] Data Science Study Program, Telkom University, Purwokerto, Indonesia
[1)]18102161@ittelkom-pwt.ac.id, [2)]ummiathiyah@telkomuniveristy.ac.id, [3)]diandraf@telkomuniversity.ac.id

Submitted :Nov 14, 2025 | **Accepted** : Dec 09, 2025 | **Published** : Jan 02, 2026

Abstract: The widespread use of gadgets and internet connectivity has become an essential aspect of daily life, especially through intensive interaction with social media platforms. Excessive usage can lead to addictive behaviors that disrupt students' academic productivity and concentration. Although research on social media addiction continues to grow, few studies specifically examine platform-level addiction (Instagram vs. TikTok) using multi-class classification approaches. Therefore, this study aims to assess the level of social media addiction among university students, focusing on users of Instagram and TikTok at Telkom University Purwokerto. The analysis employs the Naive Bayes Classifier algorithm using data collected from 100 respondents. Model performance is evaluated through a multi-class confusion matrix to compute accuracy, precision, recall, and F1-score. Separate datasets for Instagram and TikTok are used to enable platform-specific behavioral assessment. The results show that the Naive Bayes Classifier achieves strong performance, with 93% accuracy for the Instagram dataset and 90% for the TikTok dataset. Precision scores reach 95% and 91%, recall values 93% and 90%, and F1-scores 93% and 90%, respectively. These findings confirm that Naive Bayes is effective for classifying students' levels of social media addiction. Overall, this research contributes a reliable machine-learning–based approach for evaluating digital behavior and provides insights for early detection, enabling universities to design targeted interventions for students at risk of problematic usage. The methodology may also be extended to analyze engagement patterns on emerging social media platforms in future studies.

Keywords: Social Media; Addiction; Naive Bayes Classifier; Instagram; TikTok

INTRODUCTION

Social media has become an essential part of modern society, serving as a medium for communication, information exchange, and digital identity development (Watt, Mitchell, & Tuke, 2024). Platforms such as Instagram and TikTok enable users to create, share, and consume content in fast and highly interactive ways (Yang, Mousavi, Dash, Gummadi, & Weber, 2025). According to the Digital 2021: Indonesia report, many social media users in Indonesia are individuals aged 18–24 years, most of whom are university students (Riyanto, 2021). This indicates that students represent one of the most active social media user groups in the country.

High-intensity usage of social media may lead to addictive behaviors as users seek gratification from the entertainment and interactive features offered by these platforms (Zhao, 2021). Several previous studies have shown that social media addiction can negatively impact on mental health, learning concentration, and students' academic productivity (Basri, Sabri, & Rahimi, 2022) (Mou, et al., 2024). Students who experience problematic usage patterns tend to spend excessive time online, reducing time allocated to academic responsibilities. This phenomenon highlights the importance of further analysis regarding student addiction levels, particularly on widely used platforms such as Instagram and TikTok.

Machine learning approaches have increasingly been used to analyze social media behavior. Applied the Naive Bayes algorithm to classify internet addiction levels and achieved an accuracy of 93% (Fikri, Elvia, Iskandar, & Syhtia, 2024). Similarly, (Sibarani, Manalu, Hutasoit, Telaumbanua, & Elyakim, 2025) utilized Naive Bayes to detect mental health tendencies based on social media activity. Additional studies about demonstrated the

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Issue 1, Januari 2026
DOI : https://doi.org/10.33395/sinkron.v10i1.15583
e-ISSN : 2541-2019
p-ISSN : 2541-044X

effectiveness of the Naive Bayes Classifier in processing behavioral and textual data from online platforms (Ahmad, Insani, & Salim, 2024) (Satria, Hasanah, & Oktaviani, 2025). Futher confirm the algorithm's capability in recognizing emotional and behavioral patterns with high accuracy (Putri, Vitianingsih, Hamidan, Maukar, & Pratitis, 2024) (Saepudin, Widiastuti, & Irawan, 2023).

This study aims to analyze the level of social media addiction among students at Telkom University Purwokerto Campus based on their usage of Instagram and TikTok using the Naive Bayes Classifier algorithm. This method is selected due to its strength in probabilistic classification and its assumption of feature independence, which has shown effective performance in similar studies (Fikri, Elvia, Iskandar, & Syhtia, 2024) (Parinduri, Dewi, & Susiani, 2022). The dataset consists of responses from 100 students, with six behavioral indicators as input variables and addiction level categories as output.

The novelty of this research lies in applying the Naive Bayes Classifier to analyze addiction levels specifically on two of the most popular platforms among Indonesian students. Moreover, a multi-class confusion matrix is used to evaluate the model's performance in terms of accuracy, precision, recall, and F1-score (Satria, Hasanah, & Oktaviani, 2025). The findings are expected to contribute to the development of early detection models for student addictive behavior and support digital literacy initiatives within higher education settings.

LITERATURE REVIEW

The rapid development of social media platforms has encouraged extensive research on behavioral patterns, psychological tendencies, and risks associated with excessive digital engagement. Platforms such as Instagram and TikTok are highly influential among young adults, especially university students, due to their highly interactive and entertainment-driven features (Yang, Mousavi, Dash, Gummadi, & Weber, 2025). These engagement-maximizing designs increase susceptibility to compulsive use, where gratification from communication and entertainment leads to addictive behaviors (Zhao, 2021). Recent studies have highlighted multiple negative consequences of such addiction: reduced focus and academic productivity (Basri, Sabri, & Rahimi, 2022), emotional dysregulation, and lowered well-being. However, most existing surveys rely solely on self-reported questionnaires, which may not accurately reflect real-life usage behavior.

To enhance the accuracy of behavioral assessment, recent research has adopted machine learning approaches. The Naive Bayes algorithm has gained attention due to its probabilistic modeling, computational efficiency, and reliable performance even with small datasets. Naive Bayes effectively classifies internet addiction levels with high accuracy (Fikri, Elvia, Iskandar, & Syhtia, 2024), while Naive Bayes utilized it to detect mental health tendencies from social media activity (Sibarani, Manalu, Hutasoit, Telaumbanua, & Elyakim, 2025). Strong predictive ability in identifying behavioral risks, such as cyberbullying and problematic usage patterns (Ahmad, Insani, & Salim, 2024) (Satria, Hasanah, & Oktaviani, 2025). These findings confirm the suitability of Naive Bayes for analyzing digital behavior in academic contexts.

Despite ongoing progress, several research gaps remain. First, most studies investigate social media addiction in a general context rather than assessing platform-specific behavioral differences, although Instagram and TikTok feature distinct interaction styles that may influence addiction levels differently. Second, research explicitly focusing on student addiction on both platforms using machine learning remains limited, despite their popularity among university students. Third, many studies apply binary classification, while addiction severity is inherently multi-class, requiring more comprehensive evaluation indicators. As noted by (Vabalas, Gowen, Poliakoff, & Casson, 2019), insufficient model evaluation, especially in small-scale datasets, can lead to overestimated performance and reduced generalizability.

In summary, existing research demonstrates the risks of social media addiction for university students and the potential of machine learning particularly Naive Bayes to classify behavioral tendencies. However, the lack of platform-specific and multi-class analysis creates an opportunity for further investigation. To address this gap, the present study applies the Naive Bayes Classifier to analyze Instagram and TikTok-based addiction levels among university students using a multi-class confusion matrix. This framework provides a more granular understanding of digital behavior and supports early detection strategies to improve digital well-being within higher education environments.

A synthesis of previous studies indicates that the use of machine learning, particularly the Naive Bayes Classifier, is effective for detecting social media behavioral trends and mental health tendencies. Recent research has successfully utilized these algorithms to validate instruments for student attitudes and identify broad social media activity patterns. However, there remains a significant gap in platform-specific multi-class assessments that directly compare addiction levels between Instagram and TikTok using a singular multi-class confusion matrix framework. Most existing studies focus on general internet addiction or single-platform analysis without exploring how algorithmic bias stemming from class imbalance in limited sample sizes affects the misclassification of high-risk users. Consequently, this research fills that void by providing a comparative performance evaluation of

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Issue 1, Januari 2026
DOI : https://doi.org/10.33395/sinkron.v10i1.15583
e-ISSN : 2541-2019
p-ISSN : 2541-044X

addiction levels across two specific platforms, addressing the impact of class distribution on accuracy, and contributing specialized insights into behavioral assessment for university students.

METHOD

The methodology of this research consisted of several stages as presented in Figure 1, beginning with the identification of the problem, formulation of objectives, literature review, data collection, implementation of the classification model, evaluation, and interpretation of results. The main issue addressed in this study is the increasing tendency of university students to excessively use Instagram and TikTok, potentially leading to social media addiction that can disrupt academic focus and productivity. Therefore, this study aims to classify the addiction levels of students at Telkom University Purwokerto using the Naive Bayes Classifier and, subsequently, evaluate the effectiveness of this machine learning approach in identifying problematic digital behavior. The results are expected to support awareness toward digital dependency and serve as a basis for preventive actions and digital literacy enhancement on campus.

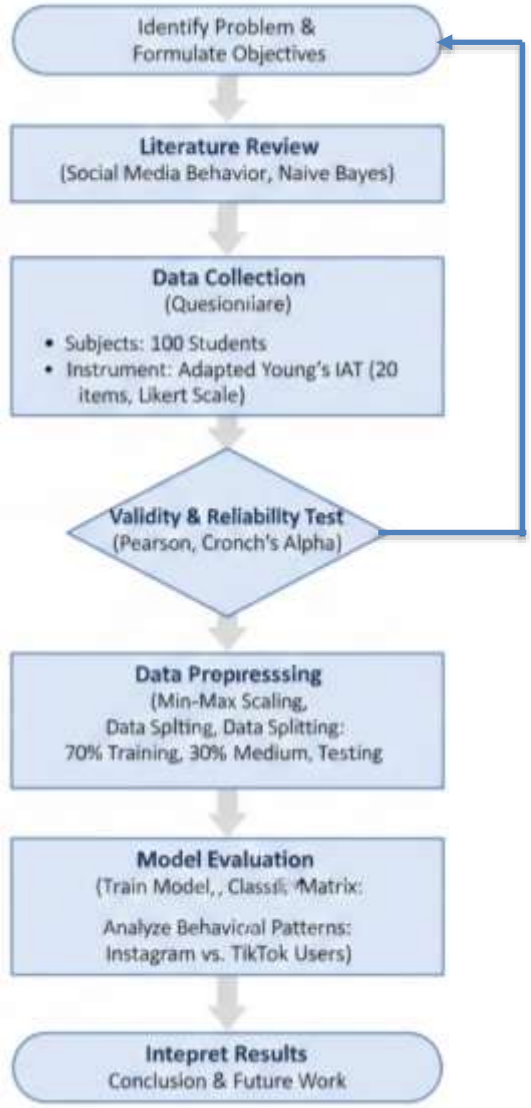

Fig. 1 Research Methodology Flowchart

The population of this study consisted of 4,012 undergraduate students from the 2018–2021 cohorts. The minimum number of respondents was calculated using Slovin's formula $n = \frac{N}{1+N(e)^2}$ with a 10% margin of error, resulting in a minimum sample size of 100 students. Data was collected using a questionnaire adapted from Young's Internet Addiction Test as well as indicators proposed by Kuss & Griffiths (Mou, et al., 2024). The

*name of corresponding author

e-ISSN : 2541-2019
p-ISSN : 2541-044X

instrument contained 20 items assessing Instagram and TikTok usage behavior using a five-point Likert scale ranging from 1 (rarely) to 5 (always). Question about Internet Addiction is:
Q1. How much Do you often feel like you 're online longer than you want to?

Q2. How often do you ignore workhouse stairs to spend more online time?

Q3. How often do you choose the excitement of the internet than intimacy / relationship with partner / friend?

Q4. How often do you make new connections with fellow internet users?

Q5. How often do other people in your life complain about how much time do you spend online?

Q6. How much often mark or work your school is declining Because too Lots how much time do you spend online?

Q7. How much how often do you check your email before doing anything else you need to do?

Q8. How often does performance or productivity decreasing because of the Internet?

Q9. How much how often do you behave defensively or closely when somebody asks what do you do online?

Q10. How often do you divert disturbing thoughts about your life with calming thoughts from the internet?

Q11. How much often you anticipate When will you be online again?

Q12. How often are you afraid that life without internet will be boring, empty, and not pleasant?

Q13. How often do you shout, scream, or behave annoyed If is there something bothering you online?

Q14. How much often you lack Sleep because login is late Evening?

Q15. How often do you feel fun with the internet when offline, or fantasizing about are online?

Q16. How often do you say ‘stay a number of minutes again " while online?

Q17. How often do you try reducing time spent online and failing?

Q18. How often do you try to hide how long you are online?

Q19. How much how often do you spend more Lots Of online time than go out with other people?

Q20. How often do you feel depression, gloom, or flustered when offline, what's missing after you get back online?

The total score obtained by each respondent was used to categorize addiction levels, where scores of 20–40 indicate low addiction, 41–60 indicate moderate addiction, and 61–100 indicate high addiction. These cut-off ranges were determined logically from the Likert-scale distribution and prior classifications in Internet addiction studies.

Instrument validity was tested using Pearson Product-Moment; items are declared valid if r-count > r-table. Reliability was assessed using Cronbach’s Alpha, where $\alpha \geq 0.70$ indicates acceptable consistency (Setiawan, Cendana, Ayres, Yuldashev, & Setyawati, 2023). All retained items met the minimum validity and reliability requirements.

The model was implemented using Python in Jupyter Notebook. The key steps include:

1. Dataset Loading: importing students’ social-media usage data
2. Attribute Selection: selecting relevant behavioral features
3. Data Preprocessing: normalization using Min-Max scaling
4. Data Splitting: 70% training and 30% testing
5. Model Training & Prediction: applying the Naive Bayes algorithm

To reduce bias and ensure robust generalization, k-fold cross-validation was applied with k = 10. The model performance was reported as the average of all folds.

Model evaluation was conducted using a multi-class confusion matrix to assess how well the Naive Bayes algorithm classified students into low, moderate, and high addiction levels. Four performance measures were calculated, including accuracy to represent the overall proportion of correct predictions, precision to indicate the correctness of predictions in each category, recall measuring the ability of the model to detect each class correctly, and the F1-score as the harmonic means of precision and recall providing a balanced performance indicator. Together, these metrics provide a comprehensive evaluation of the classifier’s ability to identify and distinguish varying levels of social media addiction among students.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Issue 1, Januari 2026
DOI : https://doi.org/10.33395/sinkron.v10i1.15583
e-ISSN : 2541-2019
p-ISSN : 2541-044X

RESULT

Implementation of Naive Bayes Classifier using Language Python programming and tools soft Jupiter Notebook. Steps its implementation is as following:

Research Dataset

The data in this study are Instagram and TikTok social media usage data among students at Telkom University, Purwokerto Campus, obtained through questionnaires. A total of 100 respondents completed the questionnaire. Furthermore, the validity and reliability of the Instagram and TikTok items in the questionnaire were tested. Validity and reliability tests in this study used the SPSS version 25 application.

Validity Test

For seem to what extent is the accuracy and precision something tool measuring in operate function measurement, used amount sample (n) = 100 respondents, so that r table value is 0.195. Validity test results Research on Instagram and TikTok questions are shown in Table 1 and Table 2.

Table 1. Results Of Instagram Validity Test

Question	amount *r*	table *r*	Information
Q1	0.662	0.195	Valid
Q2	0.726	0.195	Valid
Q3	0.808	0.195	Valid
Q4	0.712	0.195	Valid
Q 5	0.789	0.195	Valid
Q 6	0.880	0.195	Valid
Q 7	0.628	0.195	Valid
Q 8	0.806	0.195	Valid
Q 9	0.823	0.195	Valid
Q 10	0.681	0.195	Valid
Q11	0.819	0.195	Valid
Q 12	0.838	0.195	Valid
Q13	0.876	0.195	Valid
Q14	0.718	0.195	Valid
Q15	0.875	0.195	Valid
Q 16	0.845	0.195	Valid
Q 17	0.770	0.195	Valid
Q18	0.859	0.195	Valid
Q19	0.867	0.195	Valid
Q20	0.877	0.195	Valid

Table 2. Tiktok Validity Test Results

Question	amount *r*	table *r*	Information
Q1	0.785	0.195	Valid
Q2	0.857	0.195	Valid
Q3	0.872	0.195	Valid
Q4	0.837	0.195	Valid
Q 5	0.833	0.195	Valid
Q 6	0.870	0.195	Valid
Q 7	0.820	0.195	Valid
Q 8	0.896	0.195	Valid
Q 9	0.893	0.195	Valid
Q 10	0.749	0.195	Valid
Q11	0.903	0.195	Valid
Q 12	0.895	0.195	Valid
Q13	0.906	0.195	Valid
Q14	0.778	0.195	Valid
Q15	0.922	0.195	Valid
Q 16	0.815	0.195	Valid
Q 17	0.819	0.195	Valid
Q18	0.867	0.195	Valid
Q19	0.912	0.195	Valid
Q20	0.870	0.195	Valid

*name of corresponding author

e-ISSN : 2541-2019
p-ISSN : 2541-044X

The validity test results (Table 1) indicate that all questionnaire items have correlation values greater than the required threshold ($p < 0.05$), confirming that each item is valid and appropriately represents the measured variables. In addition, the reliability assessment (Table 2) shows that the instrument achieves a Cronbach's Alpha value above 0.70, demonstrating strong internal consistency. Therefore, the questionnaire is considered both valid and reliable for measuring the level of social media addiction among students at Telkom University Purwokerto Campus.

Reliability Test

Reliability test uses Cronbach's Alpha formula. Reliability test results for Instagram and TikTok questions can see in Table 3 and Table 4.

Table 3. Reliability Test Results For Instagram

Cronbach's Alpha	Information
0.969	Can reliable

Table 4. Reliability Test Results For TikTok

Cronbach's Alpha	Information
0.981	Can reliable

Table 3 presents the reliability results for the Instagram addiction questionnaire, which achieved Cronbach's Alpha of 0.969. Since this value exceeds the minimum threshold of 0.60, the instrument is considered highly reliable. Similarly, Table 4 shows that the TikTok questionnaire obtained a Cronbach's Alpha of 0.981, also above the reliability standard. These findings confirm that both instruments demonstrate excellent internal consistency and are suitable for use in further classification analysis using the Naive Bayes algorithm.

The dataset consisted of 100 respondents. Tables 5 and 6 display sample records for Instagram and TikTok, respectively. Each dataset includes students' responses to the 20 validated questionnaire items measured using a 5-point Likert scale (1 = rarely to 5 = always). Total scores were then categorized into three levels of addiction: low, moderate, and high. These labeled data served as the input for training and testing the Naive Bayes classification model.

Table 5. Research Dataset for Instagram

Name	Student ID Number	Q 1	Q 2	…	Q2 0	Information
Moh. Aminullah Al Fachri	18102166	4	2	…	2	Low
Nike Putra K. K Berutu	18102168	3	2	…	3	Low
Neo Mauizan Ali Fitrah	20102090	4	4	…	4	Heavy
Raffika Hanum	18101027	3	3	…	3	Currently
Atika Rahmadani Utami Br Ginting	18103099	2	2	…	1	Low

Next, Table 6 presents the research dataset for the TikTok platform, which has the same structure as the Instagram dataset. Respondents answered 20 questions using a five-point Likert scale to measure their intensity and motivation in using TikTok. The classification results show variations in addiction levels (low, moderate, and high). This dataset is used to compare cross-platform addiction levels and to evaluate the performance of the Naive Bayes Classifier in identifying user behavior patterns. Through this comparison, the research highlights differences in social media addiction patterns between photo-based (Instagram) and short-video-based (TikTok) platforms.

Table 6. Research Dataset for TikTok

Name	Student ID Number	Q 1	Q 2	…	Q2 0	Information
Moh. Aminullah Al Fachri	18102166	2	2	…	3	Low
Nike Putra K. K Berutu	18102168	4	4	…	4	Currently

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Issue 1, Januari 2026
DOI : https://doi.org/10.33395/sinkron.v10i1.15583
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Neo Mauizan Ali Fitrah	20102090	4	4	…	4	Heavy
Raffika Hanum	18101027	3	3	…	3	Currently
Atika Rahmadani Utami Br Ginting	18103099	2	2	…	1	Low

The data will be analyzed use Naive Bayes Classifier method. Naive Bayes Classifier analysis begins with data selection, namely election attributes that will be used for analysis. Details used in analysis level social media addiction chosen based on possibility existence patterns that are significant influence social media addiction students. Therefore that, other details such as Name, NIM, Class, and Major No used at the stage This.

Implementation Results

In the Python implementation of the Naive Bayes Classifier, the Description attribute was converted into numeric labels: 0 for high, 1 for low, and 2 for medium. The dataset was then split into 70% training data (70 samples) and 30% test data (30 samples). Model evaluation was performed using a multiclass confusion matrix to calculate accuracy, precision, recall, and F1-score, as shown in Figure 2 and 3.

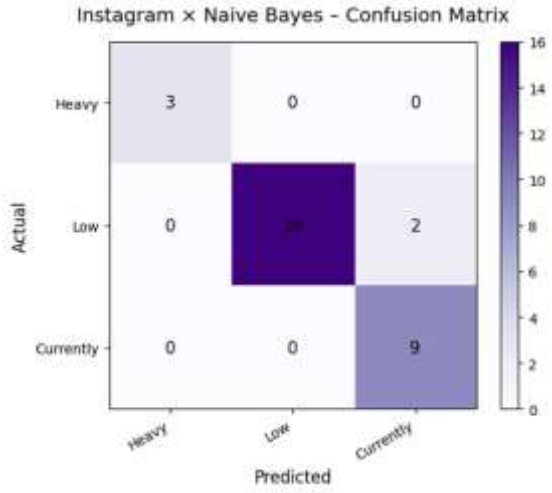

Fig 2. Confusion Matrix Results of Instagram and Naive Bayes

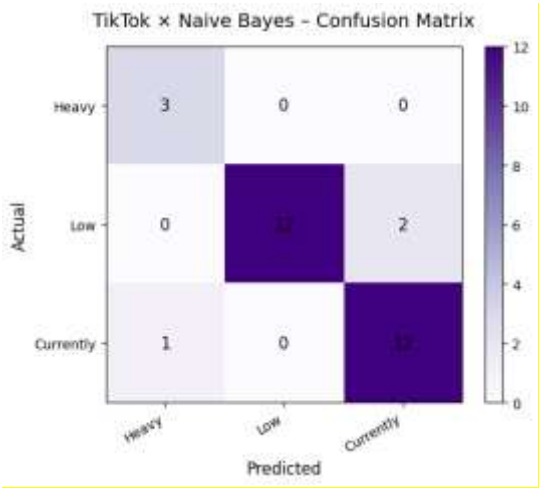

Fig 3. Confusion Matrix Results of TikTok and Naive Bayes

Figure 2 and 3 shows the classification outcomes for each class. In the Instagram dataset, many prediction errors occur in the Moderate class, where two Moderate cases were misclassified as Low. This indicates that the boundary between Low and Moderate addiction is less distinguishable. In contrast, the Heavy class was classified perfectly (TP = 3, FP = 0). For the TikTok dataset, the most errors also appear in the Heavy and Moderate classes,

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Issue 1, Januari 2026
DOI : https://doi.org/10.33395/sinkron.v10i1.15583
e-ISSN : 2541-2019
p-ISSN : 2541-044X

where one Heavy user was predicted as Moderate and two Moderate users were predicted as Heavy. This suggests that user behavior in higher addiction categories tends to overlap, making classification slightly more challenging.

The Instagram model achieved an accuracy of 93%, while the TikTok model achieved 90%. Precision values for each class are shown in Table 7.

Table 7. Precision Results for Each Class

Precision	Algorithm Naive Bayes Classification		
	Heavy	Low	Currently
Instagram	1.00	1.00	0.82
TikTok	0.75	1.00	0.86

Precision results are presented in Table 7. Instagram yields weighted precision of 95%, higher than TikTok (91%). High precision in the Heavy and Low categories on both platforms indicates that the classifier rarely assigns these labels incorrectly. However, precision for the Moderate class is lower (Instagram = 0.82; TikTok = 0.86), again confirming that this class is the most difficult to classify.

The data has a class imbalance, meaning there are many more students in the 'Low' addiction group than in the 'Heavy' group. Because Naive Bayes works by looking at how common a group is, it might prioritize predicting students as 'Low' addiction to get a higher overall score. This explains why the model is very accurate overall but occasionally misses students with 'Heavy' addiction. Future studies should use techniques like SMOTE (Synthetic Minority Over-sampling Technique) to create more balance between these groups, ensuring high-risk students are detected more effectively.

Table 8. Recall Results for Each Class

Remember	Algorithm Naive Bayes Classification		
	Heavy	Low	Currently
Instagram	1.00	0.89	1.00
TikTok	1.00	0.86	0.92

Recall values (Table 8) show a similar pattern. Instagram achieves weighted recall of 93%, while TikTok achieves 90%. Moderate TikTok users have lower recall (0.86), meaning some actual Low users are incorrectly detected as Moderate. This aligns with the confusion-matrix observations that misclassification tends to occur between adjacent addiction levels.

Table 9 shows that the weighted F1-score reaches 93% for Instagram and 90% for TikTok, indicating a good balance between precision and recall across classes. Although TikTok performance is slightly lower, both models remain consistent, confirming that Naive Bayes is stable and reliable even with a small sample size.

Table 9. Results Of F-1 scores For Each Class

F-1 score	Algorithm Naive Bayes Classification		
	Heavy	Low	Currently
Instagram	1.00	0.94	0.90
TikTok	0.86	0.92	0.89

From the final classification distribution of 100 respondents, Instagram users were grouped into 58 Low, 30 Moderate, and 12 High addiction levels. Meanwhile, TikTok users were 52 Low, 30 Moderate, and 18 High. These differences highlight distinct behavioural patterns between platforms: TikTok has more users in the High-addiction category, while Instagram users tend to fall into the Low category. This result supports the tendency that short-video platforms may trigger stronger compulsive usage than photo-based platforms.

Overall, the evaluation confirms that the Naive Bayes Classifier provides accurate, consistent, and informative classification of social media addiction, while also revealing meaningful differences in student addiction behaviour between Instagram and TikTok.

DISCUSSIONS

The findings of this study demonstrate that the Naive Bayes Classifier performs effectively in classifying social media addiction levels among Telkom University Purwokerto students, achieving 93% accuracy for Instagram

*name of corresponding author

data and 90% accuracy for TikTok data. This difference suggests that behavioral patterns on Instagram may be more consistent and predictable compared to TikTok, which provides highly dynamic and personalized content recommendations. As a result, user behavior on TikTok tends to be more variable, influencing model performance.

When compared with previous studies, strong predictive performance of Naive Bayes for detecting addictive behaviors and mental health tendencies using digital behavioral indicators (Fikri, Elvia, Iskandar, & Syhtia, 2024). The accuracy obtained in this research falls within, or slightly exceeds, the performance range observed in earlier works, reinforcing the relevance of probabilistic modeling for student behavior analysis.

Furthermore, consistently high precision, recall, and F1-scores indicate that the classifier performs well not only in overall accuracy but also in distinguishing among different addiction levels. This implies a balanced model that can serve as a practical tool for educational institutions to detect risky digital behavior in increasingly digital learning ecosystems.

The use of a multi-class confusion matrix provides a more comprehensive view of the model's classification performance compared to relying solely on accuracy. It was observed that the "high" and "moderate" addiction categories show lower misclassification rates than the "low" category. This may occur because students with intense or moderate dependency display more distinguishable behavioral characteristics than users with minimal interactions.

Additionally, behavioral attributes such as frequency of use, duration, and emotional dependence proved to be meaningful predictors of addiction tendencies. This highlights opportunities to enrich future models with psychological variables or objective digital activity logs to further improve prediction performance.

Limitations and Future Work

Despite the promising results, several limitations should be acknowledged to avoid overestimating the model's reliability:

1. Single sampling location: Data was collected from one university only, limiting generalizability to broader student populations.
2. Small sample size (n = 100): small datasets may inflate reported performance, particularly in classification tasks (Vabalas, Gowen, Poliakoff, & Casson, 2019).
3. Self-reported questionnaire: Subjective responses may introduce bias and reduce measurement reliability.
4. Lack of algorithmic comparison: Although prior studies were referenced, this work did not empirically compare Naive Bayes with other models such as Random Forest, Support Vector Machines, or deep learning methods.

Future research should therefore expand the dataset across multiple institutions, incorporate more objective behavioral metrics, and conduct benchmarking with various machine learning algorithms to identify the optimal approach for detecting social media addiction.

CONCLUSION

The results of this study confirm that the Naive Bayes Classifier is effective in classifying social media addiction levels among Telkom University Purwokerto students. The model achieved high performance, with 93% accuracy for Instagram and 90% accuracy for TikTok, supported by precision, recall, and F1-scores consistently above 90%. These findings demonstrate that the algorithm can successfully identify behavioral patterns associated with addictive use and convert subjective questionnaire responses into meaningful quantitative indicators for addiction level classification.

From a practical perspective, the findings suggest that machine learning–based approaches have strong potential to be implemented as supportive tools for educational institutions. Universities, academic advisors, and counseling services may utilize such predictive systems to assist in early detection of risky digital behavior, enabling timely monitoring and targeted interventions. By integrating this model into digital literacy programs, campuses can promote healthier technology use and enhance student well-being in increasingly online learning environments.

Future work is recommended to improve model generalization and broaden applicability. Studies should include larger and more diverse samples from multiple institutions, incorporate additional behavioral or psychological variables, and make empirical comparisons with alternative algorithms such as Random Forest, Support Vector Machines, and K-Nearest Neighbor. These enhancements are expected to produce a more robust and adaptive predictive model that can better capture evolving patterns of social media use in the rapidly developing digital era.

*name of corresponding author

e-ISSN : 2541-2019
p-ISSN : 2541-044X

REFERENCES

Ahmad, S., Insani, N., & Salim, M. (2024). Analysis of Cyberbullying on Social Media Using A Comparison of Naive Bayes, Random Forest, and SVM Algorithms. *Jurnal Teknologi Informasi dan Pendidikan, 17*(1), 75-86. doi:https://doi.org/10.24036/jtip.v17i1.807

Basri, F. N., Sabri, F., & Rahimi, M. K. (2022). Social Media Addiction and Academic Performance of University Students. *The Malaysian Journal of Islamic Sciences, 34*(2), 1-17.

Fikri, M. I., Elvia, B., Iskandar, I., & Syhtia, E. P. (2024). Klasifikasi Tingkat Kecanduan Internet Terhadap Remaja Pekanbaru Melalui Pendekatan Algoritma Naive Bayes. *ZONAsi: Jurnal Sistem Informasi, 6*(2), 424-436. doi:https://doi.org/10.31849/zn.v6i2.20191

Mou, Q., Zhuang, J., Wu, Q., Zhong, Y., Dai, Q., Cao, X., . . . Zhao, M. (2024). Social media addiction and academic engagement as serial mediators between social anxiety and academic performance among college students. *BMC Psychology, 12*(1), 190. doi:https://doi.org/10.1186/s40359-024-01635-7

Parinduri, F. Z., Dewi, R., & Susiani, S. (2022). Classification of internet addiction levels in students using the Naive Bayes algorithm. *JOMLAI: Journal of Machine Learning and Artificial Intelligence, 1*(3), 257-264. doi:https://doi.org/10.55123/jomlai.v1i3.965

Putri, V. A., Vitianingsih, A. V., Hamidan, R., Maukar, A. L., & Pratitis, N. T. (2024). Sentiment analysis on social media Instagram of depression issues using Naive Bayes method. *INOVTEK Polbeng – Informatics Series, 9*(2), 802-813. doi:https://doi.org/10.35314/spchsk42

Riyanto, A. D. (2021, August 18). *Hootsuite (We are Social): Indonesian Digital Report 2021*. Retrieved from Digital 2021 Indonesia: https://andi.link/hootsuite-we-are-social-indonesian-digital-report-2021/

Saepudin, S., Widiastuti, S., & Irawan, C. (2023). Analisis Sentimen Ulasan Platform Media Sosial Menggunakan Algoritma Naive Bayes Classifier. *Journal Sisfokom (System Information and Computers), 12*(2), 236-243. doi:10.32736/sisfokom.v12i2.1650

Satria, A. T., Hasanah, H., & Oktaviani, I. (2025). Predicting Social Media Addiction Using Machine Learning and Interactive Visualization with Streamlit. *Journal Bit-Tech, 8*(1), 778-788. doi: https://doi.org/10.32877/bt.v8i1.2715

Setiawan, A., Cendana, W., Ayres, M., Yuldashev, A. A., & Setyawati, S. P. (2023). Development and validation of a self-assessment-based instrument to measure elementary school students' attitudes in online learning. *REID (Research and Evaluation in Education), 9*(2), 184-197. doi:https://doi.org/10.21831/reid.v9i2.52083

Sibarani, J., Manalu, R., Hutasoit, D. P., Telaumbanua, W. A., & Elyakim, V. A. (2025). Detection of mental health tendencies using Naive Bayes based on social media activity. *JOMLAI: Journal of Machine Learning and Artificial Intelligence, 4*(2), 80-87. doi:https://doi.org/10.55123/jomlai.v4i2.5959

Vabalas, A., Gowen, E., Poliakoff, E., & Casson, A. J. (2019). Machine learning algorithm validation with a limited sample size. *PLoS ONE, 14*(11), e0224365. doi: https://doi.org/10.1371/journal.pone.0224365

Watt, J., Mitchell, L., & Tuke, J. (2024). Personality Profiling: How informative are social media profiles in predicting personal information? *arXiv*. doi:https://doi.org/10.48550/arXiv.2309.13065

Yang, C., Mousavi, S., Dash, A., Gummadi, K. P., & Weber, I. (2025). Studying Behavioral Addiction by Combining Surveys and Digital Traces: A Case Study of TikTok. *arXiv*. doi:https://doi.org/10.48550/arXiv.2501.15539

Zhao, L. (2021). The impact of social media use types and social media addiction on subjective well-being of college students: A comparative analysis of addicted and non-addicted students. *Computers in Human Behavior Reports, 4*(2), 100122. doi: 10.1016/j.chbr.2021.100122

*name of corresponding author