

Application of the Apriori Algorithm to Determine the Correlation of Tiktok Social Media use with the Learning Concentration of Senior High School Students

Feby Ria Syahpitri Br Ginting^{1*}, Marto Sihombing², Tioria Pasaribu³

^{1,2,3}Program Studi Sistem Informasi, Sekolah Tinggi Manajemen Informatika Dan Komputer Kaputama
febyriasyahpitribrginting@gmail.com^{1*}, martosihombing45@gmail.com², pasaributioria@gmail.com³

Abstract

Learning concentration is a crucial factor that supports the effectiveness of the learning process. Yet, the intensive use of short-video social media such as TikTok among students potentially disrupts their focus. At SMA/SMK Bina Bersaudara, students exhibit a high intensity of TikTok use, while the available student data has never been analyzed to reveal its relationship with learning concentration. This study aims to apply the Apriori algorithm, an association rule mining method in data mining, to identify patterns of relationship among student data, TikTok usage indicators, and learning concentration. The data were collected through questionnaires distributed to thirty students, transformed into binary form, and tested using RapidMiner with a minimum support of 20% and a minimum confidence of 80%. The analysis produced one valid association rule: students who open TikTok one to three times a day, most often at midday, and usually study for less than one hour per session tend to concentrate well, with a confidence level of 88.9%. The results indicate that the Apriori algorithm is able to reveal consistent relationship patterns between TikTok usage habits and learning concentration, which can serve as a data-based reference for schools, teachers, and parents in supervising social media use.

Keywords: Apriori; Association Rule; Data Mining; Learning Concentration; TikTok

1. Introduction

TikTok is a short-video social media platform whose growth has been extremely rapid and which has become highly popular among adolescents and students. Its function is not limited to entertainment; it also serves as a medium for communication, information sharing, and self-expression that contributes to shaping the interaction patterns and digital habits of its users. In the context of education, the use of TikTok among students shows potential as an informal learning medium through educational content, while at the same time strengthening digital literacy. Nevertheless, intensive use without control has the potential to cause negative impacts, such as decreased learning focus, a tendency toward addiction, disrupted sleep patterns, and imitative behavior toward viral content [1][7].

SMA/SMK Bina Bersaudara is a school located on Jl. Letjend Jamin Ginting, Kuta Parit Village, Selesai District, Langkat Regency. Based on initial observations conducted at the school, it is known that most students have a high intensity of social media use, particularly TikTok, which potentially disrupts their concentration and learning focus. However, until now, there has been no data mining-based analysis that examines the relationship between TikTok use and students' learning concentration at school, and the student data has not been utilized optimally as a basis for decision-making and educational policy [2][9].

Learning concentration is the psychological ability of an individual to focus attention continuously on learning material within a certain period of time, and it plays an important role in supporting the effectiveness of the learning process and the achievement of academic performance. When learning concentration is not well established, students tend to be easily distracted, become less productive in following the learning process, and are at risk of declining learning outcomes. In the digital era, the use of social media has become one of the external factors that potentially influences students' learning concentration, especially when used for a long duration and with high frequency [2].

Based on this need, an analytical, data-based approach is required to process and analyze data on social media use and students' learning concentration. One approach that can be used is data mining with the Apriori algorithm, namely a technique for discovering items that frequently appear together (frequent itemsets) and developing them into association rules based on support and confidence values. In the field of education, Apriori has been used to reveal the relationships among learning factors, with confidence values of around 70–80% [3][10].

Several previous studies serve as references for this research. A previous study examined the relationship between the intensity of TikTok use and learning concentration using a quantitative correlational approach and found a significant negative relationship; however, that study did not apply data mining methods to trace more detailed relationship patterns among indicators [4]. Another study explained the influence of TikTok use on learning behavior using a descriptive qualitative approach. Still, the results did not produce a quantitative pattern model that shows the combinations of conditions that most frequently occur together [2]. Meanwhile, the association rule mining approach with an Apriori-type algorithm has been shown to be effective for extracting relationship patterns from student behavior data, with confidence values ranging from 91% to 97% [5].

Therefore, this study is important to conduct as a way of combining these findings by applying the Apriori algorithm to identify more specific correlation patterns between TikTok usage indicators and the learning concentration of students at SMA/SMK Bina Bersaudara in a more objective and data-based manner. This approach not only tests the relationship in general but also maps the combinations of conditions that most frequently occur together and have the strongest association, so that the resulting association rules can be interpreted practically [6][8].

2. Page layout

This study uses the Apriori algorithm, one of the association rule mining methods in data mining, applied through a systematic and structured research procedure. The research method comprises the stages of data collection, information analysis, and the effort to achieve the research objectives so that the resulting findings are valid and can be scientifically accountable. The stages carried out in this study are shown in Figure 1.

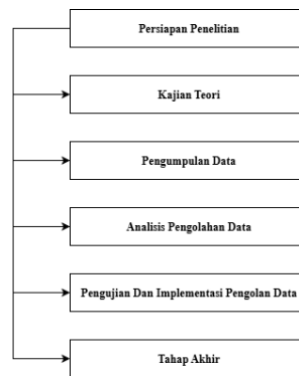


Fig. 1: Research workflow

The description of each stage of the research method used is as follows:

1. Research preparation, namely the initial planning stage that includes formulating the research questions and objectives, determining the population and sample, preparing the instrument, and arranging the research work plan.
2. Literature study, namely a review of relevant theories, previous research, and supporting findings, as the conceptual basis of the research.
3. Data collection, namely obtaining the required data through questionnaires distributed to students of SMA/SMK Bina Bersaudara.
4. Data processing and analysis, namely processing and analyzing the data to discover patterns, relationships, and findings that support the research objectives.
5. Testing and implementation of data processing, namely the evaluation and application stage of the designed method through performance testing and effectiveness analysis.
6. Preparation of the research report, namely the final stage of compiling a report that summarizes the research process, analysis results, and interpretation of the findings.

In the Apriori algorithm, two important concepts are support and confidence. The support value indicates the proportion of transactions that contain a particular combination of items. At the same time, confidence describes how strong the relationship is between items within an itemset, that is, how often item B appears when item A appears. An association rule is generally expressed in the form of *if x then y*, where x is the antecedent (initial condition) and y is the consequent (result). The support and confidence values are calculated using the following equations:

$$\text{Support}(A) = \left(\frac{\text{Number of Transactions containing } A}{\text{Total Transactions}} \right) \times 100\% \quad (1)$$

$$\text{Support}(A, B) = \left(\frac{\text{Number of Transactions containing } A \text{ and } B}{\text{Total Transactions}} \right) \times 100\% \quad (2)$$

$$\text{Confidence}(B|A) = \left(\frac{\text{Number of Transactions containing } A \text{ and } B}{\text{Number of Transactions containing } A} \right) \times 100\% \quad (3)$$

The steps for applying the Apriori algorithm in this study are illustrated in the flowchart shown in Figure 2.

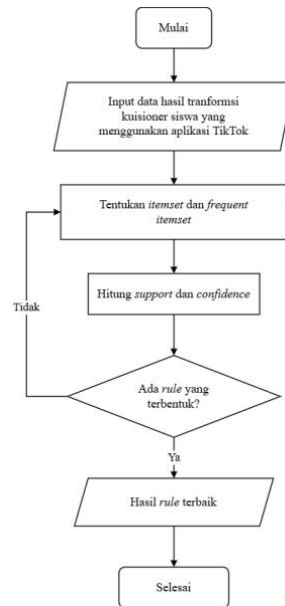


Fig. 2: Flowchart of the Apriori algorithm

The flowchart begins by inputting the transformed questionnaire data, followed by determining the itemsets and frequent itemsets, calculating the support and confidence values, and checking whether valid association rules are formed. If no rule meets the minimum threshold, the process returns to the previous stage; otherwise, the best rules are obtained, and the process ends.

3. Page style

3.1. Research supporting data

The data used in this study were obtained from questionnaires distributed to students of SMA/SMK Bina Bersaudara, grades X–XII, for the 2025–2026 academic year, with a total of thirty data records used as a sample. The data include student data, TikTok usage indicators, and learning concentration indicators. Each questionnaire answer was categorized and coded for data transformation, as presented in Table 1.

Table 1: Coding of research variables

Code	Regular	Bold
Class	Grade 10 / 11 / 12	KL1–KL3
Duration	TikTok usage duration per day	DT1–DT4
Frequency	Frequency of opening TikTok per day	FT1–FT4
Time	Most frequent time of use	WT1–WT4
While studying	Opening TikTok while studying	ST1–ST4
Content	Most frequently watched content	KT1–KT4
Education	Frequency of watching educational content	KE1–KE4
Concentration	Learning concentration indicators	KS1–KS6

After the coding process, all respondents' answers were transformed into binary form (0 and 1) to determine the number of occurrences of each item. A value of 1 indicates that the respondent selected a category, whereas 0 indicates that it was not selected. This binary form becomes the input for the calculation of support and confidence.

3.2. Formation of frequent itemsets

The initial stage of the Apriori algorithm is the formation of frequent itemsets by calculating the support value of each item combination and selecting combinations that satisfy the minimum support of 20%. The process begins with one-itemsets and continues to two-, three-, and four-itemsets. Part of the one-itemset support calculation that meets the threshold is shown in Table 2.

Table 2: Font Specifications for A4 Papers

One-itemset	Count	Support
KL2	7/30	23.3%
KL3	20/30	66.7%
DT1	8/30	26.7%
DT2	9/30	30.0%
DT3	12/30	40.0%
FT1	20/30	66.7%
FT2	8/30	26.7%

One-itemset	Count	Support
WT2	17/30	56.7%
WT3	6/30	20.0%
WT4	7/30	23.3%
ST1	10/30	33.3%
ST2	14/30	46.7%
KT1	6/30	20.0%
KT3	21/30	70.0%
KE3	14/30	46.7%
KE4	10/30	33.3%
KS4	12/30	40.0%
KS5	22/30	73.3%
KS6	16/30	53.3%

The formation of itemset combinations stopped at the fourth iteration because no five-itemset combination satisfied the minimum support. From the four-itemset combinations, the four highest-frequency patterns met the requirement and became candidate association rules, as shown in Table 3.

Table 3: Candidate highest-frequency patterns

Itemset	Count	Support
If FT1 & WT2 & KS4 → KS5	7/30	23.3%
If FT1 & ST2 & KS4 → KS5	20/30	66.7%
If FT1 & KT3 & KS4 → KS5	8/30	26.7%
If ST2 & KT3 & KE3 → KS4	9/30	30.0%

3.3. Formation of association rules

After the highest-frequency patterns were obtained, the confidence value was calculated for each candidate with a minimum confidence of 80%. From this process, one valid association rule was obtained, namely, If FT1 & WT2 & KS4 → KS5 with a confidence of 88.9% (8/9), as shown in Table 4.

Table 4: Best association rule (manual calculation)

If antecedent, then consequent	Support	Confidence
If FT1 & WT2 & KS4 → KS5	8/30	8/9 = 88.9%

The rule indicates a pattern in which, if a student opens TikTok one to three times a day (FT1), most often at midday during a break or after school (WT2), and usually studies for less than one hour per session (KS4), then the student tends to be able to concentrate well (KS5).

3.4. Implementation using RapidMiner

To validate the manual calculation results, testing was carried out using the RapidMiner application. The questionnaire data that had been converted into a matrix table format were imported into the application, and a series of operators was then arranged, namely *Discretize by Frequency*, *Numerical to Binomial*, *FP-Growth*, and *Create Association Rules*. The arrangement of these operators is shown in Figure 3.

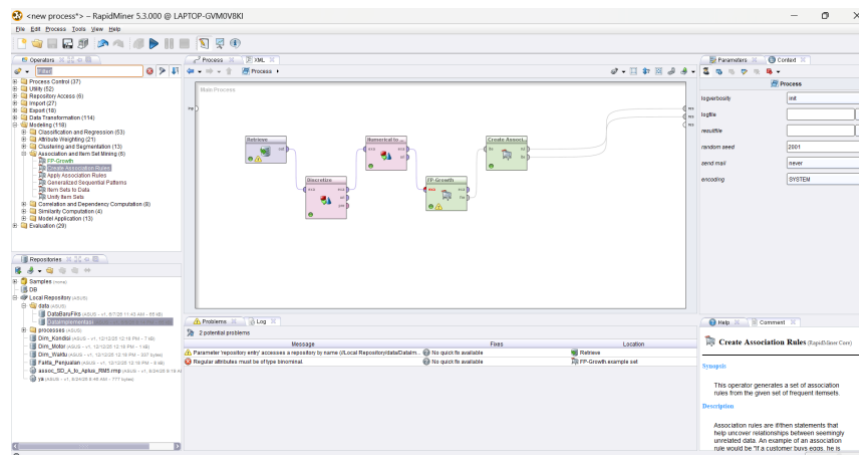


Fig. 3: Arrangement of association operators in RapidMiner

In the *FP-Growth* operator, the minimum support value was set, and in the *Create Association Rules* operator, the minimum confidence value was set. The resulting association rules, together with their support and confidence values, are shown in Figure 4.

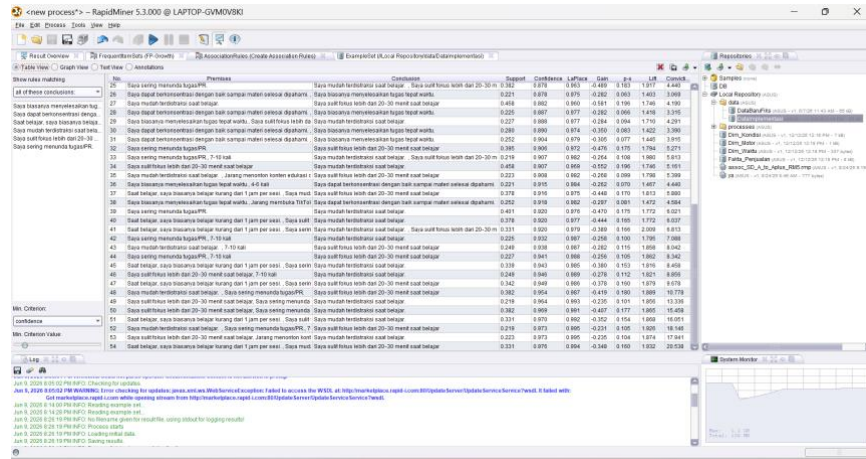


Fig. 4: Display of the resulting association rules

Based on the criterion of the product of the Support and Confidence values ($S \times C$) with a minimum support of 20% and a minimum confidence of 80%, several best rules were obtained, covering combinations of two to four itemsets, as summarized in Table 5.

Table 5: Best rules based on the $S \times C$ value

No	Rule	Sup.	Conf.	$S \times C$
1	If a student can concentrate well until the material is fully understood, then the student usually completes tasks on time (2-itemset).	54.2%	86.9%	47.1%
2	If a student has difficulty focusing for more than 20–30 minutes and often delays tasks, then the student is easily distracted while studying (3-itemset).	38.2%	96.7%	36.9%
3	If a student usually studies less than one hour per session, is easily distracted, and often delays tasks, then the student also has difficulty focusing for more than 20–30 minutes (4-itemset).	33.1%	97.6%	32.3%

The RapidMiner test results reinforce the manual calculation, showing that the various indicators of learning concentration are interrelated and form a consistent pattern. The rules with the highest $S \times C$ values confirm that short study habits and the ease with which students become distracted are closely related to a low ability to maintain learning focus. These findings are consistent with previous studies, which reported that the intensity of TikTok use tends to correlate with a decline in learning concentration.

4. Conclusion

Based on the research that has been carried out, this study concludes that there is a clear relationship between TikTok usage habits and the learning concentration of students at SMA/SMK Bina Bersaudara. Students who open TikTok at a low frequency (one to three times a day) and at particular times, such as midday, generally find it easier to focus. In contrast, students who frequently open TikTok tend to be easily distracted and have difficulty maintaining focus.

The Apriori algorithm was proven capable of discovering these relationship patterns. Using RapidMiner with thirty student records, a minimum support of 20%, and a minimum confidence of 80%, one valid rule was obtained, namely that if a student opens TikTok one to three times a day (FT1), most often at midday (WT2), and usually studies less than one hour per session (KS4), then the student tends to be able to concentrate well (KS5), with a confidence of 88.9%. Furthermore, based on the highest $S \times C$ values, the three best rules were obtained, showing that the indicators of learning concentration are interrelated and form a consistent pattern.

For future development, this research can be extended by using more data and variables, comparing the Apriori algorithm with other association methods such as Eclat or FP-Growth, or applying machine learning approaches such as clustering and classification to discover more complex patterns. The resulting findings can also be developed into an automated recommendation system that helps schools and parents monitor students' social media use and maintain their learning concentration.

Acknowledgement

The authors would like to thank STMIK Kaputama, Binjai, and SMA/SMK Bina Bersaudara for their support and permission during the data collection and completion of this research.

References

- [1] M. F. Alfatih et al., "The influence of TikTok short-form videos on attention span and study habits of students in college of vocational studies IPB University," *EduTech*, vol. 23, no. 2, pp. 202–212, 2024, doi: 10.17509/e.v23i2.69229.
- [2] C. A. Muzaqy and F. M. Suud, "The impact of social media use on students' concentration and motivation: A case study at Muhammadiyah Sambisari Elementary School," *Jurnal Kependidikan Islam*, vol. 15, no. 2, 2025.
- [3] A. Hermawan, B. Wicaksono, T. Ahmadjayadi, B. S. Prakasa, and J. D. Aruan, "Implementasi algoritma Apriori pada market basket analysis terhadap data penjualan produk supermarket," *Algoritma J. Mat. Ilmu Pengetahuan. Alam Kebumihan dan Angkasa*, vol. 2, no. 5, pp. 95–105, 2024, doi:

10.62383/algorithm.v2i5.137.

- [4] Q. Aini, "Hubungan media sosial TikTok terhadap konsentrasi belajar," *Jurnal Bimbingan Konseling Islam*, IAIH NW Pancor, 2023.
- [5] T. Wang, B. Xiao, and W. Ma, "Student behavior data analysis based on association rule mining," *Int. J. Comput. Intell. Syst.*, vol. 15, no. 1, 2022, doi: 10.1007/s44196-022-00087-4.
- [6] R. Buaton, M. Zarlis, and V. Yasin, *Konsep Data Mining dalam Implementasi*. Jakarta, Indonesia, n.d.
- [7] N. Laila, Yasmin, dan Putra, "Pengaruh penggunaan aplikasi TikTok terhadap konsentrasi dan prestasi belajar siswa/i Fakultas Dakwah dan Komunikasi UIN Sumatera Utara," vol. 4, 2022.
- [8] A. Thariq, "Implementasi market basket analysis menggunakan algoritma Apriori pada data penjualan buku," *Jurnal Kolaboratif Sains*, vol. 6, no. 3, pp. 154–163, 2023.
- [9] D. Ruswanti, D. Susilo, and R. Riani, "Implementasi CRISP-DM pada data mining untuk melakukan prediksi pendapatan dengan algoritma C4.5," *Go Infotech J. Ilm. STMIK AUB*, vol. 30, no. 1, pp. 111–121, 2024, doi: 10.36309/goi.v30i1.266.
- [10] H. Mubin dan D. Pramitha, "Pengaruh intensitas penggunaan media sosial terhadap tingkat konsentrasi peserta didik: Studi kasus di Sekolah Indonesia Kuala Lumpur," *Jurnal Multidisiplin Ilmu Akademik*, vol. 2, no. 3, pp. 69–76, 2025, doi: 10.61722/jmia.v2i3.4529.