Journal of Artificial Intelligence and Engineering Applications

Website: https://ioinformatic.org/

15th October 2023. Vol. 3. No. 1; e-ISSN : 2808-4519

The Effect of Social Media on Student Learning Motivation Using the Apriori Method

Chairmayni Pratiwi[1], Yani Maulita[2], Imeldawaty Gultom[3]

[1,2,3] Information System, STMIK Kaputama
chairmaynip@gmail.com [1], yanimaulita26@gmail.com [2], imeldagultom81@gmail.com [3]

Abstract

The success of student learning can be determined by their motivation. Students who have high learning motivation tend to have high achievement as well, otherwise their learning motivation is low, their learning achievement will also be low. student learning motivation in the subject is very low. Some students prefer to play social media rather than pay attention to the material explained by the teacher during class hours. Therefore, this study aims to explore the influence of social media on students' learning motivation. This research uses data mining method with Apriori algorithm to identify patterns related to social media usage and students' learning motivation. The Apriori algorithm is one of many algorithms in data mining that is used for frequent itemsets and association rules in databases on transactional data that are generated by identifying each item that exists, and combining larger sets of items provided that the items appear frequently enough in the database. Based on the research that has been done, the author can draw the conclusion that using the Rapid Miner 7.1 application tools in applying the apriori algorithm produces the same rules as manual calculations using 300 data on the learning motivation of Abdi Negara Binjai SMKS students and the system can generate association rules using 300 student learning motivation data with a minimum support of 12% and a minimum confidence of 75% and produce 5 association rules 3 itemsets to determine the learning motivation of Abdi Negara Binjai SMKS students. One of the rules that has the highest confidence value is, if YT and J2 then M1. Which means that every student who uses YOUTUBE Social Media with a length of use is 3-4 HOURS then INCREASES STUDY MOTIVATION. Then the less the ϕ (frequent) value is set, the more data that can be processed, as well as the minimum support value and confidence value, where the smaller the value determined, the more association results will be issued.

Keywords: *Learning Motivation, Data Mining, Apriori Algorithm, Association Rules.*

1. Introduction

Education is one of the most important sectors in a country's development. Good quality education can improve the quality of a country's human resources and help create a more qualified society. One of the factors that greatly affects the quality of education is student learning motivation. Student learning motivation is very important because it can affect students' interest in learning and their ability to understand the subject matter. Student learning motivation can also affect academic achievement. The success of student learning can be determined by their motivation. Students who have high learning motivation tend to have high achievement as well, on the contrary, low learning motivation will also have low learning achievement. High and low motivation can determine the high and low effort or enthusiasm of a person for activities, and of course the high and low enthusiasm will determine the results obtained [1]. However, in reality, there are still many students who are less motivated in learning and have difficulty in understanding the subject matter. Various factors can affect student learning motivation, such as the learning environment, learning methods, and the use of technology in learning. Some students prefer to play social media rather than pay attention to the material explained by the teacher during class hours. The reality in schools shows that low student achievement motivation is one of the obstacles in education, so an effective solution is needed to solve the problem [2]. One form of technology that is currently very popular among students is social media. Social media has become one of the most widely used platforms by teenagers and an important part of their daily lives. Social media can be used to communicate with friends, find information, or even to entertain themselves. However, it is not uncommon for students to spend too much time on social media that it interferes with their study time. This can affect students' learning motivation and academic performance. The results of this study also reveal that the highest duration of social media use occurs at the daily level, namely by updating whatever happens with social media. This is worrying, as respondents spend more time on social media than studying. In addition, they also use social media only to find friends and entertainment. No respondents used social media for education or information seeking [3]. Therefore, this study aims to explore the influence of social media on students' learning motivation. This research uses data mining method with Apriori algorithm to identify patterns related to social media usage and students' learning motivation. Apriori algorithm is one of many algorithms in data mining that is used for frequent itemsets and association rules in databases on transactional data that are generated by identifying each item, and combining larger sets of items provided that the items appear frequently enough in the database. The advantage of the Apriori algorithm is that it can make accurate predictions and can be used on very large amounts of data and can be implemented easily [4].

2. Research Methods

There are several stages in completing this research, namely:

1. Research Preparation; This stage is the initial activity, namely by determining the background of the problem, then identifying the problem and then making a problem boundary that will help the author to the next stage.
2. Formulating Problems and Objectives; After looking for identification and problem boundaries, the author will formulate problems and objectives that will provide benefits to users.
3. Data Collection; At this stage there are six ways, namely:
   a. Observation
      By conducting research and directly analyzing the condition of student learning motivation at ABDI NEGARA BINJAI SMKS, so that it can be seen the needs of the application, where this observation includes observation of software, hardware also includes searching and retrieving data related to the problems discussed related to the formation of patterns of social media influence on student motivation.
   b. Interview Technique
      The interview technique used is by asking questions and answering questions with teachers at ABDI NEGARA BINJAI SMKS in obtaining information about student learning motivation.
   c. Questionnaire (questionnaire)
      The questionnaire is conducted in each class in class X (ten), XI (eleven) and XII (twelve) in obtaining the data needed as information regarding learning motivation data on ABDI NEGARA BINJAI SMKS students.
   d. Data Processing
      At this stage, all the data obtained from the questionnaire from students at ABDI NEGARA BINJAI SMKS will be collected and selected accordingly which will be used for testing materials.
   e. Testing and Implementation
      At this stage, data validation testing and data implementation will be carried out as well as system program preparation. This stage is based on the results of data analysis that has been done before.
   f. Final Stage
      At the final stage will be discussed about the conclusions and suggestions needed for further program development.

2.1. Data Mining

Data Mining is a term used to describe the discovery of knowledge in databases. Data mining is the analysis of reviewing a data set to find unexpected relationships and summarize the data in a different way that is understandable and useful to the data owner [5]. Data Mining is an automated process on existing data. The data to be processed is in the form of very large data. Data Mining is a branch of artificial intelligence. In Data Mining there are several types of methods according to their utilization including: prediction, association, classification, clustering and estimation. In the association method, there are several techniques including the Apriori method [6].

2.2. Apriori algorithm

Apriori algorithm is a data retrieval algorithm with association rules to determine the associative relationship of a combination of items. This algorithm controls the development of candidate itemsets from frequent itemset results with support-based pruning to eliminate uninteresting itemsets by setting minsup. Association rule is one of the methods used to find relationships between data or how a group of data affects the existence of other data. Association rules are one of the methods often used to find relationships between various items. The relationship between the a priori algorithm and association is that it can find two or more attributes and two or more objects. The a priori algorithm is a type of association rule in data mining [7]. The basic methodology of association analysis is divided into two stages, including the following.

High frequency pattern analysis; At this stage look for combinations of items that meet the minimum requirements of the support value in the database. Support, is a measure that shows how much dominance an item or itemset has over the entire transaction. The support value of an item is obtained using the following formula::

$$Support\ (A) = \frac{\text{Number of Transactions Containing A}}{Total\ Transactions} X\ 100\% \quad (1)$$

The formula above explains that the support value is obtained by finding the number of transactions containing value A (one item) divided by the total number of transactions. The support value of 2 items is obtained using the formula.

$$Support\ (A, B) = P(A \cap B),\ Support\ (A,B) = \frac{\sum \text{Transactions containing A and B}}{\sum \text{Transactions}} X\ 100\%. \quad (2)$$

The formula above explains that the support value is obtained by finding the number of transactions that contain values A and B (the first item together with other items) divided by the total number of transactions.

Establishment of association rules; After all the high-frequency patterns are found, then look for association rules that meet the minimum requirements for confidence by calculating the confidence of the associative rule A to B. Confidence, is a measure that shows the relationship between two items conditionally (based on a certain condition). The confidence value of rule A to B is obtained by the following formula.

$$\text{Confidence} = P(B|A) = \frac{\sum \text{ Transaction data containing A and B}}{\sum \text{ Transaction A}} X\ 100\% \quad (3)$$

In the formula, the confidence value is obtained by finding the number of transactions containing the values of A and B (the first item together with other items) divided by the number of transactions containing A (the first item).

3. Results And Discussion

3.1. Association Analysis Methodology

This stage is divided into two parts, namely.

1. High Frequency Pattern Analysis; In this stage, the main objective is to look for item combinations that have a high frequency in the database. High frequency indicates how often the item combination appears together in the pattern. This process starts by setting a minimum support value that the combination of items must achieve to be considered significant. Support is the percentage of the number of transactions containing a particular item combination out of all transactions in the database. Item combinations that meet the minimum support requirement are called high-frequency patterns or frequent itemsets.
2. Establishment of Association Rules; This stage will be carried out after all the frequencies in the high-frequency pattern analysis stage have been found, then association rules that meet the confidence requirements are sought by calculating associativity. After the high-frequency pattern has been found, the next step is to look for association rules. An association rule is a statement of the form "if A, then B" that shows the relationship between two groups of items or variables in a transaction. In this context, A and B are two sets of items consisting of one or more items. To form an association rule, we need to calculate an associative measure, namely confidence. Confidence measures how often B appears in transactions that also contain A. The association rule that meets a certain minimum confidence requirement will be selected as the final result of the association analysis.

Tabel 1: Data Representation

No	Variabel	Sub Variabel	Code
1.	Types of Social Media	1. *Facebook*	FB
		2. *Instagram*	IG
		3. *Youtube*	YT
		4. *Tiktok*	TK
		5. *Whatsapp*	WA
		6. *Twitter*	TT
2.	Usage Time Range	1. 1-2 Hours	J1
		2. 3-4 Hours	J2
		3. 5-6 Hours	J3
		4. 7-8 Hours	J4
		5. 9-10 Hours	J5
		6. More than 10 Hours	J6
3.	Learning Motivation	1. Increase learning motivation	M1
		2. Decreases learning motivation	M2
		3. No effect on learning motivation	M3

Tabel 2: Item Representation in Matrix

p	nama	jenis media sosial						rentang waktu penggunaan						motivasi belajar		
		FB	IG	YT	TK	WA	TT	J1	J2	J3	J4	J5	J6	M1	M2	M3
1	Dinda Lorenza	0	1	0	0	0	0	0	1	0	0	0	0	1	0	0
2	Dimas Satria	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0
3	Ari Panca	0	0	1	0	0	0	0	1	0	0	0	0	1	0	0
4	Nabila Vira Ramadhani	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0
5	Safira Andina	0	1	0	0	0	0	0	1	0	0	0	0	1	0	0
6	Muhammad Rizky Fadillah	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0
7	Dimas Andreawan	0	1	0	0	0	0	0	0	1	0	0	0	1	0	0
8	Ibnu Aditya Susanto	1	0	0	0	0	0	0	0	1	0	0	0	1	0	0
9	Cindy Amelia	1	0	0	0	0	0	0	0	1	0	0	0	1	0	0
10	Dara Wuri Andriani	1	0	0	0	0	0	0	0	1	0	0	0	1	0	0
11	Ibnu Aditya	1	0	0	0	0	0	0	0	1	0	0	0	1	0	0
12	Lukman Hakim	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1
13	Mhd Nabil Ikbar	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0
14	Arya Janna Sembiring	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0
15	Sinta Puspita Sari	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0
16	Saskia Adianti	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0
17	Tria Kesumawardani	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1
18	Surya Dama Sitepu	0	0	0	0	1	0	0	0	0	0	0	1	1	0	0
19	Syahrian Priadinata	1	0	0	0	0	0	0	0	1	0	0	0	1	0	0
20	Fajar Ahmadani Napitupulu	0	0	1	0	0	0	0	1	0	0	0	0	1	0	0
21	Muhammad Riski Fadilah	1	0	0	0	0	0	1	0	0	0	0	0	0	0	1
22	Muhammad Zulfansyah	1	0	0	0	0	0	0	0	1	0	0	0	1	0	0
23	Nabila Mayang Sari	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0
24	Nisan Nurul Sakinah	0	0	0	1	0	0	0	0	0	0	1	0	0	1	0
25	Ramadhani Fitria Dewi	0	0	1	0	0	0	0	1	0	0	0	0	1	0	0
26	Dwi Sasi Mahrani	0	1	0	0	0	0	0	1	0	0	0	0	1	0	0
27	Ade Meliana	0	0	0	1	0	0	0	0	0	0	1	0	0	1	0
28	Alitto Purdana	0	0	1	0	0	0	0	1	0	0	0	0	1	0	0
29	Aidil Febriansyah	0	1	0	0	0	0	0	1	0	0	0	0	1	0	0
30	Anissa Jauh Har	0	1	0	0	0	0	0	1	0	0	0	0	0	1	0
	∑	15	6	4	2	3	0	1	19	7	0	2	1	24	3	3

4. Determine ϕ (*Frequent*); As can be seen in Table III.2. It is known that the total ϕ that can be processed with a value of ϕ (frequent) that is more than equal to 3 (three), among others, is the type of social media with codes FB, IG, YT and WA then the time span of use with codes J2 and J3 and learning motivation with code M1.
5. Define Item Set; From the results of determining ϕ (frequent) taken from table 1. Then the iteration whose value is greater than ϕ (frequent) on the variable type of social media is only the Facebook sub variable with the FB code, the Instagram sub variable with the IG code, the YouTube sub variable with the YT code and the WhatsApp sub variable with the WA code for the itemset 1 calculation pattern then the usage time span with the J2 and J3 codes for the itemset 2 pattern and the learning motivation with the M1 code for the item 3 calculation pattern. Due to the calculation of this pattern does not have the same variables, it will be determined the order of the pattern combination where for the type of social media as item 1, the time span of use as item 2 and learning motivation as item 3 in order to find the desired results with the title.
6. Test Set ϕ (Frequent); In this section will display the set formed to find the pattern of the formation of item 1 and item 2, namely 2 itemsets, as for the set pattern formed is F2 = {FB, J2}, {FB, J3}, {FB, M1}, {IG, J2}, {IG, M1}, {YT, J2}, {YT, M1}, {WA, J2}, {WA, M3}.

To make it easier to read the frequent itemset data, it is made in the form of tabulated data as can be seen in Table 3.

Tabel 3: Tabulated data of 2 itemsets

Nama *Itemset*		Jumlah
FB	J2	8
FB	J3	6
FB	M1	14
IG	J2	5
IG	M1	5
YT	J2	4
YT	M1	4
WA	M3	3
WA	J2	3

As seen in table 3. Then the data that meets the minimum requirement ϕ (frequent) is 3 (three), namely F2 = {FB, J2}, {FB, J3}, {FB, M1}, {IG, J2}, {IG, M1}, {YT, J2}, {YT, M1}, {WA, J2}, {WA, M3}. Then a combination of itemsets in F2 that meet the requirements is made followed by combining the combination items into a combination of 3 itemsets. Therefore we will calculate F3 with a pattern that is F3 = {FB, J2, M1}, {FB, J3, M1}, {IG, J2, M1}, {YT, J2, M1}, {WA, J2, M1} to find itemset pattern 3.

Then from the results obtained for F3 with ϕ (frequent) more than equal to 3 (three) are {FB, J2, M1}, {FB, J3, M1}, {IG, J2, M1}, {YT, J2, M1}, {WA, J2, M3}. After finding the itemset combination pattern, we will look for the relationship or correlation between F3 items using 2 factors, namely support and confidence obtained by the support formula, namely

$$Support(A|B\cap C) = \frac{\sum \text{Number of antecendent data that appears at once}}{\sum Total\ number\ of\ all\ data} X100\% \quad (4)$$

and the confidence value, namely with the formula

$$Confidence\ (A|B\cap C) = \frac{\sum \text{Number of antecedent and consequent data that appears at once}}{\sum \text{Total number of data in the antecedent}} X100\%. \quad (5)$$

The application of support and confidence calculations on tables that pass the minimum number of ϕ (frequent) equal to 3 (three) are {FB, J2, M1}, {FB, J3, M1}, {IG, J2, M1}, {YT, J2, M1}, {WA, J2, M3} can be seen in table 4.

Tabel 4: F3 support and confidence calculation table

If antecedent then consequent	***Support***	***Confidence***
If antecedent FB|J2 *then consequent* M1	(8/30)x100% = 26,6%	(8/8)x100% = 100%
If antecendent FB|J3 *then consequent* M1	(6/30)x100% = 20%	(6/6)x100% = 100%
If antecedent IG|J2 *then consequent* M1	(5/30)x100% = 16.6%	(4/5)x100% = 80%
If antecedent YT|J2 *then consequent* M1	(4/30)x100% = 13%	(4/4)x100% = 100%
If antecedent WA|J2 *then consequent* M3	(3/30)x100% = 10%	(3/3)x100% = 100%

After obtaining the support value and confidence value as shown in table 4. Then the multiplication between support and confidence is carried out, at this stage the minimum support is 12% and the minimum confidence taken is 75% so that the best value is obtained by multiplying the support value and confidence value as can be seen in table 5 as follows.

Tabel 5: C*S calculation table

If antecedent then consequent	***Support***	***Confidence***	**C*S**
If antecedent FB|J2 *then consequent* M1	26.6%	100%	0.26
If antecedent FB|J3 *then consequent* M1	20%	100%	0.20
If antecedent IG|J2 *then consequent* M1	16.6%	80%	0.16
If antecedent YT|J2 *then consequent* M1	13%	100%	0.13

From the data generated in Table 5. on the calculation of 3 itemsets can be concluded with the following sentence.

1. If the social media used is Facebook and the length of use is 3-4 hours, it increases student learning motivation with a support value of 26.6% and a certainty value of 100%.
2. If the social media used is Facebook and the length of use is 5-6 hours then it increases student learning motivation with a supporting value of 20% and a certainty value of 100%.
3. If the social media used is Instagram and the length of use is 3-4 hours, it increases student learning motivation with a supporting value of 16.6% and a certainty value of 100%.
4. If the social media used is Youtube and the duration of use is 3-4 hours, it increases student learning motivation with a supporting value of 13% and a certainty value of 100%

4. Testing

At the testing stage, researchers conducted experiments on all questionnaire data with different minimum support values and minimum confidence values, namely the minimum support value of 10%, 12%, and 14% then the minimum confidence of 75%, 80%, and 90%. The purpose of testing with different minimum support and minimum confidence values is to compare the test results with different values but the same data on learning motivation questionnaire data obtained from Abdi Negara Binjai SMKS students.
Based on the test results seen in Figure 1, it can be seen that at support values of 12% and 14% when combined with confidence values of 75% and 80% have the same rules results. For the combination of at least 10% support and 75% confidence, the highest number of rules is obtained from all support and confidence values combined during testing. From the graph shown in Figure 1, it can also be concluded that the higher the support and confidence values combined, the fewer rules will be generated.

Next are the test results with a support value of 12% and confidence values of 75%, 80% and 90%, the formation of itemsets and rules generated can be seen in Figure 1, Figure 2 and Figure 3.

No.	Premises	Conclusion	Support	Confidence
1	FB	M1	0.203	0.803
2	J3	M1	0.190	0.814
3	WA	J2	0.147	0.815
4	WA	J2, M3	0.147	0.815
5	J2, M3	WA	0.147	0.830
6	IG	M1	0.223	0.848
7	M3, WA	J2	0.147	0.898
8	WA	M3	0.163	0.907
9	YT	M1	0.197	1
10	J2, YT	M1	0.147	1
11	J2, WA	M3	0.147	1

Figure 1: Itemset formation results support 12% confidence 75%

No.	Premises	Conclusion	Support	Confidence
1	FB	M1	0.203	0.803
2	J3	M1	0.190	0.814
3	WA	J2	0.147	0.815
4	WA	J2, M3	0.147	0.815
5	J2, M3	WA	0.147	0.830
6	IG	M1	0.223	0.848
7	M3, WA	J2	0.147	0.898
8	WA	M3	0.163	0.907
9	YT	M1	0.197	1
10	J2, YT	M1	0.147	1
11	J2, WA	M3	0.147	1

Figure 2: Result of itemset formation support 12% confidence 80%

No.	Premises	Conclusion	Support	Confidence
1	WA	M3	0.163	0.907
2	YT	M1	0.197	1
3	J2, YT	M1	0.147	1
4	J2, WA	M3	0.147	1

Figure 3: Result of itemset formation support 12% confidence 90%

From this figure, the formation of association rules or rules generated for a minimum confidence value of 75% is 11 rules, for a minimum confidence value of 80% is 11 rules, for a minimum confidence value of 90% is 4 rules. One of the rules that has the highest confidence value is, if YT and J2 then M1. Which means that every student who uses YOUTUBE Social Media with a length of use is 3-4 HOURS then INCREASES STUDY MOTIVATION.

5. Conclusions

As for some conclusions that the author can provide from writing this thesis research, among others, are as follows.

1. Using the Rapid Miner 7.1 application tools in applying the apriori algorithm produces the same rules as manual calculations using 300 data on the learning motivation of Abdi Negara Binjai SMKS students.
2. The system can generate association rules using 300 student learning motivation data with a minimum support of 12% and a minimum confidence of 75% and produce 5 association rules of 3 itemsets to determine the learning motivation of SMKS Abdi Negara Binjai students.
3. The less the ϕ (frequent) value is set, the more data that can be processed, as well as the minimum support value and confidence value, where the smaller the value determined, the more association results will be issued.
4. With a minimum support value of 12% and a minimum support of 14%, it turns out to produce the same results as the combination of minimum confidence values of 75% and 80%.

References

[1] S. Rahman, "Pentingnya Motivasi Belajar Dalam Meningkatkan Hasil Belajar," *Merdeka Belajar dalam Menyambut Era Masy. 5.0*, no. November, pp. 289–302, 2021.
[2] A. Harso and J. Merdja, "Motivasi Belajar dan Prestasi Belajar Fisika Ditinjau dari Jenis Kelamin," *Sci. Phys. Educ. J.*, vol. 3, no. 1, pp. 11–20, 2019, doi: 10.31539/spej.v3i1.991.
[3] D. Iskandar and M. Isnaeni, "Penggunaan Internet Di Kalangan Remaja Di Jakarta," *Commun. J. Commun. Stud.*, vol. 6, no. 1, p. 57, 2019, doi: 10.37535/101006120194.
[4] Z. Abidin, A. K. Amartya, and A. Nurdin, "PENERAPAN ALGORITMA APRIORI PADA PENJUALAN SUKU CADANG KENDARAAN RODA DUA (Studi Kasus: Toko Prima Motor Sidomulyo)," *J. Teknoinfo*, vol. 16, no. 2, p. 225, 2022, doi: 10.33365/jti.v16i2.1459.
[5] D. P. Utomo and M. Mesran, "Analisis Komparasi Metode Klasifikasi Data Mining dan Reduksi Atribut Pada Data Set Penyakit Jantung," *J. Media Inform. Budidarma*, vol. 4, no. 2, p. 437, 2020, doi: 10.30865/mib.v4i2.2080.
[6] I. F. P. Ginting, D. Saripurna, and E. Fitriani, "Penerapan Data Mining Dalam Menentukan Pola Ketersediaan Stok Barang Berdasarkan Permintaan Konsumen Di Chykes Minimarket Menggunakan Algoritma Apriori," *J. SAINTIKOM (Jurnal Sains Manaj. Inform. dan Komputer)*, vol. 20, no. 1, p. 28, 2021, doi: 10.53513/jis.v20i1.2504.
[7] P. H. Simbolon, "Implementasi Data Mining Pada Sistem Persediaan Barang Menggunakan Algoritma Apriori (Studi Kasus : Srikandi Cash Credit Elektronic dan Furniture)," *J. Ris. Komput.*, vol. 6, no. 4, pp. 401–406, 2019.