

Apriori Algorithm in Social Commerce Utilization and Market Basket Analysis in Local Retail

Mahardika Inra Takaendengan^{1*}, Nolin Jufika Kabaitan², Marline Sofiana Paendong³

^{1,2}Information Systems, Universitas Sam Ratulangi, Indonesia

³Mathematics, Universitas Sam Ratulangi, Indonesia

^{1*}mahardika@unsrat.ac.id, ²jufikakabaitan@gmail.com, ³marlinepaendong@unsrat.ac.id

Abstract: Social commerce platforms provide micro-enterprises with low-cost channels to expand market reach, yet systematic evaluations of their operational effectiveness and transaction data utilization remain limited. This study examines Facebook Marketplace utilization, identifies factors affecting its effectiveness, and mines transaction patterns to formulate data-driven sales strategies. A mixed-methods descriptive exploratory design was employed. Qualitative data were collected through interviews and observations with store personnel and customers to identify operational constraints. Quantitative data comprised 200 sales transaction records, with 72 originating specifically from Facebook Marketplace, analyzed using the Apriori algorithm. The results indicate that platform adoption correlated with a 50% increase in transaction volume and a 98.43% increase in total revenue. The Apriori analysis generated six valid association rules. The strongest product association emerged between Jeans and Belts, yielding a lift ratio of 3.13. These findings demonstrate that social commerce functions as a critical discovery channel that alters operational workflows and generates structured transaction logs. Micro-enterprises, for example, could use these findings to plan targeted product mix-ups and get better synchronization of inventory. Lastly, the work done in this paper offers, among other things, a concrete solution for traditional, small-scale fashion business owners aiming at moving beyond using social media just as the platform for advertising and towards turning their streams into valuable information resources that would help them carry out strategic stock planning and online sales more efficiently.

Keywords: Social Commerce; Facebook Marketplace; Apriori Algorithm; Market Basket Analysis; SME Digitalization.

1. INTRODUCTION

Digital technology enables micro and small enterprises to expand market reach through social commerce platforms. Social commerce provides low-cost channels for product promotion and direct customer interaction. Micro-enterprises rely on these platforms to maintain competitiveness and increase sales volume without significant capital investment [1], [2], [3].

Prior research demonstrates that social media marketing increases retail website traffic and online sales [4]. Specific platforms like Facebook Marketplace positively influence consumer purchase intention by providing localized product discovery and direct messaging features [5], [6]. Concurrently, data mining techniques such as the Apriori algorithm effectively identify product associations and support market basket analysis in

traditional retail environments [7], [8]. These algorithms calculate support, confidence, and lift ratios to reveal non-obvious purchasing patterns.

Despite these advancements, current literature exhibits distinct methodological limitations. Research on social commerce adoption is mostly based on self-report surveys to measure consumer's intention to use social commerce for purchasing [2], [4], [9], [10]. This method is used to get consumer perceived willingness rather than the actual behaviour for purchasing in high-friction, and chat-based negotiation environments. Conversely, research applying the Apriori algorithm primarily focuses on large-scale traditional supermarkets or structured e-commerce platforms with automated checkout systems [6], [11], [12], [13]. These quantitative studies assume clean, pre-structured transaction databases and overlook the unstructured transaction logs and operational constraints unique to micro-enterprises operating on decentralized social media.

A distinct methodological gap remains regarding how to bridge the qualitative socio-technical friction of social commerce with the quantitative extraction of association rules from manual transaction logs. Relying solely on quantitative data mining ignores the operational realities that cause order cancellations in social commerce. Relying solely on qualitative interviews fails to capture empirical purchasing affinities. Addressing this gap requires a mixed-methods approach to contextualize algorithmic outputs within the actual operational constraints of the seller. Understanding this intersection enables micro-enterprises to translate raw sales logs into actionable business strategies.

This study examines Facebook Marketplace utilization, evaluates operational effectiveness, and extracts transaction patterns using the Apriori algorithm. The research addresses four questions. First, how do micro-enterprises utilize social commerce to support online sales? Second, what factors influence platform operational effectiveness? Third, what transaction patterns emerge from Apriori analysis? Fourth, what sales strategies can be derived from integrating qualitative operational findings with quantitative association rules? The study provides a framework for local retailers to evaluate social commerce channels and apply data mining for inventory planning.

2. METHODOLOGY

Research and Design Context

This research work takes a mixed-methods descriptive exploratory approach. The qualitative aspect assesses the degree of use of the platform and how it is used and the operational limits of its use through interviews and observation. The quantitative aspect involves running a market basket analysis on the transaction logs of the company. The area of the study is a local fashion and accessories micro-enterprise in Bitung City, Indonesia. The observation period is eight months; this period from February-March and September-December of 2025. This time span also includes a four months period of not operating Facebook Marketplace as a sales channel followed by four months period of running the Facebook Marketplace Apart from their offline stores.

Data Collection and Subjects

Qualitative data collection involves purposive sampling. The sample includes the store owner, one employee, and ten customers. The customer selection criteria required participants to have completed at least one transaction originating from a Facebook Marketplace inquiry within the observation period. Participants must also possess sufficient communication history to evaluate the negotiation and trust-building process. Researchers approached eligible customers directly after transaction completion and obtained verbal informed consent.

Quantitative data collection relies on the store transaction logs. The transaction data collection mechanism involves manual entry by store personnel into a structured Microsoft

Excel ledger at the point of sale or upon confirmation of digital payment. The ledger captures the transaction identifier, timestamp, product details, and the specific acquisition channel. Table 1 details the research subjects and data sources.

Table 1. Research Subjects and Data Sources

Category	Subject/Source	Quantity	Data Obtained
Primary	Store Owner	1	Sales strategy, platform adoption rationale, operational constraints
	Store Employee	1	Daily operations, product posting, customer response handling
	Customers	10	Purchasing experience, communication preferences, trust factors
Secondary	Transaction Logs	332 rows	Transaction ID, date, product, category, quantity, price, source
	Platform Archives	N/A	Product photos, descriptions, posting times, customer inquiries

Data Structure and Preprocessing

The Apriori algorithm requires structured transaction data. The initial dataset contains 332 raw data rows representing individual product purchases. These rows consolidate into 200 unique transactions. Table 2 defines the transaction data structure.

Table 2. Transaction Data Structure

Column Name	Data Type	Description
ID_Transaction	String	Unique identifier for each purchase event
Date	Date	Timestamp of the transaction
Product_Name	String	Standardized name of the purchased item
Category	String	Product classification (e.g., Fashion, Accessories)
Quantity	Integer	Number of units purchased
Unit_Price	Integer	Price per single unit
Total_Price	Integer	Quantity multiplied by Unit_Price
Period	Categorical	Before or After Facebook Marketplace adoption
Transaction_Source	Categorical	Physical Store, WhatsApp, or Facebook Marketplace

Data quality was confirmed through a multi-stage check that preceded the algorithmic analysis. Firstly, transaction records were matched against Facebook Messenger and WhatsApp chat logs archived for that purpose. By employing this technique, we are checking not only the correctness of the transaction source but also the product details, Because of this avoiding the classification of a walk-in (offline) sale as a digital conversion.

Finally, the data are scrutinized for missing values and duplicates as outlined by Equations (1) and (2).

$$Missing = \sum_{i=1}^n \sum_{j=1}^m I(x_{ij} = \emptyset) \quad (1)$$

$$Duplicates = N_{raw} - N_{unique} \quad (2)$$

The initial check yields zero missing values and zero duplicates. Third, product names are standardized to ensure consistent casing and spelling. Fourth, items are grouped by transaction identifier to form itemsets, defined in Equation (3).

$$T_i = \{item_1, item_2, ..., item_k\} \quad (3)$$

Fifth, the data undergoes one-hot encoding to create a binary transaction matrix suitable for algorithmic processing. The encoding logic is defined in Equation (4).

$$X_{ij} = \begin{cases} 1, & \text{if } item_j \in T_i \\ 0, & \text{if } item_j \notin T_i \end{cases} \quad (4)$$

Table 3 illustrates the one-hot encoding transformation for a subset of transactions.

Table 3. One-Hot Encoding Transformation Example

ID_Transaction	Plain T-Shirt	Cap	Jeans	Belt	Hoodie
TRX001	1	0	0	0	0
TRX002	0	0	1	0	0
TRX003	0	1	0	0	0
TRX004	0	1	0	0	1
TRX005	0	0	1	1	0

Finally, the data is filtered to isolate the 72 transactions originating specifically from Facebook Marketplace during the post-adoption period. Figure 1 illustrates the complete preprocessing and algorithmic workflow.

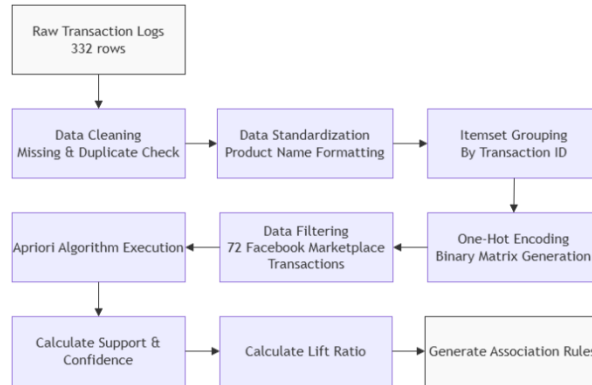


Figure 1. Preprocessing and Apriori Algorithm Workflow

Apriori Algorithm and Parameter

The Apriori algorithm identifies frequent itemsets and generates association rules based on support, confidence, and lift ratio metrics [7], [8], [9], [12]. The minimum support count threshold is calculated using Equation (5).

$$MinSupCount = MinSup \times |D| \quad (5)$$

Where $|D|$ is the total number of transactions. Support measures the frequency of an itemset across all transactions, defined in Equation (6).

$$Support(A) = \frac{\sigma(A)}{|D|} \quad (6)$$

Confidence measures the conditional probability of a consequent item given an antecedent item, defined in Equation (7).

$$Confidence(A \rightarrow B) = \frac{Support(A \cup B)}{Support(A)} \quad (7)$$

Lift ratio evaluates the strength of the association compared to random chance, defined in Equation (8).

$$Lift(A \rightarrow B) = \frac{Confidence(A \rightarrow B)}{Support(B)} \quad (8)$$

The analysis applies specific parameter thresholds to ensure the generated rules are statistically relevant and actionable for the micro-enterprise context. Table 4 details these parameters.

Table 4. Apriori Algorithm Parameters

Parameter	Value	Function in Analysis
Min Support	0.05	Filters item sets appearing in at least 5% of transactions
Min Confidence	0.40	Ensures rules have at least a 40% conditional probability
Lift Ratio	> 1.0	Confirms a positive, non-random association between products
Threshold		

Ethical Considerations

The study adheres to standard ethical protocols for micro-enterprise research. Verbal informed consent is obtained from the store owner, employee, and participating customers prior to data collection. All personal identifiers and specific store names are anonymized in the reporting phase to protect commercial privacy and participant confidentiality. Transaction logs are used strictly for aggregate pattern analysis.

3. RESULT AND DISCUSSIONS

Business Process Evaluation and Platform Utilization

The adoption of Facebook Marketplace altered the operational workflow of the micro-enterprise. The pre-adoption process relied entirely on physical store visits and direct WhatsApp inquiries. The post-adoption process introduces a digital discovery phase. Figure 2 illustrates the updated business process model.

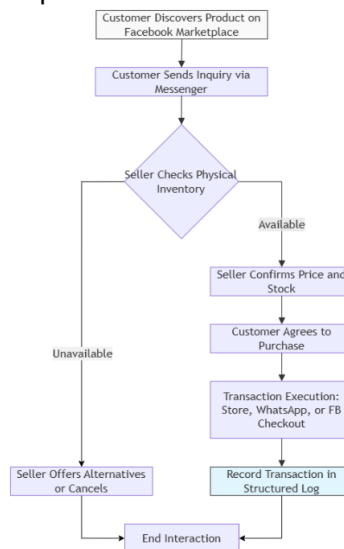


Figure 2. Post-Adoption Business Process Model

Qualitative observations indicate that platform effectiveness depends on photo quality, description clarity, and seller response speed. The primary operational constraints involve inventory synchronization and non-serious inquiries. Sellers frequently update physical stock without immediately updating digital listings, leading to order cancellations. High

volumes of unstructured messages require standardized response templates to maintain conversion rates.

Sales Performance Metrics

The dataset contains 200 unique transactions spanning eight months. The pre-adoption period recorded 80 transactions. The post-adoption period recorded 120 transactions. Table 5 details the comparative sales metrics.

Table 5. Sales Performance Comparison

Metric	Pre-Adoption	Post-Adoption	Absolute Change
Transaction Volume	80	120	+40
Total Revenue (IDR)	8,925,000	17,710,000	+8,785,000
Average Revenue per Transaction (IDR)	111,563	147,583	+36,020

The percentage increase for each metric is calculated using Equation (9).

$$\text{PercentageIncrease} = \frac{\text{Value}_{\text{post}} - \text{Value}_{\text{pre}}}{\text{Value}_{\text{pre}}} \times 100\% \quad (9)$$

Applying Equation (9) to the revenue data yields a 98.43% increase. The transaction volume increased by 50%. The average revenue per transaction increased by 32.29%. These metrics indicate that the post-adoption period correlates with higher sales activity and larger average basket sizes.

Market Basket Analysis Result

The Apriori algorithm processed the 72 transactions originating specifically from Facebook Marketplace during the post-adoption period. The algorithm applied a minimum support of 0.05 and a minimum confidence of 0.40. Table 6 presents the top frequent itemsets.

Table 6. Top Frequent Itemset

Itemset	Support Value	Transaction Count
Cap	0.444	32
Jeans	0.319	23
Plain T-Shirt	0.292	21
Jeans, Belt	0.208	15
Belt	0.208	15

The algorithm generated six valid association rules. The strongest association emerged between Jeans and Belts. First, the support for the combined itemset is calculated:

$$\text{Support}(\text{Jeans} \cup \text{Belt}) = \frac{15}{72} = 0.208$$

Next, the confidence of the rule is calculated using the support of the antecedent ($\text{Jeans} = 0.319$):

$$\text{Confidence}(\text{Jeans} \rightarrow \text{Belt}) = \frac{0.208}{0.319} = 0.652$$

Finally, the lift ratio is calculated using the support of the consequent ($\text{Belt} = 0.208$):

$$\text{Lift}(\text{Jeans} \rightarrow \text{Belt}) = \frac{0.652}{0.208}$$

Table 7 summarizes the generated association rules.

Table 7. Generated Association Rules

Antecedent	Consequent	Support	Confidence	Lift Ratio
Jeans	Belt	0.208	0.652	3.130
Belt	Jeans	0.208	1.000	3.130
Men's Shirt	Jeans	0.111	1.000	3.130
Hoodie	Cap	0.167	1.000	2.250
Sandals	Plain T-Shirt	0.083	0.500	1.714
Plain T-Shirt	Cap	0.194	0.667	1.500

A lift ratio greater than 1.0 indicates a positive association. The rule "Jeans $\rightarrow$ Belt" yields a lift of 3.130. This indicates that a customer buying Jeans is 3.13 times more likely to buy a Belt than a random customer

Discussion and Strategic Implications

The integration of qualitative operational data and quantitative transaction logs provides a comprehensive view of social commerce dynamics in micro-enterprises. The 98.43% revenue increase during the post-adoption period aligns with prior studies indicating that social media marketing expands retail reach and order volume [10], [11], [13], [14]. However, this study cannot claim strict causality. The increase may also reflect seasonal trends, general business growth, or concurrent offline promotional activities.

The qualitative insights illustrate what causes the friction in the quantitative data. As the Facebook Marketplace results appear to have high level of inquiry, the non-synchronization of real-time inventory results in cancellations of order. The operational deviation limits the conversion rate of discovery in digital into deal closing. Standardized replies to chat conversations and strict inventories synchronization rules have mitigated this friction.

The Apriori results provide empirical evidence for inventory and bundling strategies. The high lift ratio between Jeans and Belts (3.130) and Hoodies and Caps (2.250) contradicts the assumption that social commerce customers only purchase single, low-cost items. Instead, these customers purchase coordinated outfits. Micro-enterprises can leverage this by creating pre-packaged bundles. Offering a "Jeans and Belt" package at a marginal discount reduces customer decision friction and increases the average revenue per transaction, which already shows a 32.29% baseline increase.

This study has limitations. The quantitative analysis relies on a small dataset of 72 transactions. Association rules derived from small samples are sensitive to minor fluctuations in purchasing behavior and may not generalize to other retail contexts. The single case study design limits external validity. Future research should address these constraints by implementing comparative algorithmic analyses.

Subsequent studies must evaluate the computational efficiency and rule variation of FP-Growth and ECLAT algorithms against Apriori as transaction volumes scale. Researchers should also pool together transaction data from several social commerce platforms, such as TikTok Shop, Shopee, and Instagram, and compare user behavior across these different platforms. In addition, switching from association rule mining that mainly brings a picture based on description to machine learning techniques like Random Forest or sequential pattern prediction, would allow micro-businesses to not only predict customer demand but also better adapt their stock level based on their needs.

4. CONCLUSION

This study demonstrates that social commerce platforms function as effective discovery channels that correlate with expanded sales metrics for local micro-enterprises. By integrating qualitative operational observations with quantitative market basket analysis, the research identifies specific product affinities driven by digital interactions. These findings provide local retailers with a practical framework to transition social media from a basic promotional tool into a structured data source. Retailers can leverage these transaction patterns to design targeted product bundles and optimize inventory synchronization based on empirical purchasing behaviors.

The interpretation of these results is bounded by the single case study design and the analysis of a limited dataset comprising 72 social commerce transactions. Association rules derived from small samples are sensitive to minor fluctuations in consumer behavior and may not generalize to broader retail contexts. Future research should employ longitudinal designs with larger transaction volumes and compare the Apriori algorithm with FP-Growth and ECLAT to assess computational efficiency. Exploring machine learning-based predictive models across multiple social commerce platforms, such as TikTok Shop, Shopee, and Instagram, will help validate these purchasing patterns and operational constraints.

5. ACKNOWLEDGMENT

We thank the owner and personnel of the participating retail store for providing access to transaction logs and participating in interviews

6. REFERENCES

- [1] H. Arviani, D. Claretta, K. Kusnarto, N. Delinda, and S. Izzaanti, "Sosial Media Marketing : Peluang & Tantangan bagi UMKM Lokal Di Masa Pandemi Covid-19," *J. SIMBOLIKA Res. Learn. Commun. Study*, vol. 7, no. 1, pp. 47–56, Apr. 2021, doi: 10.31289/simbolika.v7i1.4356.
- [2] G. D. Rahmadiane and U. S. Utami, "ANALISIS PEMANFAATAN SOCIAL COMMERCE BAGI PENGEMBANGAN UMKM DI INDONESIA," *AdBispreneur*, vol. 6, no. 3, p. 225, Feb. 2022, doi: 10.24198/adbispreneur.v6i3.29114.
- [3] Budi Harto *et al.*, "Strategi Social Media Marketing Melalui Dukungan Teknologi Informasi dalam Kajian Kualitatif Pada UMKM Kota Bandung," *KOMVERSAL*, vol. 5, no. 2, pp. 244–261, Sep. 2023, doi: 10.38204/komversal.v5i2.1499.
- [4] L. Dolega, F. Rowe, and E. Branagan, "Going digital? The impact of social media marketing on retail website traffic, orders and sales," *J. Retail. Consum. Serv.*, vol. 60, p. 102501, May 2021, doi: 10.1016/j.jretconser.2021.102501.
- [5] S. Simatupang, E. Efendi, and D. E. Putri, "Facebook Marketplace Serta Pengaruhnya Terhadap Minat Beli," *J. EKBIS*, vol. 22, no. 1, p. 28, Mar. 2021, doi: 10.30736/je.v22i1.695.
- [6] A. Arizal, R. A. Sukmana, Y. Ulfah, S. Shaddiq, and M. Zainul, "Strategi Pemanfaatan Facebook Marketplace dalam Manajemen Periklanan," *Syntax Idea*, vol. 3, no. 6, pp. 1278–1289, Jun. 2021, doi: 10.46799/syntax-idea.v3i6.1277.
- [7] S. R. Bagaskara and D. H. Bangkalang, "Analisis dan Implementasi Market Basket Analysis (MBA) Menggunakan Algoritma Apriori dengan Dukungan Visualisasi Data," *J. Sist. Komput. Dan Inform. JSON*, vol. 4, no. 4, p. 612, Jul. 2023, doi: 10.30865/json.v4i4.6351.
- [8] K. Brighton and S. Hariyanto, "Penerapan Metode Market Basket Analisis Dengan Algoritma Apriori Pada Toko Ritel Elektronik," *Bit-Tech*, vol. 7, no. 1, pp. 37–46, Aug. 2024, doi: 10.32877/bt.v7i1.1417.

- [9] I. W. Supriana and L. A. A. Rahning Putri, "Implementasi Algoritma Apriori sebagai Association Rule Learning untuk Mengidentifikasi Pola Item Dataset Penjualan," *J. Buana Inform.*, vol. 16, no. 01, pp. 69–79, Apr. 2025, doi: 10.24002/jbi.v16i01.11384.
- [10] W. M. Lim and T. Rasul, "Customer engagement and social media: Revisiting the past to inform the future," *J. Bus. Res.*, vol. 148, pp. 325–342, Sep. 2022, doi: 10.1016/j.jbusres.2022.04.068.
- [11] M. K. E. Paendong, P. R. Rambing, C. H. Simanjuntak, A. P. Kumaat, M. R. Marentek, and N. H. J. Mandey, "Digital Marketing Ecosystem Perspective of Regional Featured Product in North Sulawesi Province, Indonesia," *Open J. Soc. Sci.*, vol. 11, no. 02, pp. 1–17, 2023, doi: 10.4236/jss.2023.112001.
- [12] F. H. Subechi, S. Surejo, and E. U. S. Utami, "Analisis Pola Pembelian Konsumen Menggunakan Algoritma Apriori pada Toko Komputer," *RIGGS J. Artif. Intell. Digit. Bus.*, vol. 4, no. 3, pp. 1504–1513, Aug. 2025, doi: 10.31004/riggs.v4i3.2172.
- [13] M. N. Assidiq, A. Bahaiddin, C. D. Octaviany, N. Ummi, and A. S. Mariawati, "Market Basket Analysis dengan Metode Association Rule Menggunakan Algoritma FP-Growth untuk Mendukung Strategi Promosi Ritel Modern," *J. Syst. Eng. Manag.*, vol. 5, no. 1, pp. 37–48, Jun. 2026.
- [14] S. A. Qalati, E. G. Vela, W. Li, S. A. Dakhan, T. T. Hong Thuy, and S. H. Merani, "Effects of perceived service quality, website quality, and reputation on purchase intention: The mediating and moderating roles of trust and perceived risk in online shopping," *Cogent Bus. Manag.*, vol. 8, no. 1, p. 1869363, Jan. 2021, doi: 10.1080/23311975.2020.1869363.