

# A Data-Driven Replication Strategy for IT Product Marketing in SMEs Using Apriori, PrefixSpan, and CloSpan Algorithms

Rendro Kasworo<sup>1</sup>, R. Rizal Isnanto<sup>2</sup>, Budi Warsito<sup>3</sup>, Shafira Ramadhina Adifa<sup>4</sup>

<sup>1,2,3</sup>Doctoral Program of Information Systems, Postgraduate School, Diponegoro University, Semarang, Indonesia

<sup>4</sup>Information System, University of Indonesia, Depok, Indonesia

<sup>1</sup>rendro@students.undip.ac.id, <sup>2</sup>rizal@ce.undip.ac.id, <sup>3</sup>budiwarsito@live.undip.ac.id,

<sup>4</sup>shafira.ramadhina@ui.ac.id

**Abstract**— Small and Medium Enterprises (SMEs) are critical to Indonesia’s economy but face significant barriers in adopting digital technologies. Despite Telkom Indonesia’s broad SME customer base, IT product adoption remains low. This study proposes a replication marketing strategy based on historical transaction data to accelerate IT product penetration in the SME segment. Utilizing Sequential Pattern Mining (SPM), we apply and compare three algorithms Apriori, PrefixSpan, and CloSpan to identify frequent and sequential purchase patterns from over 58,000 existing SME customers. These patterns are used to develop a data-driven product recommendation system tailored by business class and industry. The system aims to replicate successful customer purchase sequences for more personalized and efficient marketing. Evaluation results show improved accuracy in predicting product interest and bundling potential, supporting consultative selling and reducing sales cycle time. This research contributes both methodologically by integrating multi-algorithm SPM into a real-time marketing application and practically by offering a scalable solution for targeted IT product expansion in dynamic SME markets.

**Keywords**— IT product, Replication Strategy, SME, Sequential Pattern Mining

## I. INTRODUCTION

Small and Medium Enterprises (SMEs) are pivotal to Indonesia’s economy, contributing 61.7% to GDP and employing nearly 97% of the workforce, with over 64 million active

units driving innovation and economic stability [1]. However, their digital transformation is hindered by low digital literacy, limited financing, and insufficient tailored support, with only 37% of SMEs possessing adequate digital capabilities [2-3]. Fintech solutions and government initiatives like "SME-Go Digital" aim to enhance financial inclusion and digital adoption, yet structural barriers persist, necessitating targeted interventions to bolster SME competitiveness in the digital economy [1-2]. Similar challenges are observed in developing countries, where digital transformation among SMEs is often hindered by technological readiness and institutional support gaps [6].

To illustrate the economic significance of SMEs, Figure 1 presents the composition of Indonesian SMEs by business class. The dominance of micro-enterprises (over 98%) highlights both the scale and diversity of digitalization needs across this sector.



**Figure 1. Profile of Indonesian SMEs by business class and economic contribution**

Responding to these challenges, Telkom Indonesia has introduced Indibiz, a product series tailored for SMEs that includes digital applications, secure connectivity, business

platforms, and productivity tools. These solutions aim to address varying business maturity levels within the SME segment. However, according to Telkom's 2024 internal data, only 2.4% of its SME customers utilize more than basic connectivity services. This underutilization of IT products reveals a strategic gap between available solutions and effective customer outreach.

Figure 2 depicts the Indibiz product landscape, underscoring Telkom's effort to move beyond connectivity by offering value-added digital solutions. Yet, adoption remains stagnant, primarily due to ineffective marketing strategies that are not aligned with customer behavior.

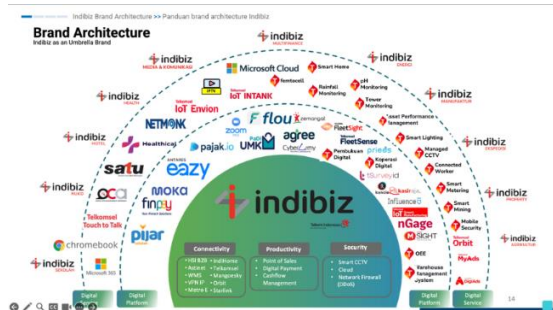


Figure 2. Telkom's Indibiz Product Lineup for the SME Segment

This mismatch points to a deeper issue: the lack of personalization and data utilization in Telkom's go-to-market approach. Despite having a rich repository of transactional history from over 58,000 SME clients, Telkom has yet to deploy data-driven insights to recommend relevant product bundles based on purchasing behavior. Early exploration of historical records reveals consistent sequential adoption patterns for instance, SMEs that first subscribe to high-speed internet often proceed to adopt digital products or content services.

To support this, Table 1 presents a sample dataset illustrating the sequential purchase patterns of SME customers over time. The data captures the chronological transaction history of individual customers, detailing their journey from the initial product purchase to subsequent ones, along with corresponding transaction dates. For instance, the table shows how customers like PT. XXXR G.I.

and XXX E.I. progressively adopt products such as High Speed Internet, Music, and Digital Products in a specific order. This sequential behavior underpins the research hypothesis that hidden patterns exist in the purchasing sequences of IT products by SMEs. Consequently, this table justifies the application of Sequential Pattern Mining (SPM) techniques, such as PrefixSpan and CloSpan, to analyze large-scale transaction data, enabling the replication of successful adoption trajectories for new SME customers with similar profiles.

Table 1. Sequential Identification

No.	Cust_Name	Date of PS	Product_Name
1	PT. XXXR G.I.	2023-01-02	High Speed Internet
		2023-02-16	High Speed Internet
		2023-02-20	Music
		2023-02-20	High Speed Internet
		2025-05-01	Digital Product
2	XXX E.I.	2021-01-05	High Speed Internet
		2021-01-05	Content
		2021-01-05	Manage Capacity Network
		2024-11-22	Produk Digital
3	PT. XXXB ER H.S.	2020-09-04	Manage Capacity Network
		2020-10-26	Music
		2020-10-28	High Speed Internet

Table 2 showcases sample data of SME customers purchasing multiple products simultaneously within a single transaction date. This phenomenon highlights the potential for bundling or cross-selling strategies, where frequently co-purchased products, such as High Speed Internet, Manage Capacity Network, and Content, can be offered as a unified package. The data, exemplified by customers like Kantor Pelayanan XXX N.K. and XXXI S OLS, supports the use of association rule mining algorithms like Apriori to identify frequent

itemsets in customer transactions. These insights will inform the development of targeted product bundling recommendations, enhancing marketing strategies by capitalizing on commonly paired product combination.

**Table 2. Product Bundling Identification**

No	Cust_ Name	Date of PS	Product (1)	Product (2)	Product (3)
1	Kantor Pelayanan XXX N.K.	2020-04-15	Manajemen Capacity Network	High Speed Internet	Music
2	XXXI S OLS	2021-01-14	High Speed Internet	Content	Manajemen Capacity Network
3	XXX AN F OLS	2020-12-24	High Speed Internet	Manajemen Capacity Network	Content

Table 3 displays a sample of SME customers who have subscribed to only a single product, primarily Digital Products, as seen with customers like XXXDI T.P. CV and XXXM. This group represents a significant untapped market opportunity, as many SME customers currently engage with only one service, typically connectivity-related. These single-product customers are the primary target for the replication strategy proposed in this research. By leveraging insights from sequential and association pattern analyses derived from multi-product customers, the study aims to transform these single-product subscribers into multi-product users. This table underscores the potential to apply data-driven recommendations to effectively expand product adoption among this customer segment.

**Table 3. Single Product Identification**

No.	Cust_Name	Product	Date of PS
1	XXXDI T.P. CV	Digital Product	2024-12-14
2	XXXM	Digital Product	2025-06-29
3	XXXPTIVE K.T. CV	Digital Product	2024-07-30

Mining these behavioral patterns using Sequential Pattern Mining (SPM) offers a promising solution. Prior research in domains such as retail and e-commerce demonstrates how algorithms like Apriori, PrefixSpan, and CloSpan can reveal temporal sequences in purchase behavior [3-4]. Yet, in the telco industry especially in B2B contexts serving SMEs such techniques remain underused.

This research addresses that gap by proposing a replication marketing strategy based on historical transaction data. Using SPM, we identify common product adoption sequences among existing SME clients and develop a recommendation model that replicates those paths for prospective customers, personalized by business class and industry. Our main contributions include:

- A comparative analysis of three SPM algorithms (Apriori, PrefixSpan, and CloSpan) using Telkom’s SME dataset.
- A personalized product recommendation framework for SME digital solutions.
- A validation of the framework’s effectiveness in enhancing marketing precision, sales cycle efficiency, and IT adoption.

Ultimately, this study bridges the gap between historical behavior and forward-looking strategy, offering a scalable approach for accelerating SME digitalization through intelligent, data-informed marketing

## II. METHOD

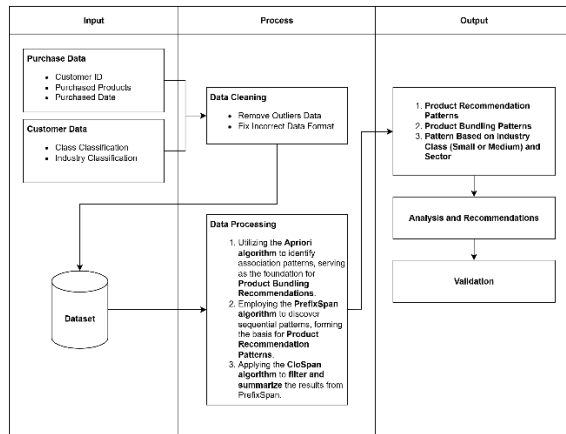
This section describes the methodological framework adopted to build a replication-based product recommendation system for SME customers. The approach integrates sequential pattern mining techniques with real transaction data to identify purchasing patterns that can inform targeted marketing strategies. Each step from data collection to algorithm implementation and model evaluation is outlined in the following subsections.

### A. RESEARCH DESIGN

This study adopts an applied quantitative research design, leveraging historical transactional data from Telkom Indonesia’s SME segment to extract sequential purchasing patterns. The goal is to develop a data-driven recommendation engine that can replicate successful product adoption paths for new customers. The overall methodology follows the data mining lifecycle: data collection, preprocessing, pattern discovery using SPM algorithms, and system evaluation.

A visual overview of the framework is shown in Fig. 3, which outlines the transformation from raw transaction logs into targeted product recommendations.

**Figure 2. Theoretical Framework: Replication-Based IT Product Recommendation for SMEs**



### B. DATA COLLECTION AND SCOPE

The dataset was sourced from Telkom Indonesia’s internal systems and contains transaction records from approximately 58,000 SME customers spanning the years 2024 and 2025. Each record includes:

- Customer ID
- Purchase date
- Subscribed IT product/service
- Customer classification (e.g., Ruko, Hotel, Multifinance)
- Industry type and geographic metadata

The study focuses on product types within Telkom's Indibiz portfolio, including connectivity, digital productivity apps, content, network solutions, and other IT-based services.

### C. SEQUENTIAL PATTERN MINING TECHNIQUES

Three SPM algorithms were implemented and compared in this study:

- Apriori: Applied for frequent itemset generation to detect commonly co-purchased services. However, it does not account for order or sequence.
- PrefixSpan: Utilized for discovering frequent subsequences using pattern-growth without candidate generation, making it efficient for large, dense datasets (Ning & Zhang, 2023).
- CloSpan: Deployed to mine closed sequential patterns, which eliminates redundancy and reduces computational overhead.

These algorithms were chosen for their complementary strengths in accuracy, efficiency, and scalability, and their relevance is established in prior literature [3-4].

### D. RECOMMENDATION ENGINE DEVELOPMENT

Based on the discovered sequential patterns, a rule-based recommendation engine was built. This engine maps each customer to a behavioral cluster (based on business class and industry), identifies their current position in the adoption sequence, and predicts the next best product to offer. Each recommendation is context-aware, considering:

- The customer's existing services
- Historical sequences from similar SMEs
- Real-time adaptability to campaign changes or product additions

The engine outputs bundled product suggestions with rank scores, allowing sales teams to execute more personalized and consultative offerings

### E. EVALUATION METRICS

To assess the effectiveness of the proposed approach, the following metrics were used:

- Support and confidence scores of sequential patterns.

- Precision and recall of recommendation outputs, benchmarked against actual future purchases (ground truth validation).
- Processing time and scalability comparison among algorithms.
- User relevance feedback, gathered from sales representatives during pilot implementation (qualitative).

## **F. ETHICAL CONSIDERATIONS**

All customer data were anonymized before processing, following Telkom's internal data governance policy. This research was approved under Telkom's internal innovation initiative and did not involve any personally identifiable information (PII).

## **III. RESULTS AND DISCUSSION**

This section presents the outcomes of applying Sequential Pattern Mining (SPM) techniques, specifically Apriori, PrefixSpan, and CloSpan, to Telkom Indonesia's SME transaction dataset, followed by a detailed discussion of their implications. The analysis focuses on identifying sequential and associative purchase patterns among SME customers to inform a replication-based product recommendation strategy. The findings validate the research hypothesis that data-driven insights can enhance marketing precision and accelerate IT product adoption in the SME segment. By comparing the performance of the three algorithms and evaluating the recommendation engine's effectiveness, this section highlights both the practical utility of the proposed framework and its contributions to personalized marketing for SMEs. The results are structured to align with the research objectives, methodology, and theoretical framework outlined in earlier sections, with each finding supported by clear data visualizations and analytical interpretations.

### **A. OVERVIEW OF PATTERN DISCOVERY**

The implementation of sequential pattern mining on historical SME transactions revealed significant behavioral sequences,

particularly in service adoption flows. For instance, SMEs categorized under "Ruko" and "Multifinance" ecosystems frequently initiated their engagement with high-speed internet packages, followed by content services and digital productivity applications. These findings support the hypothesis that SME product adoption tends to follow a replicable trajectory, influenced by business type and service dependency.

The patterns extracted through PrefixSpan and CloSpan were both more compact and semantically meaningful compared to Apriori, which often produced redundant and unordered itemsets confirming limitations cited by [9].

### **B. ALGORITHM PERFORMANCE EVALUATION**

Quantitative evaluation was conducted to compare the precision, recall, and processing time of the three algorithms. As shown in Figure 4, PrefixSpan achieved the highest precision (0.82) and recall (0.76), followed closely by CloSpan. Apriori, while effective for detecting frequent co-purchases, demonstrated lower performance metrics and significantly longer processing time (32.1 seconds), reaffirming its limitations in sequence-based contexts [7].

The superior performance of PrefixSpan is consistent with earlier retail and telco studies[8–10], suggesting its robustness in extracting high-confidence patterns in medium-to-large transactional datasets. Similarly, Apriori-based analyses have been widely applied in retail and local SME contexts to identify product association and purchase trends [11-12].

### **C. RECOMMENDATION ACCURACY AND OBSERVATION**

When applied to unseen SME profiles in Q1 2025, the recommendation engine driven by PrefixSpan-derived rules achieved an average prediction accuracy of 78%, based on comparison with actual product uptake within 60 days. This demonstrates the model's potential to replicate relevant product paths for different SME archetypes.

Sales teams who tested the engine during pilot trials noted that the generated suggestions aligned with field knowledge and customer readiness levels, particularly when recommendations were bundled with context-specific use cases (e.g., “Digital Signature” for finance SMEs post-network upgrade). These findings reinforce the growing body of research emphasizing that technology adoption among SMEs depends not only on infrastructure availability but also on managerial awareness and readiness [2], [13]. Moreover, empirical studies highlight that data-driven digital transformation models can strengthen SME competitiveness and sustainability in developing countries [4], [6]. Thus, the proposed replication-based recommendation framework aligns with global digitalization trends while addressing the unique behavioral dynamics of Indonesian SMEs.

#### IV. CONCLUSION

This study presents a data-driven approach for replicating successful IT product adoption paths among SME customers by leveraging sequential pattern mining (SPM). Using real transactional data from Telkom Indonesia, we demonstrated that SMEs follow distinct service adoption trajectories, which can be effectively extracted and generalized using algorithms such as PrefixSpan and CloSpan. The results confirm that PrefixSpan outperforms Apriori and CloSpan in both predictive accuracy and computational efficiency. The recommendation engine developed from these patterns achieved an average accuracy of 78% in real-world pilot testing, validating its practical potential for marketing personalization.

This research contributes to the growing body of literature on behavior-based product recommendations in the B2B telecommunications domain, especially for under-digitized segments like SMEs. It also highlights how internal enterprise data when ethically leveraged can be transformed into actionable insights to support strategic marketing and digital ecosystem growth.

Future work will explore integrating real-time behavioral signals, reinforcement learning for adaptive sequencing, and broader testing across varied SME sectors and geographies

#### ACKNOWLEDGMENT

The authors would like to express their gratitude to PT Telkom Indonesia for providing access to anonymized SME transactional data and for supporting this research as part of its internal innovation initiative. Special thanks are extended to the Enterprise Business Service Division for their insights on product portfolio structure and customer segmentation, and to the Data Science and Analytics Team for their technical guidance during algorithm implementation and evaluation.

This study would not have been possible without the valuable contributions of the field marketing and account manager teams who participated in the pilot testing phase and provided practical feedback to improve the recommendation framework.

The authors also acknowledge the constructive feedback from reviewers and academic advisors that helped strengthen the rigor and clarity of this research.

#### REFERENCES

- [1] M. B. Legowo, B. Indiarso, and D. Prayitno, “A Conceptualization of The Sme Digitalization Model To Support The ‘SME-Go Digital’ Program in Indonesia,” *Indones. J. Bus. Entrep.*, vol. 8, no. 3, pp. 421–429, 2022, doi: 10.17358/ijbe.8.3.421.
- [2] T. Yuwono, A. Suroso, and W. Novandari, “Information and communication technology in SMEs: a systematic literature review,” *J. Innov. Entrep.*, vol. 13, 2024, doi: 10.1186/s13731-024-00392-6.
- [3] S. Fatimah and S. K. Mukarramah, “Model Konseptual Untuk Transformasi Digital UKM Selama Pandemi Covid-19 Di Indonesia (R-Model Transformasi Digital),” *J. Soc. Soc.*, vol. 3, no. 1, pp. 11–20, 2023, doi:

- 10.54065/jss.3.1.2023.341.
- [4] Y. Nurohman, M. Kusuma, and D. Narulitasari, "Fin-Tech, Financial Inclusion, and Sustainability: a Quantitative Approach of Muslims SMEs," *Int. J. Islam. Bus. Ethics*, vol. 6, p. 54, 2021, doi: 10.30659/ijibe.6.1.54-67.
- [5] D. Prihandono, A. P. Wijaya, B. Wiratama, W. Prananta, and S. Widia, "Digital transformation to enhance Indonesian SME performance: Exploring the impact of market competition and digital strategy," *Probl. Perspect. Manag.*, vol. 22, no. 2, pp. 103–113, 2024, doi: 10.21511/ppm.22(2).2024.09.
- [6] J. Díaz-Arancibia, J. Hochstetter-Diez, A. Bustamante-Mora, S. Sepúlveda-Cuevas, I. Albayay, and J. Arango-López, "Navigating Digital Transformation and Technology Adoption: A Literature Review from Small and Medium-Sized Enterprises in Developing Countries," *Sustain.*, vol. 16, no. 14, 2024, doi: 10.3390/su16145946.
- [7] G. Aloysius and D. Binu, "An approach to products placement in supermarkets using PrefixSpan algorithm," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 25, no. 1, pp. 77–87, 2013, doi: <https://doi.org/10.1016/j.jksuci.2012.07.001>.
- [8] R. Trivonanda, R. Mahendra, I. Budi, and R. Hidayat, "Sequential Pattern Mining for e-Commerce Recommender System," 2020, pp. 393–398, doi: 10.1109/ICACSI51025.2020.9263192.
- [9] S. Ning and S. Zhang, "An improved algorithm for frequent sequence pattern mining based on PrefixSpan-ComplexPrefixSpan," in *Proceedings of the 4th International Conference on Artificial Intelligence and Computer Engineering*, 2024, pp. 48–52, doi: 10.1145/3652628.3652636.
- [10] B. Zhou, W. Hu, and T. Chen, "Pattern Extraction From Industrial Alarm Flood Sequences by a Modified CloFAST Algorithm," *IEEE Trans. Ind. Informatics*, vol. PP, p. 1, 2021, doi: 10.1109/TII.2021.3071361.
- [11] A. Setiawan and F. Putri, "Implementasi Algoritma Apriori untuk Rekomendasi Kombinasi Produk Penjualan," *Ultim. J. Tek. Inform.*, vol. 12, pp. 66–71, 2020, doi: 10.31937/ti.v12i1.1644.
- [12] S. Alya Shafira and N. Lestari Anggreini, "Analisis Pola Pembelian Konsumen Pada Data Penjualan Menggunakan Algoritma Apriori (Studi Kasus Coffee Shop Gelora Fantasi)," *J. Mhs. Tek. Inform.*, vol. 8, no. 5, pp. 10029–10034, 2024.
- [13] K. Kanojia, "Technology Adoption In Small Business Management," *Int. J. Innov. Sci. Eng. Manag.*, pp. 72–79, 2024, doi: 10.69968/ijisem.2024v3i472-79.