

Enhancing Marketing Strategies for Office Stationery Stores Using Equivalence Class Transformation Algorithms

Ega Silfa Yuliana¹, Raymond Sunardi Oetama^{2✉}

^{1,2} Department of Information Systems, Faculty of Engineering and Informatics, Universitas Multimedia Nusantara, Tangerang, Indonesia

Informasi Artikel

Riwayat Artikel

Diserahkan : 30-10-2023

Direvisi : 02-11-2023

Diterima : 10-11-2023

Kata Kunci:

Bisnis alat tulis kantor;
ECLAT; Strategi
pemasaran

Keywords :

Office supply busines;
ECLAT; Marketing
strategy,

ABSTRAK

Toko alat tulis kantor merupakan tempat yang penting bagi kantor, sekolah, dan bisnis, karena menyediakan berbagai alat dan perlengkapan yang diperlukan. Salah satu tantangan yang dihadapi oleh pemilik toko adalah bagaimana memilih produk yang tepat untuk dipromosikan agar bisnisnya berkembang dengan baik. Penelitian ini bertujuan untuk mengatasi tantangan tersebut dengan menggunakan algoritma *Equivalence Class Transformation* untuk menemukan aturan asosiasi yang berguna. Penelitian ini menggunakan beberapa metrik, seperti *support*, *confidence*, *lift ratio*, dan waktu eksekusi mengevaluasi kinerja algoritma. Hasil penelitian ini menunjukkan bahwa algoritma yang digunakan mampu menghasilkan aturan asosiasi. Dari hasil penelitian ini, ditemukan sepuluh aturan penting yang dapat membantu pemilik toko dalam membuat keputusan terkait promosi produk. Pada tahap awal pembuatan katalog promosi, prioritas diberikan pada aturan yang sangat disetujui oleh Pemilik Toko, yang pada gilirannya, memandu pemilihan item unggulan berdasarkan asosiasi produk yang kuat.

ABSTRACT

Office stationery stores facilitate operations for offices, schools, and businesses. However, the store owner must overcome the challenge of selecting the right products for promotion. This research aims to tackle this issue by comparing the performance of the *Equivalence Class Transformation* algorithms in discovering association rules using *Support*, *confidence*, and *lift ratios*. The findings reveal that the algorithm generates association rules based on *Support*, *confidence*, and *lift values*. Ten critical rules are identified, shedding light on algorithm effectiveness. Ultimately, this study underscores the significance of refining marketing approaches for brick-and-mortar stationery businesses and the value of data-driven methods in aiding decision-making. In the early promotion catalog phase, priority is assigned to rules with solid agreement by the Shop Owner, guiding the selection of featured items based on robust product associations.

Corresponding Author :

Raymond Sunardi Oetama
Information Systems Department, Faculty of Engineering and Informatics, Universitas Multimedia Nusantara, Indonesia
Jl. Scientia Boulevard, Curug Sangereng, Kec. Klp. Dua, Kabupaten Tangerang, Banten 15811
Email: raymond@umn.ac.id

INTRODUCTION

Promotion is a marketing strategy to increase awareness, interest, and sales of products or services (Mukonza and Swarts, 2020). It encompasses a range of activities such as advertising, special offers, and contests to persuade customers to make purchases and foster brand loyalty. It is a pivotal marketing mix component, conveying value and establishing differentiation from competitors. Office supply businesses typically engage in inventory management, wholesale or retail sales, marketing strategy development, and various aspects of supply chain and distribution (Imaroh et al., 2023). Some office supplies businesses may specialize in selling specific products, such as art supplies, technical equipment, or more advanced office equipment. Bundling in the office supplies industry involves offering related products (Martins et al., 2021), such as pens and notebooks or printers and ink cartridges, together at a discounted price. This strategy encourages customers to purchase multiple items as a package deal, providing added value and cost savings while boosting sales for the retailer. Customers find bundled purchases convenient and cost-effective (X. Liu & Shi, 2023), as they can acquire related items together, saving time and money. Businesses benefit from increased sales and customer satisfaction through bundling (Zhang et al., 2020), offering discounts on combined purchases. It's a win-win strategy.

Data mining is discovering and extracting valuable patterns, information, or knowledge from large datasets (Yan et al., 2020). It involves using various techniques and algorithms to analyze data and identify trends, correlations, and insights that might not be apparent through standard data analysis methods. Data mining is employed in various fields, including business, science, healthcare, and more, to make informed decisions, predict future trends, and improve processes. It can provide valuable insights for organizations to gain a competitive advantage and optimize their operations. Data mining can analyze customer transaction data and discover which products are frequently purchased together (Qisman et al., 2021). This information can help businesses identify natural pairings or complementary products that can be bundled to create attractive offers. For example, data mining may reveal that customers who purchase printers often buy ink cartridges, leading to a printer-ink cartridge bundle.

Association Rule Mining is a data mining technique that identifies patterns or associations between items in large transaction datasets (Kavitha & Subbaiah, 2020). It is especially relevant to bundling because it helps discover which products customers often purchase together. Previous research applied Association Rules for some purposes. It is instrumental in enhancing retail transactions by facilitating the strategic placement of related items, thus improving the shopping experience (Chen and Gunawan, 2023). They are crucial in detecting customer purchasing patterns (Paradkar, 2022) and providing insights into preferences and behavior (Somya et al., 2022), which informs tailored marketing strategies.

Moreover, by leveraging association rules to create bundled promotions, businesses can significantly increase profits (Wei et al., 2019), as customers are more inclined to make larger, value-driven purchases, resulting in a win-win approach that enriches the retail experience while optimizing financial outcomes. This study was conducted at Aji Shop, which sells office stationery items in Pekanbaru City. It plans to create a promotional catalog as a package of office stationery items, but the challenge lies in identifying the right products to promote. Utilizing Association Rule Mining, this promotional catalog is curated.

METHODS

Algorithms

Data mining is systematic knowledge discovery from a database (Wilim & Oetama, 2021). Association rule mining is a data mining technique first introduced in Rakesh Agrawal's research paper (Mohapatra et al., 2021). Data influences decisions in the future, as data mining techniques can extract new knowledge from existing data through processing (Santoso et al., 2021). Association rules are a powerful method to identify frequent patterns among multiple transactions (Ünvan, 2021). Each transaction comprises various items in this technique, and the algorithm's

objective is to uncover relationships and connections between these items (Aldino et al., 2021). Association rules aim to discover potential causal relationships between items within a dataset. By analyzing multiple transactions and identifying patterns, this technique helps to reveal possible associations and dependencies between different things, providing insights into the underlying structure of the data (M. Liu et al., 2022). L.H.S. (Left Hand Shake) $\rightarrow$ R.H.S. (Right Hand Shake) is an assignment form of association rule with the interpretation that if any item in L.H.S. is purchased, then the goods in R.H.S. are also purchased (Jia, 2021). The first two phases of using association rules are finding the most common combination of a data set and setting the requirements and outcomes of the association rules. The ECLAT algorithms are utilized in this investigation. ECLAT is an algorithm whose search process is a depth-first search in the database (Luna et al., 2019). The item begins by changing horizontal to vertical transactions by merging the Transaction Lists with the same item (Wu et al., 2022).

Research Flow

As illustrated in Figure 1, the research process begins by gathering data, which is then subjected to the data preparation stage, followed by the modeling phase. The ultimate step involves evaluating the performance of both generated models.

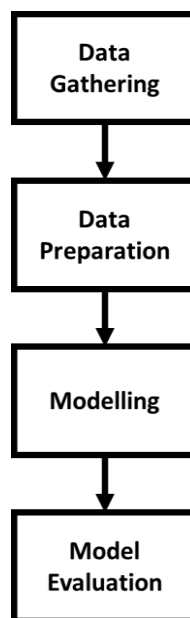


Figure 1. Research Flow

During the data gathering stage, we interviewed the shop owner to gather their specific requirements for creating association rules related to promotional procurement. The primary data source for this stage is the sales transaction reports from the shop, spanning two months from December 1, 2021, to January 31, 2022. Moving on to the data preparation stage. Data preprocessing is performed in alignment with the objectives of a Data Mining study (Hardjadinata et al., 2021). We utilize RStudio to process the raw sales transaction reports from the shop. Data cleansing is essential to address irregularities in columns and rows. Our first task involves carefully selecting the relevant attributes crucial for this study, including the transaction date, customer name, and item name. After completing the data cleansing, the refined dataset is saved in CSV format, ensuring compatibility with RStudio tools for subsequent analysis. The modeling stage involves conducting experiments to explore the performance of the ECLAT algorithm. We systematically vary vital parameters, including the Support, confidence, Length, and lift ratio. These parameters are tested using the entire transaction dataset, with minimum values set at two and a maximum length of 3. The support and confidence values are tested from 0.1 to 1, allowing

us to assess the algorithm's behavior comprehensively. Finally, we present the outcomes of the generated rules to the shop owner, seeking their valuable feedback and insights.

RESULTS AND DISCUSSION

Association Rules

Table 1 presents association rule mining results using the ECLAT algorithm, focusing on different combinations of Length, Support, and confidence thresholds. The Length is set at two, indicating that the generated association rules must involve at least two items, such as two products purchased together. The Support is started at 0.1, meaning that an item must be present in at least 10% of the total transactions to be considered a valid rule. The Confidence threshold value varies across the table from 0.1 to 0.9.

Table 1. Association Rules

Length	Support	Confidence	Number of Rules
2	0.1	0.1	25
2	0.1	0.2	25
2	0.1	0.3	25
2	0.1	0.4	18
2	0.1	0.5	17
2	0.1	0.6	7
2	0.1	0.7	3
2	0.1	0.8	1
2	0.1	0.9	0
2	0.1	0.1	25

Initially, a set of 25 association rules was generated with a confidence of 0.1, and all these rules exhibited a lift value greater than one, indicating a positive correlation between the associated items. Similar results were observed when the confidence was incrementally increased to 0.2 and 0.3, consistently yielding positively correlated rules. However, as the confidence threshold was lowered to 0.4, the number of linked rules was reduced to 18. Further decreasing the confidence to 0.5 resulted in 17 rules, while setting it at 0.6 and 0.7 led to the discovery of only 7 and 3 positively correlated rules, respectively. When a high confidence of 0.8 was applied, only one rule was obtained, and it exhibited a substantial lift value of 4.06, indicating a strong positive correlation. However, when the confidence threshold was set at 0.9, no rules met the stringent criteria, and the highest observed confidence value was only 82.35. These findings unveil the nuanced relationship between the confidence threshold and the number of positively correlated rules, highlighting the importance of setting a suitable threshold based on the specific objectives of the analysis.

The proprietor has requested that we provide the ten most exemplary rules. Our selection process for these rules is founded on their respective lift ratios. Figure 2 depicts the outcomes of applying the ECLAT algorithm to unearth association rules. The bright red highlight highlights the rule with the highest lift ratio value, a robust indicator of the correlation strength between two items. Notably, the same set of rules is identified by both algorithms, indicating their consistency in rule generation. Among these rules, the top two exhibit an identical lift value, scoring an impressive 4.06. The first of these rules asserts, "If the customer purchases Natural Paper 70gsm Kwarto (A4), then they are highly likely to purchase Natural Paper 70gsm Folio (F4)."

Conversely, the second rule states, "If the customer acquires Natural Paper 70gsm Folio (F4), it is highly likely that they will also purchase Natural Paper 70gsm Kwarto (A4)." These rules underscore a compelling and robust association between these two paper types, indicating that when customers opt for one, they are significantly inclined to choose the other. This valuable

insight can inform marketing strategies and product placement decisions, facilitating more effective and targeted approaches to drive sales and enhance customer experiences.

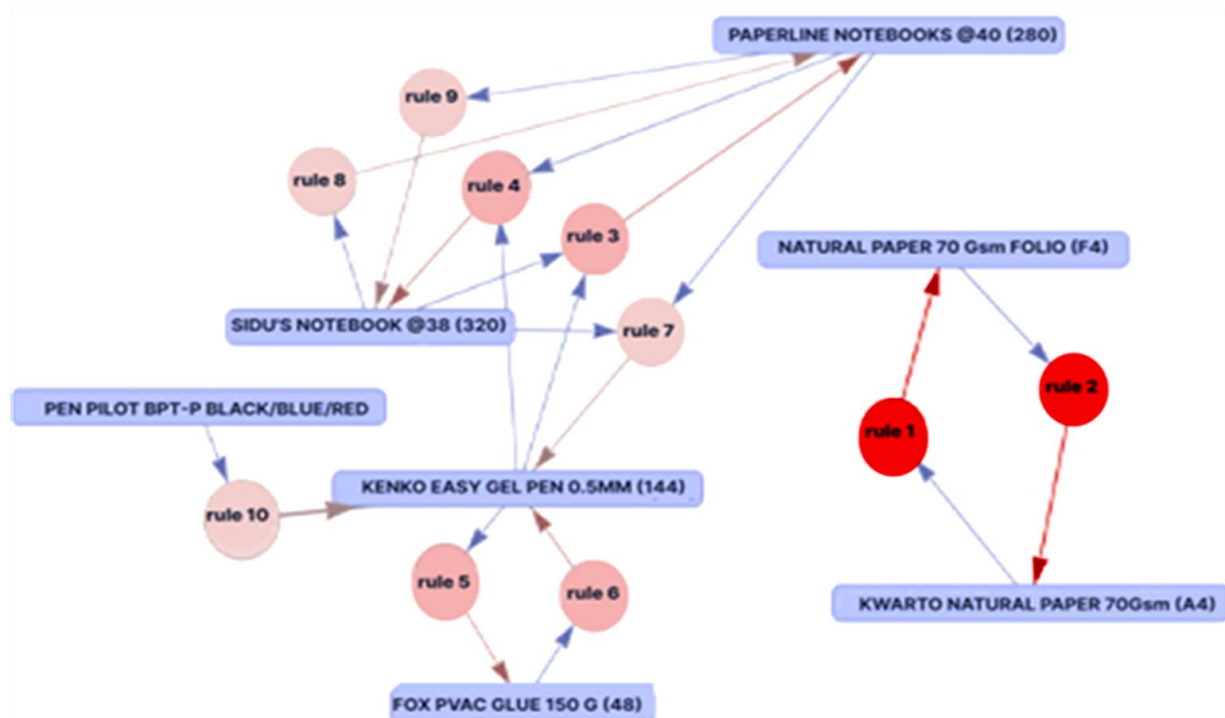


Figure 2. Rules on Highest Lift Ratio

The Shop Owner's Opinion

Table 3 presents the opinions of the shop owner on the top 10 rules with the highest lift ratio value, which were derived from the data on customer purchases. Based on the owner's experience selling these goods, four rules strongly indicate customer purchase behavior. Additionally, the owner agrees with three other rules, finds one neutral, and disagrees with two rules.

The first two rules focus on customer purchases of 70gsm natural paper in Folio and Quarto sizes. The shop owner strongly agrees with these rules as these products are popular and tend to sell out quickly. However, the owner acknowledges that some customers may only buy one size at a time based on their immediate needs. The third and fourth rules describe the purchase behavior of a large wholesaler who buys notebooks and pens from well-known brands in bulk for resale. The shop owner agrees with these rules, as these products are often purchased together in significant quantities by such buyers.

Table 2. Owner's Opinion

Rules No	Rules	Owner's Opinion	Lift Rasio
1	If 70Gsm Folio (F4) Natural Paper is purchased, then 70Gsm Quarto (A4) Natural Paper will also be purchased.	Strongly agree	4.06
2	If Natural 70Gsm Quarto (A4) Paper is purchased, then Natural 70Gsm Folio (F4) paper will also be purchased.	Strongly agree	4.06
3	If Sidu Notebook @38 (320) and Easy Gel Kenko Pen 0.5mm(144) are purchased, then Paperline Notebook @40 (280) will also be purchased.	Agree	2.41
4	If the Paperline @40 (280) Notebook and 0.5mm(144) Easy Gel Kenko Pen are purchased, then the Sidu @38 (320) Notebook will also be purchased.	Agree	2.22
5	If Fox Pvac 150 G(48) Glue is purchased, then Easy Gel Kenko 0.5mm(144) Pen will also be purchased.	Disagree	2.22
6	If Easy Gel Kenko 0.5mm(144) Pen is purchased, then Fox Pvac 150 G(48) Glue will also be purchased	Disagree	2.22
7	If Sidu Notebook @38 (320) and Paperline Notebook @40 (280) are purchased, then Easy Gel Kenko 0.5mm(144) will also be purchased.	Neutral	1.96
8	If Sidu Notebook @38 (320) is purchased, then Paperline Notebook @40 (280) will also be purchased.	Strongly agree	1.96
9	If Paperline Notebook @40 (280) is purchased, then Sidu Notebook @38 (320) will also be purchased.	Strongly agree	1.96
10	If the BPT-P Black/Blue/Red Pilot Pen is purchased, then the Easy Gel Kenko 0.5mm(144) Pen will also be purchased.	Agree	1.88

In contrast, the owner disagrees with the fifth and sixth rules, which suggest a correlation between the purchase of pens and glue. The owner notes that while large buyers purchase pens in bulk, they are likelier to buy multiple brands than glue. Thus, the owner sees no significant correlation between these two products. The shop owner takes a neutral stance towards the association rule that suggests a correlation between the purchase of Sidu Notebook @38 (320) and Paperline Notebook @40 (280) with Easy Gel Kenko 0.5mm (144). While some customers buy pens as a complement to their notebook purchase, the owner is not confident in the accuracy or reliability of this association rule.

Nonetheless, the owner believes it may still be helpful to monitor this association rule over time to see if it holds up and can be leveraged to improve sales strategies in the future. The eighth and ninth rules are simplified versions of the third, fourth, and seventh rules, focusing solely on purchasing notebooks. The shop owner strongly approves of these two rules as customers tend to buy notebooks together, but not in the same volume as paper purchases described in the first and second rules. Finally, the tenth rule has earned the shop owner's approval, highlighting an exciting trend: customers who purchase BPT-P Black/Blue/Red Pilot Pens will likely buy Kenko's Easy Gel pen, but not vice versa.

As a result, the promotion catalog is successfully made. The rules that have received solid agreement are prioritized in the initial phase of creating the promotion catalog. These rules guide the catalog's featured items based on strong associations between products. These rules are foundational for building an effective and customer-centric promotion catalog. When a customer purchases the '70Gsm Folio (F4) Natural Paper,' it is highly likely that they will also acquire the '70Gsm Quarto (A4) Natural Paper.' Conversely, if a customer chooses the 'Natural 70Gsm Quarto (A4) Paper,' it strongly indicates that they will also opt for the 'Natural 70Gsm Folio (F4)

Paper.' Similarly, the purchase of 'Sidu Notebook @38 (320)' is associated with the acquisition of 'Paperline Notebook @40 (280)' and vice versa.

CONCLUSION

Using the ECLAT algorithm, we explore the impact of varying confidence thresholds on association rule mining. Starting with a low confidence of 0.1, we generate 25 positively correlated rules, all with lift values above one. As we incrementally raise the confidence threshold to 0.4, 0.5, and beyond, the number of linked rules decreases, reflecting a more robust filter for association strength. When we reach a high confidence of 0.8, only one robust rule with a lift of 4.06 is obtained. However, no rules meet the criteria at a higher threshold of 0.9.

The ECLAT algorithm reveals ten compelling rules, shedding light on significant associations within the dataset:

1. Customers who purchase '70Gsm Folio (F4) Natural Paper' are highly likely also to buy '70Gsm Quarto (A4) Natural Paper,' and vice versa, indicating a robust correlation.
2. The acquisition of 'Natural 70Gsm Quarto (A4) Paper' suggests a substantial likelihood of purchasing 'Natural 70Gsm Folio (F4) Paper,' demonstrating a consistent association.
3. 'Sidu Notebook @38 (320)' and 'Paperline Notebook @40 (280)' are closely linked, with customers often acquiring one alongside the other, establishing a solid connection.
4. Conversely, selecting 'Paperline Notebook @40 (280)' strongly indicates customers choosing 'Sidu Notebook @38 (320),' reinforcing the strong association between these products.
5. 'Natural Paper 70gsm Folio (F4)' and '70Gsm Quarto (A4) Natural Paper' exhibit a strong and consistent association, with customers frequently opting for one when purchasing the other.
6. The purchase of '70Gsm Quarto (A4) Natural Paper' is reliably correlated with the selection of 'Natural Paper 70gsm Folio (F4),' affirming a significant connection between these paper types.
7. Customers purchasing 'Sidu Notebook @38 (320)' are also inclined to acquire 'Paperline Notebook @40 (280),' indicating a consistent association between these items.
8. Similarly, 'Paperline Notebook @40 (280)' buyers frequently acquire 'Sidu Notebook @38 (320),' reinforcing the strong relationship between these products.
9. The association between 'Natural Paper 70gsm Folio (F4)' and 'Natural Paper 70gsm Quarto (A4)' is strong, with customers often favoring one when purchasing the other.
10. The final rule underscores the significant link between '70Gsm Quarto (A4) Natural Paper' and '70Gsm Folio (F4) Natural Paper,'

The promotion catalog has been successfully made. During the initial phase of promotion catalog creation, prioritization is given to rules characterized by strong agreement, which guides the selection of featured items based on robust product associations. These rules are foundational in developing an effective, customer-centric promotion catalog.

ACKNOWLEDGEMENTS

We want to express our gratitude to Universitas Multimedia Nusantara for its Support in terms of facilities and funding for our research.

REFERENCES

- Aldino, A. A., Pratiwi, E. D., Setiawansyah, Sintaro, S., & Putra, A. D. (2021). Comparison of Market Basket Analysis to Determine Consumer Purchasing Patterns Using Fp-Growth and Apriori Algorithm. *2021 International Conference on Computer Science, Information Technology, and Electrical Engineering, ICOMITEE 2021*, 29–34. <https://doi.org/10.1109/ICOMITEE53461.2021.9650317>

- Chen, A. H.-L., & Gunawan, S. (2023). Enhancing Retail Transactions: A Data-Driven Recommendation Using Modified R.F.M. Analysis and Association Rules Mining. *Applied Sciences*, 13(18), 10057. <https://doi.org/10.3390/app131810057>
- Hardjadinata, H., Oetama, R. S., & Prasetiawan, I. (2021). Facial Expression Recognition Using Xception And DenseNet Architecture. *2021 6th International Conference on New Media Studies (CONMEDIA)*, 60–65.
- Imaroh, T. S., Widiyani, K., & Muttaqien, F. (2023). Value Chain and Supply Chain Management of Products from Women Farmer Groups in South Tangerang. *Wiga: Jurnal Penelitian Ilmu Ekonomi*, 13(2), 190–203.
- Jia, X. (2021). Research on the Role of Big Data Technology in the Reform of English Teaching in Universities. *Wireless Communications and Mobile Computing*, 2021, 1–13. <https://doi.org/10.1155/2021/9510216>
- Kavitha, & Subbaiah. (2020). Association Rule Mining using Apriori Algorithm for Extracting Product Sales Patterns in Groceries. *International Journal of Engineering Research and Technology (IJERT)*, 8(3), 5–8. www.ijert.org
- Liu, M., Gao, Y., Yang, K., Xu, J., Zheng, Q., Zhao, L., Ge, L., Li, L., Zhang, J., & Tian, J. (2022). Interpretation of STROBE-MR: a statement for strengthening the reporting of observational studies in epidemiology using Mendelian randomization. *Chinese Journal of Evidence-Based Medicine*, 22(8), 978–987. <https://doi.org/10.7507/1672-2531.202112103>
- Liu, X., & Shi, A. (2023). *Online Bundled Pricing Strategy for Agricultural Products Considering Consumers' Organic Preferences*.
- Luna, J. M., Fournier-Viger, P., & Ventura, S. (2019). Frequent itemset mining: A 25 years review. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 9(6), e1329. <https://doi.org/10.1002/widm.1329>
- Martins, P., Rodrigues, P., Martins, C., Barros, T., Duarte, N., Dong, R. K., Liao, Y., Comite, U., & Yue, X. (2021). Preference between individual products and bundles: effects of complementary, price, and discount level in Portugal. *Journal of Risk and Financial Management*, 14(5), 192.
- Mohapatra, D., Tripathy, J., Mohanty, K. K., & Nayak, D. S. K. (2021). Interpretation of Optimized Hyper Parameters in Associative Rule Learning using Eclat and Apriori. *Proceedings - 5th International Conference on Computing Methodologies and Communication, ICCMC 2021*, 879–882. <https://doi.org/10.1109/ICCMC51019.2021.9418049>
- Mukonza, C., & Swarts, I. (2020). The influence of green marketing strategies on business performance and corporate image in the retail sector. *Business Strategy and the Environment*, 29(3), 838–845. <https://doi.org/10.1002/bse.2401>
- Paradkar, R. U. (2022). *Detecting Customer Purchasing Patterns using Association Rule Mining*. Dublin, National College of Ireland.
- Qisman, M., Rosadi, R., & Abdullah, A. S. (2021). Market basket analysis using apriori algorithm to find consumer patterns in buying goods through transaction data (case study of Mizan computer retail stores). *Journal of Physics: Conference Series*, 1722(1), 12020. <https://doi.org/10.1088/1742-6596/1722/1/012020>
- Santoso, J. T., Jumini, S., Bhawika, G. W., Susilo, D., Wibowo, D., & Rahim, R. (2021). Unsupervised Data Mining Technique for Clustering Library in Indonesia. *Library Philosophy and Practice*, 4866.
- Somya, R., Winarko, E., & Priyanta, S. (2022). A hybrid recommender system based on customer behavior and transaction data using generalized sequential pattern algorithm. *Bulletin of Electrical Engineering and Informatics*, 11(6), 3422–3432.

- Ünvan, Y. A. (2021). Market basket analysis with association rules. *Communications in Statistics - Theory and Methods*, 50(7), 1615–1628. <https://doi.org/10.1080/03610926.2020.1716255>
- Wei, X., Wei, Y., Chen, P., Fan, C., Luo, H., Zhao, Q., & Kong, Y. (2019). The Analysis of "Online Silk Road" from the Perspective of Big Data. In *The New Silk Road Leads through the Arab Peninsula: Mastering Global Business and Innovation* (pp. 97–114). Emerald Publishing Limited.
- Wilim, N. N., & Oetama, R. S. (2021). Sentiment Analysis About Indonesian Lawyers Club Television Program Using K-Nearest Neighbor, Naïve Bayes Classifier, And Decision Tree. *IJNMT (International Journal of New Media Technology)*, 8(1), 50–56. <https://doi.org/10.31937/ijnmt.v8i1.1965>
- Wu, P., Niu, X., Fournier-Viger, P., Huang, C., & Wang, B. (2022). UBP-Miner: An efficient bit based high utility itemset mining algorithm. *Knowledge-Based Systems*, 248, 108865. <https://doi.org/10.1016/j.knosys.2022.108865>
- Yan, H., Yang, N., Peng, Y., & Ren, Y. (2020). Data mining in the construction industry: Present status, opportunities, and future trends. *Automation in Construction*, 119, 103331.
- Zhang, Y., Niyato, D., Wang, P., & Han, Z. (2020). Data Services Sales Design with Mixed Bundling Strategy: A Multidimensional Adverse Selection Approach. *IEEE Internet of Things Journal*, 7(9), 8826–8836. <https://doi.org/10.1109/JIOT.2020.2999824>