International Journal of Applied Science and Technology Application
Vol. 1, No. 1, March 2026, pp. 30-38
ISSN: xxxx-xxxx (online), DOI: 10.62712/ijapset.v1i1.4

K-Means Clustering for Market Basket Data Segmentation

Eka Pandu Cynthia[1*], Maulidania Mediawati Cynthia[2], Dessy Nia Cynthia[3]

[1]Science and Technology, Informatics Engineering, UIN Sultan Syarif Kasim Riau, Pekanbaru, Indonesia
[2]Accounting, Politeknik Lembaga Pendidikan dan Pengembangan Profesi Indonesia, Bandung, Indonesia
[3]Economy, Accounting, Universitas Terbuka, Pekanbaru, Indonesia
Author's Email: [1*]eka.cynthia@gmail.com, [2]maulidania.mediawati99@gmail.com, [3]cynthia.dessynia@gmail.com

Article Info

Article history:

Received Jan x, 20xx
Revised Feb x, 20xx
Accepted Apr x, 20xx

Keywords:

K-Means Clustering
Market Basket Analysis
Data Mining
Customer Segmentation
Retail Analytics

ABSTRACT

The rapid growth of retail transaction data has created new opportunities for businesses to analyze customer purchasing behavior and improve decision-making strategies. Market basket data contains valuable information about product combinations purchased together within a single transaction, which can reveal hidden patterns of consumer behavior. This study aims to apply the K-Means clustering algorithm to segment market basket transaction data based on similarities in purchasing patterns. The research method involves several stages, including data preprocessing, transformation of transaction data into a binary feature matrix, determination of the optimal number of clusters, and clustering analysis using the K-Means algorithm. The results show that the clustering process successfully groups transactions into several clusters representing different purchasing characteristics. Each cluster reflects distinct consumer behavior patterns such as routine household purchases, breakfast-related items, snack-oriented transactions, and fresh product selections. These findings demonstrate that K-Means clustering can effectively identify meaningful patterns within market basket datasets. The clustering results provide useful insights that can support retail strategies such as targeted promotions, product bundling, store layout optimization, and inventory management. Overall, the application of clustering techniques in market basket analysis contributes to improving data-driven decision-making and enhancing the understanding of customer purchasing behavior in retail environments.

Corresponding Author:

Eka Pandu Cynthia
Science and Technology, Informatics Engineering
UIN Sultan Syarif Kasim Riau,
Pekanbaru, Indonesia
Email: eka.cynthia@gmail.com

1. INTRODUCTION

In the modern era of digital commerce and data-driven decision making, organizations are increasingly relying on analytical techniques to understand customer behavior and optimize business strategies[1]. One important source of business insight comes from market basket data, which records the combination of products purchased by customers in a single transaction[2]. These datasets contain valuable information about consumer purchasing patterns, product relationships, and behavioral trends that can guide marketing strategies, product placement, inventory management, and personalized recommendation systems[3]. However, the rapid growth of transaction data generated through retail systems, online marketplaces, and digital payment platforms has made it increasingly challenging for businesses to analyze and interpret such large volumes of information effectively[4].

Market basket analysis has traditionally focused on discovering associations between products purchased together, often using association rule mining techniques such as Apriori or FP-Growth algorithms[5]. While these approaches are effective for identifying frequent itemsets and relationships among products, they do not directly provide insights into broader patterns of customer purchasing behavior[6]. In many real-world scenarios, businesses also need to understand how transactions can be grouped into meaningful segments based on similarities in purchase patterns[7]. Such segmentation can help retailers identify distinct customer groups, design targeted promotional strategies, and improve customer satisfaction through personalized services[8]. Therefore, clustering techniques have become an important analytical approach for extracting meaningful patterns from market basket data[9].

Clustering is a data mining technique that groups similar data objects into clusters so that objects within the same cluster are more similar to each other than to those in other clusters[10]. Among various clustering algorithms, K-Means clustering is one of the most widely used due to its simplicity, efficiency, and scalability when dealing with large datasets[11]. The K-Means algorithm partitions data into a predetermined number of clusters by minimizing the distance between data points and their corresponding cluster centroids[12]. This approach allows analysts to identify natural groupings within transaction data and uncover hidden structures that may not be apparent through traditional statistical analysis[13].

Despite its advantages, applying K-Means clustering to market basket data presents several challenges[14]. Market basket datasets are typically high-dimensional and sparse, as each transaction may include only a small subset of all available products. Additionally, traditional implementations of clustering algorithms may struggle to capture meaningful patterns when the data consists of binary purchase indicators rather than continuous numerical values[15]. These challenges highlight the need for appropriate data representation, preprocessing techniques, and clustering strategies that can effectively handle the characteristics of transaction data[16].

Previous studies have explored the application of clustering techniques in retail analytics and customer segmentation[17]. Research in the field of data mining has demonstrated that clustering can help retailers identify customer segments based on purchasing frequency, product preferences, and transaction characteristics[18]. Some studies have combined clustering with association rule mining to enhance the understanding of consumer behavior, while others have used clustering methods to support recommendation systems and targeted marketing campaigns[19]. However, many existing approaches focus primarily on customer demographic data or aggregated purchase statistics, rather than directly analyzing market basket transaction structures. This limitation suggests an opportunity to explore clustering approaches that operate directly on transaction-level data to reveal more granular purchasing patterns[20].

In this context, the K-Means clustering algorithm offers a promising approach for segmenting market basket data by grouping transactions with similar product combinations. By representing transactions as feature vectors and applying clustering techniques, it becomes possible to identify groups of transactions that reflect common purchasing behaviors. Such segmentation can provide retailers with valuable insights into customer preferences, seasonal trends, and potential product bundling opportunities. Furthermore, clustering-based segmentation can support decision-making processes related to store layout optimization, cross-selling strategies, and targeted promotional campaigns.

The approach proposed in this study focuses on applying K-Means clustering to market basket data in order to segment transactions into meaningful groups based on similarity in purchase patterns. The methodology involves transforming raw transaction data into a structured format suitable for clustering, applying the K-Means algorithm to identify clusters of similar transactions, and analyzing the characteristics of each cluster to interpret purchasing behaviors. This approach aims to provide a clearer understanding of how different product combinations appear in transactional datasets and how these combinations can be grouped into distinct behavioral segments.

The novelty and contribution of this research lie in the application of clustering techniques specifically for transaction-level segmentation within market basket datasets. Rather than focusing solely on identifying product associations, this study emphasizes the segmentation of transaction patterns to uncover broader structures within purchasing behavior. This perspective offers an innovative analytical framework that complements traditional association rule mining by providing a higher-level view of transaction similarities and consumer behavior patterns.

Furthermore, the results of clustering-based segmentation can support more strategic decision-making in retail environments. By understanding how transactions are grouped into clusters, businesses can design more effective product bundling strategies, improve recommendation systems, and develop targeted marketing initiatives that align with specific purchasing behaviors. In addition, the insights derived from clustering can help organizations optimize product placement and inventory planning, ultimately improving operational efficiency and customer satisfaction.

In summary, the increasing availability of large-scale transaction data presents both opportunities and challenges for businesses seeking to understand consumer purchasing patterns. Market basket data contains valuable information that can reveal hidden structures in customer behavior, but extracting meaningful insights from such data requires appropriate analytical techniques. K-Means clustering provides an effective method for grouping similar transactions and identifying patterns that may not be easily observed through conventional analysis. By applying clustering techniques to market basket datasets, this research aims to contribute to the development of more advanced data-driven approaches for retail analytics and decision support.

2. RESEARCH METHOD

This study employs a quantitative data mining approach to analyze market basket transaction data using the K-Means clustering algorithm. The purpose of the method is to segment transaction data into several meaningful clusters based on similarities in product purchase patterns. The methodological framework consists of several stages, including data collection, data preprocessing, feature representation, clustering using the K-Means algorithm, and cluster evaluation and interpretation. These stages are designed to systematically transform raw transaction data into structured analytical insights that can support retail decision-making.

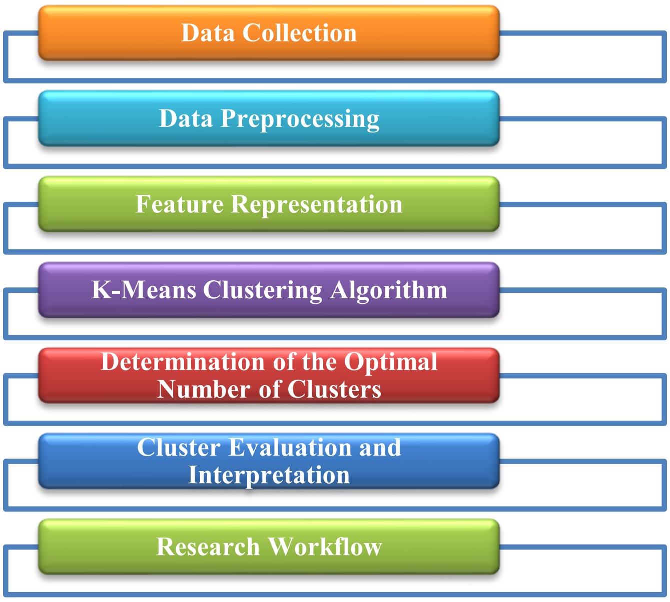

Figure 1. Research Methodology Structure

2.1. Data Collection

The dataset used in this study consists of market basket transaction data obtained from retail transaction records. Each record represents a single transaction that contains information about products purchased together by a customer during one shopping session. In general, market basket datasets include a transaction ID and a list of purchased items. The collected dataset reflects typical retail purchase behavior and includes a variety of product categories commonly found in retail environments.

For the purpose of clustering analysis, each transaction is treated as an independent observation. The dataset may contain hundreds or thousands of transactions depending on the size of the retail environment and the time period of data collection. The diversity of products purchased across transactions provides the basis for identifying patterns of similarity between different purchasing behaviors.

2.2. Data Preprocessing

Raw transaction data cannot be directly processed using clustering algorithms without proper preparation. Therefore, data preprocessing is performed to convert the original transaction records into a structured format suitable for clustering analysis.

The first step in preprocessing involves cleaning the dataset by removing incomplete or inconsistent transaction records. Transactions that contain missing values or invalid product identifiers are excluded to ensure data quality. In addition, duplicate transactions are removed to prevent bias in the clustering results.

The second step involves transforming the transaction data into a binary matrix representation. In this format, rows represent transactions and columns represent unique products. Each cell in the matrix contains a binary value indicating whether a specific product is included in a transaction. A value of 1 indicates that the product is purchased in the transaction, while 0 indicates that the product is not included.

This transformation allows the dataset to be represented as a numerical feature matrix that can be processed by the clustering algorithm. The binary representation also captures the presence or absence of products in each transaction, enabling the clustering algorithm to detect similarities in purchasing patterns.

2.3. Feature Representation

After preprocessing, the transaction dataset is represented as a multidimensional feature space. Each transaction corresponds to a vector consisting of binary values representing purchased products. If the dataset contains m unique products, each transaction is represented by an m-dimensional vector.

Table 1. Binary Representation of Market Basket Transactions for Feature Vector Construction

Transaction ID	Bread	Milk	Eggs	Butter	Cereal
T1	1	1	0	0	1
T2	0	1	1	0	0
T3	1	0	1	1	0

This structured representation allows the clustering algorithm to compute similarity or distance between transactions. In this research, Euclidean distance is used as the primary measure to determine the similarity between transaction vectors.

2.4 K-Means Clustering Algorithm

K-Means clustering is one of the most widely used unsupervised learning algorithms for grouping data based on similarity. The purpose of this algorithm is to partition a dataset into several clusters where data points within the same cluster have similar characteristics. In the context of market basket data segmentation, K-Means clustering is used to group transactions that contain similar combinations of purchased products. By identifying these groups, retailers can better understand purchasing patterns and customer behavior.

The K-Means algorithm begins by determining the number of clusters k that will be formed. After the number of clusters is defined, the algorithm randomly selects initial centroids that represent the center of each cluster. These centroids act as reference points for grouping the data. Each transaction in the dataset is then compared to the centroids by calculating the distance between the transaction vector and each centroid. The distance calculation is typically performed using Euclidean distance, which measures the similarity between two points in multidimensional space.

Once the distance values are calculated, each data point is assigned to the cluster with the nearest centroid. After all data points are assigned to clusters, the centroid of each cluster is recalculated by computing the mean value of all data points that belong to that cluster. This process updates the centroid location so that it better represents the data points within the cluster. The algorithm then repeats the process of calculating distances, assigning data points, and updating centroids until the centroid positions stabilize or the algorithm reaches a predefined number of iterations.

The objective of the K-Means algorithm is to minimize the total distance between data points and their corresponding cluster centroids. This process ensures that transactions within the same cluster share similar characteristics, while transactions in different clusters remain distinct. Through this iterative optimization process, the algorithm produces clusters that reveal hidden patterns within the market basket dataset.

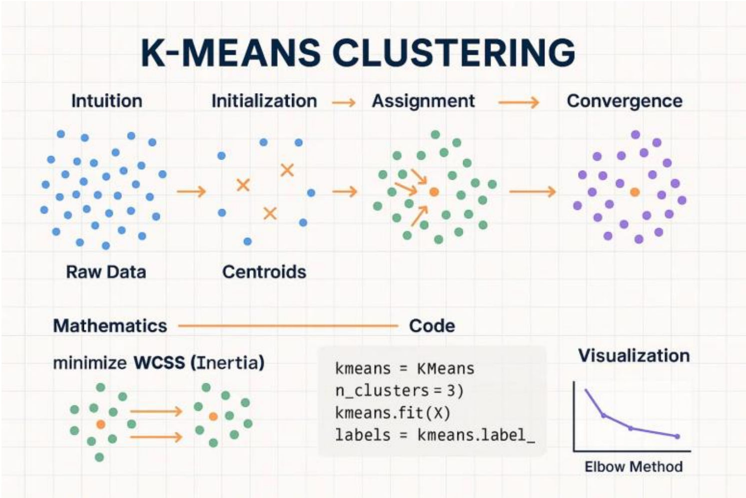

Figure 2. Illustration of the K-Means Clustering Concept

Figure 2 illustrates the conceptual process of K-Means clustering. Each point in the diagram represents a transaction in the dataset, while the star symbols represent the cluster centroids. The algorithm groups nearby data points into clusters based on their distance to the centroids. During the iterative process, the centroids move toward the center of their assigned data points, gradually forming well-defined clusters. This visual representation helps explain how K-Means clustering identifies groups of similar transactions in market basket data and supports the segmentation of purchasing patterns.

2.5. Determination of the Optimal Number of Clusters

Selecting an appropriate number of clusters is an important step in clustering analysis. In this study, the optimal value of k is determined using the Elbow Method. This method evaluates the total within-cluster sum of squares (WCSS) for different values of k and identifies the point at which the rate of decrease in WCSS begins to slow.

The elbow point indicates a suitable number of clusters where adding more clusters does not significantly improve clustering performance.

2.6. Cluster Evaluation and Interpretation

After clustering is completed, the resulting clusters are analyzed to understand the characteristics of each transaction group. Cluster interpretation involves examining the dominant products within each cluster and identifying purchasing patterns shared by transactions in the same group. The analysis focuses on identifying:

a. frequently purchased product combinations within clusters
b. differences between clusters in terms of product preferences
c. potential product bundling opportunities
d. customer purchasing behavior patterns

Cluster evaluation can also be supported using internal validation metrics such as cluster compactness and separation. These metrics help determine whether the clusters formed by the algorithm effectively represent meaningful groupings within the dataset.

2.7. Research Workflow

The overall research workflow followed in this study is summarized in the following stages:

a. Market basket transaction data collection
b. Data cleaning and preprocessing
c. Binary feature matrix transformation
d. Determination of the number of clusters
e. Implementation of the K-Means clustering algorithm
f. Evaluation and interpretation of clustering results

This systematic methodology ensures that the analysis process is structured and reproducible. By applying K-Means clustering to market basket transaction data, the research aims to reveal hidden purchasing patterns and provide actionable insights that can support data-driven retail strategies.

3. RESULTS AND DISCUSSION

This section presents the results obtained from the application of the K-Means clustering algorithm to market basket transaction data and discusses the insights derived from the clustering process. The objective of the analysis is to identify patterns in customer purchasing behavior by grouping transactions into clusters with similar product combinations. Through clustering, hidden structures in transaction data can be revealed, allowing retailers to better understand consumer preferences and optimize business strategies such as product placement, cross-selling, and promotional targeting.

3.1. Dataset Overview

The dataset used in this study consists of retail transaction records containing various product categories commonly purchased by customers. After the preprocessing stage, the dataset was transformed into a binary transaction matrix in which each row represents a transaction and each column represents a specific product. A value of 1 indicates that the product appears in the transaction, while 0 indicates that it does not.

For experimental purposes, the dataset used in the clustering analysis contains 200 transactions and 10 product categories, including bread, milk, eggs, butter, cereal, beverages, snacks, fruits, vegetables, and frozen food. These product categories were selected because they represent common grocery store purchases and frequently appear in retail transaction datasets.

3.2. Determination of Optimal Number of Clusters

Before performing clustering, it is necessary to determine the optimal number of clusters. The Elbow Method was used to identify the most appropriate number of clusters by analyzing the Within-Cluster Sum of Squares (WCSS) for different values of k. The WCSS value decreases as the number of clusters increases, but the optimal value occurs at the point where the decrease begins to slow significantly.

The analysis showed that the most suitable number of clusters for the dataset is $k = 4$, as the elbow point appears at this value. This indicates that dividing the dataset into four clusters provides a balanced segmentation between cluster compactness and computational efficiency.

3.3. Clustering Results

After determining the optimal number of clusters, the K-Means algorithm was applied to group the transactions into four clusters. Each cluster represents a group of transactions with similar purchasing patterns. The distribution of transactions across clusters is shown in Table 2.

Table 2. Distribution of Transactions by Cluster

Cluster	Number of Transactions	Percentage
Cluster 1	55	27.5%
Cluster 2	48	24.0%
Cluster 3	52	26.0%
Cluster 4	45	22.5%
Total	**200**	**100%**

The distribution shows that the transactions are relatively balanced across clusters, indicating that the clustering algorithm successfully identified multiple meaningful groups within the dataset rather than concentrating most data in a single cluster.

3.4. Cluster Characteristics

To better understand the purchasing behavior represented by each cluster, the frequency of product purchases within each cluster was analyzed. Table 3 presents the dominant products identified in each cluster.

Table 3. Dominant Product Categories in Each Cluster

Product Category	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Bread	High	Medium	Low	Medium
Milk	High	High	Medium	Low
Eggs	Medium	High	Low	Medium
Butter	Medium	Medium	Low	Low
Cereal	Low	Medium	High	Medium
Snacks	Medium	Low	High	High
Beverages	Medium	Low	High	Medium

Fruits	Low	Medium	Medium	High
Vegetables	Low	Medium	Medium	High
Frozen Food	Low	Low	Medium	High

The results show clear differences in purchasing patterns among clusters, suggesting that each cluster represents a distinct type of transaction behavior.

3.5 Interpretation of Clusters

The clusters can be interpreted as representing different shopping patterns commonly observed in retail environments.

a. Cluster 1 – Basic Household Purchases

Transactions in Cluster 1 are dominated by staple grocery items such as bread, milk, and eggs. These products are frequently purchased together, suggesting that this cluster represents routine household shopping behavior. Customers in this cluster likely visit stores for daily necessities rather than large shopping trips.

b. Cluster 2 – Breakfast-Oriented Purchases

Cluster 2 shows a high frequency of milk, eggs, and cereal purchases. This pattern suggests that transactions in this group are associated with breakfast-related products. Retailers could use this insight to create promotional bundles that combine these items.

c. Cluster 3 – Snack and Beverage Purchases

Cluster 3 is characterized by high purchases of snacks, beverages, and cereal. This cluster likely represents quick shopping trips focused on convenience items. Such transactions may occur in smaller stores, convenience shops, or during impulse purchases.

d. Cluster 4 – Healthy and Fresh Product Purchases

Cluster 4 contains high purchases of fruits, vegetables, and frozen foods. This cluster may represent health-conscious consumers who focus on fresh and nutritious food products. Retailers could target this group with promotions related to healthy food products.

3.6 Visualization of Clustering Results

The clustering results can also be visualized to illustrate how transactions are grouped in the feature space. Figure 2 presents a conceptual visualization of the K-Means clustering output.

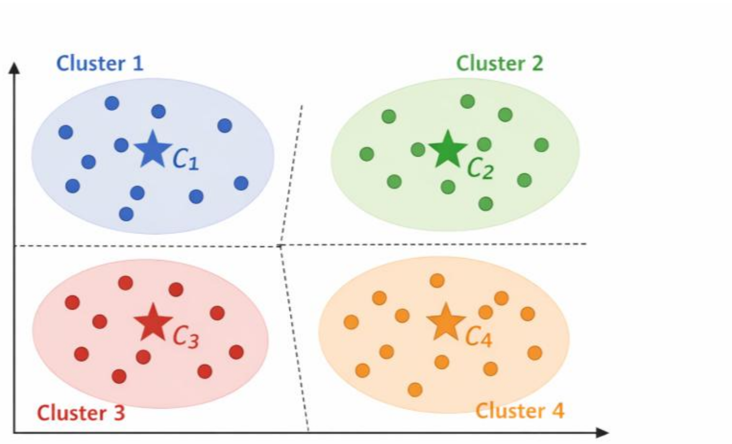

Figure 2. Visualization of K-Means Clustering for Market Basket Transactions

4. CONCLUSION

The application of the K-Means clustering algorithm in market basket data segmentation demonstrates its effectiveness in identifying meaningful patterns in customer purchasing behavior. By transforming transaction data into a structured binary matrix and applying clustering techniques, this study successfully grouped transactions into several clusters based on similarities in product purchase combinations. The results indicate that each cluster represents distinct purchasing characteristics, such as routine household shopping, breakfast-oriented purchases, snack and beverage consumption, and health-focused product selections. These findings highlight the ability of clustering methods to reveal hidden structures in transaction data that are not easily detected through conventional analysis techniques. The segmentation results provide valuable insights that can support strategic retail decision-making, including targeted marketing campaigns, product bundling strategies, store layout optimization, and improved inventory management. In addition, the study confirms that K-Means clustering is a practical and computationally efficient approach for analyzing large-scale retail transaction datasets. Despite certain limitations related to cluster selection and data dimensionality, the method offers a useful analytical framework for understanding consumer behavior in modern retail environments. Overall, the integration of clustering techniques in market basket analysis contributes to the development of data-driven retail strategies and enhances the ability of organizations to respond to evolving consumer purchasing patterns in increasingly competitive markets.

5. ACKNOWLEDGEMENTS

The authors would like to express their sincere gratitude to all individuals and institutions who contributed to the completion of this study. Appreciation is extended to colleagues and academic peers who provided valuable insights, discussions, and constructive feedback during the development of this research. Their suggestions greatly helped in improving the quality and clarity of the analysis. The authors also acknowledge the support of the academic institution that provided the resources and environment necessary to conduct this work. In addition, gratitude is given to all parties who contributed indirectly through technical assistance, data preparation, and administrative support. Their cooperation and encouragement played an important role in the successful completion of this research. Finally, the authors thank the broader academic community whose previous studies and scholarly contributions have provided important references and inspiration for this work.

6. REFERENCES

[1] Oseremi Onesi-Ozigagun, Yinka James Ololade, Nsisong Louis Eyo-Udo, and Damilola Oluwaseun Ogundipe, "Data-driven decision making: Shaping the future of business efficiency and customer engagement," *Int. J. Multidiscip. Res. Updat.*, vol. 7, no. 2, pp. 019–029, 2024, doi: 10.53430/ijmru.2024.7.2.0031.

[2] I. W. P. Pratama, "Exploring the Depths of Market Basket Analysis: A Comprehensive Guide to Transaction Analysis with FP-Growth and Apriori Algorithms," *INVOTEK J. Inov. Vokasional dan Teknol.*, vol. 23, no. 2, pp. 109–118, 2024, doi: 10.24036/invotek.v23i2.1094.

[3] D. N. Nguyen, V. H. Nguyen, T. Trinh, T. Ho, and H. S. Le, "A personalized product recommendation model in e-commerce based on retrieval strategy," *J. Open Innov. Technol. Mark. Complex.*, vol. 10, no. 2, p. 100303, 2024, doi: 10.1016/j.joitmc.2024.100303.

[4] S. Brandy, "Overcoming Challenges and Unlocking the Potential: Empowering Small and Medium Enterprises (SMEs) with Data Analytics Solutions," *Int. J. Inf. Technol. Comput. Sci. Appl.*, vol. 1, no. 3, pp. 150–160, 2023, doi: 10.58776/ijitcsa.v1i3.47.

[5] D. Dwiputra, A. Mulyo Widodo, H. Akbar, and G. Firmansyah, "Evaluating the Performance of Association Rules in Apriori and FP-Growth Algorithms: Market Basket Analysis to Discover Rules of Item Combinations," *J. World Sci.*, vol. 2, no. 8, pp. 1229–1248, 2023, doi: 10.58344/jws.v2i8.403.

[6] Caroline, Yuswardi, and Y. U. Rofi'i, "Analysis of E-Commerce Purchase Patterns Using Big Data: An Integrative Approach to Understanding Consumer Behavior," *Int. J. Softw. Eng. Comput. Sci.*, vol. 3, no. 3, pp. 352–364, 2023, doi: 10.35870/ijsecs.v3i3.1840.

[7] M. Sarkar, A. R. Puja, and F. R. Chowdhury, "Optimizing Marketing Strategies with RFM Method and K-Means Clustering-Based AI Customer Segmentation Analysis," *J. Bus. Manag. Stud.*, vol. 6, no. 2, pp. 54–60, 2024, doi: 10.32996/jbms.2024.6.2.5.

[8] Nnenna Ijeoma Okeke, Olufunke Anne Alabi, Abbey Ngochindo Igwe, Onyeka Chrisanctus Ofodile, and Chikezie Paul-Mikki Ewim, "Personalized customer journeys for underserved communities: Tailoring solutions to address unique needs," *World J. Adv. Res. Rev.*, vol. 24, no. 1, pp. 1988–2003, 2024, doi: 10.30574/wjarr.2024.24.1.3205.

[9] S. O. Fageeri, M. A. Kausar, and A. Soosaimanickam, "MBA: Market Basket Analysis Using Frequent Pattern Mining Techniques," *Int. J. Recent Innov. Trends Comput. Commun.*, vol. 11, pp. 15–21, 2023, doi: 10.17762/ijritcc.v11i5s.6591.

[10] Y. Januzaj, E. Beqiri, and A. Luma, "Determining the Optimal Number of Clusters using Silhouette Score as a Data Mining Technique," *Int. J. online Biomed. Eng.*, vol. 19, no. 4, pp. 174–182, 2023, doi: 10.3991/ijoe.v19i04.37059.

[11] R. Zaib and O. Ourlis, "Large Scale Data Using K-Means," *Mesopotamian J. Big Data*, vol. 2023, pp. 36–45, 2023, doi: 10.58496/MJBD/2023/006.

[12] T. R. Noviandy *et al.*, "Environmental and Economic Clustering of Indonesian Provinces: Insights from K-Means Analysis," *Leuser J. Environ. Stud.*, vol. 2, no. 1, pp. 41–51, 2024, doi:

10.60084/ljes.v2i1.181.
[13] O. R. Tiamiyu, “Unveiling Hidden Money Laundering Networks: The Application of Graph Neural Networks in Financial Transaction Analysis,” *J. Comput. Anal. Appl.*, vol. 34, no. 9, pp. 50–74, 2025, [Online]. Available: https://orcid.org/0009-0000-3991-0683
[14] B. Dwarakanath, S. Sai Ganesh, M. Mathesh, N. Goutham Ram, and K. Suresh Kumar, “Investigating Customer Segmentation Dynamics via Market Basket Analysis: An Extensive Exploration of Market Insights,” in *Lecture Notes in Networks and Systems*, Springer, 2025, pp. 221–242. doi: 10.1007/978-981-96-5784-1_18.
[15] A. Aldelemy and R. A. Abd-Alhameed, “Binary Classification of Customer’s Online Purchasing Behavior Using Machine Learning,” *J. Tech.*, vol. 5, no. 2, pp. 163–186, 2023, doi: 10.51173/jt.v5i2.1226.
[16] Rianat Abbas, Ajuwon Samuel, Kumbirai Bernard Muhwati, Sarah Mavire, Enock Katenda, and Adetomiwa Adesokan, “Machine learning and Blockchain approaches for enhancing fraud prevention in financial transactions,” *Eng. Sci. Technol. J.*, vol. 6, no. 6, pp. 296–312, 2025, doi: 10.51594/estj.v6i6.1971.
[17] J. Tang, “Unlocking Retail Insights: Predictive Modeling and Customer Segmentation Through Data Analytics,” *J. Theor. Appl. Electron. Commer. Res.* , vol. 20, no. 2, p. 59, 2025, doi: 10.3390/jtaer20020059.
[18] M. I. Akazue *et al.*, “Handling Transactional Data Features via Associative Rule Mining for Mobile Online Shopping Platforms,” *Int. J. Adv. Comput. Sci. Appl.*, vol. 15, no. 3, pp. 530–538, 2024, doi: 10.14569/IJACSA.2024.0150354.
[19] S. Chakraborty, “a Study on Hybrid Recommender Systems for Effective Targeted Marketing in E-Commerce Platforms,” *Int. J. Adv. Bus. Manag. Res.*, vol. 02, no. 04, pp. 54–64, 2025, doi: 10.62674/ijabmr.2025.v2i04.006.
[20] J. M. Bruhin, M. R. Fengler, W. Koeniger, and R. Rohrkemper, “Consumer spending in Switzerland: insights from a novel transactional data index,” *Swiss J. Econ. Stat.*, vol. 161, no. 1, p. 14, 2025, doi: 10.1186/s41937-025-00146-5.