

[Click here](#) and write your Article Category

Implementation of K-Means Clustering for Inventory Item Classification in Inventory Management

Febry Yusmita

Department of Information System, Faculty of Computer Science, Universitas Dharmawangsa, Medan, 20221, North Sumatra, Indonesia

ARTICLE INFORMATION

Received: February 00, 00
Revised: March 00, 00
Available Online: April 00, 00

KEYWORDS

K-Means Clustering; Data Mining; Inventory Management; RapidMiner; Inventory Classification

CORRESPONDENCE

Phone: +62 896-9267-4646
E-mail: febiyusnita2702@gmail.com

A B S T R A C T

Effective inventory management is essential for maintaining stock availability, minimizing storage costs, and improving operational efficiency. However, managing large numbers of inventory items can be challenging when classification is performed manually or based on subjective criteria. Therefore, a data-driven approach is needed to group inventory items according to their characteristics and usage patterns. This study aims to implement the K-Means clustering algorithm to classify inventory items and support more systematic inventory management. This study employs a quantitative approach using historical inventory data as the primary dataset. The data undergo preprocessing, including data cleaning, handling missing values, and feature selection. Relevant inventory attributes, such as stock quantity, item demand, and transaction frequency, are used as clustering variables, depending on data availability. The K-Means algorithm is applied to group inventory items into clusters based on similarities in their characteristics. The resulting clusters are analyzed and interpreted to identify inventory patterns and support stock management decisions. Cluster quality can be evaluated using the Silhouette Score or other appropriate clustering evaluation metrics. The implementation of K-Means clustering is expected to produce distinct groups of inventory items based on their characteristics and usage patterns. These groups can help identify items with different levels of demand, stock movement, and inventory management priorities. The clustering results can also provide useful information for determining replenishment priorities, monitoring stock levels, and improving the allocation of storage resources. The application of K-Means clustering provides a data-driven approach to inventory item classification that can support more structured and efficient inventory management. By identifying similarities among inventory items, the proposed method can assist inventory managers in making informed decisions regarding stock monitoring and replenishment planning. Nevertheless, the effectiveness of the classification depends on data quality, the selection of relevant attributes, and the appropriate determination of the number of clusters. Further evaluation using actual inventory data is necessary to establish the practical benefits of the proposed approach.

INTRODUCTION

Inventory management is one of the most important operational activities in a company because it directly affects product availability, customer satisfaction, and operational efficiency. Proper inventory management enables companies to maintain stock availability according to demand while minimizing storage costs and the risk of overstock or stock shortages. However, many companies still rely on conventional inventory recording systems, making inventory control less effective and prone to data inconsistencies[1][2][3].

PT. Deli Farm Indonesia is a company engaged in the distribution of fresh food and daily necessities. The company conducts product sales through online platforms such as WhatsApp and Shopee. Although digital sales have been

implemented, inventory management is still unable to provide comprehensive information regarding product sales patterns and inventory movement. Consequently, management often experiences difficulties in determining appropriate stock levels, resulting in excess inventory for some products while other products experience shortages. These conditions reduce operational efficiency and increase inventory holding costs[4][5][6].

Advances in information technology have encouraged organizations to utilize data mining techniques for extracting useful knowledge from large volumes of data. Data mining provides analytical methods capable of identifying hidden patterns, predicting future conditions, and supporting strategic decision-making. One of the most widely used unsupervised learning techniques is the K-Means Clustering algorithm, which groups data objects according to their similarity based on distance calculations. This method has the advantages of simple implementation, fast computation, and the ability to process large datasets efficiently[7][8].

In inventory management, K-Means Clustering can be applied to classify products according to sales characteristics and stock movement. Products with similar transaction patterns can be grouped into the same cluster, allowing companies to identify fast-moving, medium-moving, and slow-moving inventory items. Such information is valuable for determining inventory policies, warehouse arrangement, purchasing strategies, and production planning. Compared with manual analysis, clustering provides a more objective and data-driven approach for inventory decision-making.

Several previous studies have demonstrated that the K-Means algorithm has been successfully implemented in various clustering problems, including customer segmentation, student classification, and product grouping. Nevertheless, studies focusing on inventory classification in fresh product distribution companies remain relatively limited. Therefore, this study attempts to fill this gap by implementing the K-Means Clustering algorithm on inventory transaction data from PT. Deli Farm Indonesia using the Knowledge Discovery in Databases (KDD) framework and RapidMiner as the data mining platform.

The objective of this research is to implement the K-Means Clustering algorithm to classify inventory items based on sales characteristics and inventory movement. The clustering results are expected to provide useful information for inventory control, improve warehouse management efficiency, and support managerial decision-making in determining inventory policies.

METHOD

This study employed a quantitative approach using data mining techniques to classify inventory items at PT. Deli Farm Indonesia. The research adopted the Knowledge Discovery in Databases (KDD) framework, which consists of several systematic stages, namely data collection, data selection, preprocessing, data transformation, data mining, and interpretation of the results. The KDD framework was selected because it provides a structured process for extracting meaningful knowledge from large datasets and has been widely implemented in data mining research.

The data used in this study were obtained from inventory transaction records at PT. Deli Farm Indonesia. The dataset contains information related to inventory items, including product codes, inventory quantities, and sales transaction data during the observation period. Prior to the clustering process, the collected data underwent preprocessing to eliminate duplicate records, correct inconsistent data, and remove incomplete information that could affect clustering performance. Data transformation was subsequently performed to convert the dataset into a format suitable for analysis using the K-Means algorithm.

The clustering process was conducted using the K-Means Clustering algorithm. Initially, the number of clusters (k) was determined based on the research objectives. The algorithm then randomly selected initial centroids and calculated the Euclidean distance between each data object and the centroids. Each inventory item was assigned to the nearest cluster according to the minimum distance value. After all data objects were assigned, new centroid positions were recalculated based on the average values of each cluster. This iterative process continued until no further changes occurred in cluster membership, indicating that the clustering process had converged.

The Euclidean Distance formula was applied to measure the similarity between data objects and cluster centroids, as expressed in Equation (1).

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

where $d(x, y)$ represents the Euclidean distance between two data objects, x_i denotes the value of the observed attribute, y_i represents the centroid value of the corresponding attribute, and n indicates the total number of attributes used in the clustering process.

The implementation of the K-Means algorithm was performed using RapidMiner software. RapidMiner was selected because it provides an integrated environment for data preprocessing, clustering, visualization, and evaluation without requiring extensive programming knowledge. The clustering results generated by RapidMiner were analyzed to identify inventory groups with similar sales characteristics and stock movement patterns. These results were subsequently interpreted to provide recommendations for improving inventory management, warehouse arrangement, and inventory planning at PT. Deli Farm Indonesia.

RESULTS AND DISCUSSION

This study utilized inventory sales transaction data obtained from PT. Deli Farm Indonesia during the January–February 2025 observation period. A total of 40 inventory items representing fresh food products and staple commodities were analyzed using the K-Means Clustering algorithm. The sales quantity recorded for each product during the observation period served as the clustering variables. Before the clustering process was performed, the dataset underwent preprocessing to remove duplicate records, correct inconsistent data, and ensure data completeness. The cleaned dataset was subsequently imported into RapidMiner to perform the clustering analysis.

Table 1. Sample of Inventory Sales Dataset

Code	Product	January (kg)	February (kg)
KD_1_TSB	Tepung Segitiga Biru	12	19
KD_2_TMN	Timun	48	44
KD_3_MS	Minyak Sunco	60	58
KD_4_KP	Kopi	450	448
KD_5_BRS	Beras	60	70
KD_6_AF	Ayam Filet	115	200
KD_7_BB	Blue Band	69	75
KD_8_AKF	Ayam Katsu Frozen	105	115
KD_9_LMN	Lemon	30	29
KD_10_BM	Bawang Merah	65	73

Source: PT. Deli Farm Indonesia (processed by the authors).

Table 1 presents a sample of the inventory sales dataset used in this research. The complete dataset consisted of 40 inventory items with monthly sales quantities that became the basis for clustering analysis. These sales values represent product movement and were selected because they reflect inventory demand during the observation period. Higher sales values indicate products with greater inventory turnover, while lower values indicate products with slower movement. The clustering analysis was conducted using the K-Means algorithm with the number of clusters set to two ($k = 2$). The algorithm initialized the centroid positions randomly, calculated the Euclidean distance between each inventory item and the cluster centroids, reassigned cluster membership based on the minimum distance, and recalculated the centroid positions until convergence was achieved. The iterative process stopped after the second iteration because no further changes in cluster membership occurred.

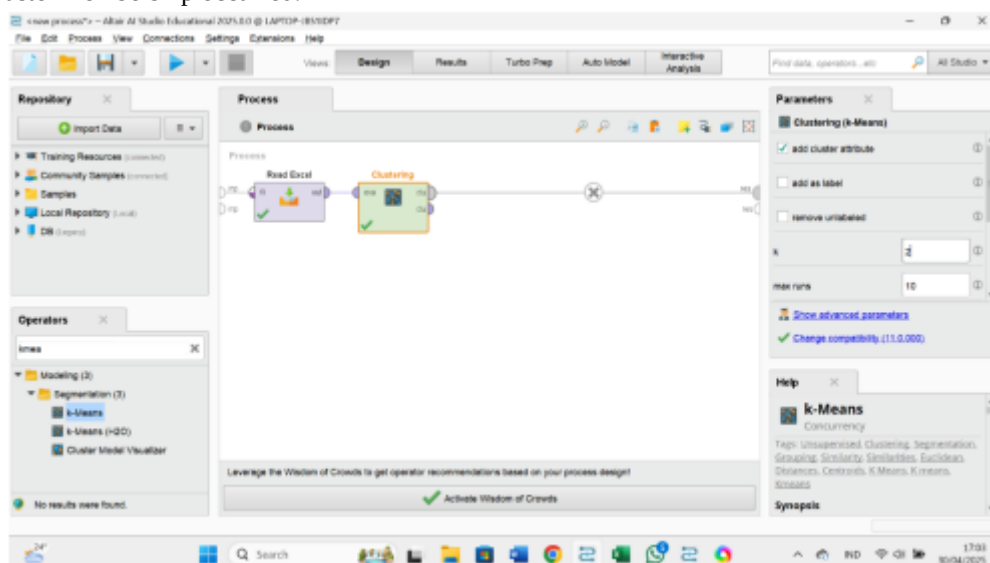


Figure 1. Implementation of the K-Means Clustering process in RapidMiner.

Figure 1 illustrates the implementation of the K-Means algorithm using RapidMiner. The process begins with importing the inventory dataset, followed by connecting the dataset to the K-Means operator, specifying the number of clusters, and executing the clustering process. RapidMiner automatically performs centroid calculation, cluster assignment, and iteration until stable clustering results are obtained. This implementation significantly simplifies the clustering procedure compared with manual computation.

The final clustering results classified all inventory items into two distinct groups according to their sales characteristics. The classification produced meaningful information regarding inventory demand patterns that can be utilized as a basis for inventory planning.

Table 2. Summary of Final Clustering Results

Cluster	Characteristics	Inventory Recommendation
Cluster 1	Low to moderate sales	Maintain minimum inventory level
Cluster 2	High sales	Increase inventory availability

Source: Research Results.

Based on Table 2, Cluster 1 consists of products with relatively low inventory turnover, including vegetables, fish, meat, spices, and several processed food products. These products require more conservative inventory management because excessive stock may increase storage costs and the risk of product deterioration. In contrast, Cluster 2 contains products with consistently high sales volumes, primarily staple commodities that experience continuous customer demand. Consequently, products assigned to Cluster 2 should receive higher replenishment priority to ensure product availability and prevent stock shortages.

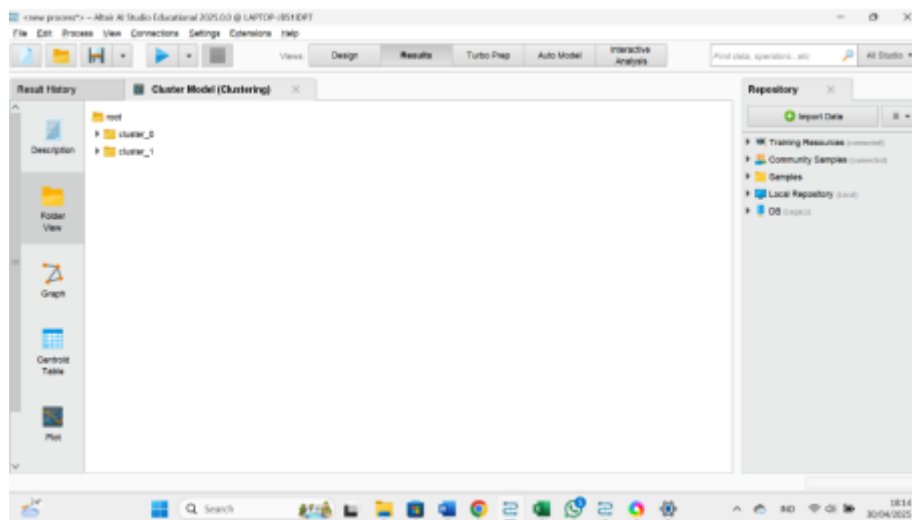


Figure 2. Visualization of inventory clustering generated by RapidMiner.

Figure 2 presents the visualization of the final clustering results generated by RapidMiner. The software displays the distribution of inventory items according to the similarity of their sales characteristics, making it easier to distinguish between low-demand and high-demand products. The graphical visualization also confirms that the clustering process successfully separated the inventory data into two well-defined groups.

Overall, the implementation of the K-Means Clustering algorithm successfully identified inventory groups with different sales characteristics. The resulting clusters provide valuable decision support for inventory management, warehouse organization, procurement planning, and stock control at PT. Deli Farm Indonesia. These findings are consistent with previous studies demonstrating that K-Means Clustering is an effective approach for inventory classification. Nevertheless, this study is limited by the use of a two-month observation period and only two clustering groups. Future research may improve the model by incorporating longer observation periods, additional inventory attributes, and comparative clustering techniques such as Fuzzy C-Means, DBSCAN, or Hierarchical Clustering to obtain more comprehensive classification results.

CONCLUSION

This study successfully implemented the K-Means Clustering algorithm to classify inventory items at PT. Deli Farm Indonesia based on sales transaction data. Using the Knowledge Discovery in Databases (KDD) framework and RapidMiner software, the inventory data were successfully grouped into two clusters representing different sales characteristics. Cluster 1 consists of products with relatively low to moderate sales volumes, whereas Cluster 2 contains

products with consistently high sales volumes and inventory turnover. These clustering results provide meaningful information regarding inventory demand patterns and can support more effective inventory planning and warehouse management.

The implementation of the K-Means algorithm demonstrates that data mining techniques can effectively transform inventory transaction data into useful knowledge for managerial decision-making. The resulting product classification enables the company to prioritize inventory replenishment for high-demand products while minimizing excessive stock for low-demand products. Consequently, the proposed approach contributes to improving inventory management efficiency, reducing storage costs, and supporting more accurate inventory control strategies at PT. Deli Farm Indonesia. Although the proposed approach produced satisfactory clustering results, this study is limited by the use of sales transaction data collected over only a two-month observation period and by employing a fixed number of two clusters. Future research is recommended to incorporate longer observation periods, additional inventory attributes, and comparative clustering algorithms, such as Fuzzy C-Means, DBSCAN, or Hierarchical Clustering, to improve clustering accuracy and provide more comprehensive decision support for inventory management.

REFERENCES

Buku

- [1] SULTAN ZAKY / HELLO WORLD JURNAL ILMU KOMPUTER - VOL. 5 NO. 2 (2026) EDISI JULI ISSN: 2829-8616 (ONLINE) <https://doi.org/10.30743/xxxxx Sultan Zaky177>
- [2] Andi Zulherry, Muhammad Basri, Muhammad Haris, Ferdy Riza, Zuli Agustina Gultom, Farid Akbar Siregar, Okvi
- [3] Nugroho, Mahardika Abdi Prawira Tanjung. Komunikasi Data dan Jaringan Komputer. Medan: UMSU Press, 2025, pp. 202.
- [4] Indah Purnama Sari. Algoritma dan Pemrograman. Medan: UMSU Press, 2023, pp. 290.
- [5] Indah Purnama Sari. Buku Ajar Pemrograman Internet Dasar. Medan: UMSU Press, 2022, pp. 300.
- [6] Indah Purnama Sari. Buku Ajar Rekayasa Perangkat Lunak. Medan: UMSU Press, 2021, pp. 228.
- [7] Janner Simarmata Arsan Kumala Jaya, Syarifah Fitrah Ramadhani, Niel Ananto, Abdul Karim, Betrisandi,
- [8] Muhammad Ilham Alhari, Cucut Susanto, Suardinata, Indah Purnama Sari, Edson Yahuda Putra. Komputer dan Masyarakat. Medan: Yayasan Kita Menulis, 2024, pp.162.
- [9] Mahdianta Pandia, Indah Purnama Sari, Alexander Wirapraja Fergie Joanda Kaunang, Syarifah Fitrah Ramadhani
- [10] Stenly Richard Pungus, Sudirman, Suardinata Jimmy Herawan Moedjahedy, Elly Warni, Debby Erce Sondakh. Pengantar Bahasa Pemrograman Python. Medan : Yayasan Kita Menulis, 2024, pp.180
- [11] Zelvi Gustiana Arif Dwinanto, Indah Purnama Sari, Janner Simarmata Mahdianta Pandia, Supriadi Syam, Semmy Wellem Taju Fitrah Eka Susilawati, Asmah Akhriana, Rolly Junius Lontaan Fergie Joanda Kaunang. Perkembangan Teknologi Informatika. Medan: Yayasan Kita Menulis, 2024, pp.158

Jurnal

- [12] A. Ariyanto, R. Bangun, M. R. M. Indillah, A. F. M. Trenggana, D. R. Sholihah, M. Ariyanti, et al., Manajemen Pemasaran. Medan, Indonesia: Yayasan Kita Menulis, 2023.
- [13] A. Selay, G. D. Andigha, A. Alfarizi, M. I. B. Wahyudi, M. N. Falah, M. Encep, and M. Khaira, "Sistem Informasi Penjualan," Karimah Tauhid, vol. 2, no. 1, pp. 232–237, 2023.
- [14] S. Aulia, "Klasterisasi Pola Penjualan Pestisida Menggunakan Metode K-Means Clustering (Studi Kasus di Toko Juanda Tani Kecamatan Hutabayu Raja)," Djtechno: Jurnal Teknologi Informasi, vol. 1, no. 1, pp. 1–5, 2020.
- [15] F. Handayani, "Aplikasi Data Mining Menggunakan Algoritma K-Means Clustering untuk Mengelompokkan Mahasiswa Berdasarkan Gaya Belajar," Jurnal Teknologi dan Informasi, vol. 12, no. 1, pp. 46–63, 2022.
- [16] P. Apriyani, A. R. Dikananda, and I. Ali, "Knowledge Discovery in Database pada Data Mining," 2023.
- [17] Z. Zulham et al., "Implementasi Knowledge Discovery in Database Menggunakan RapidMiner," 2023.
- [18] N. Rahmawita et al., "Database sebagai Sistem Penyimpanan Data Terintegrasi," 2023.
- [19] A. Gunawan, "Hubungan Persediaan Dengan Penjualan Pada Laporan Keuangan Koperasi Bukit Muria Jaya Tahun 2021–2022," Jurnal PBM: Pengembangan Bisnis dan Manajemen, vol. 23, pp. 95–109, 2023.
- [20] N. Sari, "Perencanaan dan Pengendalian Persediaan Barang dalam Upaya Meningkatkan Efektivitas Gudang," Jurnal Bisnis, Logistik dan Supply Chain (BLOGCHAIN), vol. 2, no. 2, pp. 85–91, 2022.
- [21] A. M. D. Putri, "Analisis Pengendalian Stok Beras Komersial dengan Metode Economic Order Quantity (EOQ) dan Reorder Point (ROP) di Perum Bulog Cabang Surakarta," 2025.

- [22] I. Darmansyah, "Sistem Informasi Persediaan Barang pada Rumah Sakit Umum Daerah Kabupaten Batu Bara," *SABER: Jurnal Teknik Informatika, Sains dan Ilmu Komunikasi*, vol. 3, no. 1, pp. 169–177, 2025.
- [23] A Zulherry, FA Siregar, ZA Gultom, EA Raihan (2023). Optimalisasi Website untuk Monitoring Jaringan OPD di Dinas Kominfo Kota Medan dengan Metode Triangulasi. *Bulletin of Computer Science Research* 3 (5), 357-363
- [24] Sari, I.P., Hariani, P.P., Al-Khowarizmi, A., Ramadhani, F., Sulaiman, O.K., Satria, A., & Manurung, A.A. (2024). CLUSTERING HIV/AIDS DISEASE USING K-MEANS CLUSTERING ALGORITHM. *Proceeding International Seminar on Islamic Studies* 5 (1), 1668-1676
- [25] Sari, I.P., Ramadhani, F., Satria, A., & Sulaiman, O.K. Leukocoria Identification: A 5-Fold Cross Validation CNN and Adaboost Hybrid Approach. *2023 6th International Seminar on Research of Information Technology and Intelligent Systems (ISRITI)*, 486-491
- [26] Manurung, A.A., Nasution, M.D., & Sari, I.P. (2023). Implementation of Fuzzy K-Nearest Neighbor Method in Dengue Disease Classification. *2023 11th International Conference on Cyber and IT Service Management (CITSM)*, 1-4
- [27] A Zulherry, TS Gunawan, W Wanayumini (2021). Analisis Hasil Pendukung Keputusan Mendapatkan Rumah Dinas Perusahaan Menggunakan Metode Analytical Hierarchy Process (AHP) dan Technique for Order Preference by Similarity to Ideal Solution (TOPSIS). *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 2021 I. Dan And K. Tik, "Pemanfaatan Metode Clustering Untuk Memiliki Keterampilan Teknologi," Vol. 02, No. 02, Pp. 36–41, 2023.
- [28] A Ichsan, A Zulherry, TA Lubis, BAZ Shahnaz (2025). Utilization of Mobile Applications to Speed Up The Search for Android-Based Index Places. *IJATCoS: Indonesian Journal of Applied Technology, Computer and Science* 2 (1)
- [29] Sari, I.P., Batubara, I.H., & Al-Khowarizmi, A. (2021). Sensitivity Of Obtaining Errors In The Combination Of Fuzzy And Neural Networks For Conducting Student Assessment On E-Learning. *International Journal of Economic, Technology and Social Sciences (Injects)* 2 (1), 331-338
- [30] Sari, I.P., Al-Khowarizmi, A., & Batubara, I.H. (2021). Cluster Analysis Using K-Means Algorithm and Fuzzy CMeans Clustering For Grouping Students' Abilities In Online Learning Process. *Journal of Computer Science, Information Technology and Telecommunication Engineering* 2 (1), 139-144
- [31] Apdilah, D., & Sari, I.P. (2021). Optimization Of The Fuzzy C-Means Cluster Center For Credit Data Grouping Using Genetic Algorithms. *Al'adzkiya International of Computer Science and Information Technology (AIoCSIT) Journal* 2 (2), 156-163
- [32] W. Budiono, I. Nugroho, N. A. Santoso, And G. Gunawan, "Penerapan Metode Fuzzy K-Means Clustering Untuk Pengelompokan Konten Halaman Web Secara Otomatis," *INNOVATIVE: Journal Of Social Science Research*, Vol. 4, No. 3, Pp. 13232–13241, 2024.
- [33] A Zulherry (2023) Decision making for network security with simple additive weighting method. *Journal of Intelligent Decision Support System (IDSS)* 6 (3), 155-159
- [34] M Basri, A Zulherry (2025). Analysis of the Impact of Gambling and Online Loans in the Perspective of Informatics, Islam, and Kemuhammadiyah. *AR-RASYID: Jurnal Pendidikan Agama Islam* 5 (1) Vol. 15, No. 2, Pp. 160–174, 2015.