
ANALYSIS OF THE APPLICATION OF THE K-MEANS ALGORITHM TO THE CLUSTERING METHOD APPROACH FOR GROUPING CONSUMER PURCHASING TRENDS IN ONE OF THE TEXTILE COMPANIES

Debby Kurniawan¹, Hidayat Anwari², Christina Juliane³

Sekolah Tinggi Manajemen Informatika dan Komputer LIKMI, Bandung, Indonesia

valentinoanwar99@gmail.com¹, hidayatanwari22@gmail.com², Christina.juliane@likmi.ac.id³

ABSTRACT

In Indonesia, the regulation of sexual abuse crimes is a critical aspect of ensuring justice and protection for victims. However, challenges remain in the effectiveness and comprehensiveness of these regulations. This study aims to analyze and evaluate the current legal framework addressing sexual abuse in Indonesia, identifying gaps and proposing improvements to enhance legal protections for victims. The research employs a qualitative approach, utilizing legal analysis and case studies to assess the application of existing laws. Data collection involves reviewing legal documents, court cases, and expert interviews to gather comprehensive insights into the regulatory landscape. The findings indicate significant shortcomings in the legal framework, including inconsistencies in legal definitions, procedural delays, and inadequate victim support mechanisms. The study discusses the implications of these findings, emphasizing the need for a more cohesive and victim-centered approach in legal reforms. This research underscores the necessity for legislative improvements to address the identified gaps in the regulation of sexual abuse crimes. Recommendations include clearer legal definitions, expedited legal processes, and enhanced victim support services. These measures are essential for ensuring justice and effective protection for victims of sexual abuse in Indonesia.

Keywords: Clustering, K-Means Algorithm, Data Mining, Consumer Purchasing Trends, Textile Industry, Business Strategy.

Corresponding Author: Debby Kurniawan

E-mail: valentinoanwar99@gmail.com



INTRODUCTION

In an era of increasingly free business climate development, companies face increasingly complex challenges and fierce competition (Mujiati, 2023). There are often obstacles for companies where demand or demand is enormous. However, the company cannot meet the needs of this demand because one of the factors causing this is having a large stock of goods (Eunike et al., 2021). However, the type or type differs from the demand (need) consumers need. Therefore, a stock of goods fills the storage area (Wijaya et al., 2016).

On the other hand, the company needs land to store or restock goods in high consumer demand. One of the ways to survive and thrive while changing market dynamics is for companies to sharpen their business strategies (Yunus, 2016). The right strategy is needed so that the company can provide added value to consumers and remain relevant in an increasingly competitive market (Fatyandri et al., 2023). In addition to a sense of trust, the company also needs to pay attention to a sense of comfort for consumers (Darmawan, 2017).

According to Rofaida et al. (2019), one effective strategic approach is to focus on product development that provides added value to consumers, both in terms of benefits and quality. However, challenges come not only from inter-company competition but also from the demands of society and the environment.

In this context, a company is not only faced with competition with business competitors who do the same business and produce similar products (Sanawiri & Iqbal, 2018). However, it must also consider its responsibilities to society and the environment to meet the rapidly changing market needs in this increasingly complex era (April, 2018). The products produced not only have to be of high quality by existing standards, but the service to customer satisfaction must also be considered and be environmentally friendly and pay attention to social aspects such as human rights by the provisions of the International Labor Organization (ILO) (Toriq & Martoatmodjo, 2014).

Therefore, this research aims to utilize data mining with a clustering approach to examine strategies that companies can apply in producing products that are not only from the quality aspect but also pay attention to aspects of trust and comfort for consumers and social commitment. Thus, this research is expected to contribute to understanding modern business dynamics and guide companies in developing sustainable and responsible business strategies. The existence of research can be adopted and become an evaluation for companies in other fields.

METHOD

This section will explain the research design, which should include the selection of data samples, data collection methods, and data analysis techniques. For example, this research may use a secondary data analysis approach by using historical data on textile companies' consumer purchases over time. In addition, it will explain how the clustering analysis process will be carried out by separating the data into groups with similar purchasing patterns using the K-Means algorithm.

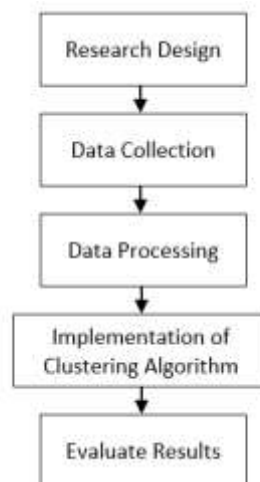


Figure 1. Research Stage

Data collection will describe the data sources used in this research. The data source may come from the textile company's internal database or data on an analytics platform. The data collection process will be described in detail, including data preprocessing processes such as cleaning, removing missing values, and variable coding if required.

Algorithm implementation will explain how to implement the K-Means algorithm in the clustering analysis of consumer purchase trends. The explanation will include stages such as centroid initialization, distance calculation, data clustering, centroid update, and iteration until convergence. In addition, it will explain how this algorithm is applied to the collected consumer purchase dataset.

In this result evaluation, we will describe the metrics used to evaluate the quality of clustering produced by the K-Means algorithm (Sari et al., 2018). Examples may include metrics such as inertia,

silhouette score, or Davies-Bouldin index. In addition, we will explain how the clustering results are evaluated and how the findings from the analysis are interpreted.

Data mining involves finding interesting patterns or information in selected data using specific techniques or methods (Mardi, 2017). Data mining is a process that involves discovering functional patterns, information, or knowledge from large and complex data sets. It involves analyzing data from multiple points of view and using various techniques or methods to uncover patterns hidden in the data (Juwita & Ali, 2024). Data mining mainly aims to gain valuable insights from existing data. One of the techniques used in data mining is classification. Classification divides data into predetermined categories or classes based on specific attributes (Haryati et al., 2015). The goal is to group data into the most appropriate class based on its characteristics.

One of the commonly used classification algorithms in data mining is K-Means. K-Means is a clustering algorithm used to group data into clusters based on the similarity of their characteristics. (Zuhal, 2022). This algorithm works by grouping data into k, called groups, where each data observation belongs to a group with the closest average. The stages of the K-means clustering algorithm, which will be discussed at a later stage in this study

RESULTS AND DISCUSSION

The data we use is data obtained from one of the companies that focuses on the textile category. For the dataset we use is as follows:

Table 1. Data Source

Customer's Name	Month	Motif Name	Qty (Mtr)	Sales		COGS		
				Price /M	Total	OGS/M	Total	
TIO	Rayon	Jan-22	APTKT-2 K	1,25	45.000	56.250	39.000	48.750
TIO	Cotton	Jan-22	Cotton W.	12,00	45.000	540.000	38.500	462.000
TIO	Rayon	Jan-22	RAYON V	32,00	45.000	#####	39.000	#####
TIO	Cotton	Jan-22	Cotton W.	106,00	45.000	#####	38.500	#####
SYA	Silk	Jan-22	CARTAGE	52,82	42.500	#####	41.000	#####
SYA	Rayon	Jan-22	Pueblos	160,00	42.500	#####	39.000	#####

The data will be evaluated in rows and columns, producing 2 axes, the x-axis, and y-axis, which are used to group the data into how many parts.

Clustering

At this stage, the clustering process data is carried out by selecting the x and y axes with QTY (Mtr) combined with the type of fabric. It will produce the total number of QTY consumers need, and then we will group it into 3 clusters. The data mining clustering formula that we use is the clustering formula or clustering calculating the Euclidean distance between two data points XX and YY with coordinates $X = (x_1, x_2, \dots, x_n)$ $X = (x_1, x_2, \dots, x_n)$ and $Y = (y_1, y_2, \dots, y_n)$. Euclidean distance $= \sum_{i=1}^n (x_i - y_i)^2$ Euclidean distance $= \sum_{i=1}^n (x_i - y_i)^2$. The column or column data obtained from the following data will become a table or data shown in Figure 2 data grouping.

Table 2. Data grouping

No	Customer Name	Sales				Meter			
		Qty (Mtr)	Rayon	Cotton	Linen	Viscose	Silk	Tencel	Rami
1	TIO	1,25	1,25	-	-	-	-	-	-
2	TIO	12,00	-	12,00	-	-	-	-	-
3	TIO	32,00	32,00	-	-	-	-	-	-
4	TIO	106,00	-	106,00	-	-	-	-	-
5	SYA	52,82	-	-	-	-	52,82	-	-

No	Customer Name	Sales Qty (Mtr)	Rayon	Cotton	Linen	Meter Viscose	Silk	Tencel	Rami
6	SYA	160,00	160,00	-	-	-	-	-	-
7	SYA	120,00	120,00	-	-	-	-	-	-
8	SYA	160,00	160,00	-	-	-	-	-	-
9	SYA	123,00	123,00	-	-	-	-	-	-
10	SYA	2,00	-	-	2,00	-	-	-	-
11	SYA	2,00	-	-	2,00	-	-	-	-
12	SYA	49,65	-	-	-	-	49,65	-	-

The previous process then divided the class into 3 groupings so that a group will be formed, which we call priority one, priority 2, and priority 3. It can be explained that group one is the highest priority group for companies to have priority 2 fabric stock, which is a priority after priority one, and the least priority is priority three, which is to stock goods. The data is obtained by referring to the calculation after grouping through the calculation of the nearest average; on this occasion, we were assisted with orange tools. After using the orange tools, we use or implement the k-means algorithm, as seen in Figure 1.

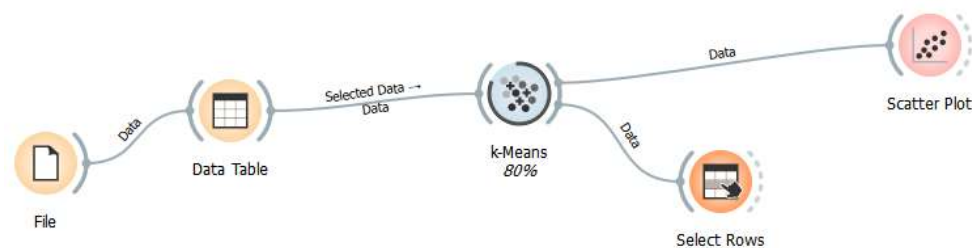


Figure 1. Implement The K-Means Algorithm

Our data focus on fabric type, where we calculate the average order of the fabric. There are seven types of fabrics that we process, as can be seen in Figure 2.

	Feature 1	Feature 2	Feature 3	Feature 4	Feature 5	Feature 6	Feature 7
1	1.25	0	0	0	0	0.00	0
2	0	12	0	0	0	0.00	0
3	32	0	0	0	0	0.00	0
4	0	106	0	0	0	0.00	0
5	0	0	0	0	52.52	0.00	0
6	160	0	0	0	0	0.00	0
7	120	0	0	0	0	0.00	0
8	160	0	0	0	0	0.00	0
9	123	0	0	0	0	0.00	0
10	0	0	2	0	0	0.00	0
11	0	0	2	0	0	0.00	0
12	0	0	0	0	49.65	0.00	0
13	0	0	0	0	49.86	0.00	0
14	0	0	0	0	49.57	0.00	0
15	0	0	0	0	50.19	0.00	0
16	0	0	0	0	40.65	0.00	0
17	0	0	0	0	20.3	0.00	0
18	0	3	0	0	0	0.00	0
19	0	3	0	0	0	0.00	0
20	150	0	0	0	0	0.00	0
21	150	0	0	0	0	0.00	0
22	150	0	0	0	0	0.00	0

Figure 2. Seven Types Of Fabrics

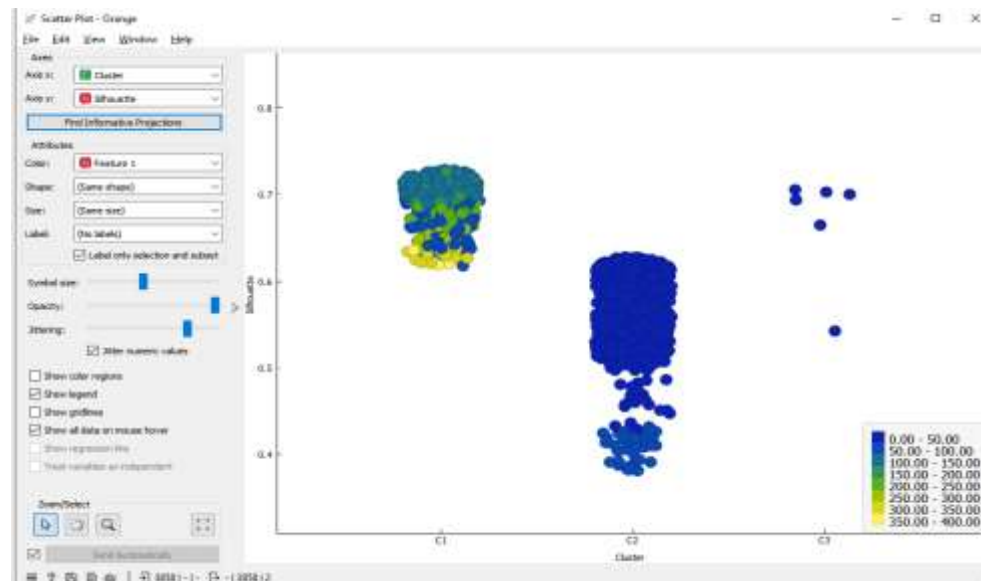


Figure 3. Cluster Results

CONCLUSION

Clustering analysis using k-means helps identify different customer segments based on fabric preferences, allowing companies to prioritize fabric inventory appropriately. This approach not only facilitates better inventory management and effectively meets customer demands, but also provides important implications for more targeted marketing strategies and product development tailored to diverse market needs.

REFERENCES

- Aprizal, S. (2018). *Orientasi Pasar dan Keunggulan Bersaing: Studi Kasus Penjualan Komputer* (Vol. 1). CELEBES MEDIA PERKASA.
- Darmawan, D. (2017). Hubungan Antara Inovasi Produk, Kepuasan dan Kepercayaan Pelanggan. *Jurnal Ekonomi Dan Bisnis*, 7(1), 36–42.
- Eunike, A., Setyanto, N. W., Yuniarti, R., Hamdala, I., Lukodono, R. P., & Fanani, A. A. (2021). *Perencanaan Produksi dan Pengendalian Persediaan: Edisi Revisi*. Universitas Brawijaya Press.
- Fatyandri, A. N., Fiona, F., Fernando, R., Wijaya, R. C., Alexandro, W., & Agustian, W. (2023). Meningkatkan Keunggulan Kompetitif Dengan Memilih Strategi Tingkat Bisnis Yang Tepat. *Jurnal Mirai Management*, 8(2), 230–236.
- Haryati, S., Sudarsono, A., & Suryana, E. (2015). implementasi data mining untuk memprediksi masa studi mahasiswa menggunakan algoritma c4. 5 (studi kasus: universitas dehasen bengkulu). *Jurnal Media Infotama*, 11(2).
- Juwita, I., & Ali, I. (2024). Penerapan Pola Penjualan Dengan Menggunakan Metode Algoritma Asosiasi Fp-Growth Bertujuan Untuk Meningkatkan Penjualan Kopi Di Point Coffee. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 8(2), 1600–1607.
- Mardi, Y. (2017). Data mining: Klasifikasi menggunakan algoritma c4. 5. *Jurnal Edik Informatika Penelitian Bidang Komputer Sains Dan Pendidikan Informatika*, 2(2), 213–219.
- Mujiati, N. W. (2023). Perkembangan, metode pendekatan dan tantangan manajemen sumber daya manusia. *Forum Manajemen*, 21(1), 68–86.
- Rofaida, R., Aryanti, A. N., & Perdana, Y. (2019). Strategi inovasi pada industri kreatif digital: Upaya memperoleh keunggulan bersaing pada era Revolusi Industri 4.0. *Jurnal Manajemen Dan Keuangan*, 8(3), 402–414.
- Sanawiri, B., & Iqbal, M. (2018). *Kewirausahaan*. Universitas Brawijaya Press.

- Sari, V. N., Yupianti, Y., & Maharani, D. (2018). Penerapan Metode K-Means Clustering Dalam Menentukan Predikat Kelulusan Mahasiswa Untuk Menganalisa Kualitas Lulusan. *JURTEKSI (Jurnal Teknologi Dan Sistem Informasi)*, 4(2), 133–140.
- Toriq, M., & Martoatmodjo, S. (2014). Pengaruh fasilitas dan kualitas layanan terhadap kepuasan pelanggan pada SPBU Pertamina 54.612. 64 di Sidoarjo. *Jurnal Ilmu & Riset Manajemen*, 3(8).
- Wijaya, D., Mandey, S., & Sumarauw, J. S. B. (2016). Analisis pengendalian persediaan bahan baku ikan pada PT. Celebes minapratama bitung. *Jurnal EMBA: Jurnal Riset Ekonomi, Manajemen, Bisnis Dan Akuntansi*, 4(2).
- Yunus, E. (2016). *Manajemen strategis*. Penerbit Andi.
- Zuhal, N. K. (2022). Study Comparison K-Means Clustering Dengan Algoritma Hierarchical Clustering: AHC, K-Means Clustering, Study Comparison. *Seminar Nasional Teknologi & Sains*, 1(1), 200–205.



© 2024 by the authors. This work has been submitted for possible open-access publication under the terms and conditions of the Creative Commons Attribution (CC BY SA) license (<https://creativecommons.org/licenses/by-sa/4.0/>).