JURTEKSI (Jurnal Teknologi dan Sistem Informasi) ISSN 2407-1811 (Print)
Vol. X No 2, Maret 2024, hlm. 395 - 402 ISSN 2550- 020 (Online)
DOI: http://dx.doi.org/10.33330/jurteksi.v10i2.3064
Available online at http://jurnal.stmikroyal.ac.id/index.php/jurteksi

K-MEANS ALGORITHM TO DETERMINE MARKETING STRATEGY AT CODEVERSE COMPUTER ACCESSORIES STORE

Muhamad Naufal Burhanuddin Balit [1*], Fandy Setyo Utomo [1]
[1]Informatics, Universitas Amikom Purwokerto
email: *burhannaufal808@gmail.com

Abstract: Artificial Intelligence (AI) is currently gaining popularity across various industries, including healthcare, finance, and others. In this study, AI technology is employed to devise an optimal marketing strategy for Code Verse Computer Accessories Store using the K-Means algorithm. As part of machine learning, the K-Means algorithm, categorized under unsupervised learning, is implemented to cluster sales data for computer accessory products over the last three months of 2023. The results of the K-Means analysis identify two main clusters. Cluster one (Cluster 1) comprises products such as Mouse, Keyboard, Monitor, Headset, and Speaker, indicating consistent purchasing patterns and high consumer interest. Recommendations are made to increase stock for Cluster 1. Meanwhile, Cluster two (Cluster 2) consists of Mic products with lower interest, and it is not advisable to increase stock. The implementation of K-Means provides insights into purchasing patterns, enabling Code Verse to develop more effective marketing and inventory management strategies.

Keywords: K-Means algorithm; artificial intelligence; clustering

Abstract: Kecerdasan Buatan (AI) kini meraih popularitas dalam berbagai industri, termasuk sektor kesehatan, keuangan, dan lainnya. Pada penelitian ini, teknologi AI digunakan untuk merancang strategi pemasaran optimal bagi Toko Aksesoris Komputer CodeVerse dengan menggunakan Algoritma K-Means. Sebagai bagian dari machine learning, Algoritma K-Means, yang termasuk dalam kategori unsupervised learning, diimplementasikan untuk mengelompokkan data penjualan produk selama tiga bulan terakhir tahun 2023. Hasil dari analisis K-Means mengidentifikasi dua cluster utama. Cluster pertama (Cluster 1) terdiri dari produk Mouse, Keyboard, Monitor, Headset, dan Speaker, menunjukkan pola pembelian yang konsisten dan tingginya minat konsumen. Rekomendasi untuk menambah stok diberikan. Sementara itu, Cluster kedua (Cluster 2) terdiri dari produk Mic dengan minat lebih rendah, dan tidak disarankan untuk menambah stok. Implementasi K-Means memberikan wawasan tentang pola pembelian, memungkinkan CodeVerse mengembangkan strategi pemasaran dan manajemen persediaan yang lebih efektif.

Keywords: Algoritma k-means; kecerdasan buatan; clustering

INTRODUCTION

Artificial Intelligence (AI) is currently a very popular technology, and it has been adopted by various industries, ranging from the healthcare sector to finance, and so on [1]. The use of AI in the healthcare industry encompasses a wide range of applications, such as diagnostic assistance through medical image analysis, prediction of potential health problems through patient data

JURTEKSI (Jurnal Teknologi dan Sistem Informasi) ISSN 2407-1811 (Print)
Vol. X No 2, Maret 2024, hlm. 395 - 402 ISSN 2550- 020 (Online)
DOI: http://dx.doi.org/10.33330/jurteksi.v10i2.3064
Available online at http://jurnal.stmikroyal.ac.id/index.php/jurteksi

analysis algorithms, accelerated drug discovery, personalized treatment through genetic data analysis, administrative support with AI systems, and AI-based virtual health assistant services to provide 24/7 access to patients [2]. All of these optimize patient care, improve operational efficiency, and provide innovative solutions in the healthcare industry [2].

The application of artificial intelligence in finance enhances operational efficiency and decision-making processes, including identifying customer investments, improving accuracy, and reducing associated risks. It also encompasses management functions, rapid data analysis, and increased productivity, leading to improved management and cost reduction in the financial sector [3]. The application of artificial intelligence in finance promises enhanced efficiency, decision-making, and risk management, with marketing strategies leveraging techniques like the K-Means algorithm experiencing a surge in popularity.[4].

The algorithm is able to perform this task without requiring category labels for the given data [5]. These algorithms not only provide deep insight into customer behavior but also help identify potential market segments [6]. Using the power of AI data analysis, the use of the K-Means algorithm can be optimized to identify effective marketing strategies in the CodeVerse computer accessories store [7].

The study "Using K-Means Algorithm in Clustering Cases of Infant Stunting in Tegal Perfume Village" aimed to identify growth patterns of infants in Tegal village through K-Means algorithm. Results revealed two distinct groups: Group 0 with 392 infants and Group 1 with 3 infants, categorized based on age, weight, and height, where 287 infants were normal and 108 experienced growth spurts. The Davies Bouldin Index (DBI) indicated effective cluster quality, suggesting validation through alternative methods like Silhouette and exploring differences between manual DBI assessments and tools like RapidMiner is recommended [8].

The research employs the K-Means algorithm to analyze sales data, predict product demand, and optimize inventory management for companies. By clustering products based on sales performance and using preprocessing techniques like normalization, it aids in understanding customer demand variations and adjusting inventory levels effectively. The segmentation approach derived from cluster analysis results helps companies anticipate demand fluctuations, avoid stockouts, and optimize production processes by emphasizing the importance of understanding sales patterns. [9].

A study conducted at Pulo Brayan Community Health Center utilized the K-Means Clustering Algorithm to analyze patient visit data from March 1st to May 31st, 2023, categorizing 949 patient records based on gender, age, and diagnosis. The algorithm identified three clusters, with Cluster 1, comprising mainly males aged 40 and above with a high prevalence of Acute Respiratory Tract Infections (ARTIs), being the largest with 524 patients. This research aids in targeted prevention, treatment, and education programs, as well as informing health policies and resource allocation for better healthcare management, thus enhancing overall public health.[10].

JURTEKSI (Jurnal Teknologi dan Sistem Informasi) ISSN 2407-1811 (Print)
Vol. X No 2, Maret 2024, hlm. 395 - 402 ISSN 2550- 020 (Online)
DOI: http://dx.doi.org/10.33330/jurteksi.v10i2.3064
Available online at http://jurnal.stmikroyal.ac.id/index.php/jurteksi

The Pulo Brayan Community Health Center conducted a comprehensive investigation analyzing patient visit data from March 1st to May 31st, 2023, using the K-Means Clustering Algorithm on 949 records. The analysis revealed three clusters: low, moderate, and high. Cluster 1, predominantly comprising males aged 40 and above, exhibited a significant incidence of Acute Respiratory Tract Infections (ARTIs), indicating potential health disparities in the community. [11].

The essence of this study is to use the K-Means clustering algorithm to analyze the sample of Riau University scholarship recipients in 2020, 2021 and 2022. The research results show that group 0 is the majority of students from the D3 Commercial Shipping Manageme nt research program at level 5. Currently, most Cluster 1 students from the seventh semester of the accounting and management studies program, with a GPA greater than or equal to 3.51, will graduate. In addition, group 2 is dominated by students of the nursing education program of the fifth week with the lowest GPA of 3.51 [12].

Implementing the K-Means algori thm strategically enhances CodeVerse software store's marketing efforts, consolidating its leading position as a one-stop destination for computer users seeking integrated solutions, spanning high-quality hardware sales to professio nal computer repair services.

METHOD

In this research method, resear chers use several stages. These stages will be described in Image 1.

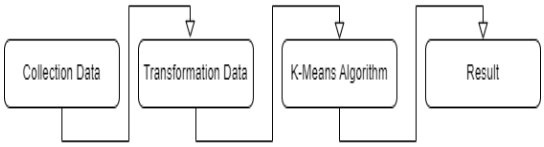

Image 1. Research Flow

Image 1 explains the flow of the research, starting from data collection, data transformation, the use of the K-Means algorithm, then the results of the research appear.

Data Collection

The dataset consists of 1000 entries of sales data from October to December 2023 at CodeVerse Computer.

$$d_{(x,y)} = \sqrt{\sum_i (xiyi)^2} \quad (1)$$

Then for the explanation of the above formula is as follows:

$d_{(x-y)}$: the distance between the data at points x and y

x: object data point

y: centroid data point

I: number of data attributes

The K-Means method, widely applicable in tasks like image and text analysis, exhibits drawbacks due to its initial random centroid distribution, often resulting in varied outcomes and non-uniform classification, leading to inconsistencies in results.

Data Mining

The basic principle of data mining is to discover hidden information in databases, which is an important part of Knowledge Discovery in Databases (KDD), which aims to find useful information and patterns in data [13].

JURTEKSI (Jurnal Teknologi dan Sistem Informasi) ISSN 2407-1811 (Print)
Vol. X No 2, Maret 2024, hlm. 395 - 402 ISSN 2550- 020 (Online)
DOI: http://dx.doi.org/10.33330/jurteksi.v10i2.3064
Available online at http://jurnal.stmikroyal.ac.id/index.php/jurteksi

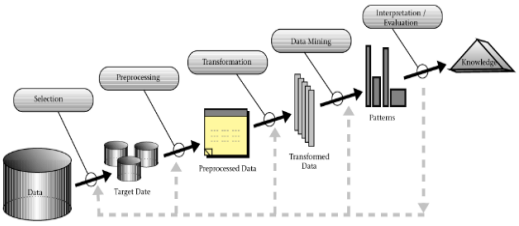

Image 2. Data mining working order

In Image 2, the data mining process begins with understanding the relevant domain or application, followed by careful selection, cleaning, and conversion of datasets. Subsequently, the data is reduced and subjected to prediction processes to extract pertinent components, which are then formulated into regression variables for specific analysis. [14].

K-Means Algorithm

K-Means is an unsupervised method that selects a subset of the population as the centers of the first group [15].

Clustering

Basically, clustering is an attempt to find and group data that has something in common [16]. Clustering technique is an approach in the field of data mining that groups a number of data into clusters based on similar characteristics or objects [17].

RESULT AND DISCUSSION

Cluster Number (K) Selection

At this point, the researchers identify the cluster centers, in this case represented by statements 1 and 5. Table 1. data from October to December. Based on table 1 data, the number of clusters was determined. The cluster data used can then be seen in table 2.

Centroid Initialization

Subsequently, centroid calcula tion will be performed using the formula, with Table 3 providing the initial centroids K1 and K2, where K1 denotes the selling products and K2 represents products that are not selling well.

Table 2. Sales For Data Cluster Materials

Product	Oct		Nov		Dec	
	Free	Sold	Free	Sold	Free	Sold
Mouse	100	50	100	79	100	89
Keyboard	100	70	100	32	100	55
Monitor	100	48	100	45	100	76
Headset	100	34	100	68	100	67
Speaker	100	79	100	54	100	98
Mic	100	67	100	88	100	35

JURTEKSI (Jurnal Teknologi dan Sistem Informasi) ISSN 2407-1811 (Print)
Vol. X No 2, Maret 2024, hlm. 395 - 402 ISSN 2550- 020 (Online)
DOI: http://dx.doi.org/10.33330/jurteksi.v10i2.3064
Available online at http://jurnal.stmikroyal.ac.id/index.php/jurteksi

Table 3. Determination of K1 and K2

Iteration 1	Cluster Center (4.5)			
Data 1	C1	79	54	98
Data 2	C2	67	88	35

After K1 and K2 are determined randomly, the next step is to carry out calculations to get the values of C1 and C2. In this calculation, the Euclidean distance formula is used. Further information can be found in Table 4.

Table 4. Cluster data calculations

Product	C1	C2
Mouse	39.3	57.3
Keyboards	49.1	59.5
Monitor	39.1	62.4
Headset	56.4	50.1
Speaker	0	72.6
Mic	72.6	0

Data Grouping

The results of data grouping can be seen in table 5.

Table 5. grouping data results

Product	C1	C2	Closest Distance	Data Groups
Mouse	39.3	57.3	39.3	Cluster 1
Keyboards	49.1	59.5	49.1	Cluster 1
Monitor	39.1	62.4	39.1	Cluster 1
Headset	56.4	50.1	56.4	Cluster 1
Speaker	0	72.6	0	Cluster 1
Mic	72.6	0	0	Cluster 2

From the data in table 5, it can be concluded that data 1, 2, 3, 4, and 5 are included in Cluster 1. Meanwhile, data number 6 is included in Cluster 2.

Centroid Update

In this research, centroids were recalculated based on existing data points in cluster table 5. Then, for conclusions the centroid update results can be seen in table 6.

Table 6. data cluster after update

Iteration 1	Cluster Center (4.5)			
Data 1	C1	79	54	98
Data 2	C2	67	88	35

To find the values in table 6, the calculations can be seen in manual calculations as follows.

$$C_{x1\,baru}=50+70+48+34+79=\frac{281}{5}=56,2 \quad (2)$$

$$C_{y1\,baru}=79+32+45+68+54=\frac{278}{5}=55,6 \quad (3)$$

$$C_{z1\,baru}=89+55+76+67+98=\frac{385}{5}=77 \quad (4)$$

$$C_{x2\,baru}=\frac{67}{1}=67 \quad (5)$$

$$C_{y2\,baru}=\frac{88}{1}=88 \quad (6)$$

$$C_{z2\,baru}=\frac{35}{1}=35 \quad (7)$$

Reflection is carried out until each cluster is the same as the previous iteration, in this case, those in Iteration 1. For the new data grouping, it can be seen in table 7.

Table 7. Grouping Data Results Literation 2

Product	C1	C2	Closest Distance	Data Groups
Mouse	27.0	57.3	27.0	Cluster 1
Keyboards	35.1	59.5	35.1	Cluster 1
Monitor	13.4	62.4	13.4	Cluster 1
Headset	27.3	50.1	27.3	Cluster 1
Speaker	31.0	72.6	31.0	Cluster 1
Mic	54.1	0	0	Cluster 2

JURTEKSI (Jurnal Teknologi dan Sistem Informasi) ISSN 2407-1811 (Print)
Vol. X No 2, Maret 2024, hlm. 395 - 402 ISSN 2550- 020 (Online)
DOI: http://dx.doi.org/10.33330/jurteksi.v10i2.3064
Available online at http://jurnal.stmikroyal.ac.id/index.php/jurteksi

Table 8. Calculate The Result From K-Means

Product name	Iteration 1				Iteration 2			
	C1	C2	Closest Distance	Data Groups	C1	C2	Closest Distance	Data Groups
Mouse	39.3	57.3	39.3	Cluster 1	27.0	57.3	27.0	Cluster 1
Keyboards	49.1	59.5	49.1	Cluster 1	35.1	59.5	35.1	Cluster 1
Monitors	39.1	62.4	39.1	Cluster 1	13.4	62.4	13.4	Cluster 1
Headsets	56.4	50.1	56.4	Cluster 1	27.3	50.1	27.3	Cluster 1
Speaker	0	72.6	0	Cluster 1	31.0	72.6	31.0	Cluster 1
Mic	72.6	0	0	Cluster 2	54.1	0	0	Cluster 2

Iteration

Researchers iteratively refined data grouping until they consistently identified stable clusters across iterations, noting that in Iterations 1 and 2, Cluster 1 encompassed data 1 through 5, while Cluster 2 consisted of data 6, thus concluding the process.

CONCLUSION

The K-Means algorithm analysis of body oils sales data consistently identified two clusters across Iterations 1 and 2, with stable average values. Cluster 1 signifies high consumer demand, while Cluster 2 remains the most popular. In contrast, there has been a notable decline in sales of third-party premium soap products, prompting the decision to cease additional supplies in the coming months, as summarized in Table 8.

BIBLIOGRAPHY

[1] H. Gunawan and V. Purwayoga, "Data Mining Menggunakan Algoritma K-Means Clustering Untuk Mengetahui Potensi Penyebaran Virus Corona Di Kota Cirebon," *J. Sisfokom (Sistem Inf. dan Komputer)*, vol. 11, no. 1, pp. 1–8, 2022, doi: 10.32736/sisfokom.v11i1.1316.

[2] E. M. Fitri, R. R. Suryono, and A. Wantoro, "Klasterisasi Data Penjualan Berdasarkan Wilayah Menggunakan Metode K-Means Pada Pt Xyz," *J. Komputasi*, vol. 11, no. 2, pp. 157–168, 2023, doi: 10.23960/komputasi.v11i2.12582.

[3] U. A. Nasron and M. Habibi, "Analysis of Marketplace Conversation Trends on Twitter Platform Using K-Means," *Compiler*, vol. 9, no. 1, pp. 51–62, 2020, doi: 10.28989/compiler.v9i1.579.

[4] T. Hardiani, "Analisis Clustering Kasus Covid 19 di Indonesia Menggunakan Algoritma K-Means," *J. Nas. Pendidik. Tek. Inform.*, vol. 11, no. 2, pp. 156–165, 2022, doi: 10.23887/janapati.v11i2.45376.

[5] M. Hermansyah, R. A. Hamdan, F. Sidik, and A. Wibowo, "Klasterisasi Data Travel Umroh di Marketplace Umroh.com Menggunakan Metode K-Means," *J. Ilmu Komput.*, vol. 13, no. 2, p. 8, 2020, doi: 10.24843/jik.2020.v13.i02.p06.

[6] R. Nainggolan and E. Purba, "Perbaikan Performa Cluster K-Means Menggunakan Sum

Squared Error (Sse) Pada Analisis Online Customer Review Terhadap Produk Toko Online," *J. TIMES*, vol. VIII, no. 2, pp. 1–8, 2019.

[7] A. Aditya, I. Jovian, and B. N. Sari, "Implementasi K-Means Clustering Ujian Nasional Sekolah Menengah Pertama di Indonesia Tahun 2018/2019," *J. Media Inform. Budidarma*, vol. 4, no. 1, p. 51, 2020, doi: 10.30865/mib.v4i1.1784.

[8] P. Apriyani, A. R. Dikananda, and I. Ali, "Penerapan Algoritma K-Means dalam Klasterisasi Kasus Stunting Balita Desa Tegalwangi," *Hello World J. Ilmu Komput.*, vol. 2, no. 1, pp. 20–33, 2023, doi: 10.56211/helloworld.v2i1.230.

[9] D. Anggarwati, O. Nurdiawan, I. Ali, and D. A. Kurnia, "Penerapan Algoritma K-Means Dalam Prediksi Penjualan Karoseri," *J. Data Sci. Inform.*, vol. 1, no. 2, pp. 58–62, 2021.

[10] Okta Jaya Harmaja, Hadirat Halawa, Wicarda Sandi Hulu, and Seriani Loi, "Implementasi Algoritma K-Means Clustering Untuk Pengelompokkan Penyakit Pasien Pada Puskesmas Pulo Brayan," *J. Sains dan Teknol.*, vol. Volume 5 N, no. 1, pp. 150–157, 2023.

[11] W. Sirait, S. Defit, and G. W. Nurcahyo, "Algoritma K-Means Untuk Klasterisasi Tugas Akhir Mahasiswa Berdasarkan Keahlian," *J. Sistim Inf. dan Teknol.*, vol. 1, no. 3, pp. 25–30, 2019, doi: 10.35134/jsisfotek.v1i3.6.

[12] Q. A'yuni, A. Nazir, L. Handayani, and I. Afrianty, "Penerapan Algoritma K-Means Clustering untuk Mengetahui Pola Penerima Beasiswa Bank Indonesia (BI)," *J. Comput. Syst. Informatics*, vol. 4, no. 3, pp. 530–539, 2023, doi: 10.47065/josyc.v4i3.3343.

[13] S. Wulandari, "Clustering Kecamatan Di Kota Bandung Berdasarkan Indikator Jumlah Penduduk Dengan Menggunakan Algoritma K-Means," *Semnas Ristek (Seminar Nas. Ris. …*, pp. 128–132, 2020, [Online]. Available: https://www.proceeding.unindra.ac.id/index.php/semnasristek/article/view/1688%0Ahttps://www.proceeding.unindra.ac.id/index.php/semnasristek/article/download/1688/235

[14] M. Luqmanul Hakim, Junervin, S. Rambe, S. Azmi, and A. Kurniasi Alu, "ANALISIS DAN DESAIN PEMETAAN UMKM AGROINDUSTRI JAWA TENGAH DENGAN METODE K-MEANS CLUSTER UNTUK PENINGKATAN EKSPOR," *JATI (Jurnal Agroindustri Terap. Indones.*, vol. 1, no. 02, 2024, doi: 10.17605/OSF.IO/U4CJF.

[15] I. N. M. Adiputra, "Clustering Penyakit Dbd Pada Rumah Sakit Dharma Kerti Menggunakan Algoritma K-Means," *Inser. Inf. Syst. Emerg. Technol. J.*, vol. 2, no. 2, p. 99, 2022, doi: 10.23887/insert.v2i2.41673.

[16] L. Suriani, "Pengelompokan Data Kriminal Pada Poldasu Menentukan Pola Daerah Rawan Tindak Kriminal Menggunakan Data Mining Algoritma K-Means Clustering," *J. Sist. Komput. dan*

Inform., vol. 1, no. 2, p. 151, 2020, doi: 10.30865/json.v1i2.1955.

[17] A. Rohmah, F. Sembiring, and A. Erfina, "Implementasi Algoritma K-Means Clustering Analysis untuk Menentukan Hambatan Pembelajaran Daring (Studi Kasus: SMK Yaspim Gegerbitung)," *Semin. Nas. Sist. Inf. dan Manaj. Inform.*, pp. 290–298, 2021, [Online]. Available: https://sismatik.nusaputra.ac.id/index.php/sismatik/article/view/32