

Penerapan Algoritma K-Means Dalam Aplikasi Business Intelligence Untuk Segmentasi Pelanggan

Christopher Michael Lauw¹, Husain^{2**}, Muhamad Wisnu Alfiansyah³, Felita Nathania⁴,
Anyas Asmina Nurul Adha⁵, Ayu Nursila⁶

^{1,2,3,4,5,6} Fakultas Teknik, Universitas Bumigora, Mataram, Indonesia
¹2001020015@universitasbumigora.ac.id ²husain@universitasbumigora.ac.id
³wisnu@universitasbumigora.ac.id ⁴2301020067@universitasbumigora.ac.id
⁵2201020041@universitasbumigora.ac.id ⁶2301020003@universitasbumigora.ac.id

ABSTRAK

Business Intelligence (BI) adalah suatu konsep dan teknologi yang digunakan untuk mengumpulkan, menganalisis, mengelola, dan menyajikan informasi bisnis yang dapat digunakan oleh organisasi untuk mendukung pengambilan keputusan yang lebih baik. Dalam mengimplementasikan Business Intelligence juga dapat dikomparasikan dengan algoritma data mining seperti pengelompokan dengan menggunakan K-Means. Dalam penelitian ini K-Means digunakan untuk melakukan segmentasi pelanggan sehingga perusahaan dapat mengetahui apa yang dibutuhkan oleh pelanggan mereka. Penelitian ini menggambarkan penerapan algoritma K-Means dalam konteks aplikasi Business Intelligence untuk tujuan segmentasi pelanggan dan pembuatan dashboard. Pembuatan dashboard interaktif sering terlepas dari algoritma yang digunakan sehingga dashboard tersebut menjadi pasif atau hanya dapat menampilkan visualisasi data saja. Tahapan penelitian yang digunakan yaitu pengumpulan data yang dilakukan melalui internet, kemudian dilakukan data preprocessing dengan beberapa tahapan didalamnya yaitu extract, transform dan load, serta data cleaning untuk membersihkan data setelah dilakukan data preprocessing dilakukan clustering dengan menggunakan algoritma K-Means, setelah dilakukan clustering, yaitu pembuatan dashboard. Hasil dari penelitian ini berupa dashboard dengan menggunakan power BI.

Kata kunci: Business Intelligence, k-means, data mining, customer segmentation

ABSTRACT

Business Intelligence (BI) is a concept and technology used to collect, analyze, manage and present business information that can be used by organizations to support better decision making. In implementing Business Intelligence, it can also be compared with data mining algorithms such as grouping using K-Means. In this research, K-Means is used to segment customers so that companies can find out what their customers need. This research describes the application of the K-Means algorithm in the context of Business Intelligence applications for the purpose of customer segmentation and dashboard creation. Creating interactive dashboards is often independent of the algorithm used so that the dashboard becomes passive or can only display data visualizations. The research stages used were data collection carried out via the internet, then data preprocessing was carried out with several stages in it, namely extract, transform and load, as well as data cleaning to clean the data after the data preprocessing was carried out, clustering was carried out using the K-Means algorithm, after clustering was carried out, namely creating a dashboard. The results of this research are in the form of a dashboard using Power BI.

Keywords: Business Intelligence, k-means, data mining, customer segmentation

Info Artikel :

Disubmit: 09 September

Direview: 11 November

Diterima : 20 Desember

Copyright © 2023 - Journal UPU. All rights reserved

I. PENDAHULUAN

Di Era industry 4.0 kebutuhan perusahaan akan teknologi semakin besar, salah satunya seperti Business Intelligence[1]. *Business Intelligence (BI)* adalah suatu konsep dan teknologi yang digunakan dalam dunia bisnis untuk mengumpulkan, menganalisis, mengelola, dan menyajikan informasi bisnis yang dapat digunakan oleh organisasi untuk mendukung pengambilan keputusan yang lebih baik[2]. Tujuan utama dari BI adalah menyediakan wawasan yang relevan dan akurat kepada pemangku kepentingan bisnis sehingga mereka dapat mengambil keputusan yang lebih baik dan strategis[3]. *Business Intelligence* juga dapat dikomparasikan dengan berbagai algoritma data mining seperti K-Means yang digunakan untuk pengelompokan atau *clustering*, dalam penelitian ini algoritma K-Means digunakan untuk segmentasi pelanggan[4].

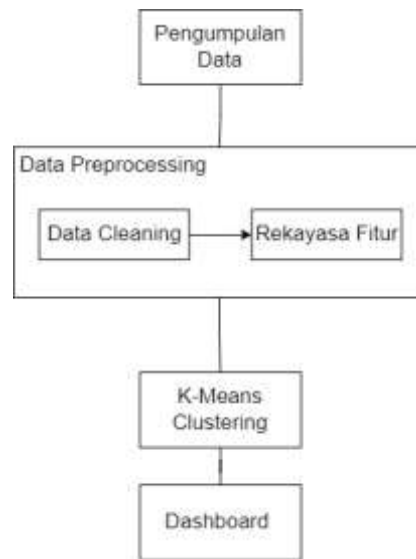
Segmentasi pelanggan merupakan proses membagi pasar atau basis pelanggan suatu bisnis atau organisasi[5] menjadi kelompok-kelompok yang lebih kecil atau segmen-segmen berdasarkan karakteristik atau perilaku yang serupa. Tujuan dari segmentasi pelanggan adalah untuk memahami kebutuhan[6], preferensi, dan perilaku pelanggan dengan lebih baik sehingga bisnis dapat mengarahkan upaya pemasaran, penjualan, dan pelayanan kepada setiap segmen dengan lebih efektif[7]. Dalam melakukan segmentasi pelanggan dibutuhkan suatu algoritma, dalam penelitian ini menggunakan algoritma K-Means yang merupakan algoritma data mining untuk *clustering* atau pengelompokan[8].

Algoritma K-Means merupakan salah satu algoritma clustering yang digunakan dalam analisis data dan pembelajaran mesin[9]. Tujuannya adalah untuk mengelompokkan data ke dalam beberapa kelompok atau cluster berdasarkan kemiripan karakteristik antara data[10]. Algoritma K-Means bekerja dengan cara mengidentifikasi pusat cluster (centroid) dan mengelompokkan setiap data ke dalam cluster yang memiliki centroid terdekat[11]. Salah satu penerapan algoritma K-Means yaitu segmentasi pelanggan dalam penerapan business intelligence. Dalam melakukan segmentasi terdapat empat kategori yaitu unggul[12], elite menengah dan rendah. *Elite customer* Yang mempunyai pendapatan tertinggi dan pengeluaran tertinggi pula[13]. *Unggul customer* Yang mempunyai pendapatan tinggi dan pengeluaran tinggi tetapi tidak setinggi *Elite*. *Menengah customer* Yang mempunyai gaji rendah dan pengeluaran lebih sedikit[14]. *Rendah customer* Yang mempunyai gaji paling rendah dan pengeluaran paling sedikit[15].

Adapun permasalahan yang terjadi yaitu suatu perusahaan sangat susah untuk melakukan pengelompokan sistem BI sebelumnya menggunakan grafik yang ditampilkan dalam dashboard tanpa algoritma analisa apapun[16]. Sehingga mempersulit *stackholder* dalam melihat ciri-ciri *customer* melalui data historical yang dimiliki dan telah disajikan dalam *dashboard*[17]. Oleh karena itu solusi yang ditawarkan dalam permasalahan tersebut yaitu dengan menggunakan algoritma *K-Means Clustering*[18].

Terdapat beberapa penelitian terdahulu yang membahas terkait algoritma K-Means dalam Business Intelligence[19]. Yang pertama penelitian yang dilakukan oleh [20] yang membahas terkait prediksi pergantian pelanggan menggunakan analisis teks untuk Business Intelligence dengan algoritma evolusioner, pada penelitian tersebut belum terdapat *dashboard* yang digunakan untuk menyajikan data[21]. Penelitian berfokus pada pengimplementasian algoritma K-Means pada analisa dashboard *Business Intelligence*[22]. Penelitian ini akan menghasilkan sebuah visualisasi data dalam bentuk penyajian dalam *dashboar* guna untuk menyajikan data agar lebih mudah dianalisa[23].

II. METODE PENELITIAN.



Gambar 1 Tahapan Penelitian

Adapun tahapan penelitian yang ditunjukkan pada gambar 1 diatas yang pertama yaitu dilakukan pengumpulan data[24]. Pengumpulan data dilakukan melalui situs kaggle, dimana data tersebut terdiri dari 2205 record data historikal pelanggan. Kemudian data yang telah dikumpulkan dilakukan preprocessing, terdiri dari pembersihan data atau *data cleaning* kemudian dilakukan rekayasa fitur, setelah data telah di *preprocessing* akan dilakukan pengelompokan atau *clustering* dengan menggunakan algoritma K-Means, tahap terakhir yaitu berbentuk dashboard analisis[25].

III. HASIL DAN PEMBAHASAN

A. Pengumpulan Data

Data diambil dan dikumpulkan melalui situs kaggle. Data yang dikumpulkan berbentuk format *xlsx* dengan jumlah sebanyak 2205 baris data, dan terdiri dari 33 atribut[26], tetapi dalam penelitian ini atribut yang digunakan hanya 10 atribut antara lain yaitu Income, AgeGroup, NumCatalogPurchases, NumCatalogPurchases, TotalSpending, Education, Complain, cluster, MntSweetProducts[27]. Kidhome, TeenHome Berikut merupakan data yang digunakan terdapat pada gambar 2 dan 3 dibawah ini.

	ID	Year_Birth	Education	Marital_Status	Income	Kidhome	Teenhome	Dt_Customer	Recency	MntWines
0	5524	1957	Graduation	Single	58138.0	0	0	04-09-2012	58	635
1	2174	1954	Graduation	Single	46344.0	1	1	08-03-2014	38	11
2	4141	1965	Graduation	Partner	71613.0	0	0	21-08-2013	26	426
3	6182	1984	Graduation	Partner	26646.0	1	0	10-02-2014	26	11
4	5324	1981	PhD	Partner	58293.0	1	0	19-01-2014	94	173

Gambar 2 Hasil Pengumpulan Data

AcceptedCmp2	Complain	Z_CostContact	Z_Revenue	Response	Age	AgeGroup	anak	TotalSpending	clusters
0	0	3	11	1	58	Dewasa Usia Pertengahan	0	1617	Elite
0	0	3	11	0	61	Tua	2	27	Rendah
0	0	3	11	0	50	Dewasa Usia Pertengahan	0	776	Unggul
0	0	3	11	0	31	Dewasa	1	53	Menengah
0	0	3	11	0	34	Dewasa	1	422	Elite

Gambar 3 Hasil Pengumpulan Data

B. Data Preprocessing.

Data preprocessing merupakan langkah atau teknik yang digunakan pada data mentah sebelum diproses [28] atau digunakan untuk melakukan segmentasi dengan menggunakan algoritma K-Means dalam konteks analisis data [29]. Pada penelitian ini pemrosesan data terdiri dari dua langkah yaitu pembersihan data (*data cleaning*) dan rekayasa fitur [30].

- Data Cleaning.

Melakukan data cleaning sehingga data yang telah dikumpulkan tidak terdapat data yang kosong atau *missing value* sehingga data tersebut informatif untuk digunakan sebagai bahan pembuatan dashboard [31].

```
[5]: missing_value = pelanggan.isnull().sum().to_dict()
[6]: pelanggan = pelanggan.dropna()
    pelanggan.isnull().sum()
```

Gambar 4 Data Cleaning

- Rekayasa Fitur.

Melakukan rekayasa fitur pada atribut usia, dimana akan dikelompokkan berdasarkan jarak usia dan akan diubah menjadi kategori [32]. Dimana pelanggan yang berusia lebih dari atau samadengan 13 tahun akan dikategorikan "Remaja", pelanggan yang berusia lebih dari 20 tahun dan lebih dari atau sama dengan 39 tahun, akan dilategorikan Dewasa [33]. Kemudian pelanggan yang berusia lebih dari atau samadengan 40 tahun atau sama dengan lebih dari 59 tahun maka akan dikategorikan sebagai "Dewasa Usia Pertengahan", pelanggan yang memiliki usia lebih dari 60 tahun akan dikategorikan sebagai "Tua" [34].

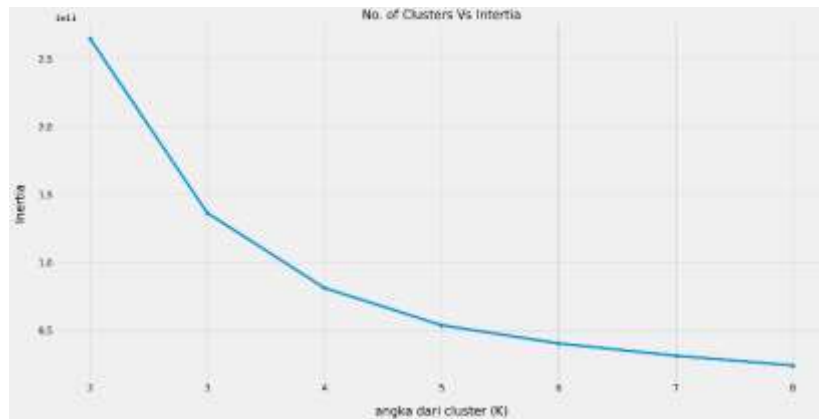
```
pelanggan.loc[(pelanggan['Age'] >= 13) & (pelanggan['Age'] <= 19), 'AgeGroup'] = 'Remaja'
pelanggan.loc[(pelanggan['Age'] >= 20) & (pelanggan['Age'] <= 39), 'AgeGroup'] = 'Dewasa'
pelanggan.loc[(pelanggan['Age'] >= 40) & (pelanggan['Age'] <= 59), 'AgeGroup'] = 'Dewasa Usia Pertengahan'
pelanggan.loc[(pelanggan['Age'] >= 60), 'AgeGroup'] = 'Tua'

C:\Users\MICHAEL\AppData\Local\Temp\ipykernel_20156124\79595107.py:1: FutureWarning: Setting an item of incompatible dtype is deprecated and will raise in a future error of pandas. Value 'Remaja' has dtype incompatible with float64, please explicitly cast to a compatible dtype first.
  pelanggan.loc[(pelanggan['Age'] >= 13) & (pelanggan['Age'] <= 19), 'AgeGroup'] = 'Remaja'
```

Gambar 5 Rekayasa Fitur

C. K-Means Clustering.

Pertama harus terlebih dahulu menganalisa jumlah cluster yang akan digunakan agar lebih optimal. Seperti yang terlihat pada gambar dibawah ini. Gambar dibawah menjelaskan bahwa pengelompokan atau segmen pelanggan menjadi 4 *cluster* dikarenakan nilai inersia pada angka lebih dari empat sangat rendah. Dimana inersia merupakan suatu metrik yang mengukur sejauh apa titik-titik data dalam suatu pengelompokan dari pusat (*centroid*)



Gambar 6 Grafik Penentu Jumlah Cluster

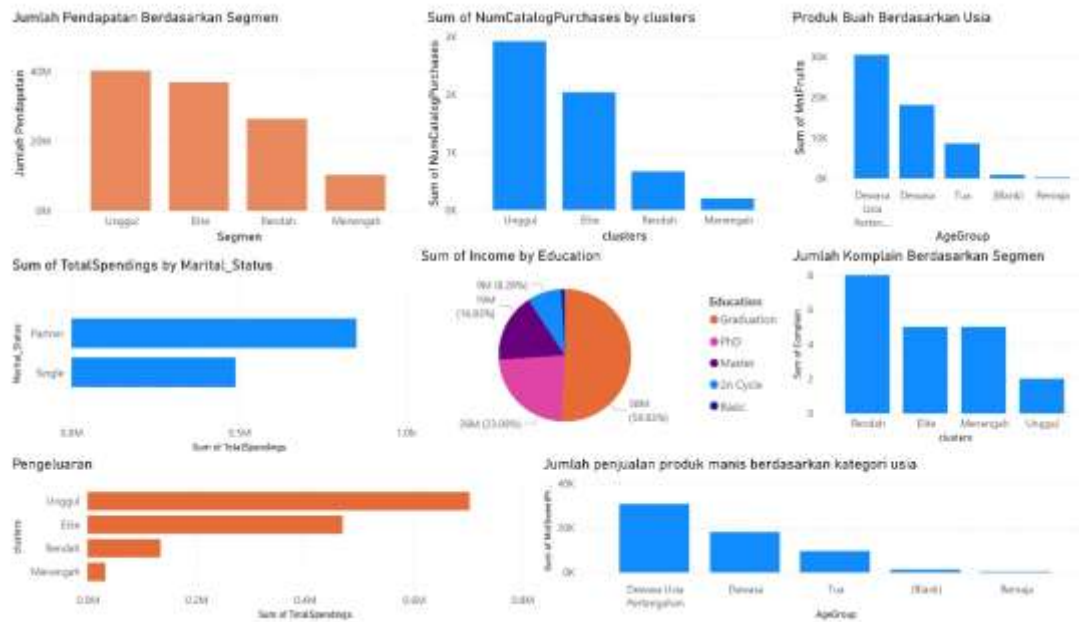
Kemudian dilakukan clustering dengan algoritma *K-Means* dimana jumlah *cluster* yang digunakan yaitu 4 *cluster*. Hasil proses clustering dapat terlihat pada gambar dibawah ini. Atribut yang digunakan dalam *clustering* yaitu 5 atribut sebagai dasar untuk melakukan *clustering*, Adapun atribut tersebut yaitu *Income*, *Age*, *anak*, *TotalSpending*, *clusters*. Dimana *Income* yaitu pendapatan dari pelanggan atau gaji, *Age* yaitu usia dari pelanggan, *totalSpending*, yaitu total pengeluaran saat pelanggan tersebut berbelanja. Dan setelah keempat atribut digunakan maka akan dilakukan pengelompokan berdasarkan keempat atribut tersebut pada atribut cluster, terdapat rentang yaitu 0-4 dimana 0 merupakan kategori Elite yaitu pelanggan Yang mempunyai pendapatan tertinggi dan pengeluaran tertinggi pula[13]. 3 merupakan kategori Unggul yaitu pelanggan Yang mempunyai pendapatan tinggi dan pengeluaran tinggi tetapi tidak setinggi Elite. 2 merupakan kategori Menengah yaitu pelanggan Yang mempunyai gaji rendah dan pengeluaran lebih sedikit[14]. Dan 0 yaitu kategori Rendah pelanggan Yang mempunyai gaji paling rendah dan pengeluaran paling sedikit[15].

	Income	Age	anak	TotalSpending	clusters
0	58138.0	58	0	1617	2
1	46344.0	61	2	27	1
2	71613.0	50	0	776	0
3	26646.0	31	1	53	3
4	58293.0	34	1	422	2

Gambar 7 Hasil K-Means Clustering

D. Dashboard Analisis.

Data yang telah dilakukan pengelompokan dengan algoritma *K-Means Clustering* akan disajikan dalam bentuk dashboard analisis dengan menggunakan *software microsoft PowerBI* yang merupakan suatu program atau aplikasi yang biasa digunakan untuk memvisualisasikan data.



Gambar 8 Hasil Output Dashboard Analisis

Data dalam dashboard tersebut berisi visualisasi data yang menjelaskan jumlah pendapatan berdasarkan segmen dimana terdapat empat segmen yaitu unggul, elite, rendah, dan menengah. Dari grafik jumlah pendapatan berdasarkan segmen yaitu, kategori pelanggan atau *customer* yang terdapat pada grafik diatas memberi wawasan terkait pendapatan perbulan yang diperoleh dari masing-masing kategori. Pada grafik "sum of NumCatalogPurchases by cluster" yaitu jumlah pembelian dari katalog berdasarkan pengelompokan (*cluster*) dalam hal ini perusahaan dapat memahami katetgori produk yang paling unggul dan dapat meningkatkan kualitas produksi dari produk tersebut. Pada grafik produk buah berdasarkan usia, yaitu grafik yang mempresentasikan jumlah peminat buah dari usia dewasa pertengahan, dewasa, tua dan remaja sehingga dapat diketahui bahwa peminat produk buah yaitu tertinggi pada usia dewasa pertengahan. Pada grafik "sum of total spendings by martial status" yaitu total pengeluaran berdasarkan status pernikahan, dimana perusahaan dapat memahami mana lebih banyak pengeluaran pelanggan yang telah menikah atau yang masih belum menikah, sehingga perusahaan dapat memutuskan untuk meningkatkan penjualan produk rumah tangga. Pada grafik "sum of income by education" grafik yang mempresentasikan data pendapatan berdasarkan tingkat Pendidikan. Pada grafik jumlah komplain berdasarkan segmen, dimana grafik tersebut mempresentasikan pelanggan yang melakukan komplain berdasarkan segmen pelanggan, sehingga perusahaan dapat mengetahui jumlah komplain tertinggi berdasarkan segmen pelanggan. Pada grafik pengeluaran mempresentasikan pengeluaran berdasarkan segmen pelanggan, sehingga perusahaan dapat mengambil keputusan untuk memberikan diskon atau bonus ke pelanggan yang kategori unggul, dikarenakan pelanggan yang kategori unggul paling tinggi pengeluarannya ketika berbelanja.

IV SIMPULAN DAN SARAN

Penerapan algoritma *K-Means Clustering* dalam *business intelligent* memiliki dampak yang sangat baik sehingga membantu suatu organisasi atau perusahaan dalam menyajikan data secara visual yang menarik dan informatif. Peran algoritma K-Means yaitu melakukan pengelompokan sehingga pelanggan dibagi menjadi beberapa kelompok dan dapat dianalisa sesuai dengan kelompok atau clusternya. Dengan adanya algoritma pengelompokan pemanfaatan *Dashboard* yang Lebih Efisien Dashboard BI dapat disusun dengan lebih baik untuk menyajikan informasi segmentasi pelanggan secara visual, mempermudah stakeholder dalam memahami

karakteristik pelanggan melalui data historis. Saran pada penelitian selanjutnya yaitu, menggunakan model pengelompokan lainnya dan membandingkan hasil dari model tersebut kepada *K-Means Clustering*. Selain itu dibutuhkan pengembangan sistem sehingga *dashboard* dapat terhubung langsung ke sumber data atau *database*.

DAFTAR PUSTAKA

- [1] Berlin and Y. C. Giap, "Penerapan Business Intelligence Pada Cv. Tanggamas Chemichal Dengan Metode Olap," *Algor*, vol. 2, no. 1, pp. 57–65, 2020.
- [2] J. O. Ong and M. Mahazan, "Strategi pengelolaan sdm dalam peningkatan kinerja perusahaan berkelanjutan di era industri 4.0," *Bus. Econ. Commun. Soc. Sci. J.*, vol. 2, no. 1, pp. 159–168, 2020.
- [3] N. Renaldo, D. Jollyta, S. Suhardjo, L. Fransisca, and M. Rosyadi, "Pengaruh Fungsi Sistem Intelijen Bisnis terhadap Manfaat Sistem Pendukung Keputusan dan Organisasi," in *SEMINAR NASIONAL INFORMATIKA (SENATIKA)*, 2022, pp. 61–78.
- [4] A. Supriyadi, A. Triayudi, and I. D. Sholihati, "Perbandingan algoritma k-means dengan k-medoids pada pengelompokan armada kendaraan truk berdasarkan produktivitas," *JIPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform.)*, vol. 6, no. 2, pp. 229–240, 2021.
- [5] H. Mawarni, G. Testiana, and M. L. Dalafranka, "IMPLEMENTASI ALGORITMA K-MEANS UNTUK SEGMENTASI PELANGGAN PADA PT. BINTANG MULTI SARANA CABANG TUGUMULYO," *J-Icon J. Komput. dan Inform.*, vol. 11, no. 2, pp. 227–236, 2023.
- [6] A. B. Setiawan, D. R. Putera, and T. Wirawan, "Representasi Makna Budaya Pada Iklan Redbus" Pendekar Tatang Tapasti", in *Prosiding Seminar Nasional Desain dan Media*, 2023, pp. 1014–1023.
- [7] S. Adha, "Faktor revolusi perilaku konsumen era digital: sebuah tinjauan literatur," *JIPIS*, vol. 31, no. 2, pp. 134–148, 2022.
- [8] A. P. Pramudiansyah, "Segmentasi Pelanggan Menggunakan Algoritma K-Means Berdasarkan Model Recency Frequency Monetary," *J. Ilm. Ilmu Komput. Fak. Ilmu Komput. Univ. Al Asyariah Mandar*, vol. 7, no. 2, pp. 6–19, 2021.
- [9] D. F. Pasiribu, I. S. Damanik, E. Irawan, and H. S. Tambunan, "Memanfaatkan Algoritma K-Means Dalam Memetakan Potensi Hasil Produksi Kelapa Sawit PTPN IV Marihat," *BIOS J. Teknol. Inf. Dan Rekayasa Komput.*, vol. 2, no. 1, pp. 11–20, 2021.
- [10] T. Hartati, O. Nurdian, and E. Wiyandi, "Analisis Dan Penerapan Algoritma K-Means Dalam Strategi Promosi Kampus Akademik Maritim Suaka Bahari," *J. Sains Teknol. Transp. Marit.*, vol. 3, no. 1, pp. 1–7, 2021.
- [11] Z. Nabila, A. R. Isnain, P. Permata, and Z. Abidin, "Analisis data mining untuk clustering kasus covid-19 di Provinsi Lampung dengan algoritma k-means," *J. Teknol. Dan Sist. Inf.*, vol. 2, no. 2, pp. 100–108, 2021.
- [12] A. B. H. Kiat, Y. Azhar, and V. Rahmayanti, "Penerapan Metode K-Means Dengan Metode Elbow Untuk Segmentasi Pelanggan Menggunakan Model RFM (Recency, Frequency, & Monetary)," *J. Repos*, vol. 2, no. 7, p. 945, 2020.
- [13] M. Z. Piliang, "PENGLOLAAN EKONOMI INDONESIA: PENGALAMAN ORDE BARU," *At-Tanmiyah J. Ekon. dan Bisnis Islam*, vol. 2, no. 1, pp. 207–238, 2023.
- [14] N. K. S. Julyantari, I. K. Budiarta, and N. M. D. K. Putri, "Implementasi K-Means Untuk Pengelompokan Status Gizi Balita (Studi Kasus Banjar Titih)," *J. Janitra Inform. dan Sist. Inf.*, vol. 1, no. 2, pp. 92–101, 2021.
- [15] C. J. Silalahi, A. Situmorang, and J. F. Naibaho, "Implementasi Metode K-Means Clustering Untuk Memetakan Daerah Potensial Penghasil Padi di Provinsi Sumatera Utara," *J. Ilm. Tek. Inform. METHOTIKA*, vol. 2, no. 2, pp. 49–57, 2022.
- [16] D. Ananda, S. Rohimah, B. Susilo, D. Wulan, and A. Mustofa, "Implementasi K-Means Dalam Pengelompokan Data Akta Kelahiran di Indonesia: Implementation of K-Means in Grouping Birth Certificate Data in Indonesia," in *SENTIMAS: Seminar Nasional Penelitian dan Pengabdian Masyarakat*, 2022, pp. 66–71.
- [17] I. Satibi, "Manajemen Publik Dalam Perspektif Teoritik dan Empirik." Unpas Press, 2023.
- [18] P. Subekti, T. D. Andini, and M. Islamiyah, "Sistem Penentuan Konsentrasi Jurusan Bagi Mahasiswa Informatika Menggunakan Metode K-Means Di Institut Asia Malang," *J. Manaj. Inform.*, vol. 12, no. 1, pp. 25–39, 2022.
- [19] I. Romli, "Penerapan Data Mining Menggunakan Algoritma K-Means Untuk Klasifikasi Penyakit Ispa," *Indones. J. Bus. Intell.*, vol. 4, no. 1, pp. 10–15, 2021.
- [20] I. V. Pustokhina *et al.*, "Dynamic customer churn prediction strategy for business intelligence using text analytics with evolutionary optimization algorithms," *Inf. Process. Manag.*, vol. 58, no. 6, p. 102706, 2021, doi: 10.1016/j.ipm.2021.102706.
- [21] H. Santoso, H. Magdalena, and H. Wardhana, "Aplikasi Dynamic Cluster pada K-Means Berbasis Web untuk Klasifikasi Data Industri Rumah," *MATRIK J. Manajemen, Tek. Inform. dan Rekayasa Komput.*, vol. 21, no. 3, pp. 541–554, 2022.
- [22] F. Yenila, S. Wahyuni, E. Rianti, H. Marfalino, and D. Gusmita, "Sistem Pakar Deteksi Hemangioma pada Batita menggunakan Metode Hybrid," *J. Inf. dan Teknol.*, pp. 265–270, 2022.
- [23] K. Haryono, "Dashboard E-Procurement Berbasis Website Guna Monitoring dan Preventing Studi Kasus PT HJ Glove Indonesia," 2020.
- [24] S. U. Putri, E. Irawan, and F. Rizky, "Implementasi Data Mining Untuk Prediksi Penyakit Diabetes Dengan Algoritma C4.5," *Kesatria J. Penerapan Sist. Inf. (Komputer dan Manajemen)*, vol. 2, no. 1, pp. 39–46, 2021.
- [25] A. Zein and E. Suharyanto, "SENTIMEN ANALISIS PADA KOMENTAR PENDEK EVALUASI DOSEN OLEH MAHASISWA (EDOM) PROGRAM STUDI SISTEM INFORMASI UNIVERSITAS PAMULANG," *J. Ilmu Komput.*, vol. 5, no. 1, p. 17, 2022.
- [26] S. A. FADHILAH, "Analisis Perilaku Pembelian Konsumen Menggunakan Metode Association Rule-Market Basket Analysis Dan Clustering Analysis (Studi Kasus: Jore Coffee & Eatery)," 2022.
- [27] M. A. N. Siregar, "Klasifikasi Penerimaan Nasabah Kredit Pensiun & Pra Pensiun Untuk Mengurangi Resiko Kredit Macet Menggunakan Metode Naïve Bayes," *SkripsiKu-2022*, vol. 1, no. 1, 2022.
- [28] S. Y. Nursyi'ah, A. Erfina, and C. Warman, "Analisis Sentimen Pembelajaran Daring Pada Masa Pandemi Covid-19 Di Twitter Menggunakan Algoritma Naïve Bayes," in *Prosiding Seminar Nasional Sistem Informasi dan Manajemen Informatika Universitas Nusa Putra*, 2021, pp. 117–123.
- [29] N. H. Harani, C. Prianto, and F. A. Nugraha, "Segmentasi Pelanggan Produk Digital Service Indihome Menggunakan Algoritma K-Means Berbasis Python," *J. Manaj. Inform.*, vol. 10, no. 2, pp. 133–146, 2020.
- [30] D. I. Putri and M. Y. Putra, "Komparasi Algoritma dalam Memprediksi Perubahan Harga Saham GOTO Menggunakan

- RapidMiner,” *J. Khatulistiwa Inform.*, vol. 11, no. 1, pp. 14–20, 2023.
- [31] N. AISYAH, “IMPLEMENTASI ALGORITMA K-NEAREST NEIGHBOR UNTUK SISTEM REKOMENDASI TOPIK PENELITIAN PADA JURUSAN ILMU KOMPUTER UNIVERSITAS LAMPUNG BERBASIS KLASIFIKASI,” 2021.
- [32] A. S. Yaumi, Z. Zulfiqar, and A. Nugroho, “Klasterisasi Karakter Konsumen Terhadap Kecenderungan Pemilihan Produk Menggunakan K-Means,” *JOINTECS (Journal Inf. Technol. Comput. Sci.)*, vol. 5, no. 3, pp. 195–202, 2020.
- [33] N. Khoiruzzaman, R. D. Ramadhani, and A. Junaidi, “Hasil Klasifikasi Algoritma Backpropagation dan K-Nearest Neighbor pada Cardiovascular Disease,” *J. Dinda Data Sci. Inf. Technol. Data Anal.*, vol. 1, no. 1, pp. 17–27, 2021.
- [34] R. Efendi, D. Andreswari, and A. Mukhtadin, “Pengelompokkan dan Pemetaan Derajat Kesehatan Masyarakat Pada Tingkat Kelurahan Kota Lubuklinggau Dengan Metode FUZZY C-MEANS,” *Rekursif J. Inform.*, vol. 10, no. 2, pp. 153–166, 2022.