

Terbit online pada laman web jurnal: <https://jurnal.plb.ac.id/index.php/tematik/index>

T E M A T I K

Jurnal Teknologi Informasi Komunikasi (e-Journal)

Vol. 10 No. 2 (2022) 204 - 209

ISSN Media Elektronik: 2443-3640

## Penerapan Algoritma *K-Means* Untuk Penentuan Wilayah Penjualan Potensial Pada Perusahaan Jasa *Cleaning Service* *Implementation of K-Means Algorithm for Determining Areas Potential Sales to Cleaning Service Companies*

Reynalda Vonna Syalwa<sup>1</sup>, Vivine Nurcahyawati<sup>2</sup>, Tutut Wuriyanto<sup>3</sup>  
<sup>1,2,3</sup>Sistem Informasi, Fakultas Teknologi Informatika, Universitas Dinamika  
<sup>1</sup>vonnasyalwareynalda@gmail.com, <sup>2</sup>vivine@dinamika.ac.id<sup>2</sup>, <sup>3</sup>tutut@dinamika.ac.id<sup>3</sup>

### Abstract

Promotion is important for companies in encouraging sales and helping consumers recognize and remember the company brand. Without proper promotion, it is impossible for a business to be known to consumers, which can be detrimental to brand growth and sales. This is the same as the condition that occurs in cleaning service companies located in East Java, where the company has problems when receiving customers in several different places in East Java, namely that the company does not have information about potential areas to promote its service offerings, so this has an impact on the company cannot plan promotions that are right on target and the company can miss valuable business opportunities. The aim of this research is to process company sales data from 2021 to 2023 with data mining using the *k-means* clustering approach to group 3 clusters, namely Cluster 0 (C0) with less potential, Cluster 1 (C1) with sufficient potential, Cluster 2 (C2) has great potential. Based on the results of data processing obtained, Cluster 0 contains 4 less potential areas, namely the Mojokerto, Jombang, Pasuruan, Malang areas. Cluster 1 contains 3 areas with potential, namely Madura, Gresik, Lamongan. In cluster 2 there are 2 areas with great potential, namely the Sidoarjo area and Surabaya.

Keywords: *cleaning service, data mining, k-means, clustering, sales*

### Abstrak

Promosi merupakan hal penting bagi perusahaan dalam mendorong penjualan serta membantu konsumen mengenal dan mengingat merek perusahaan. Apabila tanpa promosi yang tepat, suatu bisnis tidak mungkin dikenal konsumen sehingga dapat merugikan pertumbuhan dan penjualan merek. Hal tersebut sama dengan kondisi terjadi di perusahaan *cleaning service* yang berlokasi di Jawa Timur dimana perusahaan memiliki permasalahan pada saat menerima pelanggan di beberapa tempat yang berbeda di Jawa Timur perusahaan yakni tidak mempunyai informasi tentang area potensial untuk melakukan promosi penawaran jasanya sehingga hal tersebut berdampak bagi perusahaan yaitu tidak dapat merancang promosi yang tepat sasaran dan perusahaan bisa melewatkan peluang bisnis yang berharga. Tujuan dari penelitian ini adalah untuk melakukan pengolahan data penjualan perusahaan dari tahun 2021 sampai dengan tahun 2023 dengan data mining menggunakan pendekatan *clustering k-means* mengelompokkan 3 cluster yakni Cluster 0 (C0) wilayah kurang berpotensi, Cluster 1 (C1) cukup berpotensi, Cluster 2 (C2) sangat berpotensi. Berdasarkan hasil dari pengolahan data didapatkan Cluster 0 berjumlah 4 wilayah kurang berpotensi yakni wilayah Mojokerto, Jombang, Pasuruan, Malang. Cluster 1 berjumlah 3 wilayah cukup berpotensi yakni wilayah Madura, Gresik, Lamongan. Pada cluster 2 berjumlah 2 wilayah sangat berpotensi, yakni wilayah Sidoarjo, dan Surabaya.

Kata kunci: *cleaning service, data mining, k-means, clustering, penjualan*

### 1. Pendahuluan

PT.XYZ adalah sebuah perusahaan yang bergerak dibidang penjualan jasa *cleaning service*. Usaha *cleaning service* merupakan pekerjaan dalam memelihara kebersihan serta memberikan layanan kebersihan di sebuah tempat [1]. PT.XYZ berfokus

pada jasa membersihkan kasur, sofa, jok mobil, dan kursi kantor, untuk membersihkannya menggunakan beberapa alat canggih seperti *vacuum cleaner*. PT.XYZ ini berlokasi di Jawa Timur berdiri sejak tahun 2021 bertahan hingga sampai saat ini tahun 2023. Perkembangan usaha *cleaning service* pada saat ini merupakan usaha yang mendatangkan keuntungan

yang besar dikarenakan banyaknya perumahan dan apartemen membutuhkan manusia pada perawatan [2].

Sejauh ini penyebaran dalam perkembangan usaha *cleaning service* pada PT.XYZ tersebar hingga beberapa kabupaten di Jawa Timur. Proses transaksi dilakukan menggunakan sistem *online* dan untuk meningkatkan pendapatan pada perusahaan perlu dilakukan kegiatan promosi. Namun, pada saat ini perusahaan belum punya informasi tentang mana daerah berpotensi untuk melakukan kegiatan promosi.

Pada permasalahan tersebut pihak perusahaan PT.XYZ merencanakan solusi dengan melakukan pengolahan data untuk menghasilkan informasi terkait area potensial dalam memasarkan jasanya. dimana data yang diolah adalah data area pelanggan pada tahun 2021 hingga 2023 untuk dijadikan objek penelitian.

Tujuan dari penelitian adalah mengolah data dengan menerapkan data mining khususnya analisis *clustering* untuk memberikan rekomendasi daerah potensial dalam penjualan jasa *cleaning service*, pengolahan ini didukung menggunakan data mining. Data mining adalah pemrosesan data menggunakan teknik statistika, matematika untuk mengidentifikasi informasi yang berhubungan dengan database [3][4]. Pengelompokan data transaksi pelanggan dibagi menjadi wilayah kabupaten potensial, cukup potensial, kurang potensial terhadap penjualan jasa *cleaning service* dengan menggunakan metode *Clustering*, *clustering* adalah metode analisis data yang melakukan proses pemodelan dan pengelompokan data [5]. Setelah dilakukannya pengelompokan data sesuai kebutuhan, maka proses *cluster* diawali tanpa mengikuti proses hirarki metode itu disebut dengan metode *k-means clustering* [6][7]. Data yang telah dikelompokkan mampu memudahkan instansi mengetahui wilayah kabupaten yang potensial, cukup potensial, dan kurang potensial terhadap pemilihan lokasi kantor operasional yang dibuka.

Beberapa penelitian terdahulu menggunakan *algoritma k-means clustering* yakni Randi P. dan Wadisman (2018) penggunaan *algoritma k-means clustering* dalam memilih pelanggan potensial data bertitik pusat centroid terbesar yang akan menjadi pelanggan potensial [8]. Zulrahmadi (2020) implementasi *algoritma k-means* dalam pengelompokan wilayah potensial terhadap penjualan sepeda motor dimana pusat centroid menggunakan *euclidean distance* didalam menentukan wilayah [9]. Penelitian Fatmawati (2018) *algoritma k-means clustering* pengelompokan wilayah yang terjangkau penyakit demam berdarah, hasil Cluster 1 dengan kategori wilayah tertinggi berjumlah 4, Cluster 2 tingkatan sedang jumlah 13 provinsi, dan Cluster 3 tingkatan rendah berjumlah 17 provinsi [10]. Penelitian Maulida (2018) *algoritma k-means* dalam pengelompokan kunjungan wisatawan ke objek wisata unggulan pada provinsi Jakarta dengan

menemukan *Cluster 1* = Ancol, *Cluster 2* = Taman Mini Indonesia, *Cluster 3* = Museum Sejarah [11]. Penelitian Triansyah (2018) *algoritma k-means clustering* dalam menentukan strategi marketing diperoleh 2 *cluster* untuk informasi tipe sepatu paling laris[12].

Berdasarkan beberapa penelitian yang sudah ada kelebihan dari *k-means* yaitu bersifat sederhana dalam konsep, *k-means* menghasilkan kelompok yang mudah diinterpretasikan karena setiap *cluster* diwakili oleh pusat yang sudah ditentukan. *K-Means* cocok digunakan pada penelitian saat ini karena pada segmentasi wilayah *k-means* dapat membantu penulis mengelompokkan area geografis berdasarkan karakteristik tertentu yang relevan untuk penjualan jasa *cleaning service*.

## 2. Metode Penelitian

Objek penelitian menggunakan data yang bersumber dari penjualan jasa *cleaning service* tahun 2021 hingga 2023 pada PT.Xyz. Dalam penelitian ini metode yang digunakan yaitu metode *clustering* menggunakan *k-means*. Seperti pada Gambar 1.



Gambar.1 Rancangan Penelitian

Tahap awal yang dilakukan adalah mengumpulkan data dimana berisi menentukan data yang akan diolah nantinya sesuai dengan studi kasus yang dibahas dan data yang dikumpulkan selanjutnya akan disimpan.

Tanggal	Nama pelanggan	Alamat	Item
3-01-2021	Hazzah	Kecamatan.Wates, Mojokerto	Cuci Kasur
3-01-2021	Hajrulrafi	Kecamatan.Cerme, Gresik	Cuci Sofa

Pada tahap kedua yang dilakukan yakni menentukan data set. tahapan menentukan data set merupakan proses kedua sebelum terjadinya pengolahan data. Pada saat menentukan data set hal yang perlu dilakukan ialah mencari apa dibahas pada penelitian, dan mencari data set yang akurat sesuai dengan studi kasus penelitian. Tahapan ketiga tahapan *cleansing* data merupakan proses perbaikan data, pembersihan data, membuang sebagian data yang ganda, serta memverifikasi data yang dianggap tidak konsisten [13], [14]. Hal ini berguna untuk memudahkan proses perhitungan *cluster k-means*. Tahapan menentukan *cluster* dimana tahapan ini dilakukan dengan mengkategorikan atau pengelompokan sekelompok objek berdasarkan atribut dari objek tersebut kedalam (K) kluster/partisi [15]. Tahapan keempat menentukan titik *centroid*. *Centroid* merupakan nilai pusat (*center*) dari sebuah *cluster*. Pada tahap ini menentukan titik centroid dapat dilakukan dengan cara memilih secara acak. Tahapan menghitung jarak *centroid* terdekat pada penelitian ini dapat dilakukan dengan melakukan perhitungan antara titik centroid dengan data. Perhitungan dilakukan dengan beberapa kali iterasi untuk mendapatkan hasil akhir. Perhitungannya menggunakan rumus korelasi antar dua objek, yang terdapat pada Persamaan 1 *Eulidance Distance*.

$$D_{(w,x)} = \sum_{i=0}^n (w_i - x_i)^2 \quad (1)$$

n adalah Jumlah data, w adalah Nilai objek, x adalah Nilai *centroid*.

Tahapan kelima merupakan tahapan penjabaran hasil dari perhitungan jarak *centroid* yang sudah dilakukan pada tahap sebelumnya dan dikelompokkan menjadi beberapa *cluster*.

Tahapan hasil cluster merupakan tahapan akhir dimana setelah dilakukannya beberapa kali iterasi menghasilkan data *cluster* yang tidak mengalami perubahan maka, iterasi berhenti dilakukan dan hasil akhir didapat dari data *cluster* terbaru dan menghasilkan sebuah informasi dari tiap *cluster*.

### 3. Hasil dan Pembahasan

#### 3.1. Menentukan Data set

Pada tahapan menentukan data set dimana data yang diambil merupakan data mentah dari penjualan jasa *cleaning service* yang berkisaran 3 tahun pada PT.Xyz yakni pada tahun 2020, tahun 2022, hingga tahun 2023 seperti terlihat pada Tabel 1 - 3.

Tabel 1. Data Set 2021

Tanggal	Nama pelanggan	Alamat	Item
3-01-2022	Dewa	Kecamatan.Balong bendo, Sidoarjo	Cuci Sofa
4-01-2022	Putri	Kecamatan.Cerme, Gresik	Cuci Kasur

Tanggal	Nama pelanggan	Alamat	Item
2-01-2023	Nadya	Kecamatan.Sukolilo, Surabaya	Cuci Kasur
3-01-2023	Dewi Anggraini	Kecamatan.Waru, Sidoarjo	Cuci Kasur

#### 3.1 Melakukan *Cleansing Data*

Pada tahapan *cleansing* data, data yang akan dibersihkan adalah data set dari penjualan jasa *cleaning service* pada PT. Xyz di tahun 2021 hingga tahun 2023. Tabel 4 adalah hasil data yang sudah di-*cleansing*.

Tabel 4. Hasil dari *cleansing* data

Kabupaten	Penjualan Tahun 2021	Penjualan Tahun 2022	Penjualan Tahun 2023
Surabaya	28	38	41
Madura	29	19	34
Sidoarjo	40	46	66
Mojokerto	12	22	17
Gresik	29	34	41
Lamongan	12	15	53
Jombang	10	7	0
Pasuran	9	9	0
Malang	10	14	0

#### 3.2. Menentukan jumlah *cluster*

Jumlah *cluster* pada penelitian ini ditentukan menjadi 3, dimana nantinya ke-3 *cluster* menjadi penentu hasil penelitian yang terkait wilayah yang masuk didalam pengelompokan sangat berpotensi, cukup berpotensi, dan kurang berpotensi terhadap penjualan jasa *cleaning service* pada provinsi Jawa Timur.

#### 3.3. Menentukan titik *centroid* dan menghitung jarak

Penentuan titik *centroid* bisa dilakukan dengan cara acak, pada penelitian ini penulis mengambil titik *centroid* berdasarkan angka terendah, angka menengah, dan angka tertinggi mengenai total penjualan jasa *cleaning service* yang berkisaran 3 tahun pada PT. XYZ, yakni tahun 2020, tahun 2022, hingga pada tahun 2023 seperti terlihat pada Tabel 5..

Pada tahapan menghitung jarak *centroid* terdekat pada saat proses perhitungan dari rumus *Eulidance Distance* akan dihitung dengan data *cleansing*.

Tabel 5. Sample *Centroid* Acak

<i>Centroid</i>	Kabupaten	Penjualan Tahun 2021	Penjualan Tahun 2022	Penjualan Tahun 2023
C0	Mojokerto	12	22	17
C1	Surabaya	28	38	41
C2	Sidoarjo	40	46	66

Proses iterasi 1, menghitung *cluster* 0 (C0), proses perhitungan diambil dari *cleansing* data pada kabupaten nomor urut 1 dan nilai *centroid* C0.

$$D_1 = \sqrt{(12 - 28)^2 + (22 - 38)^2 + (17 - 41)^2}$$

$$D_1 = 32,98.$$

Menghitung *cluster* 1 (C1), proses perhitungan diambil dari data *cleansing* pada kabupaten nomor urut 1 dan nilai *centroid* C1.

$$D_1 = \sqrt{(28 - 28)^2 + (38 - 38)^2 + (41 - 41)^2}$$

$$D_1 = 0$$

Menghitung *cluster* 2 (C2), proses perhitungan diambil dari data *cleansing* pada kabupaten nomor urut 1 dan nilai *centroid* C2.

$$D_1 = \sqrt{(40 - 28)^2 + (46 - 38)^2 + (66 - 41)^2}$$

$$D_1 = 28,86$$

### 3.4. Mengelompokkan data masing - masing *cluster*.

Setelah perhitungan jarak *centroid* terdekat, dilanjutkan data dikelompokkan sesuai dengan *cluster* masing - masing data. Hasil pengelompokan terlihat pada Tabel 6.

Dilihat dari tabel 6 didapatkan *cluster* C0 terdapat pada wilayah kabupaten Mojokerto, Jombang, Pasuruan, Malang. *Cluster* C1 pada wilayah kabupaten Surabaya, Madura, Lamongan. *Cluster* C2 pada wilayah kabupaten Sidoarjo, Gresik dimana hasil perhitungan pada tiap *cluster* sudah dijabarkan detail sehingga mendapatkan informasi bahwa tiap kabupaten yang memiliki hasil angka paling kecil pada *cluster* hal tersebut akan dipilih menjadi *cluster* dengan jarak terdekat.

Tabel 6. Data Sample Perhitungan Iterasi 1

Kabupaten	C0	C1	C2	Jarak Terdekat
Surabaya	32,98	0	28,86	C1
Madura	24,22	20,27	43,28	C1
Sidoarjo	61,32	28,86	0	C2
Mojokerto	0	32,98	61,32	C0
Gresik	31,76	4,12	29,83	C2
Lamongan	36,67	30,47	43,74	C1
Jombang	22,75	54,46	82,32	C0
Pasuran	21,61	53,69	81,76	C0
Malang	18,89	50,80	79,24	C0

Selanjutnya mulai menentukan titik *centroid* baru untuk melakukan iterasi 2 dengan menghitung rata-rata hasil dari kelompok masing *cluster* pada iterasi 1. Perhitungan menentukan *centroid* baru berikut:

*Cluster* 0 (C0) berjumlah 4 data

$$C0(X1) = (12 + 10 + 9 + 10) / 4$$

$$C0(X1) = 10$$

$$C0(X2) = (22 + 7 + 9 + 14) / 4$$

$$C0(X2) = 13$$

$$C0(X3) = (17 + 0 + 0 + 0) / 4$$

$$C0(X3) = 17$$

*Cluster* 1 (C1) Berjumlah 3 data

$$C0(X1) = (28 + 29 + 12) / 3$$

$$C0(X1) = 23$$

$$C0(X2) = (38 + 19 + 15) / 3$$

$$C0(X2) = 24$$

$$C0(X3) = (41 + 34 + 53) / 3$$

$$C0(X3) = 42$$

*Cluster* 2 (C2) berjumlah 2 data

$$C0(X1) = (40 + 29) / 2$$

$$C0(X1) = 34$$

$$C0(X2) = (46 + 34) / 2$$

$$C0(X2) = 40$$

$$C0(X3) = (66 + 41) / 2$$

$$C0(X3) = 53$$

Setelah proses perhitungan dilakukan, akan didapati data *centroid* baru untuk iterasi 2. Data *centroid* baru bisa dilihat pada tabel 7.

Tabel 7. *Centroid* baru

<i>Centroid</i>	Penjualan Tahun 2021	Penjualan Tahun 2022	Penjualan Tahun 2023
C0	10	13	17
C1	23	24	42
C2	34	40	53

Setelah *centroid* baru akan dilakukan *cluster* terhadap data sample. dengan cara mengulang perhitungan sesuai dengan iterasi 1, maka pada iterasi 2 didapatkan *cluster* baru data wilayah yang termasuk dari kategori pengelompokan sesuai dengan *cluster* masing-masing seperti pada Tabel 8 data sample perhitungan iterasi 2.

Tabel 8. Data Sample Perhitungan Iterasi 2

Kabupaten	<i>Cluster</i>
Surabaya	C2
Madura	C1
Sidoarjo	C2
Mojokerto	C0
Gresik	C1
Lamongan	C1
Jombang	C0
Pasuran	C0
Malang	C0

Iterasi ke 3 setelah dilakukan sebuah perhitungan didapatkan hasil yang sama dengan Tabel 9.

Tabel 9. Data *Sample* Perhitungan Iterasi 3

Kabupaten	Cluster
Surabaya	C2
Madura	C1
Sidoarjo	C2
Mojokerto	C0
Gresik	C1
Lamongan	C1
Jombang	C0
Pasuruan	C0
Malang	C0

Iterasi akan berhenti dilakukan bila data *cluster* tidak mengalami perubahan. Melihat tabel 8 dan tabel 9 isi *cluster* tidak mengalami perubahan sehingga iterasi diberhentikan dengan hasil akhir *cluster* C0 terdapat pada wilayah kabupaten Mojokerto, Jombang, Pasuruan, Malang. *Cluster* C1 pada wilayah kabupaten Gresik, Madura, Lamongan. *Cluster* C2 pada wilayah kabupaten Sidoarjo, Surabaya.

Perhitungan penelitian *cluster* data set penjualan jasa *cleaning service* hanya sampai pada iterasi 2.

Tabel 10. *Cluster* 0

Kabupaten	Penjualan Tahun 2021	Penjualan Tahun 2022	Penjualan Tahun 2023
Mojokerto	12	22	17
Jombang	10	7	0
Pasuruan	9	9	0
Malang	10	14	0

Tabel 10 menunjukkan kategori *cluster* 0 dimana terlihat pada hasil penjualan pada tahun 2021 hingga 2023 memiliki total penjualan rendah atau sedikit. Hal tersebut menandai bahwa *cluster* 0 termasuk dalam kategori wilayah kurang berpotensi.

Tabel 11. *Cluster* 1

Kabupaten	Penjualan Tahun 2021	Penjualan Tahun 2022	Penjualan Tahun 2023
Madura	29	19	34
Gresik	29	34	41
Lamongan	12	15	53

Tabel 11 menunjukkan kategori *cluster* 1 dimana terlihat pada hasil penjualan pada tahun 2021 hingga 2023 memiliki total penjualan yang cukup lumayan. Hal tersebut menandai bahwa *cluster* 1 termasuk dalam kategori wilayah cukup berpotensi.

Tabel 12 menunjukkan kategori *cluster* 2 dimana terlihat pada hasil penjualan pada tahun 2021 hingga 2023 memiliki total penjualan yang banyak. Hal

tersebut menandai bahwa *cluster* 2 termasuk dalam kategori wilayah sangat berpotensi.

Tabel 12. *Cluster* 2

Kabupaten	Penjualan Tahun 2021	Penjualan Tahun 2022	Penjualan Tahun 2023
Surabaya			1
Sidoarjo			6

#### 4. Kesimpulan

Setelah melakukan perhitungan dan pembahasan *Algoritma K-Means clustering* penulis dapat menarik kesimpulan bahwa *K-Means Clustering* dapat digunakan sebagai metode untuk mengelompokkan wilayah potensial terhadap penjualan jasa *cleaning service*. Hal ini dapat dilihat dari 9 data penjualan jasa yang diolah didapatkan hasil kategori wilayah kurang berpotensi (C0) sebanyak 4 wilayah yaitu wilayah Mojokerto, Jombang, Pasuruan, Malang. Pada kategori cukup berpotensi (C1) berjumlah 3 wilayah yaitu wilayah Madura, Gresik, Lamongan, dan kategori sangat berpotensi (C2) berjumlah 2 wilayah yaitu Surabaya, Sidoarjo

#### Daftar Rujukan

- [1] D. Septianawati and Yamtana, "Penerapan Peregangan Untuk Menurunkan Keluhan Nyeri Otot Rangka (MSDs) Pada Petugas Kebersihan di RSUD dr. Tjitrowardoyo," Politeknik Kesehatan Jogjakarta, 2019. [Online]. Available: <https://medium.com/@arifwicaksanaa/pengertian-use-case-a7e576e1b6bf>
- [2] Ardianto, "Peran Cleaning Service Terhadap Kebersihan Fakultas Dakwah Dan Komunikasi UIN Alauddin Makassar," Islam Negeri Alauddin Makassar, 2019. [Online]. Available: [http://www.scopus.com/inward/record.url?eid=2-s2.0-84865607390&partnerID=tZOtx3y1%0Ahttp://books.google.com/books?hl=en&lr=&id=2LIMMD9FVXkC&oi=fnd&pg=PR5&dq=Principles+of+Digital+Image+Processing+fundamental+techniques&ots=HjrHeuS\\_](http://www.scopus.com/inward/record.url?eid=2-s2.0-84865607390&partnerID=tZOtx3y1%0Ahttp://books.google.com/books?hl=en&lr=&id=2LIMMD9FVXkC&oi=fnd&pg=PR5&dq=Principles+of+Digital+Image+Processing+fundamental+techniques&ots=HjrHeuS_)
- [3] Nugroho Arif Sudibyo, Ardymulya Iswardani, Kartika Sari, and Siti Suprihatiningsih, "Penerapan Data Mining Pada Jumlah Penduduk Miskin Di Indonesia," *J. Lebesgue J. Ilm. Pendidik. Mat. Mat. dan Stat.*, vol. 1, no. 3, pp. 199–207, 2020, doi: 10.46306/lb.v1i3.42.
- [4] H. Prastiwi, Jeny Pricilia, and Errissya Rasywir, "Implementasi Data Mining Untuk Menentukan Persediaan Stok Barang Di Mini Market Menggunakan Metode K-Means Clustering," *J. Inform. Dan Rekayasa Komputer (JAKAKOM)*, vol. 2, no. 1, pp. 141–148, 2022, doi: 10.33998/jakakom.2022.2.1.34.
- [5] H. Sibarani, Solikhun, W. Saputra, I. Gunawan, and Z. M. Nasution, "Penerapan Metode K-Means Untuk Pengelompokan Kabupaten/Kota Di Provinsi Sumatera Utara Berdasarkan Indikator Indeks Pembangunan Manusia," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 6, no. 1, pp. 154–161, 2022, doi: 10.36040/jati.v6i1.4590.
- [6] Y. Lase and E. Panggabean, "Implementasi Metode K-Means Clustering Dalam Sistem Pemilihan Jurusan Di SMK Swasta Harapan Baru," *J. Teknol. dan Ilmu Komput. Prima*, vol. 2, no. 2, p. 43, 2019, doi: 10.34012/jutikomp.v2i2.723.
- [7] F. Nasari and C. J. M. Sianturi, "Penerapan Algoritma K-Means Clustering Untuk Pengelompokan Penyebaran Diare

- Di Kabupaten Langkat,” *CogITO Smart J.*, vol. 2, no. 2, pp. 108–119, 2016, doi: 10.31154/cogito.v2i2.19.108-119.
- [8] R. R. Putra and C. Wadisman, “Implementasi Data Mining Pemilihan Pelanggan Potensial Menggunakan Algoritma K-Means Implementation of Data Mining for Potential Customer Selection Using K-Means Algorithm,” *J. Inf. Technol. Comput. Sci.*, vol. 1, no. 1, pp. 72–77, 2018.
- [9] Zulrahmadi, S. Defit, and Y. Yunus, “Pemetaan Wilayah Potensial Terhadap Penjualan Sepeda Motor Honda Menggunakan K-Means Clustering,” *J. Inform. Ekon. Bisnis*, vol. 2, no. 2018, pp. 53–59, 2020, doi: 10.37034/infek.v2i2.41.
- [10] K. Fatmawati and A. P. Windarto, “Data Mining: Penerapan Rapidminer Dengan K-Means Cluster Pada Daerah Terjangkit Demam Berdarah Dengue (Dbd) Berdasarkan Provinsi,” *Comput. Eng. Sci. Syst. J.*, vol. 3, no. 2, p. 173, 2018, doi: 10.24114/cess.v3i2.9661.
- [11] L. Maulida, “Penerapan Datamining Dalam Mengelompokkan Kunjungan Wisatawan Ke Objek Wisata Unggulan Di Prov. Dki Jakarta Dengan K-Means,” *JISKA (Jurnal Inform. Sunan Kalijaga)*, vol. 2, no. 3, p. 167, 2018, doi: 10.14421/jjska.2018.23-06.
- [12] D. Triyansyah and D. Fitriana, “Analisis Data Mining Menggunakan Algoritma K-Means Clustering Untuk Menentukan Strategi Marketing,” *J. Telekomun. dan Komput.*, vol. 8, no. 3, p. 163, 2018, doi: 10.22441/incomtech.v8i3.4174.
- [13] W. D. Budimulia and F. Ridho, “Penerapan Komputasi Paralel Pada Aplikasi Data Cleaning Multiple Data Edit,” *Semin. Nas. Off. Stat.*, vol. 2019, no. 1, pp. 7–14, 2020, doi: 10.34123/semnasoffstat.v2019i1.120.
- [14] I. K. Juni Arta, G. Indrawan, and G. R. Dantes, “Data Mining Rekomendasi Calon Mahasiswa Berprestasi Di Stmik Denpasar Menggunakan Metode Technique for Others Reference By Similarity To Ideal Solution,” *JST (Jurnal Sains dan Teknol.*, vol. 5, no. 2, pp. 11–21, 2017, doi: 10.23887/jstundiksha.v5i2.8549.
- [15] B. Harahap, “Penerapan Algoritma K-Means Untuk Menentukan Bahan Bangunan Laris (Studi Kasus Pada UD. Toko Bangunan YD Indarung),” *Reg. Dev. Ind. Heal. Sci. Technol. Art Life*, pp. 394–403, 2019, [Online]. Available: <https://ptki.ac.id/jurnal/index.php/readystar/article/view/82>