



PEMETAAN PENDUDUK PENERIMA BANTUAN SOSIAL DESA WARU JAYA MENGGUNAKAN ALGORITMA K-MEANS CLUSTERING

Deni Bahtiar

Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Pamulang

Email: dheni810@gmail.com

Kata kunci:

Pemetaan Penerima
Bantuan Dana Sosial
Desa Waru Jaya

Abstrak

Bagi pemerintah Indonesia masalah kemiskinan merupakan masalah lama yang belum dan sulit untuk diselesaikan. Pemerintah sendiri telah melakukan beberapa upaya dalam melakukan pengentasan kemiskinan melalui program bantuan sosial diantaranya Bantuan Langsung Tunai (BLT), Program Keluarga Harapan (PKH) dll. Berdasarkan hasil survey yang dilakukan bantuan yang di berikan pemerintah kepada penduduk tidak sesuai dengan apa yang mereka butuhkan. Satu kesulitan yang terkadang dihadapi oleh pemerintah dalam proses penanganan kemiskinan adalah proses pembagian bantuan sosial yang tidak merata dan tidak tepat sasaran. Penyaluran bantuan sosial oleh pemerintah yang disalurkan kepada Desa Waru Jaya diberikan kepada penduduk miskin/rentan miskin. Namun terdapat beberapa kendala. Kendala tersebut dikarenakan banyak data yang belum diolah sehingga memungkinkan proses penentuan penerima bantuan sosial tidak tepat sasaran. Untuk mengatasi masalah tersebut, dibutuhkan metode penerapan data mining agar penyaluran bantuan ini tepat sasaran. Oleh karena itu dibutuhkan metode klasterisasi (clustering) yang merupakan pengelompokan item data. Banyak metode klasterisasi yang diusulkan oleh para ahli, salah satunya adalah k-means method atau metode k-means yang termasuk dalam metode berbasis partisi, dimana metode ini membagi data kedalam sejumlah kelompok. Metode k-means merupakan algoritma klasterisasi yang paling sering digunakan karena prosesnya sederhana untuk diimplementasikan dan dijalankan. Selain mampu melakukan klasterisasi dan mudah diadaptasi, operasi matematis dalam metode k-Means juga relative lebih sederhana

PENDAHULUAN

Kemiskinan merupakan salah satu persoalan mendasar yang menjadi pusat perhatian suatu negara baik di perkotaan maupun desa. Salah satu aspek penting untuk mendukung strategi penanggulangan kemiskinan adalah dengan tersedianya data kemiskinan yang akurat dan tepat sasaran. Kategori masyarakat miskin merupakan suatu kondisi dimana masyarakat tidak memiliki akses sarana dan prasarana dasar lingkungan yang memadai, serta

mata pencaharian yang tidak menentu yang mencakup seluruh multidimensi (Hasymi et al., 2021).

Bagi pemerintah Indonesia masalah kemiskinan merupakan masalah lama yang belum dan sulit untuk diselesaikan. Pemerintah sendiri telah melakukan beberapa upaya dalam melakukan pengentasan kemiskinan melalui program bantuan sosial diantaranya Bantuan Langsung Tunai (BLT), Program Keluarga Harapan (PKH) dll. Berdasarkan hasil survey yang dilakukan bantuan yang di berikan pemerintah kepada penduduk tidak sesuai dengan apa yang mereka butuhkan. Satu kesulitan yang terkadang dihadapi oleh pemerintah dalam proses penanganan kemiskinan adalah proses pembagian bantuan sosial yang tidak merata dan tidak tepat sasaran. Ini disebabkan karena validasi data sering diabaikan sehingga menimbulkan data yang tidak akurat (Sunia et al., 2017).

Penyaluran bantuan sosial oleh pemerintah yang disalurkan kepada Desa Waru Jaya diberikan kepada penduduk miskin/rentan miskin. Namun terdapat beberapa kendala. Kendala tersebut dikarenakan banyak data yang belum diolah sehingga memungkinkan proses penentuan penerima bantuan sosial tidak tepat sasaran. Untuk mengatasi masalah tersebut, dibutuhkan metode penerapan data mining agar penyaluran bantuan ini tepat sasaran. Oleh karena itu dibutuhkan metode klasterisasi (*clustering*) yang merupakan pengelompokan item data.

Banyak metode klasterisasi yang diusulkan oleh para ahli, salah satunya adalah *k-means method* atau metode *k-means* yang termasuk dalam metode berbasis partisi, dimana metode ini membagi data kedalam sejumlah kelompok. Metode *k-means* merupakan algoritma klasterisasi yang paling sering digunakan karena prosesnya sederhana untuk diimplementasikan dan dijalankan. Selain mampu melakukan klasterisasi dan mudah diadaptasi, operasi matematis dalam metode *k-Means* juga relatif lebih sederhana

Algoritma

Algoritma adalah urutan logis langkah langkah penyelesaian masalah yang disusun secara sistematis. Metode pengurutan digambarkan dalam sejumlah langkah terbatas yang mengarah pada solusi permasalahan.

Data Mining

(Ridwan, Suyono, & Saroso, 2013) Data mining adalah proses yang menggunakan teknik statistic, matematika, kecerdasan buatan, dan machine learning untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar. Istilah data mining memiliki hakikat sebagai disiplin ilmu yang tujuan utamanya adalah untuk menemukan, menggali, atau menambang pengetahuan dari data atau informasi yang kita miliki. Data mining, sering juga disebut sebagai Knowledge

Discovery Database (KDD). KDD adalah kegiatan yang meliputi pengumpulan, pemakaian data, historis untuk menemukan keteraturan, pola atau hubungan dalam kumpulan data berukuran besar (Astuti et al., 2019)

Data mining adalah sebagai bahan keterangan tentang kejadian nyata atau fakta-fakta yang dirumuskan dalam sekelompok lambing tertentu yang tidak acak yang menunjukkan jumlah, tindakan, atau hal. Data adalah fakta dari suatu pernyataan yang berasal dari kenyataan, dimana pernyataan tersebut merupakan hasil pengukuran atau pengamatan. Jadi dapat disimpulkan bahwa data adalah sekumpulan fakta-fakta hasil pengukuran dan pengamatan yang belum memiliki arti dan dapat diolah menjadi suatu informasi tertentu (Sunia et al., 2017)

Fungsi Data Mining

Data mining mengidentifikasi fakta – fakta atau kesimpulan yang disarankan berdasarkan penyaringan melalui data untuk menjelajah pola – pola atau anomali – anomali data. Data mining mempunyai lima fungsi, yaitu:

a) *Classification*

Menyimpulkan definisi-definisi karakteristik dari sebuah grup. Contoh: pelanggan-pelanggan perusahaan yang telah berpindah ke saingan perusahaan lain.

b) *Clustering*

Mengidentifikasi kelompok-kelompok dari barang-barang dengan produk-produk dengan berbagai karakteristik khusus, clustering berbedda dengan classification, dimana pada *clustering* tidak terdapat definisi-definisi karakteristik awal yang diberikan pada waktu *classification*.

c) *Association*

Mengidentifikasi hubungan antara kejadian-kejadian yang terjadi pada suatu waktu, seperti isi-isi dari keranjang belanja.

d) *Sequencing*

Hampir sama dengan, mengidentifikasi hubungan-hubungan yang mengunjungi *supermarket* secara berulang.

e) *Forecasting*

Memperkirakan nilai pada masa yang akan datang berdasarkan pola- pola dengan sekumpulan data yang besar, seperti peramalan permintaan pasar.

Rapidminer

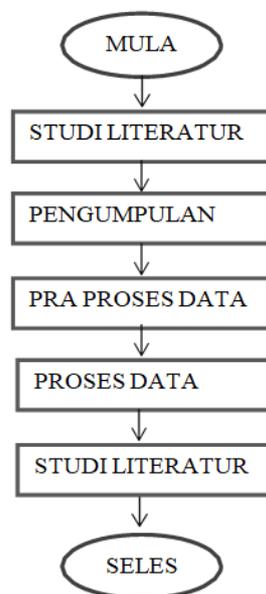
RapidMiner merupakan perangkat lunak yang bersifat terbuka (open source). RapidMiner adalah sebuah solusi untuk melakukan analisis prediksi. RapidMiner menggunakan berbagai teknik deskriptif dan prediksi dalam memberikan wawasan kepada

pengguna sehingga dapat membuat keputusan yang paling baik. RapidMiner memiliki kurang lebih 500 operator data mining, termasuk operator untuk input, output, data preprocessing dan visualisasi. RapidMiner merupakan software yang berdiri sendiri untuk analisis data dan sebagai mesin data mining yang dapat diintegrasikan pada produknya sendiri. RapidMiner ditulis dengan menggunakan bahasa java sehingga dapat bekerja di semua system operasi.

RapidMiner sebelumnya bernama YALE (Yet Another Learning Environment), dimana versi awalnya mulai dikembangkan pada tahun 2001 oleh Ralfklinkenberg, Ingo Mierswa, dan Simon Fischer di Artificial Intelligence Unit dari University of Dortmund. RapidMiner didistribusikan di bawah lisensi AGPL (GNU Afero General Public Licence) versi 3. RapidMiner sebagai software open source untuk data mining tidak perlu diragukan lagi karena software ini sudah terkemuka didunia.

METODOLOGI PENELITIAN

Tahapan Penelitian



Gambar 3.2 Flowchart Tahapan Penelitian

Pra Proses Data (Pre-Processing Data)

Data mentah yang telah diperoleh tidak serta merta dapat langsung diolah. Agar dapat diolah menggunakan metode *clustering K-Means* diperlukan pra proses data guna memperoleh atribut dan tipe data yang cocok dan berguna untuk diolah dalam proses *data mining*.

Visi

“Religius, Berpendidikan, Insfrastruktur yang baik dan Terselenggaranya Pelayanan Prima”

Misi

Misi pemerintah kelurahan Desa Waru Jaya Kecamatan Parung, Bogor adalah:

1. Menciptakan Kondisi lingkungan yang kondusif dan agamis, ditunjang olehtersedianya sarana peribadatan dan Sukmber Daya Manusia yang kompeten.
2. Membangun pola hidup sehat melalui pemberdayaan kader kesehatan dan optimalisasi Desa Siaga.
3. Menyelenggaran Pemerintah yang transparan, akuntabilitas, partisipatif dan responsive.
4. Meningkatkan dan memperdayakan peranan perempuan dan pemuda dengan tetap memelihara adat istiadat dan budaya local.
5. Membangun sarana dan prasarana yang berbasis pada Agamis, berpendidikan yang produktif, insfratuktur perdesaan, dalam upaya peningkatan keimanan dan pendidikan.

HASIL DAN PEMBAHASAN

Proses Clustering

Pada tahap ini dilakukan proses utama yaitu segmentasi pengelompokan data kependudukan. Berikut merupakan penerapan algoritma k-means dengan asumsi bahwa parameter input adalah jumlah dataset sebanyak n data dan jumlah inisialisasi *centroid* k adalah 2. Data yang diambil untuk penelitian berjumlah 823 data. Percobaan dilakukan dengan menggunakan parameter-parameter berikut:

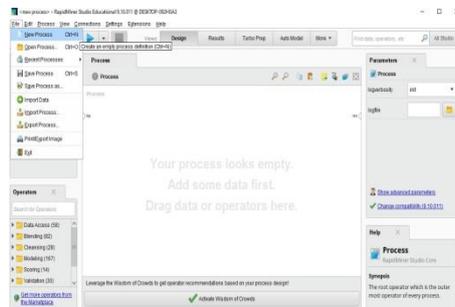
Jumlah cluster = 3
Jumlah data = 823
Jumlah atribut = 3

Pengujian RapidMiner

Pada pengujian ini penulis menggunakan *tool rapidminer* sebagai alat pengujian dataset. Langkah awal yang dilakukan adalah membuka aplikasi *RapidMiner Studio Versi 9.10*.



Gambar 4.1 Gambar Tampilan Awal RapidMiner



Gambar 4.2 Gambar Tampilan Menu Utama RapidMiner

Disini penulis memilih menu *open process* karena sebelumnya telah melakukan percobaan menggunakan aplikasi *RapidMiner*. Bila belum pernah melakukan percobaan, maka dapat memilih menu *new process* dan memulai dengan pengolahan data yang baru. Dan menampilkan tampilan awal dari *new process* yang kemudian dilakukan tahapan proses *clustering k-means*.

Berikut adalah tahapan proses pengujian yang dilakukan dalam *RapidMiner*:

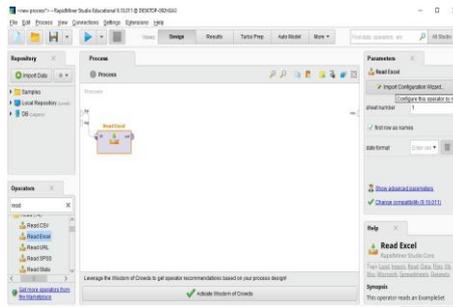


Gambar 4.3 Gambar Design Process

Gambar diatas merupakan gambaran keseluruhan proses *clustering*. Terdapat 5 tahapan yang terdapat dalam gambar *design process clustering*, yaitu:

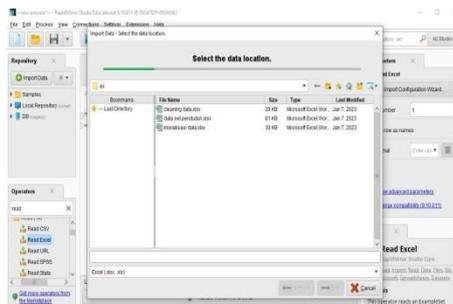
1. Read Excel

Read Excel merupakan tahapan pertama saat akan melakukan proses *clustering*. Dalam tahap ini dilakukan penginputan dataset berupa file berekstensi *.xls* ke dalam *RapidMiner*. Data tersebut adalah data yang telah melalui tahap *preprocessing* sebelumnya dan siap untuk diolah menggunakan metode *clustering k-means*.



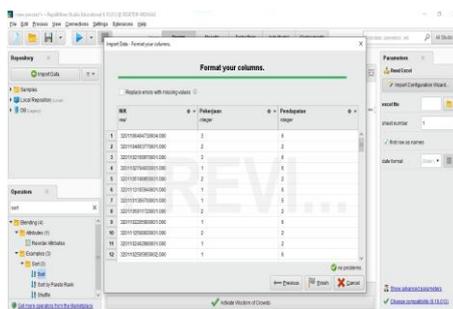
Gambar 4.4 Gambar Process Read Excel

Agar data yang berekstensi .xls dapat dibaca maka dilakukan import data dari computer dengan menggunakan *Parameter Import Configuration Wizard*.



Gambar 4.5 Gambar Data Import Wizard

Sebelum proses import data selesai terlebih dahulu dilakukan pemilihan tipe data yang sesuai dengan atribut yang akan diolah.



Gambar 4.6 Gambar Pemilihan Atribut

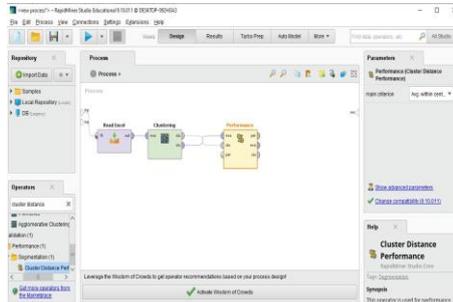
2. *Clustering* Tahapan ini adalah tahapan dimana dilakukannya pengoperasian proses *clustering* menggunakan algoritma *k-means*. Dalam proses ini ditentukan dalam parameter berapa jumlah *cluster* yaitu 3.



Gambar 4.7 Process Clustering K-Means

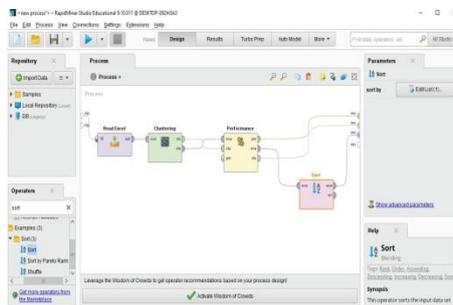
3. Performance

Tahapan *performance* ini, bertujuan untuk mencari nilai DBI (*Davies Bouldin Index*). Dimana parameter main criterion disetting Davies Bouldin dan maximize guna memaksimalkan nilai hasil dari DBI.



Gambar 4.8 Gambar Performance

4. *Sort* Tahapan ini dilakukan operasi pengurutan anggota cluster 0 sampai dengan 2. Dengan parameter atribut name adalah cluster.



Gambar 4.9 Gambar Process Sort

Setelah semua proses telah dijalankan, maka ditampilkan hasil sebagaiberikut :

ID	Row No.	id	cluster	DBI	Paketan	Produk
1	1	nama_0	0	1.1	2	
2	2	nama_0	0	1.1	2	
3	3	nama_0	0	1.1	2	
4	4	nama_0	0	1.1	2	
5	5	nama_0	0	1.1	2	
6	6	nama_0	0	1.1	2	
7	7	nama_0	0	1.1	2	
8	8	nama_0	0	1.1	2	
9	9	nama_0	0	1.1	2	
10	10	nama_0	0	1.1	2	
11	11	nama_0	0	1.1	2	
12	12	nama_0	0	1.1	2	
13	13	nama_0	0	1.1	2	
14	14	nama_0	0	1.1	2	
15	15	nama_0	0	1.1	2	
16	16	nama_0	0	1.1	2	
17	17	nama_0	0	1.1	2	
18	18	nama_0	0	1.1	2	
19	19	nama_0	0	1.1	2	
20	20	nama_0	0	1.1	2	
21	21	nama_0	0	1.1	2	
22	22	nama_0	0	1.1	2	
23	23	nama_0	0	1.1	2	
24	24	nama_0	0	1.1	2	
25	25	nama_0	0	1.1	2	
26	26	nama_0	0	1.1	2	
27	27	nama_0	0	1.1	2	
28	28	nama_0	0	1.1	2	
29	29	nama_0	0	1.1	2	
30	30	nama_0	0	1.1	2	
31	31	nama_0	0	1.1	2	
32	32	nama_0	0	1.1	2	
33	33	nama_0	0	1.1	2	
34	34	nama_0	0	1.1	2	
35	35	nama_0	0	1.1	2	
36	36	nama_0	0	1.1	2	
37	37	nama_0	0	1.1	2	
38	38	nama_0	0	1.1	2	
39	39	nama_0	0	1.1	2	
40	40	nama_0	0	1.1	2	
41	41	nama_0	0	1.1	2	
42	42	nama_0	0	1.1	2	
43	43	nama_0	0	1.1	2	
44	44	nama_0	0	1.1	2	
45	45	nama_0	0	1.1	2	
46	46	nama_0	0	1.1	2	
47	47	nama_0	0	1.1	2	
48	48	nama_0	0	1.1	2	
49	49	nama_0	0	1.1	2	
50	50	nama_0	0	1.1	2	
51	51	nama_0	0	1.1	2	
52	52	nama_0	0	1.1	2	
53	53	nama_0	0	1.1	2	
54	54	nama_0	0	1.1	2	
55	55	nama_0	0	1.1	2	
56	56	nama_0	0	1.1	2	
57	57	nama_0	0	1.1	2	
58	58	nama_0	0	1.1	2	
59	59	nama_0	0	1.1	2	
60	60	nama_0	0	1.1	2	
61	61	nama_0	0	1.1	2	
62	62	nama_0	0	1.1	2	
63	63	nama_0	0	1.1	2	
64	64	nama_0	0	1.1	2	
65	65	nama_0	0	1.1	2	
66	66	nama_0	0	1.1	2	
67	67	nama_0	0	1.1	2	
68	68	nama_0	0	1.1	2	
69	69	nama_0	0	1.1	2	
70	70	nama_0	0	1.1	2	
71	71	nama_0	0	1.1	2	
72	72	nama_0	0	1.1	2	
73	73	nama_0	0	1.1	2	
74	74	nama_0	0	1.1	2	
75	75	nama_0	0	1.1	2	
76	76	nama_0	0	1.1	2	
77	77	nama_0	0	1.1	2	
78	78	nama_0	0	1.1	2	
79	79	nama_0	0	1.1	2	
80	80	nama_0	0	1.1	2	
81	81	nama_0	0	1.1	2	
82	82	nama_0	0	1.1	2	
83	83	nama_0	0	1.1	2	
84	84	nama_0	0	1.1	2	
85	85	nama_0	0	1.1	2	
86	86	nama_0	0	1.1	2	
87	87	nama_0	0	1.1	2	
88	88	nama_0	0	1.1	2	
89	89	nama_0	0	1.1	2	
90	90	nama_0	0	1.1	2	
91	91	nama_0	0	1.1	2	
92	92	nama_0	0	1.1	2	
93	93	nama_0	0	1.1	2	
94	94	nama_0	0	1.1	2	
95	95	nama_0	0	1.1	2	
96	96	nama_0	0	1.1	2	
97	97	nama_0	0	1.1	2	
98	98	nama_0	0	1.1	2	
99	99	nama_0	0	1.1	2	
100	100	nama_0	0	1.1	2	

Gambar 4.10 Gambar Example Set Result

Example Set Result menampilkan data hasil dari klasterisasi data. *Label cluster* terbagi menjadi tiga kelompok yaitu *cluster 0*, *cluster 1* dan *cluster 2*. Pembagian ini berdasarkan hasil kedekatan tiap masing-masing data dengan jarak terdekat.

Name	Type	Missing	Status	File ID	Instances
id	Integer	0	T	id	823
cluster	Nominal	0		cluster_1 (71), cluster_2 (776), cluster_3	
NIK	Text	0		201101010103000, 201104010102000, 201040201041402	
Pekerjaan	Integer	0	T	1	1301
Pendidikan	Integer	0	T	6	2277

Gambar 4.11 Gambar Example Set Statistic

Example Set Statistic menampilkan hasil statistic dari proses *clustering k- means*. Terdapat nilai minimal dan maksimal pada dua atribut yang diolah dan rata-rata nilai keluar.

Cluster Model

Description
Cluster 0: 713 cases
Cluster 1: 1 cases
Cluster 2: 319 cases
Total number of cases: 823

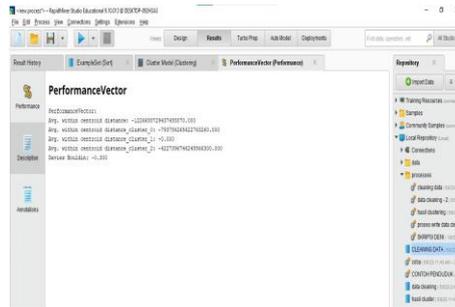
Gambar 4.12 Gambar Cluster Model

Gambar diatas menampilkan hasil pembagian data terhadap tiap cluster. Cluster 0 memiliki 104 anggota, cluster 1 memiliki 1 anggota dan cluster 2 memiliki 718 anggota. Total dari semua data set yang diuji adalah 823 anggota.

Atribut	cluster_1	cluster_2	cluster_3
NIK	201101010103000	201040201041402	201101010103000
Pekerjaan	1301	2	1301
Pendidikan	2277	2	2348

Gambar 4.13 Gambar Centroid Table

Gambar ini menampilkan nilai *centroid* masing-masing atribut pada tiap *cluster*. Nilai tersebut menjadikan acuan perhitungan pada setiap dataset dengan cara mengukur kedekatan nilai dengan masing-masing titik pusat *cluster*.



Gambar 4.15 Gambar Description Performance Vektor

Tampilan diatas merupakan detail hasil dari proses Performance pada Rapidminer yang menampilkan rata-rata jarak titik pusat pada masing-masingcluster dan juga menampilkan nilai DBI.

Kesimpulan

Dari penelitian yang sudah dilakukan maka dapat ditarik kesimpulan sebagai berikut:

1. Pemetaan penduduk penerima bantuan sosial menggunakan algoritma clustering k-means berdasarkan cluster yaitu penduduk yang layak menerima bantuan, penduduk yang kurang layak menerima bantuan dan penduduk yang tidak layak menerima bantuan.
2. Hasil pengujian mendapatka nilai DBI (Davies Bouldin Index) sebesar 0,300 yang berarti memiliki jarak antar anggota *cluster* yang cukup baik karena mendekati angka nol.

Saran

Mengingat banyaknya hal yang belum dapat diterapkan dari penelitian ini, maka penulis mempertimbangkan beberapa saran sebagai berikut:

3. Hasil cluster yang terbentuk dapat dikembangkan dengan metode lain agar dapat mendapatkan informasi lain yang belum diketahui.
4. Penelitian ini dapat dikembangkan lagi menjadi basis pengetahuan untuk sisitem penduduk keputusan pemetaan penduduk calon penerima bantuan.

DAFTAR PUSTAKA

- Andean, R., Fendy, S., & Nugroho, A. (2019). Klasterisasi Pengendalian Persediaan Aki Menggunakan Metode K-Means. *JOINTECS (Journal of Information Technology and Computer Science)*, 4(1), 5. <https://doi.org/10.31328/jointecs.v4i1.998>
- Astuti, S. I., Arso, S. P., & Wigati, P. A. (2019). Pemetaan Penduduk Calon Penerima Bantuan Renovasi Rumah Desa Pesangkalan Menggunakan Algoritma Clustering K-Means. *Analisis Standar Pelayanan Minimal Pada Instalasi Rawat Jalan Di RSUD Kota Semarang*, 3, 103–111.
- Hasymi, M. A., Faisol, A., & Ariwibisono, F. X. (2021). *SISTEM INFORMASI GEOGRAFIS PEMETAAN WARGA KURANG MAMPU DI KELURAHAN KARANGBESUKI MENGGUNAKAN METODE K-MEANS*. 5(1), 284–290.

- Nst, A. H., Munthe, I. R., & Juledi, A. P. (2021). *Implementasi Data Mining Algoritma Apriori untuk Meningkatkan Penjualan*. 06, 188–197.
- Pemetaan, K., Penerima, P., Renovasi, B., & Means, M. A. K.-. (2022). *Jurnal Pendidikan dan Konseling*. 4, 637–646.
- Rahmawati, L., Widya Sihwi, S., & Suryani, E. (2016). Analisa Clustering Menggunakan Metode K-Means Dan Hierarchical Clustering (Studi Kasus :Dokumen Skripsi Jurusan Kimia, Fmipa, Universitas Sebelas Maret). *Jurnal Teknologi & Informasi ITSmart*, 3(2), 66. <https://doi.org/10.20961/its.v3i2.654>
- Sari, D. A., Ginting, B. S., & Maulita, Y. (2022). *Pengelompokan Data Penduduk Penerima BSTP (Bantuan Sosial Tunai Pandemic) Menggunakan Metode Algoritma K-means Clustering (Kantor Desa Padang Brahrang)*. 6(3).
- Sunia, D., Jusia, P. A., Informatika, T., Informatika, T., Subdistrict, J., & Selatan, J. (2017). *PENERAPAN DATA MINING UNTUK CLUSTERING DATA PENDUDUK MISKIN MENGGUNAKAN ALGORITMA K-MEANS*. 2016,121–134.
- Wardhani, A. K. (2016). K-Means Algorithm Implementation for Clustering of Patients Disease in Kajen Clinic of Pekalongan. *Jurnal Transformatika*, 14(1), 30. <https://doi.org/10.26623/transformatika.v14i1.387>