



KLASIFIKASI PENDUDUK KURANG MAMPU DENGAN METODE K-MEANS UNTUK OPTIMALISASI PROGRAM BANTUAN SOSIAL

Bayu Baskoro¹⁾, Aris Gunaryati, Albaar Rubhasy

Program Studi Sistem Informasi, Universitas Nasional, Indonesia

Abstrak: Kemiskinan merupakan permasalahan yang selalu terjadi pada negara berkembang, khususnya Indonesia. Kemiskinan merupakan kondisi seseorang ketika tidak sanggup untuk menyanggupi kebutuhan sehari-hari. Selama pandemi Covid-19 kemiskinan di DKI Jakarta cenderung meningkat. Pada wilayah Jakarta Selatan jumlah penduduk miskin terdapat 61,77 ribu jiwa (2,73%) per 2019, lalu meningkat menjadi 78,09 ribu jiwa (3,34%) ditahun 2020 dan menjadi 81,50 ribu jiwa (3,56%) pada tahun 2021. Adapun tujuan dari penelitian ini yaitu untuk mengoptimalkan program bantuan sosial yang dinilai masih kurang tepat sasaran. Dengan menggunakan metode K-Means penelitian ini mencoba untuk membagi data penduduk kurang mampu kedalam 3 cluster berdasarkan variabel pendidikan, pekerjaan, tanggungan dan pengeluaran. Dari hasil proses K-Means didapatkan kelompok yang setelah dianalisis menghasilkan kelompok layak mendapat bantuan, sangat layak mendapat bantuan dan kurang layak mendapat bantuan dengan hasil pengujian mendapatkan skor DBI (Davies Bouldin Index) yang cukup baik, yaitu -0,996.

Kata kunci: K-Means, kemiskinan, bantuan sosial, clustering, DBI

I. PENDAHULUAN

Kemiskinan merupakan permasalahan yang sedang dihadapi hampir diseluruh dunia. Berdasarkan statistik, taraf kemiskinan setiap tahun selalu mengalami penurunan ataupun peningkatan di masing-masing daerah (Sepriyanti et al., 2022). Dalam negara berkembang, khususnya Indonesia. Kemiskinan menjadi satu dari banyaknya persoalan dasar yang menjadi inti perhatian pemerintah masa kini, dari beberapa sumber dikatakan bahwa kemiskinan ini adalah suatu penyakit ekonomi yang patut diperbaiki atau paling tidak diturunkan. Kemiskinan

merupakan kondisi seseorang ketika tidak sanggup untuk menyanggupi kebutuhan pokok sehari-hari. Salah satu penyebab kemiskinan adalah kurangnya pendidikan yang disandang oleh seseorang (Seran, 2012).

Salah satu upaya dalam mengatasi kemiskinan yaitu dengan program bantuan sosial, bantuan sosial bisa datang dari siapa saja baik itu perorangan, kelompok, organisasi dan juga pemerintah, bantuan sosial yang diberikanpun dapat berupa uang tunai maupun kebutuhan makanan.

Program bantuan sosial banyak membantu masyarakat kurang mampu dan juga masyarakat yang terkena musibah seperti bencana alam, wabah penyakit dan juga pandemi. Selama pandemi Covid-19

¹⁾ bayubaskoro2019@student.unas.ac.id

masyarakat banyak terbantu berkat program bantuan sosial. Namun meskipun begitu penyaluran program bantuan sosial yang diberikan dinilai masih kurang tepat sasaran karena terdapat masyarakat kurang mampu atau penduduk yang lebih patut mendapat bantuan sosial malah tidak mendapat bantuan sosial tersebut.

Salah satu contohnya, beberapa berita melaporkan bahwa sebagian penduduk di DKI Jakarta telah mengembalikan bantuan sosial yang telah diterima. Hal ini dilakukan karena mereka mengetahui bahwa ada penerima bantuan yang seharusnya tidak layak menerima, sementara masih banyak warga yang membutuhkan bantuan sosial tetapi tidak terdaftar sebagai penerima (Pratama & Akbar, 2021).

Program bantuan sosial tidak tepat sasaran dikarenakan beberapa alasan. Pertama, validasi data penduduk kurang mampu tidak berjalan dengan baik sehingga data yang digunakan untuk menentukan penerima bantuan sosial tidak akurat. Kedua, dalam menentukan penerima bantuan sosial masih dilakukan secara konvensional dengan memonitor langsung kediaman keluarga calon penerima bansos yang berarti penyelenggaraan bantuan sosial memerlukan waktu yang lama dan memakan banyak biaya karena bagian panitia harus menyeleksi secara langsung siapa yang berhak menerima bantuan sosial (Hasymi et al., 2021).

Maka dari itu, demi mengatasi permasalahan tersebut lantas diperlukan sistem yang bisa memberitahukan informasi tentang penduduk kurang mampu dan dapat mengelola data penduduk kurang mampu serta dapat mengklasifikasikan data tersebut untuk menentukan prioritas masyarakat sebagai penerima bantuan sosial guna membantu panitia bantuan sosial dalam memilih

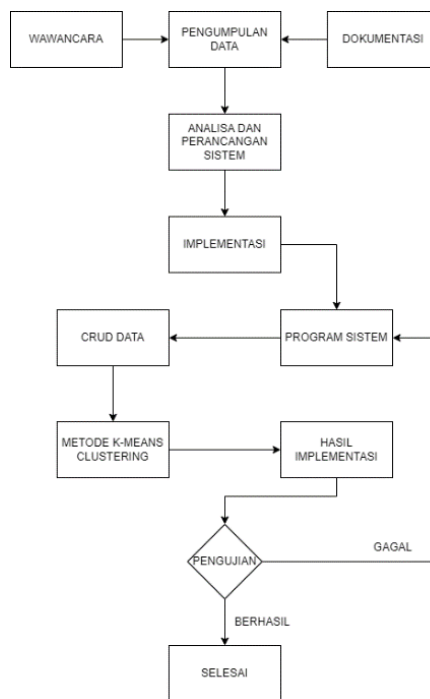
penduduk kurang mampu dan benar-benar layak untuk menerima bantuan sosial.

Adapun tujuan dari penelitian ini antara lain untuk membantu dan mempermudah penyelenggara bantuan sosial dalam memilih penduduk kurang mampu sebagai penerima bantuan sosial dan juga diharapkan mampu membantu pemerintahan daerah dalam mengatasi laju kemiskinan yang terjadi.

II. METODOLOGI

2.1. Tahapan Penelitian

Penelitian umumnya melalui beberapa tahapan yang saling terkait untuk mencapai tujuan penelitian yang diinginkan. Pada penelitian ini, tahapan penelitian dapat dilihat pada Gambar 1.



Gambar 1. Tahapan Penelitian

Dalam tahapan penelitian ini dilakukan pengumpulan dengan metode wawancara yaitu proses tanya jawab secara langsung antara peneliti dan responden kepada

narasumber selain itu juga menggunakan dokumentasi dari dokumen atau catatan-catatan yang sudah ada. Dokumen atau catatan tersebut berupa dokumen mengenai orang atau kelompok dan juga kejadian-kejadian yang terjadi di lingkup sosial (Alhamid & Anufia, 2019).

Analisa dan Perancangan Sistem

Pada tahap ini dilakukan analisa tentang kebutuhan sistem dan juga pengumpulan data terhadap kebutuhan dalam pembuatan sistem seperti perancangan database, perancangan website dan juga transformasi data.

Implementasi

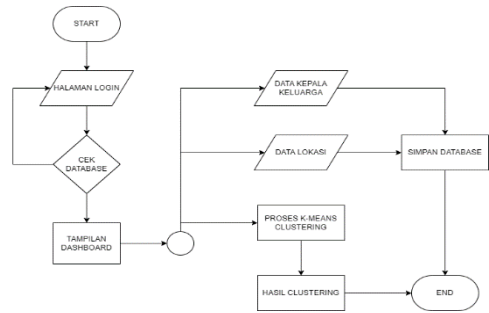
Pada tahap ini yang dilakukan adalah menuliskan program untuk membuat sistem yang berbasis website, penulisan program untuk membuat CRUD data dan juga implementasi metode K-Means Clustering kedalam bahasa pemograman PHP. Metode K-Means ini sering digunakan untuk permasalahan klasteriasi (Novi & Mubarak, 2021)

Pengujian

Tahap ini dilakukan pengujian pada pemrosesan metode K-Means Clustering apakah berjalan dengan baik atau tidak, pada proses pengujian penulis menggunakan software Rapidminer sebagai bahan perbandingan hasil yang didapat pada sistem dengan hasil yang dihasilkan oleh software Rapidminer.

2.2. Perancangan Sistem

Pada gambar 2 merupakan flowchart dari sistem klasifikasi yang akan dibuat.



Gambar 2 Flowchart Sistem

Sistem dimulai dari halaman login. Setelah login, sistem akan mengecek data login di database. Jika data login ada, maka akan diarahkan ke halaman dashboard. Pada halaman dashboard, admin dapat melakukan penginputan data kepala keluarga dan data lokasi. Selain itu, ada juga halaman proses k-means clustering yang digunakan untuk mengelompokkan data kepala keluarga berdasarkan pusat kluster yang ditentukan.

2.3. Transformasi Data

Dalam penelitian ini menggunakan data sample penduduk kurang mampu yang berjumlah 50 data dan menggunakan 4 variabel antara lain pendidikan (PD), pekerjaan (PK), tanggungan (TG) dan pengeluaran perbulan (PG).

Adapun dalam proses perhitungan menggunakan metode K-Means hanya dapat mengelola data bertipe numerik, karena pada sampel data penduduk kurang mampu yang akan digunakan memiliki tipe data char, maka data tersebut akan ditransformasikan atau diinisialisasikan terlebih dahulu menjadi tipe data numerik seperti pada Tabel 1.

Tabel 1 Transformasi Data

Variabel	Kriteria	Inisialisasi
Pendidikan	Tidak bersekolah	1
	SD	2
	SMP	3
	SMA	4
	Tidak Bekerja	1
Pekerjaan	Buruh	2
	Pedagang	3
	Karyawan swasta	4
	0 s/d 1.000.000	1
	1.000.000 s/d 1.500.000	2
Pengeluaran	1.500.000 s/d 2.000.000	3
	2.000.000 s/d 2.500.000	4

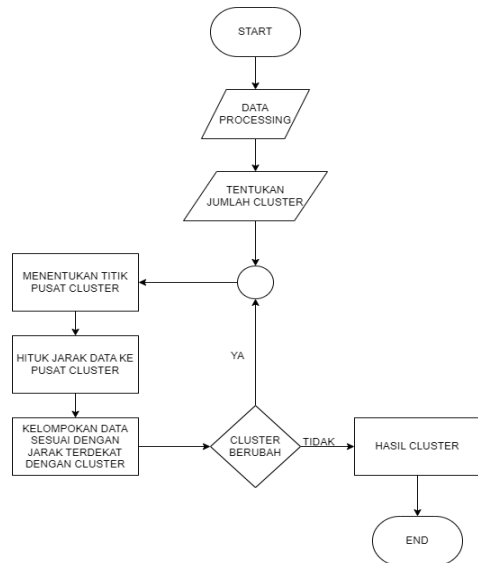
Setelah diinisialisasi maka hasil dari transformasi data akan menjadi seperti pada Tabel 2.

Tabel 2 Hasil Transformasi Data

ID	PD	PK	TG	PG
1	1	2	4	1
2	2	3	3	2
3	3	4	4	3
4	2	3	2	2
5	4	4	4	4
6	3	2	4	4
7	3	4	2	4
8	2	2	3	2
9	3	3	1	3
10	4	4	2	3
11	1	3	3	1
12	1	3	2	2
13	2	2	3	3
14	2	3	2	2
15	3	3	3	3
16	2	2	4	4
17	2	3	3	4
18	4	4	2	4
19	3	3	1	3

ID	PD	PK	TG	PG
20	3	4	3	4
21	1	2	3	1
22	2	2	3	2
23	3	3	3	2
24	2	3	4	2
25	3	3	4	3

2.4. Flowchart K-Means



Gambar 3 Flowchart K-Means

K-Means merupakan metode pengelompokan data numerik berdasarkan jarak antar data, yang membagi data ke dalam kelompok dengan jumlah tertentu. K-Means sangat populer karena kemampuannya untuk mengelompokan data dalam jumlah besar dan cepat. Selama proses clustering, setiap data akan masuk ke kelompok tertentu dan kemungkinan besar dapat pindah ke kelompok lain pada proses selanjutnya (Nasari et al., 2016).

Pada gambar 3 merupakan flowchart dari metode k-means, adapun algoritma dari metode k-means adalah sebagai berikut :

- 1) Tentukan jumlah cluster

- 2) Tentukan pusat centeroid awal, pada dasarnya tahap ini dipilih dari data secara random
- 3) Hitung jarak terdekat dari setiap data dengan centeroid awal yang sudah ditentukan menggunakan rumus *Euclidean Distance*.

$$d(x_i, x_j) = \sqrt{\sum_{l=1}^n (x_{il} - y_{il})^2} \quad (1)$$

- 4) Menentukan jarak terdekat antara data dengan centeroid.
- 5) Tentukan centeroid baru dengan menghitung rata-rata dari setiap cluster dengan rumus

$$C_{kl} = \frac{x_{1l} + x_{2l} + x_{3l} + \dots + x_{pl}}{p} \quad (2)$$

- 6) Tempatkan data kedalam cluster yang sesuai dengan jarak terdekat dengan cluster. Proses akan selesai jika objek tidak berpindah ke cluster lain, namun jika objek masih berpindah ke cluster lain, maka langkah ke-3 akan diulang kembali.

III. HASIL DAN PEMBAHASAN

3.1. Perhitungan K-Means

Dalam penelitian ini data yang digunakan akan dikelompokkan kedalam 3 cluster dengan pusat masing-masing centeroid seperti pada tabel 3.

Tabel 3 Pusat Centeroid

Cluster	PD	PK	TG	PG
1	2	3	3	2
2	2	3	2	2
3	1	2	4	1

Tahapan selanjutnya yaitu menghitung jarak setiap data dengan cluster menggunakan rumus *Euclidean Distance*.

Jarak data ke-1 dengan pusat cluster 1 :

$$\sqrt{(1+2)^2 + (2+3)^2 + (4+3)^2 + (1+2)^2} = 2,00$$

Jarak data ke-1 dengan pusat cluster 2 :

$$\sqrt{(1+2)^2 + (2+3)^2 + (4+2)^2 + (1+2)^2} = 2,65$$

Jarak data ke-1 dengan pusat cluster 3 :

$$\sqrt{(1+1)^2 + (2+2)^2 + (4+4)^2 + (1+1)^2} = 0,00$$

Setelah diketahui jarak antara data dengan masing-masing cluster, didapatkan hasil iterasi 1 seperti pada tabel 4.

Tabel 4 Hasil Iterasi 1

Data-i	Jarak ke -			Cluster
	C1	C2	C3	
1	2,00	2,65	0,00	C3
2	0,00	1,00	2,00	C1
3	1,00	0,00	2,65	C2
4	2,65	3,46	3,61	C1
5	3,16	3,61	4,69	C1
6	2,65	3,16	3,61	C1
7	2,65	2,45	4,58	C1
8	1,00	1,41	1,73	C1
9	2,45	1,73	4,24	C2
10	2,65	2,45	4,58	C2
11	1,41	1,73	1,41	C1
12	1,41	1,00	2,45	C2
13	1,41	1,73	2,45	C1
14	1,00	0,00	2,65	C2
15	1,41	1,73	3,16	C1
16	2,45	3,00	3,16	C1
17	2,00	2,24	3,46	C1
18	3,16	3,00	5,10	C2
19	2,45	2,65	4,24	C2
20	2,45	2,65	4,24	C1
21	1,73	2,00	1,00	C3
22	1,00	1,41	1,73	C1
23	1,00	1,41	1,73	C1
24	1,00	2,00	1,73	C1
25	1,73	2,45	3,00	C1

Setelah proses perhitungan dan pengelompokan sesuai dengan cluster maka selanjutnya mencari pusat cluster baru dengan cara menghitung rata-rata dari tiap anggota cluster, maka hasilnya akan seperti pada Tabel 5.

Tabel 5 Pusat Centeroid Baru

Cluster	PD	PK	TG	PG
1	2,47	2,87	3,47	2,87
2	2,75	3,38	1,75	2,88
3	1	2	3,50	1

Selanjutnya menghitung kembali jarak data dengan centeroid menggunakan rumus Euclidean Distance seperti proses sebelumnya. Jika data tidak berpindah cluster maka proses selesai, namun jika masih berpindah cluster ulangi langkah-langkah seperti sebelumnya.

Pada penelitian ini, iterasi berhasil dilakukan sebanyak 8 kali, hasil akhir pada perhitungan iterasi ke-8 dapat dilihat pada Tabel 6.

Tabel 6. Perhitungan Jarak Terdekat dan Cluster

ID	PD	PK	TG	PG	Jarak terdekat	Cluster
1	1	2	4	1	1.66	C3
2	2	3	3	2	0.52	C3
3	3	4	5	3	1.05	C3
4	2	3	2	2	1.65	C1
5	4	4	4	4	1.49	C1
6	3	2	4	4	1.22	C1
7	2	4	2	4	0.92	C2
8	2	2	3	2	0.66	C3
9	3	3	1	3	1.02	C3
10	4	4	2	3	0.93	C2
11	1	3	3	1	1.20	C3
12	1	3	2	2	1.27	C3
13	2	2	3	3	1.33	C3
14	2	3	2	2	1.05	C3
15	3	3	3	3	0.99	C1
16	2	2	4	4	1.49	C1
17	2	3	3	4	1.22	C1
18	4	4	2	4	1.02	C2
19	3	3	1	3	1.02	C2
20	3	4	3	4	1.22	C1
21	1	2	3	1	1.27	C3
22	2	2	3	2	0.66	C3

ID	PD	PK	TG	PG	Jarak terdekat	Cluster
23	3	3	3	2	1.33	C3
24	2	3	4	2	1.20	C3
25	3	3	4	3	0.70	C1

3.2. Analisa Hasil

Dari pemrosesan K-Means Clustering pada 50 data sample yang digunakan, proses berhenti pada iterasi ke-8 dan didapat 3 kelompok atau cluster yaitu cluster 1 dengan 10 anggota cluster, cluster 2 dengan 16 anggota cluster dan cluster 3 dengan 24 anggota cluster.

Tabel 7 Hasil Pembagian Cluster

Cluster 1	Cluster 2	Cluster 3
10 Keluarga	16 Keluarga	24 Keluarga

Cluster 1 dengan pusat cluster (2-3-3-2) yang dimaksudkan bahwa anggota dalam cluster ini memiliki kepala keluarga dengan pendidikan terakhir antara SMP-SMA, pekerjaannya rata-rata karyawan swasta dengan tanggungan rata-rata 4 anggota keluarga dan pengeluaran rata-rata perbulannya sekitar Rp. 1.650.000 s/d Rp. 2.500.000,- yang dapat diartikan dalam cluster ini merupakan kelompok yang kurang layak mendapat bantuan sosial.

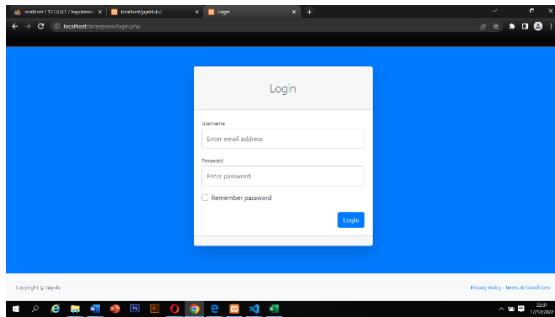
Cluster 2 dengan pusat cluster (2-3-2-2) yang dalam cluster ini beranggotakan kepala keluarga dengan pendidikan terakhir SMP-SMA, pekerjaannya rata-rata sebagai pedagang dengan tanggungan rata-rata 2 anggota keluarga dan pengeluaran perbulan sekitar Rp. 1.150.000 s/d Rp. 2.000.000,- yang dapat diartikan dalam cluster ini merupakan kelompok yang layak menerima bantuan sosial.

Cluster 3 dengan pusat cluster (1-2-4-1) dalam cluster ini beranggotakan kepala keluarga dengan pendidikan terakhir rata-rata tamat SD, pekerjaan rata-rata sebagai buruh dengan

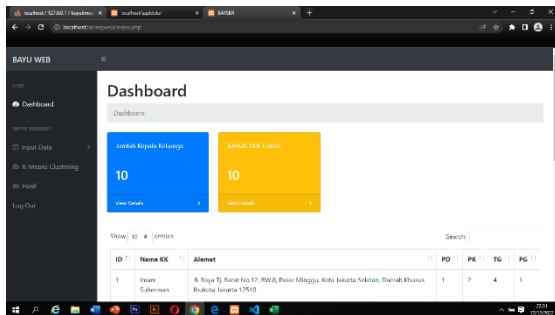
tanggungannya 3 anggota keluarga dan pengeluaran perbulan rata-rata Rp. 850.000 s/d Rp. 1.000.000,- yang berarti pada cluster ini merupakan kelompok yang sangat layak dan menjadi prioritas untuk menerima bantuan sosial.

3.3. Implementasi Sistem

Gambar 4 merupakan implementasi sistem pada halaman login. User dalam sistem ini hanya ada admin untuk mengelola data.

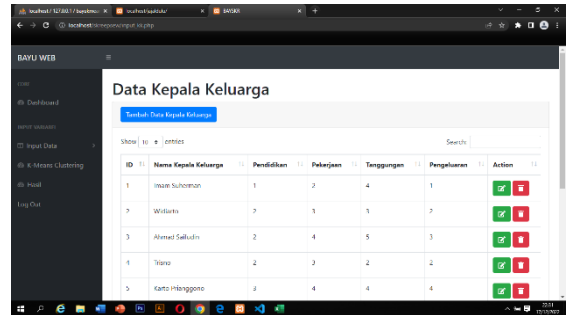


Gambar 4 Implementasi Halaman Login



Gambar 5 Halaman Dashboard

Pada halaman dashboard ditampilkan beberapa informasi seperti jumlah data kepala keluarga yang ada dalam database dan juga titik lokasi yang disimpan. Dalam halaman dashboard juga menampilkan tabel berupa data kepala keluarga.

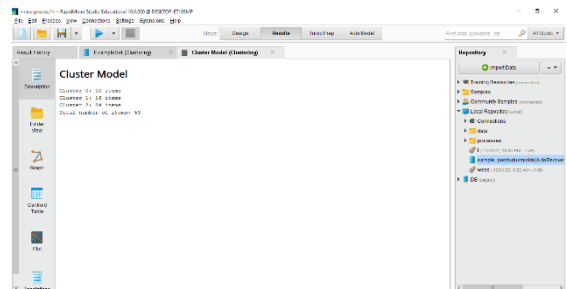


Gambar 6 Halaman Input Data

Pada gambar 6 merupakan tampilan dari halaman untuk mengelola data seperti menginput, edit dan delete.

3.4. Pengujian

Pada tahap pengujian, peneliti menggunakan software Rapidminer sebagai membandingkan hasil yang diperoleh dari sistem yang sudah dibangun.



Gambar 7 Pengujian dengan Rapidminer

Dari Gambar 7 diketahui bahwa jumlah anggota dari masing-masing cluster sama dengan proses yang dilakukan pada sistem yang dibuat dan juga pada pengujian performance didapatkan skor DBI (*Davies Bouldin Index*) dengan nilai skor 0.996.

V. KESIMPULAN

Berdasarkan kajian ini, proses pengklasifikasian penerima bantuan sosial dengan menggunakan metode K-Means telah menghasilkan 3 cluster yaitu penduduk yang kurang layak mendapat bantuan, layak mendapat bantuan dan sangat layak

mendapat bantuan sehingga dapat mengetahui penduduk mana yang akan menjadi prioritas sebagai penerima bantuan sosial. Kriteria yang digunakan dalam penelitian, seperti pendidikan, pekerjaan, tanggungan, dan pengeluaran per bulan, telah diproses oleh algoritma K-Means untuk menentukan penerima bantuan. Hasil dari proses K-Means telah menyediakan alat yang berguna untuk pengelolaan program bantuan sosial, membantu memastikan bahwa bantuan didistribusikan secara adil dan efektif. Namun, penting untuk dicatat bahwa hasil penelitian ini harus dievaluasi lebih lanjut dan ditingkatkan, dan faktor-faktor tambahan mungkin perlu dipertimbangkan untuk mengoptimalkan proses pengklasifikasian penerima bantuan sosial.

DAFTAR PUSTAKA

- Alhamid, T., & Anufia, B. (2019). *RESUME: INSTRUMEN PENGUMPULAN DATA*.
- Hasymi, M. A., Faisol, A., & Ariwibisono, F. X. (2021). SISTEM INFORMASI GEOGRAFIS PEMETAAN WARGA KURANG MAMPU DI KELURAHAN KARANGBESUKI MENGGUNAKAN METODE K-MEANS CLUSTERING. In *Jurnal Mahasiswa Teknik Informatika* (Vol. 5, Issue 1).
- Nasari, F., Jhony, C., & Sianturi, M. (2016). *Penerapan Algoritma K-Means Clustering Untuk Pengelompokan Penyebaran Diare Di Kabupaten Langkat*.
- Novi, N. & Mubarak, A. (2021). Penerapan Algoritma K-Means Clustering untuk Menentukan Kelas Unggulan di SMP Pelita Bandung, *Infomatek: Jurnal Informatika, Manajemen dan Teknologi*, 23(2), 97-106.
- Pratama, B., & Akbar, A. S. (2021). KLASIFIKASI PENENTUAN WARGA PENERIMA BANTUAN SOSIAL DI MASA PANDEMI MENGGUNAKAN METODE NAIVE BAYES. *Jurnal Nasional Informatika*, 2(2), 143–157.
- Sepriyanti, N., Sani Nahampun, R., Zikri, M. H., Ambarani, I., & Rahmadayan, A. (2022). *Penerapan K-Means Clustering Untuk Mengelompokkan Tingkat Kemiskinan di Provinsi Riau*. <https://journal.irpi.or.id/index.php/sentimas>
- Seran, S. (2012). DETERMINAN FAKTOR SOSIAL DAN EKONOMI TERHADAP KEMISKINAN PENDUDUK. *Jurnal Ekonomi Pembangunan*, 13(1), 62–78.