

PENERAPAN ALGORITMA K-MEANS PADA KLASTERISASI DATA PENERIMA PKH DI KECAMATAN MEDAN TIMUR

Reihan Hidayat Putra¹, M Fakhriza²

^{1,2} Universitas Islam Negeri Sumatera Utara, Medan, Indonesia

Email : reeihaan02@gmail.com

Abstrak

Dalam era digital dan teknologi informasi saat ini, pengelolaan data menjadi krusial untuk mendukung efektivitas program pemerintah, terutama dalam penyaluran bantuan sosial. Program Keluarga Harapan (PKH) merupakan salah satu program yang signifikan dalam upaya penanggulangan kemiskinan di Indonesia. Di Kantor Kecamatan Medan Timur, ketersediaan data penerima PKH menjadi kunci untuk mengoptimalkan penyaluran bantuan dan memastikan pemerataan. Namun, pengolahan data mengalami kesulitan dalam menentukan daerah dengan peningkatan atau penurunan penerima bantuan karena perubahan jumlah penduduk, penerima bantuan, dan anggaran setiap tahun. Untuk mengatasi masalah ini, data mining dengan proses klasterisasi diterapkan untuk mengamati perkembangan kemiskinan di Kecamatan Medan Timur. Klasterisasi data, yang merupakan teknik analisis data untuk mengelompokkan data berdasarkan kesamaan karakteristik, menjadi penting untuk memahami pola-pola penerima manfaat PKH. Algoritma K-Means digunakan untuk membagi data penerima PKH ke dalam beberapa klaster, dengan tujuan meminimalkan varians intra-klaster dan memaksimalkan varians antar-klaster. Penerapan algoritma K-Means diharapkan dapat memberikan pemahaman mendalam tentang pola-pola atau kelompok penerima bantuan di wilayah tersebut, sehingga dapat menjadi dasar dalam menyusun strategi yang lebih efektif untuk memahami perkembangan setiap kelurahan di Kecamatan Medan Timur.

Kata Kunci: Program Keluarga Harapan, Klasterisasi Data, K-Means, Bantuan Sosial, Sistem Informasi.

Abstract

In the current era of digital and information technology, data management has become crucial to support the effectiveness of government programs, especially in the distribution of social assistance. The Family Hope Program (PKH) is a significant initiative in the effort to alleviate poverty in Indonesia. At the Medan Timur District Office, the availability of PKH recipient data is key to optimizing aid distribution and ensuring equity. However, data processing faces difficulties in determining areas with increased or decreased aid recipients due to changes in population, aid recipients, and budget allocations each year. To address this issue, data mining with clustering processes is applied to observe poverty trends in the Medan Timur District. Data clustering, a data analysis technique for grouping data based on similar characteristics, is important for understanding patterns among PKH beneficiaries. The K-Means algorithm is used to divide PKH recipient data into several clusters, aiming to minimize intra-cluster variance and maximize inter-cluster variance. The application of the K-Means algorithm is expected to provide deeper insights into patterns or groups of aid recipients in the region, thereby serving as a foundation for developing more effective strategies to understand the development of each sub-district in the Medan Timur District.

Keywords: Family Hope Program, Data Clustering, K-Means Algorithm, Social Assistance Distribution, Information System.



Pendahuluan

Dalam era digital dan teknologi informasi saat ini, pengelolaan data menjadi aspek yang sangat penting dalam mendukung efektivitas program-program pemerintah, terutama dalam hal penyaluran bantuan sosial kepada masyarakat yang membutuhkan. Salah satu program yang memiliki dampak signifikan dalam upaya penanggulangan kemiskinan di Indonesia adalah Program Keluarga Harapan (PKH). Di Kantor Kecamatan Medan Timur, sebagai bagian dari upaya pemerintah dalam mengelola program tersebut, ketersediaan data penerima PKH menjadi kunci dalam mengoptimalkan penyaluran bantuan dan memastikan bahwa bantuan yang diberikan menyebar rata [1].

Pada Kantor Kecamatan Medan Timur terdapat kerumitan dalam pengolahan data selama ini yaitu menentukan daerah yang mengalami peningkatan atau penurunan penerima bantuan di antara sekian banyak data penduduk miskin setiap kelurahan di Kecamatan Medan Timur, Kesulitan itu muncul akibat adanya perubahan jumlah penduduk, jumlah penerima bantuan, dan jumlah anggaran yang diberikan setiap tahun. Dengan mencermati permasalahan di atas data mining menggunakan proses clustering untuk mengamati perkembangan kemiskinan setiap wilayah di Kecamatan Medan Timur, pentingnya masing-masing parameter tersebut menjadi tolak ukur untuk memilih populasi yang akan menjadi pertimbangan untuk menerima bantuan Program Keluarga Harapan (PKH).

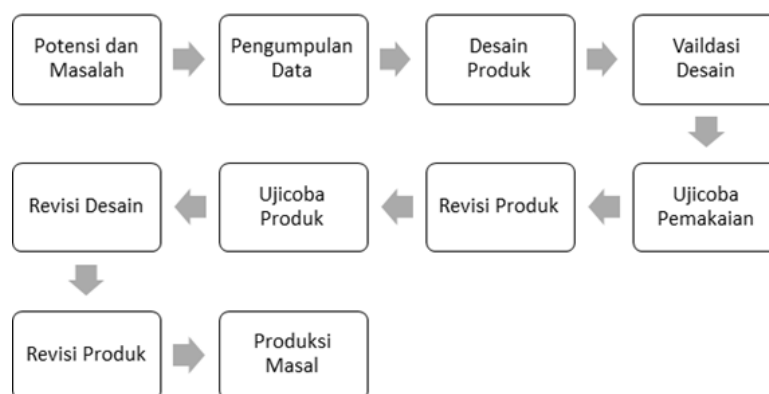
Klasterisasi data merupakan salah satu teknik dalam analisis data yang bertujuan untuk mengelompokkan data ke dalam kelompok-kelompok atau klaster-klaster berdasarkan kesamaan karakteristik tertentu [2]. Pada konteks penerima Program Keluarga Harapan (PKH) di Kantor Kecamatan Medan Timur, klasterisasi data menjadi penting untuk memahami pola-pola yang mungkin ada di antara penerima manfaat PKH. Salah satu algoritma yang sering digunakan dalam klasterisasi data adalah algoritma K-Means. Algoritma ini bekerja dengan cara membagi data ke dalam K klaster, di mana K merupakan jumlah klaster yang diinginkan [3]. Tujuan utama dari algoritma ini adalah untuk meminimalkan varians intra-klaster, yaitu varians antara titik data dalam satu klaster, dan memaksimalkan varians antara klaster [4].

Dengan menerapkan algoritma K-Means pada data penerima PKH di Kantor Kecamatan Medan Timur, diharapkan dapat menghasilkan pemahaman yang lebih mendalam tentang pola-pola atau kelompok-kelompok yang ada di antara wilayah penerima bantuan tersebut. Hal ini diharapkan dapat menjadi landasan dalam menyusun strategi yang lebih efektif dalam memahami perkembangan setiap kelurahan.

Metodologi Penelitian

Metode yang diterapkan dalam penelitian ini adalah menggunakan metode *Research and Development*. Metode penelitian dan pengembangan atau *Research and Development* (R&D) adalah metode penelitian yang digunakan untuk menghasilkan produk tertentu, dan menguji keefektifan produk tersebut. Untuk dapat menghasilkan produk tertentu digunakan penelitian yang bersifat analisis kebutuhan untuk menguji keefektifan produk supaya dapat berfungsi di masyarakat luas, maka diperlukan penelitian untuk menguji keefektifan produk [5].

Dalam melakukan penelitian menggunakan metode R&D (*Research and Development*), metode ini memiliki langkah-langkah sebagai membantu penulis untuk pengerjaan penelitian ini. Adapun langkah-langkah metode R&D yaitu sebagai berikut :



Gambar 1. Metode R&D.

Hasil dan Pembahasan

1. Algoritma K-Means

Algoritma K- Means merupakan salah satu algoritma dengan kemampuan melakukan partisi pada klaster yang diinginkan dengan melakukan pendefinisian centroid terlebih dahulu [6]. Digunakannya metode ini dengan alasan bahwa pengelompokan dataset harus dilakukan dengan membentuk kelompok–kelompok atau kelas–kelas taksonomi [7].

Berikut proses penerapan algoritma K-Means:

- 1) Tentukan k sebagai jumlah cluster yang akan dibentuk.
- 2) Tentukan k titik pusat cluster (centroid) awal yang dilakukan secara random. Penentuan pada centroid awal dilakukan secara acak atau random dari objek yang tersedia sebanyak k cluster, untuk menghitung centroid cluster ke-i berikutnya, digunakan rumus sebagai berikut:

$$v = \frac{\sum_{i=1}^n x_i}{n} : i = 1, 2, 3, \dots, n \quad (2)$$

Dimana v: Centroid pada Cluster

xi: Objek Ke-i

n: banyaknya objek pada anggota cluster

- 3) Menghitung jarak dari setiap objek ke masing–masing centroid dari masing–masing cluster menggunakan Euclidean Distance, dengan rumus sebagai berikut:

$$d(x, y) = ||x - y|| = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (3)$$

Keterangan:

xi: objek x ke i

yi: daya y ke I

n: banyaknya objek

- 4) Alokasikan dari masing–masing objek kedalam centroid paling dekat. Untuk pengalokasian objek kedalam masing–masing cluster pada saat iterasi secara umum dilakukan dengan cara hard k-means dimana secara tegas setiap objek dinyatakan sebagai anggota cluster dengan mengukur jarak kedekatan sifatnya terhadap titik pusat cluster tersebut.
- 5) Lakukan iterasi dan kemudian tentukan posisi centroid baru dengan menggunakan persamaan.
- 6) Ulangi langkah nomor 3 jika posisi centroid baru tidak sama [8].

2. Penerapan Algoritma K-Means pada Klasterisasi Data Penerima PKH di Kecamatan Medan Timur

Penerapan algoritma K-Means pada klasterisasi data penerima Program Keluarga Harapan (PKH) di Kecamatan Medan Timur bertujuan untuk mengelompokkan penerima PKH berdasarkan karakteristik tertentu sehingga pemerintah atau lembaga terkait dapat membuat keputusan yang lebih tepat dalam mendistribusikan bantuan [9]. Terdapat 11 kelurahan di kecamatan medan timur. Berikut adalah data yang diambil beserta deskripsi umum mengenai penerapan algoritma K-Means pada klasterisasi data tersebut:

Tabel 1. Data Penerima PKH Tahun 2021.

No	Penerima	Anggaran	Penduduk
1	137	164.975.000	11022
2	318	384.275.000	10351
3	3	3.575.000	6237
4	198	238.300.000	9931
5	255	308.675.000	11104
6	391	476.550.000	10579

7	487	589.125.000	14037
8	268	332.700.000	11813
9	302	367.900.000	9217
10	68	84.950.000	4822
11	30	34.275.000	7140

Tabel 2. Data Penerima PKH Tahun 2022

No	Penerima	Anggaran	Penduduk
1	112	134.848.000	10083
2	342	411.768.000	9789
3	9	10.836.000	4455
4	120	144.480.000	13054
5	149	179.396.000	11539
6	240	288.960.000	15491
7	302	363.608.000	20785
8	155	186.620.000	15508
9	259	311.836.000	11558
10	20	24.080.000	4227
11	9	10.836.000	6372

- 1) Menentukan jumlah cluster
Pada penelitian ini penulis menggunakan 3 cluster. Maka dari itu nilai $K = 3$.
- 2) Tentukan k titik cluster atau centeroid awal

Tabel 3. Centeroid Awal Data Tahun 2021.

No	centeroid	X	Y	Z
2	1	318	384.275.000	10351
4	2	198	238.300.000	9931
6	3	391	476.550.000	10579

Tabel 4. Centeroid Awal Data Tahun 2022.

No	centeroid	X	Y	Z
2	1	342	411.768.000	9789
4	2	120	144.480.000	13054
6	3	240	288.960.000	15491

- 3) Menghitung jarak dari setiap objek ke masing–masing centroid dari masing–masing cluster

Tabel 5. Jarak Cluster Data Tahun 2021

	C1	C2	C3
219300000	73325000,01	311575000	
0	145975000	92275000	
380700000	234725000	472975000	
145975000	0	238250000	

75600000	70375000,01	167875000
92275000	238250000	0
204850000	350825000	112575000,1
51575000,02	94400000,02	143850000
16375000,04	129600000	108650000
299325000,1	153350000,1	391600000
350000000	204025000	442275000

4) Alokasikan dari masing–masing objek kedalam centroid paling dekat

Tabel 6. Hasil Cluster

Minimum	Cluster
73325000,01	2
0	1
234725000	2
0	2
70375000,01	2
0	3
112575000,1	3
51575000,02	1
16375000,04	1
153350000,1	2
204025000	2

5) Lakukan iterasi dan kemudian tentukan posisi centroid baru dengan menggunakan persamaan

Tabel 7. Hasil Akhir Iterasi data 2021

No	Penerima	Anggran	Penduduk	C1	C2	C3	Minimum	Cluster
1	137	164.975.000	11022	183412500	59760000,09	367862500	59760000,09	2
2	318	384.275.000	10351	35887500	279060000	148562500	35887500	1
3	3	3.575.000	6237	344812500	101640000	529262500	101640000	2
4	198	238.300.000	9931	110087500	133085000	294537500	110087500	1
5	255	308.675.000	11104	39712500	203460000	224162500	39712500	1
6	391	476.550.000	10579	128162500	371335000	56287500,03	56287500,03	3
7	487	589.125.000	14037	240737500	483910000	56287500,03	56287500,03	3
8	268	332.700.000	11813	15687500,05	227485000	200137500	15687500,05	1
9	302	367.900.000	9217	19512500,05	262685000	164937500	19512500,05	1
10	68	84.950.000	4822	263437500,1	20265000,22	447887500,1	20265000,22	2
11	30	34.275.000	7140	314112500	70940000	498562500	70940000	2

Tabel 8. Hasil Akhir Iterasi data 2022

No	Penerima	Anggaran	Penduduk	C1	C2	C3	Minimum	Cluster
1	112	134.848.000	10083	252840000,1	36120000,01	165550000	36120000	2
2	342	411.768.000	9789	24080000,63	313040000	111370000,1	24080001	1
3	9	10.836.000	4455	376852000,2	87892000,13	289562000,1	87892000	2
4	120	144.480.000	13054	243208000	45752000,15	155918000	45752000	2

5	149	179.396.000	11539	208292000	80668000,03	121002000	80668000	2
6	240	288.960.000	15491	98728000	190232000,1	11438000,17	11438000	3
7	302	363.608.000	20785	24080000,63	264880000,2	63210000,42	24080001	1
8	155	186.620.000	15508	201068000	87892000,22	113778000	87892000	2
9	259	311.836.000	11558	75852000,09	213108000	11438000,17	11438000	3
10	20	24.080.000	4227	363608000,2	74648000,17	276318000,2	74648000	2
11	9	10.836.000	6372	376852000,1	87892000,05	289562000,1	87892000	2

Berdasarkan hasil clustering, maka didapat cluster data tahun 2021 dan tahun 2022 dapat dilihat di table 9.

Tabel 9. Hasil Clustering Data

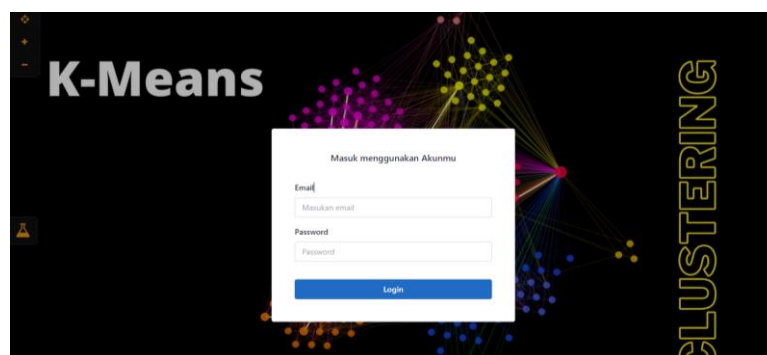
No	Kelurahan	2021	2022	Status
1	DURIAN	Cluster 2	Cluster 2	Tetap
2	GAHARU	Cluster 1	Cluster 1	Tetap
3	GANG BUNTU	Cluster 2	Cluster 2	Tetap
4	GLUGUR DARAT I	Cluster 1	Cluster 2	Meningkat
5	GLUGUR DARAT II	Cluster 1	Cluster 2	Meningkat
6	PULO BERAYAN BENGKEL	Cluster 3	Cluster 3	Tetap
7	PULO BERAYAN DARAT I	Cluster 3	Cluster 1	Menurun
8	PULO BERAYAN DARAT II	Cluster 1	Cluster 2	Meningkat
9	PULO BERAYAN BENGKEL BARU	Cluster 1	Cluster 3	Meningkat
10	PERINTIS	Cluster 2	Cluster 2	Tetap
11	SIDODADI	Cluster 2	Cluster 2	Tetap

3. Implementasi

Implementasi adalah fase di mana rancangan sistem diwujudkan menjadi sebuah aplikasi atau perangkat lunak yang fungsional [10]. Pada tahap ini, tampilan antarmuka pengguna dengan berbagai elemen dan tata letaknya direalisasikan sesuai dengan perancangan, di mana setiap elemen tidak hanya terlihat secara visual tetapi juga menjalankan fungsi-fungsi spesifik yang telah direncanakan [11].

a. Tampilan *Form Login*

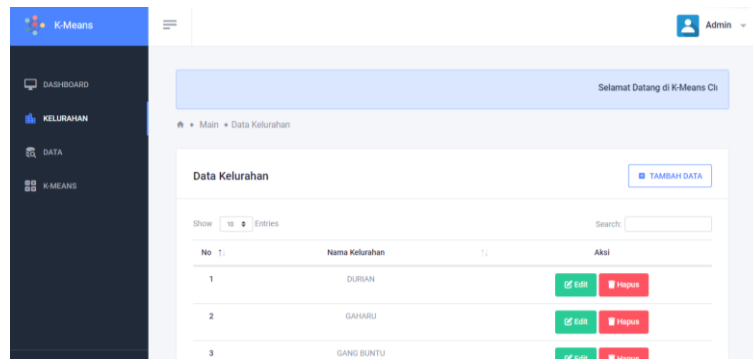
Form login adalah *form* yang berfungsi sebagai gerbang autentikasi untuk mengidentifikasi dan memverifikasi pengguna yang sah. Dalam implementasinya, form login melibatkan beberapa komponen utama: field untuk nama pengguna dan kata sandi, tombol untuk mengirimkan data.



Gambar 2. *Form Login*

b. Tampilan Menu Kelurahan

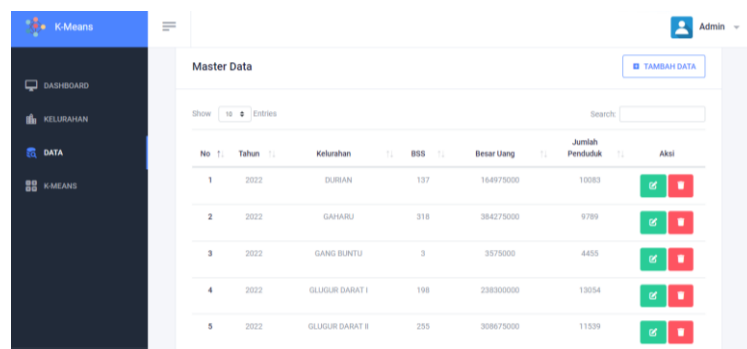
Menu Kelurahan adalah menu yang berisi data kelurahan yang ada di kecamatan medan timur, admin dapat menambah, mengubah. Serta menghapus data pada halaman ini.



Gambar 3. Tampilan Menu Kelurahan

c. Tampilan Menu Data

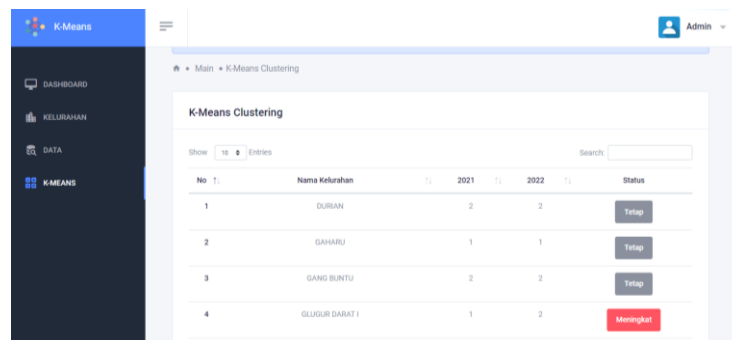
Menu Data adalah menu yang berisi data penerima PKH masing-masing kelurahan, menu ini terintegrasi dengan data kelurahan, admin dapat menambah, mengubah dan menghapus data di halaman ini.



Gambar 4. Tampilan Menu Data

d. Tampilan Menu K-Means

Menu K-Means adalah menu yang menampilkan hasil dari perhitungan algoritma K-Means, menu ini menampilkan hasil perkembangan masing-masing kelurahan di kecamatan medan timur.



Gambar 5. K-Means

Kesimpulan

Hasil penelitian menunjukkan bahwa penerapan algoritma K-Means berhasil memberikan pemahaman mendalam tentang pola-pola atau kelompok penerima bantuan di wilayah tersebut. Penemuan ini dapat menjadi dasar dalam menyusun strategi yang lebih efektif untuk memahami perkembangan setiap kelurahan di Kecamatan Medan Timur, sehingga mendukung penyaluran bantuan yang lebih tepat sasaran dan efisien. Dengan klusterisasi ini, lembaga terkait dapat membuat strategi yang lebih efektif dalam mendistribusikan bantuan.

Daftar Pustaka

[1] G. Gustientiedina, M. H. Adiya and Y. Desnelita, "Penerapan Algoritma K-Means untuk Clustering Data Obat-Obatan," *Jurnal Nasional Teknologi Dan Sistem Informasi*, vol. 5, no. 1, pp.

- 17-24, 2019.
- [2] S. M. Sinaga, J. T. Hardinata and M. Fauzan, "Implementasi Data Mining Clustering Tingkat Kepuasan Konsumen Terhadap Pelayanan Go-Jek," *Kesatria: Jurnal Penerapan Sistem Informasi (Komputer dan Manajemen)*, vol. 2, no. 2, pp. 118-124, 2021.
 - [3] S. Astuti, S. samsudin and T. triase, "Penerapan Data Mining Dalam Menentukan Penerima Beasiswa UPZ (Unit Pengumpulan Zakat) Menggunakan Algoritma K-Means," *JSI: Jurnal Sistem Informasi (E-Journal)*, vol. 13, no. 2, pp. 2288-2297, 2021.
 - [4] A. Solichin and K. Khairunnisa, "Klasterisasi persebaran virus Corona (Covid-19) di DKI Jakarta menggunakan metode K-Means," *Fountain of Informatics Journal*, vol. 5, no. 2, pp. 52-59, 2020.
 - [5] Sugiyono, *Metode Penelitian Pendidikan (Pendekatan Kuantitatif, Kualitatif, dan R&D)*, Bandung: CV. Alfabeta, 2017.
 - [6] D. L. Aulia and R. A. Putri, "Sistem Informasi Monitoring Industri Kecil Menengah menggunakan Algoritma K-Means pada Dinas Perindustrian Perdagangan PROVSU Berbasis Website," *ZONAsi: Jurnal Sistem Informasi*, vol. 5, no. 3, pp. 406-418, 2023.
 - [7] Y. Mardi, "Data mining: Klasifikasi menggunakan algoritma c4. 5," *Jurnal Edik Informatika Penelitian Bidang Komputer Sains dan Pendidikan Informatika*, vol. 2, no. 2, pp. 213-219, 2017.
 - [8] A. Purwanto, A. Primajaya and A. Voutama, "Penerapan Algoritma C4. 5 dalam Prediksi Potensi Tingkat Kasus Pneumonia di Kabupaten Karawang," *JUSTIN (Jurnal Sistem dan Teknologi Informasi)*, vol. 8, no. 4, pp. 390-396, 2020.
 - [9] M. R. Ramadhan, Y. R. Nasution and A. M. Harahap, "Management Information System Implementation Doyan Drinks Using K-Means," *Journal of Applied Engineering and Technological Science (JAETS)*, vol. 4, no. 1, pp. 490-504, 2022.
 - [10] Y. R. Nasution and M. Eka, "Penerapan Algoritma K-Means Clustering Pada Aplikasi Menentukan Berat Badan Ideal," *ALGORITMA: Jurnal Ilmu Komputer Dan Informatika*, vol. 2, no. 1, pp. 77-81, 2018.
 - [11] M. R. Fahrezi and M. Fakhriza, "Sistem Pendukung Keputusan Pemilihan Lokasi Camping berdasarkan Kesulitan Track di Sumatera Utara dengan Menggunakan Metode Topsis," *JISTech (Journal of Islamic Science and Technology)*, vol. 8, no. 2, pp. 116-133, 2023.