

ANALISIS DATA PROGRAM KELUARGA HARAPAN PADA TINGKAT KEMISKINAN DENGAN MENGGUNAKAN FUZZY C-MEANS DI KECAMATAN PABEDILAN

Titan Septian Nugraha, Rudi Kurniawan, Yudhistira Arie Wijaya

Teknik Informatika, STMIK IKMI CIREBON

Jl. Perjuangan No. 10 B Majasem Kec. Kesambi Kota Cirebon

Titannugraha21@gmail.com

ABSTRAK

Kemiskinan, sebagai isu mendasar di Indonesia, tetap menjadi tantangan yang belum teratasi sepenuhnya oleh pemerintah pusat maupun daerah. Oleh karena itu penelitian ini bertujuan untuk menganalisis data guna mengurangi kemiskinan dan meningkatkan kesejahteraan masyarakat dengan kondisi ekonomi rendah di Kecamatan Pabedilan. Fokusnya adalah Program Keluarga Harapan (PKH), sebuah inisiatif pemerintah Indonesia untuk memberikan bantuan sosial bersyarat kepada Keluarga Miskin (KM) yang diidentifikasi sebagai penerima manfaat PKH. Penelitian ini menggunakan Metode Algoritma Fuzzy C-Means sebagai teknik pengelompokan data. Metode ini terpilih karena sederhana namun efektif dalam mengatasi atribut dengan variasi nilai yang beragam. Dengan memfokuskan analisis pada Kecamatan Pabedilan, penelitian ini diharapkan dapat memastikan bahwa Program Keluarga Harapan memberikan manfaat optimal bagi masyarakat sasaran, serta memberikan kontribusi positif dalam upaya pengurangan kemiskinan di wilayah tersebut. Teknik kluster dipilih sebagai solusi untuk menghadapi variasi nilai atribut, menunjukkan keunggulan dalam konsistensi, akurasi, dan efisiensi waktu komputasi. Hasil penelitian ini Algoritma Fuzzy C-Means pada data penerima PKH Kecamatan Pabedilan, ditemukan tiga cluster, Cluster 0: 473 items Karakteristik nominal: 633,333-1,600,000 (masuk dalam nominal besar). Cluster 1: 656 items Karakteristik nominal: 18,666-250,000 (masuk dalam nominal kecil). Cluster 2: 1,572 items Karakteristik nominal: 300,000-583,333 (masuk dalam nominal sedang).

Kata kunci : Bansos, PKH, Data mining, Fuzzy C-Means, Kemiskinan

1. PENDAHULUAN

Program Keluarga Harapan (PKH) merupakan inisiatif pemerintah Indonesia untuk memberikan bantuan sosial bersyarat kepada Keluarga Miskin (KM) yang telah diidentifikasi sebagai penerima manfaat PKH. Dalam upaya percepatan penanggulangan kemiskinan, Pemerintah Indonesia telah mengimplementasikan program ini Program Keluarga Harapan (PKH) sejak tahun 2007. Program perlindungan sosial ini, yang juga dikenal di dunia internasional dengan istilah Conditional Cash Transfers (CCT), Misi besar PKH untuk menurunkan tingkat kemiskinan semakin mendesak, mengingat bahwa pada Maret 2016, Persentase penduduk miskin di Indonesia saat ini mencapai 10,86% dari keseluruhan populasi, setara dengan sekitar 28,01 juta orang (BPS 2016).

Kemiskinan merupakan salah satu isu mendasar yang menjadi fokus perhatian suatu negara, baik di perkotaan maupun di desa. [1]. Di Indonesia sendiri kemiskinan telah menjadi sebuah fenomena dan fakta, salah satu masalah yang sejak dulu hingga sekarang masih juga belum bisa teratasi baik oleh pemerintah pusat maupun oleh pemerintah daerah [2]. Kemiskinan dapat diartikan sebagai kondisi seseorang yang tidak mampu memenuhi kebutuhan pokoknya [3]. Sehingga, pemerintah berusaha memberikan bantuan melalui Program Keluarga Harapan (PKH) yang bertujuan untuk meningkatkan kesejahteraan masyarakat. Program Keluarga

Harapan adalah inisiatif pemberian uang tunai dengan syarat bagi keluarga yang berada dalam kondisi sangat miskin, memenuhi kriteria peserta, dan telah ditetapkan oleh Kementerian Sosial [4].

Sejumlah penelitian yang memanfaatkan Metode Algoritma Fuzzy C-Means [5] Dalam menerapkan analisis data, peneliti seringkali dihadapkan pada sejumlah atribut dengan variasi nilai yang beragam. Peneliti memilih teknik kluster sebagai solusi untuk mengatasi permasalahan yang muncul, Dikarenakan teknik ini menunjukkan keunggulan dalam konsistensi, akurasi, dan efisiensi waktu komputasi. Metode kluster dipilih sebagai solusi karena mampu meningkatkan efektivitas identifikasi sel-sel tumor dan mengukur area tumor pada gambar otak melalui proses segmentasi citra MRI otak.

Penelitian ini bertujuan untuk menganalisis data untuk mengurangi kemiskinan dan meningkatkan kesejahteraan masyarakat yang berada dalam kondisi ekonomi rendah di Kecamatan Pabedilan dan penting untuk memastikan bahwa program Hal ini dapat memberikan manfaat optimal bagi masyarakat sasaran dan memberikan kontribusi positif terhadap upaya pengurangan kemiskinan di wilayah Kecamatan Pabedilan. Metode Algoritma Fuzzy C-Means merupakan salah satu teknik yang sering diterapkan dan relatif sederhana dalam pengelompokan data karena mampu menghasilkan

estimasi yang efektif tanpa memerlukan parameter yang berlebihan. [6].

Hasil dari pengolahan data PKH yaitu Pemetaan wilayah berdasarkan tingkat pengangguran terbuka dan tingkat kemiskinan dapat memberikan informasi yang bermanfaat bagi pemerintah untuk memahami permasalahan dan mendukung proses pengambilan kebijakan di tingkat wilayah [7].

2. TINJAUAN PUSTAKA

2.1. Data mining

Data mining merupakan suatu proses penambangan informasi krusial dari suatu kumpulan data. Penggalan informasi yang signifikan ini melibatkan proses yang kompleks, seperti pemanfaatan kecerdasan buatan, teknik statistik, ilmu matematika, machine learning, dan metode lainnya. Teknik-teknik canggih ini digunakan untuk mengidentifikasi dan mengekstraksi informasi yang bernilai dari database yang besar. Disiplin ilmu data mining telah berkembang dalam ranah kecerdasan buatan (AI) dan rekayasa pengetahuan (KE), memiliki akar dalam machine learning dan statistika, dan merambah ke berbagai bidang ilmu komputer serta disiplin ilmu lainnya seperti biologi, lingkungan, keuangan, jaringan, dan sebagainya [8].

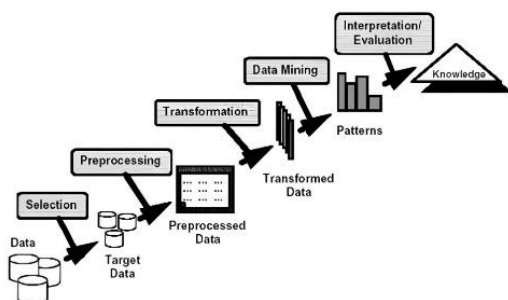
2.2. Analisis

Analisis cluster adalah suatu cabang dalam statistik analisis multivariat yang digunakan secara khusus untuk mengelompokkan objek yang memiliki karakteristik serupa [9].

2.3. Clustering

Clustering, yang merujuk pada pengelompokan atau pembagian data, adalah salah satu metode dalam data mining yang digunakan untuk mengelompokkan sejumlah besar data. Tujuannya adalah untuk mengamati kemiripan antara nilai-nilai yang bersifat homogen dari setiap data. Proses clustering melibatkan pengamatan terhadap objek-objek dan data yang ingin dikelompokkan dari setiap data. Dalam proses clustering, data dibantu menggunakan pendekatan algoritma yang bertujuan untuk menyajikan informasi lebih rinci dan memfokuskan pada target pengelompokan data dalam suatu sistem matematis yang terstruktur [10].

2.4. Fuzzy C-Means



Fuzzy C-Means Clustering (FCM), atau dikenal juga sebagai Fuzzy Isodata merupakan salah satu metode clustering yang merupakan bagian dari metode Hard K-Means. FCM menggunakan model pengelompokan fuzzy sehingga data dapat menjadi anggota dari semua kelas atau cluster terbentuk dengan derajat atau tingkat keanggotaan yang berbeda antara 0 hingga 1. Tingkat keberadaan data dalam suatu kelas atau cluster ditentukan oleh derajat keanggotaannya [11]. Algoritma FCM menjadi bagian integral dari teknik klaster. FCM digunakan untuk mengelompokkan data berdasarkan rentang nilai 0 hingga 1, di mana nilai keanggotaan menentukan sejauh mana data termasuk dalam suatu klaster. Semakin tinggi nilai keanggotaan, semakin tinggi derajat keanggotaannya, dan sebaliknya, semakin kecil nilai keanggotaan, semakin rendah derajat keanggotaannya. FCM merupakan varian Fuzzy dari K-Means yang mengalami beberapa modifikasi untuk membedakannya dari K-Means [12].

2.1. Cluster model from data

Tabel 1. Hasil cluster model from data

No.	Cluster	Jumlah items
1.	Cluster 0	473
2.	Cluster 1	656
	Cluster 2	1572
Total number of items : 2701		

Setelah melakukan proses klasterisasi pada data uji menggunakan Algoritma Fuzzy C-Means dengan optimalisasi Cluster Distance Performance, hasilnya terdiri dari 3 klaster. Klaster 0 memiliki 473 item, Klaster 1 memiliki 656 item, dan Klaster 2 memiliki 1572 item, dari total 2701 dataset. Dalam konteks nominal penerima PKH, karakteristik tiap klaster adalah sebagai berikut: Klaster 0, berdasarkan nominal yang diperoleh, termasuk dalam kategori besar (633333-1600000); Klaster 2, berdasarkan nominal yang diperoleh, termasuk dalam kategori sedang (300000-583333); dan Klaster 1, berdasarkan nominal yang diperoleh, termasuk dalam kategori kecil (18666-250000).

Jika *heading* anda melebihi satu, gunakan level kedua *heading* seperti di contoh dibawah. *Heading* dituliskan dengan **boldface** dengan menggunakan huruf besar. **Pengutipan/sitasi WAJIB menggunakan Mendeley Style IEEE** contoh [1], atau jika lebih dari satu contoh [2] [3] [4], Untuk referensi gunakan journal atau proceeding, minimal 10 buah.

3. METODE PENELITIAN

Dalam melakukan analisis data PKH yang akan di teliti sesuai pemodelan data yang akan Dalam melakukan analisis data PKH yang akan diteliti sesuai pemodelan data yang akan diambil agar mempermudah penelitian dan berjalan sesuai di

inginkan maka dibuat alur atau tahapan dalam penelitian ini sebagai berikut:

3.1. Selection

Selection merupakan suatu tahap di mana informasi diambil, dikumpulkan, dipilih, dan diperiksa. Pada fase ini, data yang digunakan sebagai materi penelitian berasal dari tahun 2023, kemudian dipilih atribut sesuai dengan kebutuhan.

3.2. Preprocessing

Preprocessing merupakan bagian dari fase persiapan informasi sebelum melibatkan proses data mining. Dalam tahap preprocessing, data yang digunakan umumnya melewati beberapa sesi, seperti cleaning, reduction, dan integration. Sebelum melakukan proses prediksi, perlu dilakukan proses preprocessing terlebih dahulu, yang mencakup penghapusan data yang tidak relevan, penanganan data yang hilang, dan menangani data yang missing.

3.3. Transformasi

Dalam tahap ini, dilakukan transformasi data ke dalam format yang sesuai untuk diproses oleh algoritma dan alat data mining. Tujuannya adalah untuk mempermudah koordinasi data yang akan diolah oleh algoritma dan perangkat yang digunakan dalam penelitian.

3.4. Data mining

Data Mining merupakan proses pengolahan data yang mengikuti algoritma dengan konsistensi sesuai alur data mining. Algoritma yang digunakan dalam penelitian ini adalah Regresi Linear, yang tergolong dalam kategori algoritma untuk peramalan atau prediksi.

3.5. Evaluation

Tahap evaluasi digunakan untuk mengevaluasi hasil dari Fuzzy C-Means guna mendapatkan kesimpulan dari proses data mining. Penilaian performa dilakukan dengan menggunakan DBI (Davies Bouldin Index).

3.6. Knowledge

Tahapan Knowledge ini menampilkan data yang dihasilkan dari proses prediksi. Proses ini bertujuan untuk menyajikan hasil sehingga informasinya menjadi lebih mudah dipahami.

4. HASIL DAN PEMBAHASAN

4.1. Selection

Langkah pertama adalah mencari lokasi file Itu sebelumnya dibuat dalam format *xlsx* dengan operator *Read Excel* yang berfungsi baca data *xlsx*. Kemudian menggunakan operator *Set Role* Untuk mengidentifikasi id dalam dataset, sebagaimana terlihat dalam gambar di bawah ini.



Gambar 1. Operator Read Excel

Parameter operator *Set Role* yang digunakan sebagaimana terlihat dalam tabel di bawah ini.



Gambar 2. Operator Set Role

Tahap selection pada model proses dalam RapidMiner sebagaimana terlihat dalam gambar di bawah ini.



Gambar 3. Model proses tahap Selection di Rapidminer

4.2. Pre-processing

Pada tahap pre-processing atau proses pembersihan data, yang mencakup penanganan nilai yang missing atau tidak konsisten, dilakukan setelah melakukan analisis terhadap atribut dataset yang dipilih. Sebelumnya, dilakukan evaluasi untuk memastikan apakah atribut tersebut mengandung nilai yang hilang atau tidak konsisten. Hasil dari analisis statistik dataset, seperti yang terlihat pada gambar 4 di bawah ini, menunjukkan bahwa tidak ada atribut yang memiliki nilai yang missing. Selain itu, pemeriksaan langsung per-record menunjukkan bahwa dataset ini konsisten dalam hal nilai-nilainya.

NO	Integer	0	47379	278831	178763.655
NOMINAL	Integer	0	18966	1800000	417960.691
KODE_WILAYAH	Nominal	0	3209042011 (125)	3209042008 (309)	3209042003 (263) ... [1 more]
ALAMAT	Nominal	0	Dusun Karanganyar (1)	DUSUN WAGI (263)	DUSUN WAGI (263), DUSUN WAGI (258) ... [51 more]
NO_RT	Nominal	0	29 (16)	3 (520)	3 (520), 1 (519) ... [19 more]
NO_RW	Nominal	0	9 (1)	2 (647)	2 (647), 1 (616) ... [9 more]
DUSUN	Nominal	0	9 (1)	- (2672)	- (2672), MANIS (6) ... [15 more]
DIR	Nominal	0	Dir Dapason (370)	Dir JS (1477)	Dir JS (1477), Dir Retusa (854) ... [1 more]
GELOMBANG	Nominal	0	QEL 7 (370)	QEL 3 (1477)	QEL 3 (1477), QEL 1 (854) ... [1 more]

Gambar 4. Hasil dari Analisis Statistik Dataset

Kemudian menggunakan operator *Select Attributes*, Operator ini memilih subset Atribut dari ExampleSet dan menghapus atribut lainnya, sebagaimana terlihat dalam gambar di bawah ini.



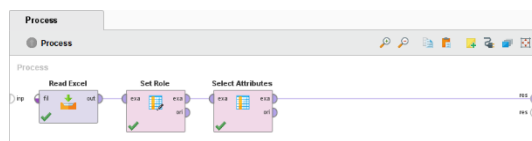
Gambar 5. Operator Select Attributes

Dari hasil pembacaan operator *Select Attributes* didapat informasi sebagai berikut.

Tabel 2. Hasil Select Attributes

No.	Uraian	Keterangan
1.	Record	2701
2.	Special Attribute	2
3.	Regular Attribute	5
4.	Attribute :	
	NO	Integer
	Cluser	Numeric
	nominal	Integer
	Kode_wilayah	Numeric
	Alamat	Numeric
	Dir	Numeric
	gelombang	Numeric

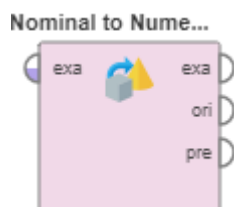
Tahap pre-processing pada model proses dalam RapidMiner sebagaimana terlihat dalam Gambar 6 di bawah ini



Gambar 6. Model proses tahap Pre-processing di Rapidminer

4.3. Transformation

Dikarenakan dataset ada yang menggunakan tipe data nominal sementara algoritma Fuzzy C-Means memerlukan tipe data numeric, maka transformasi diperlukan untuk mengubah data nominal menjadi Numerical. Hal ini dapat dilakukan dengan menggunakan operator *Nominal to Numerical*, sebagaimana terlihat dalam gambar di bawah ini.



Gambar 7. Operator Nominal to Numerical

Parameter dan Seselect Attributes pada operator *Nominal to Numerical* yang digunakan sebagaimana terlihat dalam Tabel 3 di bawah ini.

Tabel 3. Parameter Nominal to Numerical

No.	Parameter	Isi
1.	attribute filter type	subset
2.	coding type	unique integers

Hasil dari penggunaan operator *Nominal to Numerical* sebagaimana terlihat dalam gambar di bawah ini.

Tabel 4. Hasil atribut yang di pilih Nominal to Numerical

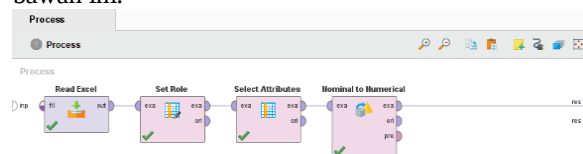
No.	Parameter	Isi
1.	attribut filter type	subset
2.	attributes	Select Attributes ...
		Kode wilayah
		Alamat
		Dir
		Gelombang

Hasil dari penggunaan operator *Nominal to Numerical* tampak pada gambar dibawah ini.

KODE_WILAYAH	Numeric	0	Min	0	Max	12	Average	5.968
ALAMAT	Numeric	0	Min	0	Max	52	Average	19.802
DIR	Numeric	0	Min	0	Max	2	Average	0.821
GELOMBANG	Numeric	0	Min	0	Max	2	Average	0.821
NOMINAL	Integer	0	Min	19565	Max	1600000	Average	417960.691

Gambar 8. Hasil Nominal to Numerical

Tahap transformation pada model proses dalam RapidMiner sebagaimana terlihat dalam gambar di bawah ini.



Gambar 9. Model proses tahap Transformation di Rapidminer

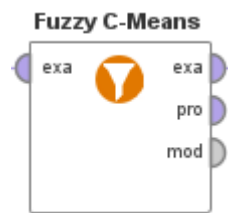
4.4. Data Mining

Penentuan jumlah cluster adalah langkah yang harus dilakukan. Nilai K yang digunakan melibatkan identifikasi jumlah kluster optimal serta jumlah anggota optimal dalam setiap kluster. Pengelompokan data dilakukan dengan menggunakan algoritma Fuzzy C-Means dengan Cluster Distance Performance dan di optimasi dengan Davies Bouldin Index pada Tools RapidMiner versi 10.3.

Desain model proses hingga tahap Data Mining sebagaimana terlihat dalam gambar di bawah ini.



Gambar 10. Desain Fuzzy C-Means



Gambar 11. Operator Fuzzy C-Means

pengelompokan (clustering) yang digunakan dalam analisis data. FCM adalah metode klasterisasi fuzzy yang memungkinkan suatu objek data masuk ke lebih dari satu kelompok dengan tingkat keanggotaan yang berbeda-beda. Ini berbeda dengan pendekatan klasterisasi klasik yang bersifat keras (hard clustering), di mana suatu objek hanya dapat termasuk dalam satu kelompok. Fuzzy C-Means digunakan untuk membentuk model klaster fuzzy dari data. Operator ini memungkinkan pengguna untuk menentukan jumlah klaster yang diinginkan dan menghasilkan kelompok-kelompok yang memiliki tingkat keanggotaan untuk setiap objek data. FCM bekerja dengan mengoptimalkan pusat klaster dan tingkat keanggotaan setiap objek terhadap setiap klaster.

5. KESIMPULAN DAN SARAN

Berdasarkan hasil implementasi Algoritma Fuzzy C-Means pada data penerima PKH Kecamatan Pabedilan, ditemukan tiga cluster dengan jumlah anggota masing-masing sebagai berikut: Cluster 0: 473 items Karakteristik nominal: 633,333-1,600,000 (masuk dalam nominal besar). Cluster 1: 656 items Karakteristik nominal: 18,666-250,000 (masuk dalam nominal kecil). Cluster 2: 1,572 items Karakteristik nominal: 300,000-583,333 (masuk dalam nominal sedang).

Berdasarkan hasil penelitian ini menyadari adanya kekurangan pada penelitian ini, karena keterbatasan baik dalam perihal waktu atau pengetahuan. Dalam rangka memperbaiki kekurangan serta guna penyempurnaan penelitian ini maka memberikan beberapa saran sebagai berikut: Penelitian ini dapat di kembangkan lagi dengan menggunakan algoritma *forecasting* yang lainnya, seperti *Neural Network*, *Naïve Bayes*, *Decision Tree*, dll. Pada penelitian selanjutnya bisa menggunakan dataset dengan periode yang lebih banyak.

DAFTAR PUSTAKA

- [1] D. Bahtiar *et al.*, "Pemetaan Penduduk Penerima Bantuan Sosial Desa Waru Jaya Menggunakan Algoritma K-Means Clustering," *Sci. Sacra J. Sains*, vol. 3, no. 2, pp. 29–39, 2023, [Online]. Available: <http://pijarpemikiran.com/index.php/Scientia>
- [2] A. Indriyanti, A. Nugroho, and I. Romli, "Penerapan Algoritma Fuzzy C-Means Untuk Pengelompokan Data Penduduk Miskin Di Indonesia Berdasarkan Kabupaten dan Kota," *Pros. Sains dan Teknol.*, vol. 1, no. 1, pp. 141–146, 2022.
- [3] S. Abdul Azis, S. Defit, and Y. Yunus, "Klasterisasi Dana Bantuan Pada Program Keluarga Harapan (PKH) Menggunakan Metode K-Means," *J. Inform. Ekon. Bisnis*, vol. 3, no. 2, pp. 53–59, 2021, doi: 10.37034/infkeb.v3i2.66.
- [4] F. I. Putri, R. Damayanti, and Kismiantini, "Penerapan Algoritma K-Means Untuk Mengelompokan Kecamatan Di Kabupaten Gunungkidul Berdasarkan Program Keluarga Harapan," *Pros. Semin. Nas. Mat. Stat. dan Apl.*, vol. 2, pp. 408–418, 2022.
- [5] V. Herlinda, D. Darwis, and Dartono, "Analisis Clustering Untuk Recrederialing Fasilitas Kesehatan Menggunakan Metode Fuzzy C-Means," *JTSI J. Teknol. dan Sist. Inf.*, vol. 2, no. 2, pp. 94–99, 2021, [Online]. Available: <http://jim.teknokrat.ac.id/index.php/JTSI>
- [6] A. Surya Maulana, A. Nazir, L. Handayani, and I. Afrianty, "Kajian Ilmiah Informatika dan Komputer Penerapan Algoritma Fuzzy C-Means untuk Melihat Pola Penerima Beasiswa Bank Indonesia," *Media Online*, vol. 3, no. 6, pp. 670–679, 2023, doi: 10.30865/klik.v3i6.788.
- [7] A. E. Wardoyo and N. Triuspita, "Penentuan Cluster Optimum pada Tingkat Pengangguran dan Tingkat Kemiskinan di Jawa Timur Menggunakan Algoritma Fuzzy C-Means," *BIOS J. Teknol. Inf. dan Rekayasa Komput.*, vol. 1, no. 2, pp. 40–47, 2021, doi: 10.37148/bios.v1i2.10.
- [8] B. G. Sudarsono, M. I. Leo, A. Santoso, and F. Hendrawan, "Analisis Data Mining Data Netflix Menggunakan Aplikasi Rapid Miner," *JBASE - J. Bus. Audit Inf. Syst.*, vol. 4, no. 1, pp. 13–21, 2021, doi: 10.30813/jbase.v4i1.2729.
- [9] R. A. Ningtyas, Y. N. Nasution, and Syaripuddin, "Pengelompokan Kabupaten/Kota Di Pulau Kalimantan Dengan Fuzzy C-Means Berdasarkan Indikator Kemiskinan," *J. EKSPONENSIAL*, vol. 13, no. 2, pp. 141–146, 2022.
- [10] B. G. Sudarsono and S. P. Lestari, "Clustering Penerima Beasiswa Yayasan Untuk Mahasiswa Menggunakan Metode K-Means," *J. Media Inform. Budidarma*, vol. 5, no. 1, p. 258, 2021, doi: 10.30865/mib.v5i1.2670.
- [11] D. L. Rahakbauw, V. Y. I. Ilwaru, and M. H. Hahury, "Implementasi Fuzzy C-Means Clustering Dalam Penentuan Beasiswa," *BAREKENG J. Ilmu Mat. dan Terap.*, vol. 11, no. 1, pp. 1–12, 2017, doi: 10.30598/barekengvol11iss1pp1-12.
- [12] A. Yudhistira, A. A. Aldino, and D. Darwis, "Analisis Klasterisasi Penilaian Kinerja Pegawai Menggunakan Metode Fuzzy C-Means (Studi Kasus: Pengadilan Tinggi Agama bandar lampung)," *J. Ilm. Edutic Pendidik. dan Inform.*, vol. 9, no. 1, pp. 77–82, 2022, doi: 10.21107/edutic.v9i1.17134.