



Klasifikasi Pencari Kerja pada Disnaker Menggunakan Metode K-Means Clustering

M. Aqshal Al Fachrizy, Hendri*

Fakultas Sains Dan Teknologi, Program Studi Sistem Komputer, Universitas Pembangunan Panca Budi, Medan, Indonesia

Email: ¹maqshalalfachrizy@email.com, ²*hendry@dosen.pancabudi.ac.id

Email Korespondensi: hendry@dosen.pancabudi.ac.id

Abstrak—Dinas Tenaga Kerja (Disnaker) merupakan institusi pemerintah yang bertugas mendukung, mengontrol, dan mengawasi sektor ketenagakerjaan. Selain itu, Disnaker juga bertanggung jawab dalam memberikan pelatihan keterampilan khusus kepada calon karyawan sesuai dengan kebutuhan pasar kerja serta memberikan akses yang luas terhadap kesempatan kerja. Penelitian ini bertujuan untuk melakukan pengelompokan data pencari kerja berdasarkan tingkat pendidikan di wilayah Medan menggunakan Algoritma K-Means melalui pendekatan metode clustering dengan penerapan Teknik Data Mining menggunakan perangkat lunak RapidMiner guna memperoleh data yang akurat dan relevan. Hasil implementasi Algoritma K-Means pada data pencari kerja di wilayah Medan menunjukkan terbentuknya 6 kelompok (k6) dengan nilai Davies Bouldin Index (DBI) sebesar 0.185. Penelitian ini menunjukkan bahwa penggunaan cluster (k6) sebagai k optimal memberikan nilai DBI yang terbaik, yaitu 0.185, menandakan tingkat kemiripan data dalam setiap kelompok semakin berdekatan.

Kata Kunci: K-Means; Clustering; Rapid Miner; Pencari Kerja; Klasifikasi

Abstract—The Manpower Office (Disnaker) is a government institution tasked with supporting, controlling, and supervising the employment sector. In addition, the Manpower Office is also responsible for providing specialized skills training to prospective employees to the needs of the job market and providing broad access to employment opportunities. This research aims to cluster job seeker data based on education level in the Medan area using the K-Means Algorithm through a clustering method approach with the application of Data Mining Techniques using RapidMiner software to obtain accurate and relevant data. The results of the implementation of the K-Means Algorithm on job seeker data in the Medan area show the formation of 6 groups (k6) with a Davies Bouldin Index (DBI) value of 0.185. This study shows that the use of cluster (k6) as the optimal k provides the best DBI value, which is 0.185, indicating the level of similarity of data in each group is getting closer.

Keywords: K-Means; Clustering; Rapid Miner; Worker; Classification

1. PENDAHULUAN

Undang undang Nomor 7 Tahun 1981 tentang wajib lapor ketenagakerjaan (Lembaran Negara Republik Indonesia Tahun 1981 Nomor 39, Tambahan Lembaran Negara Republik Indonesia Nomor 3201) setiap pemberi kerja di daerah wajib menyampaikan informasi lowongan pekerjaan kepada Dinas Tenaga Kerja dan Transmigrasi dan Instansi yang bertanggung jawab di bidang ketenagakerjaan Kabupaten / Kota. K- Means clustering merupakan algoritma yang mudah dan sederhana. Algoritma Clustering memiliki daya tarik yang luas dan digunakan untuk analisis data eksplorasi [1]. Tahapan tahapan Data Mining yaitu, Pembersihan data, Integrasi data, seleksi data, transformasi data, proses mining, Evaluasi Pola dan presentasi pengetahuan[2]. Peningkatan algoritma Clustering K-Means menggunakan waktu yang lebih baik dan akurat. Analisis cluster merupakan pengelompokan satu set objek dengan benda yang sama yang lebih mirip satu sama lain [3].

Penggunaan Data Mining dalam pengelompokkan para pencari kerja terdaftar dikarenakan Data Mining sudah banyak diterapkan dalam berbagai bidang. Salah satunya metode clustering khususnya metode K-Means. Data Mining. Hasil pengolahan dengan Data Mining berupa sebuah pengetahuan (knowledge) baru yaitu pengelompokan data jumlah pencari kerja terdaftar, sehingga dapat. K-Means clustering merupakan salah satu metode data clustering non-hirarki yang mengelompokkan data dalam bentuk satu atau lebih cluster/kelompok. Data-data yang memiliki karakteristik yang sama dikelompokkan dalam satu cluster/kelompok dan data yang memiliki karakteristik yang berbeda dikelompokkan dengan cluster/kelompok yang lain sehingga data yang berada dalam satu cluster/kelompok memiliki tingkat variasi yang kecil [4].

Data Mining merupakan bagian dari proses Knowledge Discovery in Database (KDD) Dimana proses tersebut mengacu pada proses yang luas untuk menemukan pengetahuan dalam data dan menekankan aplikasi khusus pada metode Data Mining [5]. Pengelompokan (clustering) merupakan bagian dari ilmu Data Mining yang bersifat tanpa arahan (unsupervised). Clustering adalah proses pembagian data ke dalam kelas atau cluster berdasarkan tingkat kesamaannya[6]. Dinas Tenaga Kerja (Disnaker) adalah lembaga pemerintah yang mendukung, mengontrol dan mengawasi sektor ketenagakerjaan dan memberikan pelatihan keterampilan khusus kepada calon karyawan seperti yang dipersyaratkan oleh para pencari kerja dan memberikan kesempatan kerja yang luas, meningkatkan peluang perekrutan layanan dan memberikan informasi pasar tenaga kerja dan pasar tenaga kerja.

Irfan Sofi, 2023 melakukan penelitian, dan terungkap bahwa BLT Desa terutama dilaksanakan di daerah-daerah dengan tingkat kejadian anggaran yang tinggi. Temuan lain termasuk perbedaan antara jumlah BLT Desa dan KPM Desa, inkonsistensi antara DTKS Harga non-tunai tidak terlalu tinggi di kepulauan dan daerah pegunungan[7]. Ada penelitian lain yang dilakukan oleh Eti Yonika, Sri Ritno, 2019 mengatakan Data tentang populasi dapat diklasifikasikan dengan menggunakan algoritma Classification and Regression (Cart) dengan melihat data yang melebihi tingkat kesesuaian tertinggi berdasarkan data tentang populasi sebelumnya. Namun dari segi pemekaran calon dan cabang total, sangat



bermasalah jika data yang diungkapkan dalam jumlah yang cukup besar. [8]. Pada penelitian lain yang dilakukan oleh S Ramadani dkk, 2019 Dari 20 data yang didistribusikan di antara ketiga kelompok tersebut, Klaster 1 memiliki enam data untuk penjualan, Klaster 2 berisi delapan untuk kemiskinan, dan Klaster 3 berisi enam untuk kemiskinan. Klaster 2 adalah grup yang paling sering digunakan dan ada 8 Data Kemiskinan (1,37, 3,5, dan 3). Jelas bahwa klaster 2 memiliki konsentrasi kemiskinan yang tinggi dalam Jumlah Tanggungan = 1 (0-1) dengan Ijazah Tertinggi Suami = 3 (SMP/Sederajat), dan Lapangan Usaha Dari Pekerjaan Utama Seorang Suami = 4. (Perikanan Tangkap) [9].

Menurut Badan Pusat Statistik (BPS) Provinsi Sumatera Utara mencatat ada sebanyak 1,23 juta orang di Sumut yang terdampak Covid-19. Dari jumlah tersebut, terbanyak atau mencapai sekitar 1,02 juta merupakan kalangan pekerja yang mengalami pengurangan jam kerja. Jumlah ini terdiri atas pengangguran akibat Covid-19 sebanyak 107.000 orang. Kemudian bukan angkatan kerja karena pandemi sebesar 39.000 orang, sedangkan yang tidak bekerja 64.000 orang. Tingkat pengangguran terbuka (TPT) Agustus 2020 juga meningkat sebesar 1,52 persen atau sebanyak 508.000 orang. Penduduk usia kerja didefinisikan sebagai penduduk yang berusia 15 tahun ke atas. Itu adalah pekerjaan, bukan pekerjaan. Proporsi pendukung yang ikut dalam angkatan kerja adalah pendukung yang melakukan kegiatan ekonomi. Pencari kerja berdasarkan tingkat pendidikan merupakan peran utama dalam dunia kerja. Pendidikan merupakan kunci utama untuk mencari pekerjaan yang lebih cocok bagi para pencari kerja profesional, agar tidak menimbulkan kegagalan dalam bidang pekerjaan. Provinsi Sumatera Utara masih kekurangan kualitas lapangan kerja, menjadi peningkatan pertumbuhan perekonomian yang tidak dapat tercapai sehingga terjadi pekerja yang gagal dan tidak profesional.

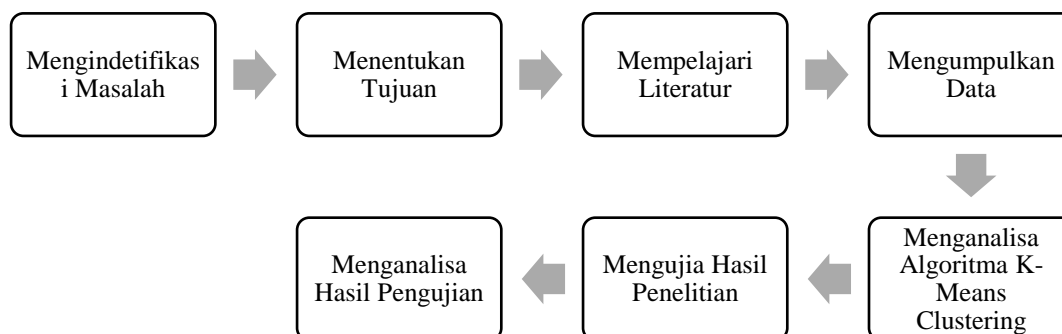
Tujuan penelitian ini bertujuan untuk mengelompokkan data pencari kerja berdasarkan tingkat Pendidikan khususnya di wilayah Medan menggunakan Algoritma K-Means melalui metode clustering dengan menerapkan Data Mining menggunakan software RapidMiner agar menghasilkan data yang akurat dan relevan.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Clustering K-Means adalah teknik data mining yang digunakan dalam penelitian ini. Proses pembersihan, integrasi, seleksi, transformasi, dan penggalian termasuk pengujian pola dan penampilan hasil. K-means clustering adalah teknik clustering data non-hirarki yang mengelompokkan data dalam satu atau lebih cluster atau kelompok. Satu cluster atau kelompok terdiri dari data dengan karakteristik yang sama, dan kelompok lain memiliki karakteristik yang berbeda, sehingga ada tingkat variasi yang kecil antara data yang tergabung dalam satu cluster atau kelompok. Ilmu Data Mining menggunakan clustering. Clustering adalah pembagian data ke dalam kelas atau cluster berdasarkan tingkat kesamaannya. Gambar 1 menunjukkan langkah-langkah yang dilakukan.

Data mining adalah proses untuk menemukan pola dan teknik statistik matematika, kecerdasan buatan, dan pembelajaran mesin untuk mengekstrak dan mengidentifikasi pengetahuan dari database [10][11][12]. Analisis Klaster mengklasifikasi objek sehingga setiap objek yang paling dekat kesamaannya dengan objek lain berada dalam klaster yang sama [13][14][15][16]. Salah satu metode analisis klaster non-hirarki yang dikenal sebagai klasifikasi K-Means bertujuan untuk membagi semua objek yang ada ke dalam satu atau lebih klaster atau kelompok objek berdasarkan atributnya. Dengan demikian, objek dengan atribut yang sama dikelompokkan dalam satu klaster, dan objek dengan atribut yang berbeda dikelompokkan dalam klaster yang berbeda [17][18][19][20].



Gambar 1. Tahapan Penelitian

Pada tahapan Penelitian yang terlihat pada gambar 1 dapat dijelaskan:

1. Identifikasi

Tujuan penelitian ditentukan melalui identifikasi masalah. Bagaimana memanfaatkan algoritma K-Means untuk menerapkan data mining pencari kerja di Kota Medan adalah masalah yang diidentifikasi dalam penelitian ini.



2. Studi Literatur

Pada tahap ini, analisis dilakukan terhadap kumpulan data yang akan diolah dan digunakan. Studi literatur dilakukan untuk mempelajari data atau sistem yang akan dibangun, berdasarkan referensi jurnal dan penelitian yang telah dilakukan sebelumnya.

3. Pengumpulan Data

Pengumpulan data adalah proses untuk mengumpulkan data pencari kerja yang terdaftar di Dinas Ketenagakerjaan Kota Medan dari tahun 2020 sampai 2023. Data yang dikumpulkan kemudian disortir menjadi data asli dan dibagi menjadi beberapa kategori yang akan di klastering

4. Adapun langkah-langkah pada proses clustering dengan menggunakan algoritma K-Means, adalah sebagai berikut:

a. Menentukan Jumlah Cluster Jumlah cluster yang ditentukan untuk mengelompokan data pada penelitian ini sebanyak 3 cluster. Dengan rumus sebagai berikut:

$$SSE = \sum_k^K = 1 \sum_{X_i} = S_k ||N_i - C_k || \tag{1}$$

b. Menentukan Centroid Pusat awal cluster (centroid) ditentukan secara random atau acak.

$$V = \frac{\sum_{i=1}^n X_i}{n} : i = 1,2,3, \dots \dots n \tag{2}$$

c. Menghitung Jarak dari Centroid Menghitung jarak antara titik centroid dengan titik tiap objek dengan menggunakan Euclidian Distance.

$$d(x, y) = ||x - y|| = \sqrt{\sum_{i=1}^n (X_i - Y_i)^2} : 1 = 1,2,3, \dots \dots n \tag{3}$$

d. Alokasikan masing-masing objek ke centroid terdekat. Untuk mengalokasikan objek ke dalam masing-masing cluster dengan cara mengelompokkan berdasarkan jarak minimum objek ke pusat cluster. Dalam penelitian ada urutan kerangka kerja yang harus diikuti, urutan kerangka kerja ini merupakan langkah-langkah yang dilakukan dalam penulisan.

5. Implementasi

Implementasi adalah tahap penerapan sistem dan pengujian. Saat aplikasi siap digunakan dalam keadaan sebenarnya, semua orang akan tahu apakah sistem baru berfungsi dengan baik.

3. HASIL DAN PEMBAHASAN

3.1 Hasil

1. Data Selection

Pemilihan data adalah langkah pertama dalam proses penyelesaian pengelompokan data pencari kerja. Data ini diambil dari situs web Disnaker pada tahun 2023, dengan atribut seperti nama kabupaten/kota, tingkat pendidikan tertinggi, jumlah pencari kerja, dan tahun. Dataset pencari kerja tersebut disajikan dalam tabel 1 berikut:

Tabel 1. Dataset Awal

Kode provinsi	Nama provinsi	Kode kabupaten kota	Nama kabupaten kota	Jenis kelamin	Pendidikan tertinggi	Jumlah pencari kerja	Satuan	Tahun
12	SUMATERA UTARA	1275	MEDAN	LAKI-LAKI	DIPLOMA I/II/III/AKADEMI	217	ORANG	2023
12	SUMATERA UTARA	1275	MEDAN	PEREMPUAN	DIPLOMA I/II/III/AKADEMI	229	ORANG	2023
12	SUMATERA UTARA	1275	MEDAN	LAKI-LAKI	S3	1	ORANG	2023
12	SUMATERA UTARA	1275	MEDAN	PEREMPUAN	S3	2	ORANG	2023
.....
12	SUMATERA UTARA	1275	MEDAN	LAKI-LAKI	SMP	177	ORANG	2023
12	SUMATERA UTARA	1275	MEDAN	PEREMPUAN	SMP	216	ORANG	2023
12	JAWA BARAT	1275	MEDAN	LAKI-LAKI	UNIVERSITAS	84	ORANG	2023
12	SUMATERA UTARA	1275	MEDAN	PEREMPUAN	UNIVERSITAS	49	ORANG	2023



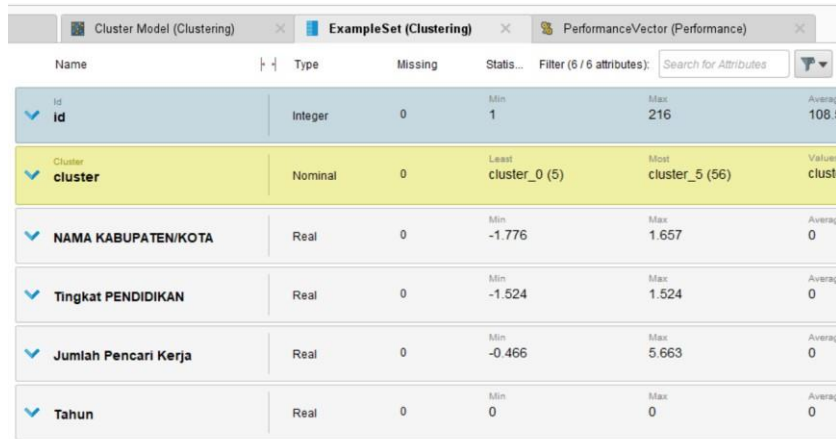
Setelah memilih data pencari kerja untuk digunakan, alat RapidMiner akan digunakan untuk memproses kembali data untuk menghilangkan atribut yang tidak digunakan. Attribute yang akan diproses termasuk nama kabupaten atau kota, tingkat pendidikan tertinggi, jumlah pencari kerja, dan tahun, dapat dilihat pada tabel 2 dibawah ini:

Tabel 2. Pemilihan Data

Nama kabupaten/kota	Pendidikan tertinggi	Jumlah pencari kerja	Tahun
MEDAN	SMA	4454	2023
MEDAN	DIPLOMA I/II/III/AKADEMI	229	2023
MEDAN	SD	16	2023
MEDAN	SMP	139	2023
....
MEDAN	SMK	0	2023
MEDAN	TIDAK/BELUM TAMAT SD	0	2023
MEDAN	PROFESI	0	2023
MEDAN	UNIVERSITAS	84	2023

2. Preprocessing Data

Tahap kedua adalah preprocessing data pada data pencari kerja. Tahap ini menghapus atribut yang tidak digunakan agar proses Data Mining tidak terhambat. Atribut yang tidak ada nilai ditunjukkan pada gambar 2 di bawah ini.



Gambar 2. Missing Value Check

3. Transformation Data

Tahap ketiga yakni transformasi data, data yang akan ditransformasi yakni atribut pendidikan tertinggi yang semula masih berbentuk polynominal dilakukan inisialisasi menjadi numerik agar dapat dilakukan proses Data Mining. Berikut ini tabel 3, attribute yang sudah di inisialisasi terlebih dahulu.

Tabel 3 Inisialisasi pada attribute pendidikan tertinggi

Pendidikan tertinggi	Inisialisasi
DIPLOMA I/II/III/AKADEMI	1
PROFESI	2
TIDAK/BELUM TAMAT SD	3
SD	4
SMA	5
SMK	6
SMP	7
UNIVERSITAS	8

Tabel 4. Transformasi data berdasarkan tingkat pendidikan

Nama Kabupaten/Kota	Pendidikan Tertinggi	Jumlah Pencari Kerja	Tahun
MEDAN	1	446	2023
MEDAN	2	0	2023
MEDAN	3	0	2023
MEDAN	4	30	2023
MEDAN	5	8298	2023
MEDAN	6	0	2023

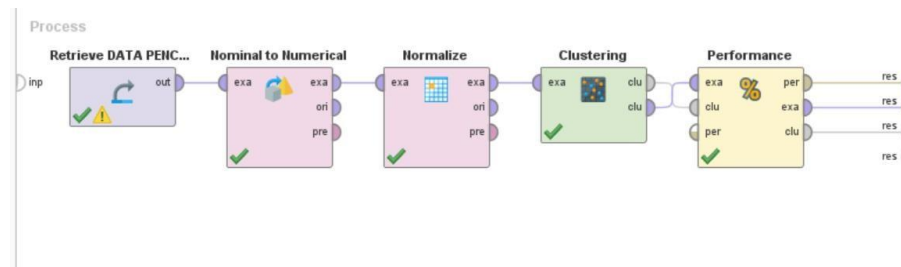


Nama Kabupaten/Kota	Pendidikan Tertinggi	Jumlah Pencari Kerja	Tahun
MEDAN	7	266	2023
MEDAN	8	890	2023
....
MEDAN	1	120	2023
MEDAN	2	0	2023
MEDAN	3	0	2023
MEDAN	4	154	2023
MEDAN	5	1108	2023
MEDAN	6	0	2023
MEDAN	7	393	2023
MEDAN	8	133	2023

Pada tabel 4 tersebut merupakan tranformasi data berdasarkan Tingkat Pendidikan, Dimana menunjukkan jumlah pencari kerja berdasarkan Pendidikan tinggia yang dimiliki oleh para pencari kerja.

4. Data Mining

Pada langkah selanjutnya, proses data mining, pengelompokan data pencari kerja berdasarkan tingkat pendidikan dilakukan menggunakan alat RapidMiner. Berikut ini adalah susunan proses clustering yang dilakukan oleh Algoritma K-Means di alat RapidMiner, dapat dilihat pada gambar 3 dibawah ini:

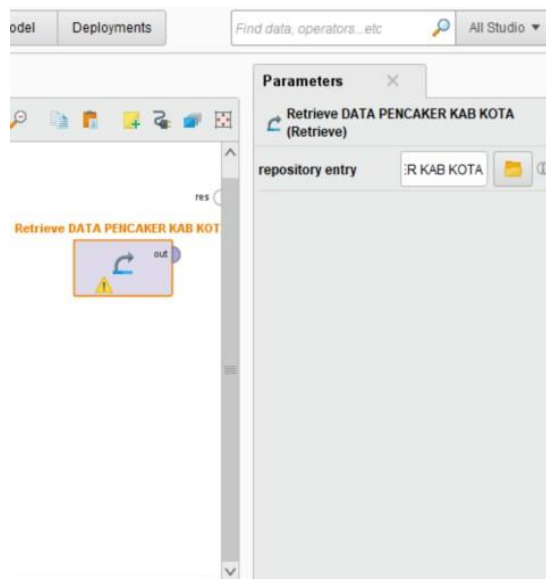


Gambar 3. Proses clustering pada tools RapidMiner

Sebelum memulai proses data mining, susunan operator yang digunakan oleh alat RapidMiner untuk pengelompokan data pencari kerja menggunakan Algoritma K-Means dijelaskan sebagai berikut:

a. Retreive

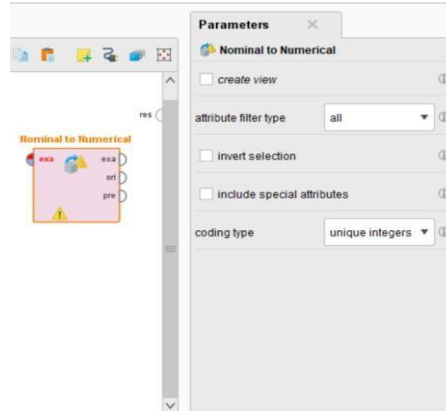
Pada gambar di bawah ini, operator retrieve mengakses data pencari kerja yang disimpan di repositori lokal dan memasukkannya ke dalam proses. Operator retrieve memulai proses data mining dan dapat menilai apakah data sudah layak untuk tahap selanjutnya dari proses data mining. Objek RapidMiner adalah sekumpulan data, yang dapat berupa koleksi atau model. Pada gambar 4 dibawah ini merupakan objek retrieve



Gambar 4. Operator Retrieve

b. Nominal to Numerical

Operator Nominal to Numerical digunakan untuk mengubah jenis atribut non-numerik, seperti nama kota dan sekolah tertinggi, agar proses data mining tidak terhambat. Operator ini juga mengubah jenis atribut dan memetakan semua nilai atribut menjadi nilai numerik.

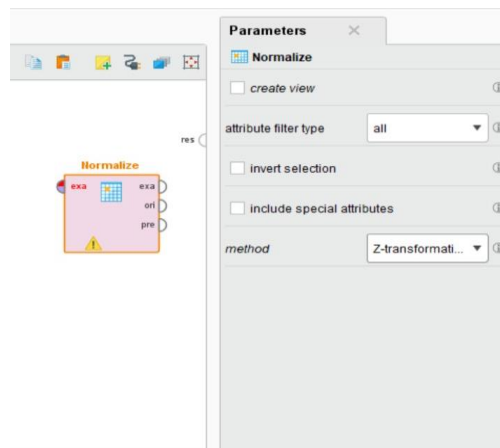


Gambar 5. Operator nominal to numerical

Gambar 5 merupakan operator nominal to numerical yang mana operator ini mengubah jenis atribut yang non numerik. Hal tersebut dilakukan agar proses data mining tidak terhambat.

c. Normalize

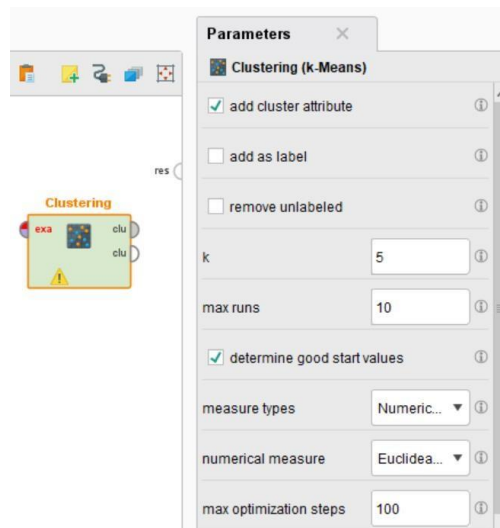
Nilai atribut yang dipilih dinormalkan oleh operator ini; normalisasi digunakan untuk menskalakan nilai sehingga menghasilkan rentang tertentu. Ini juga berguna untuk membandingkan atribut dengan jarak Eucdlian, yaitu semua atribut dengan skala yang sama, untuk perbandingan yang adil. Karena data pencari kerja ini merupakan standarisasi penggunaan range antar, metode transformasi range Z digunakan untuk normalisasi data. Berikut gambar 6 menjelaskan bagian dari Normalize:



Gambar 6. Operator Normalize

d. Clustering

Salah satu operator penting dalam proses pengelompokan data pencari kerja adalah operator pengelompokan.



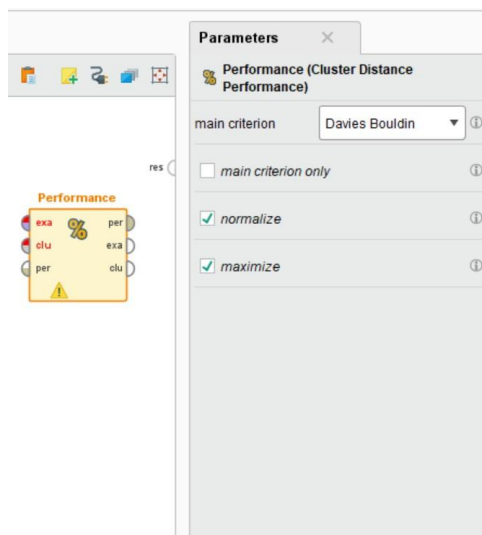
Gambar 7. Operator clustering (k-means)



Pada gambar 7, Ini melakukan pengelompokan menggunakan Algoritma K-Means dan menentukan cluster dari k2 hingga k10 pada data pencari kerja. Untuk mengetahui cluster mana yang memberikan hasil yang paling akurat, lakukan running sebanyak sepuluh kali. Tentukan cluster mana yang paling dekat dengan 0 hasilnya.

e. Performance

Operator ini menilai kinerja metode clustering berbasis centroid dengan menggunakan Davies Bouldin Index (DBI). Nilai maksimal DBI adalah 1, dan semakin kecil nilainya, semakin mirip data dalam kelompok. Untuk mengetahui kinerja data pencari kerja yang paling dekat, Davies Bouldin Index digunakan.



Gambar 8. Operator performance

Pada gambar 8, menentukan main criteriation dari performance (Cluster Distance Performance), dimana yang dipilih menjadi main criteriation adalah Davies Bouldin Index

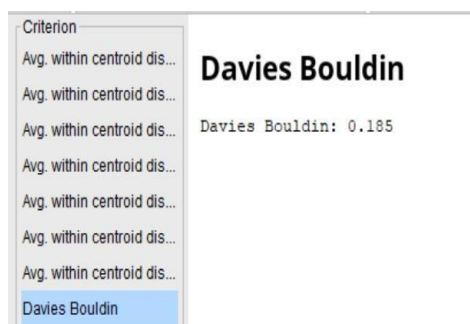
5. Evaluasi

Evaluasi hasil pengelompokan menggunakan Davies Bouldin Index (DBI) adalah tahap terakhir dari proses data mining menggunakan alat RapidMiner. Hal ini dilakukan untuk menentukan hasil pengelompokan terbaik, atau cluster, yang dihasilkan setelah menjalankan operator pengelompokan sampai sepuluh kali. DBI terbaik yang dihasilkan ditunjukkan pada tabel 5 di bawah ini.

Tabel 5 Hasil cluster DBI

Perulangan	Cluster	DBI
1	K2	0.326
2	K3	0.257
3	K4	0.212
4	K5	0.195
5	K6	0.185
6	K7	0.208
7	K8	0.221
8	K9	0.219
9	K10	0.213

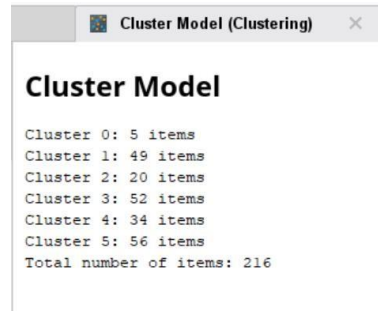
Untuk menemukan k ideal, yang berada pada (k6) dengan nilai DBI 0.185, algoritma K-Means digunakan pada data pemcari kerja berdasarkan tingkat pendidikan. Hasilnya menunjukkan bahwa k ideal dilakukan running sebanyak sepuluh kali, diperlihatkan pada gambar 9 dibawah ini:



Gambar 9. Nilai Davies Bouldin

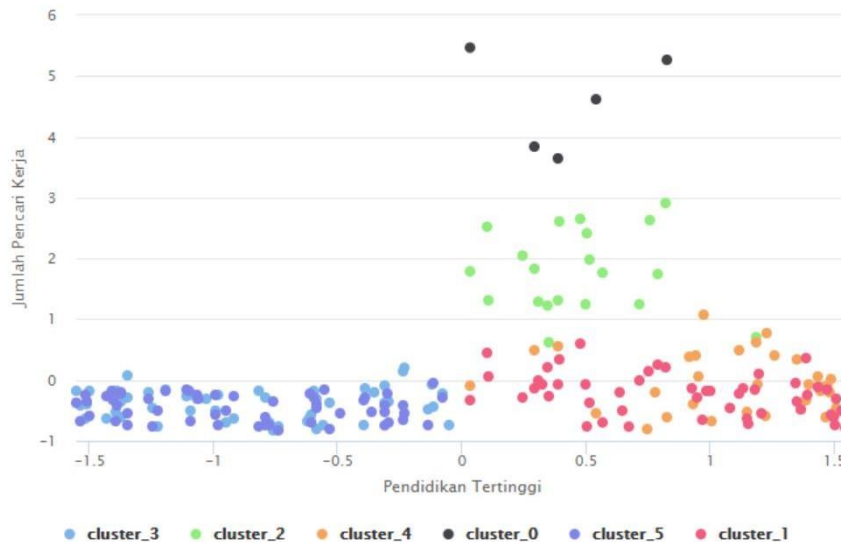


Setelah mengetahui nilai Davies Bouldin dibawah ini merupakan gambar cluster model pengelompokan datapencari kerja berdasarkan tingkat pendidikan yakni sebagai berikut.



Gambar 10. Cluster model pengelompokan data pencari kerja berdasarkan tingkat pendidikan

Dalam cluster model gambar 10 di atas, setelah melakukan running sebanyak sepuluh kali menggunakan perhitungan Davies Bouldin Index (DBI), ditemukan cluster (k6) dengan nilai DBI 0.185. Cluster ini terdiri dari cluster 0 (C0), cluster 1 (C1), cluster 2 (C2), cluster 3 (C3), cluster 4 (C4), dan cluster 5 (C5). Selanjutnya, hasil klasterisasi data pencari kerja berdasarkan tingkat pendidikan disajikan.



Gambar 11. Visualizations dengan tipe plot scatter/bubble pada data pencari kerja

Cluster_0 yang berwarna hitam terdiri dari 5 anggota, Cluster_1 yang berwarna pink terdiri dari 49 anggota, Cluster_2 yang berwarna hijau terdiri dari 20 anggota, Cluster_3 yang berwarna biru terdiri dari 52 anggota, Cluster_4 yang berwarna jingga terdiri dari 34 anggota dan Cluster_5 yang berwarna ungu terdiri dari 56 anggota. Cluster_5 merupakan jumlah pencari kerja terbanyak yang terdiri dari 56 anggota, masing-masing anggota terdiri dari DIPLOMA I/II/III/AKADEMI dengan jumlah pencari kerja terbanyak berjumlah 1062 orang, PROFESI dengan jumlah pencari kerja terbanyak berjumlah 20 orang, SD dengan jumlah pencari kerja terbanyak berjumlah 714 orang dan TIDAK/BELUM TAMAT SD dengan jumlah pencari kerja terbanyak berjumlah 2 orang. Sehingga dapat disimpulkan jumlah pencari kerja yang mendominasi yakni pencari kerja dengan tingkat pendidikan DIPLOMA I/II/III/AKADEMI yang berjumlah 1062 orang. Berikut ini merupakan tabel 6 hasil klasterisasi pencari kerja berdasarkan Tingkat Pendidikan.

Tabel 6. Hasil klasterisasi pencari kerja yang mendominasi pendidikan SMA

No.	Cluster	Nama Kabupaten/Kota	Pendidikan Tertinggi	Jumlah Pencari Kerja	Tahun
1.	Cluster_0	Medan	SMA	22819	2023
2.	Cluster_0	Medan	SMK	19133	2023
3.	Cluster_0	Medan	SMA	16990	2023
4.	Cluster_0	Medan	SMK	20476	2023
5.	Cluster_0	Medan	SMA	16333	2023

Tabel 6 menunjukkan hasil klasterisasi pencari kerja yang memiliki pendidikan SMA sebagai yang dominan. Data tersebut mencantumkan nomor cluster, nama Kabupaten/Kota (Medan), pendidikan tertinggi (SMA), jumlah pencari kerja, dan tahun pengumpulan data (2023). Klasterisasi ini menunjukkan bahwa di wilayah Medan, jumlah pencari kerja dengan latar belakang pendidikan SMA mendominasi dalam beberapa kelompok (Cluster_0). Jumlah pencari



kerja dengan pendidikan SMA yang tercatat dalam tabel menunjukkan variasi jumlah yang signifikan di berbagai kelompok, yang menunjukkan kompleksitas pasar kerja di wilayah tersebut.

Tabel 7. Hasil klasterisasi pencari kerja yang mendominasi pendidikan UNIVERSITAS

No.	Cluster	Nama Kabupaten/Kota	Pendidikan Tertinggi	Jumlah Pencari Kerja	Tahun
1.	Cluster_1	Medan	SMK	0	2023
2.	Cluster_1	Medan	SMP	1409	2023
3.	Cluster_1	Medan	UNIVERSITAS	2400	2023
4.	Cluster_1	Medan	SMP	446	2023
5.	Cluster_1	Medan	UNIVERSITAS	883	2023
....
48.	Cluster_1	Medan	SMP	393	2023
49.	Cluster_1	Medan	UNIVERSITAS	133	2023

Tabel 7 menampilkan hasil klasterisasi pencari kerja yang memiliki pendidikan UNIVERSITAS sebagai yang dominan. Data tersebut mencakup nomor cluster, nama Kabupaten/Kota (Medan), jenis pendidikan tertinggi (UNIVERSITAS), jumlah pencari kerja, dan tahun pengumpulan data (2023). Dalam klasterisasi ini, terlihat bahwa jumlah pencari kerja dengan latar belakang pendidikan UNIVERSITAS mendominasi dalam Cluster_1 di wilayah Medan. Meskipun terdapat variasi dalam jumlah pencari kerja di dalam kelompok ini, kehadiran individu dengan pendidikan universitas menonjol. Hal ini menunjukkan adanya tren pencarian pekerjaan yang lebih berkualifikasi di wilayah tersebut, yang dapat menggambarkan permintaan pasar kerja terhadap tenaga kerja yang memiliki pendidikan tinggi.

Tabel 8. Hasil klasterisasi pencari kerja yang mendominasi pendidikan SMK

No.	Cluster	Nama Kabupaten/Kota	Pendidikan Tertinggi	Jumlah Pencari Kerja	Tahun
1.	Cluster_2	Medan	SMA	8298	2023
2.	Cluster_2	Medan	SMA	8728	2023
3.	Cluster_2	Medan	SMK	12057	2023
4.	Cluster_2	Medan	SMA	6622	2023
5.	Cluster_2	Medan	SMK	11036	2023
6.	Cluster_2	Medan	SMK	7681	2023
....
19.	Cluster_2	Medan	SMA	11576	2023
20.	Cluster_2	Medan	SMK	9613	2023

Tabel 8 menunjukkan hasil klasterisasi pencari kerja yang memiliki pendidikan SMK sebagai yang dominan. Data ini mencakup nomor cluster, nama Kabupaten/Kota (Medan), jenis pendidikan tertinggi (SMK), jumlah pencari kerja, dan tahun pengumpulan data (2023). Dari tabel tersebut, dapat dilihat bahwa dalam klasterisasi ini, jumlah pencari kerja yang memiliki latar belakang pendidikan SMK mendominasi dalam Cluster_2 di wilayah Medan. Terdapat variasi jumlah yang signifikan dalam jumlah pencari kerja dengan pendidikan SMK di dalam kelompok ini, yang menunjukkan kompleksitas pasar kerja di wilayah tersebut. Hal ini mungkin mencerminkan permintaan pasar kerja terhadap tenaga kerja dengan keterampilan dan keahlian yang spesifik yang diajarkan di sekolah menengah kejuruan (SMK).

Tabel 9. Hasil klasterisasi pencari kerja yang mendominasi pendidikan SD

No.	Cluster	Nama Kabupaten/Kota	Pendidikan Tertinggi	Jumlah Pencari Kerja	Tahun
1.	Cluster_3	Medan	DIPLOMA I/II/III/AKADEMI	446	2023
2.	Cluster_3	Medan	SD	511	2023
3.	Cluster_3	Medan	PROFESI	3	2023
4.	Cluster_3	Medan	DIPLOMA I/II/III/AKADEMI	950	2023
5.	Cluster_3	Medan	SD	1828	2023
....
51.	Cluster_3	Medan	TIDAK/BELUM TAMAT SD	0	2023
52.	Cluster_3	Medan	SD	1504	2023



Tabel 9 menunjukkan hasil klasterisasi pencari kerja yang memiliki pendidikan SD sebagai yang dominan. Data ini mencakup nomor cluster, nama Kabupaten/Kota (Medan), jenis pendidikan tertinggi (SD), jumlah pencari kerja, dan tahun pengumpulan data (2023). Dari tabel tersebut, terlihat bahwa dalam klasterisasi ini, jumlah pencari kerja yang memiliki latar belakang pendidikan SD mendominasi dalam Cluster_3 di wilayah Medan. Terdapat variasi jumlah yang signifikan dalam jumlah pencari kerja dengan pendidikan SD di dalam kelompok ini. Selain itu, data juga mencantumkan beberapa entri yang menunjukkan adanya variasi tingkat pendidikan, seperti "DIPLOMA I/II/III/AKADEMI" dan "TIDAK/BELUM TAMAT SD", yang menunjukkan kompleksitas tingkat pendidikan di antara pencari kerja di wilayah tersebut. Hal ini menyoroti pentingnya pemahaman terhadap kebutuhan pasar kerja serta pendekatan yang sesuai dalam menyediakan pelatihan dan kesempatan kerja bagi individu dengan latar belakang pendidikan yang beragam.

Tabel 10. Hasil klasterisasi pencari kerja yang mendominasi pendidikan SMP

No.	Cluster	Nama Kabupaten/Kota	Pendidikan Tertinggi	Jumlah Pencari Kerja	Tahun
1.	Cluster_4	Medan	SMP	266	2023
2.	Cluster_4	Medan	UNIVERSITAS	890	2023
3.	Cluster_4	Medan	SMP	5035	2023
4.	Cluster_4	Medan	UNIVERSITAS	1308	2023
5.	Cluster_4	Medan	SMP	1987	2023
....
33.	Cluster_4	Medan	SMP	3355	2023
34.	Cluster_4	Medan	UNIVERSITAS	792	2023

Tabel 10 menampilkan hasil klasterisasi pencari kerja yang memiliki pendidikan SMP sebagai yang dominan. Data ini mencakup nomor cluster, nama Kabupaten/Kota (Medan), jenis pendidikan tertinggi (SMP), jumlah pencari kerja, dan tahun pengumpulan data (2023). Dari tabel tersebut, terlihat bahwa dalam klasterisasi ini, jumlah pencari kerja yang memiliki latar belakang pendidikan SMP mendominasi dalam Cluster_4 di wilayah Medan. Terdapat variasi jumlah yang signifikan dalam jumlah pencari kerja dengan pendidikan SMP di dalam kelompok ini. Selain itu, terdapat juga entri yang menunjukkan adanya individu dengan pendidikan UNIVERSITAS, menandakan keberagaman latar belakang pendidikan di antara pencari kerja dalam cluster ini. Hal ini mencerminkan kompleksitas pasar kerja di wilayah tersebut, di mana terdapat permintaan akan tenaga kerja dengan berbagai tingkat pendidikan dan keterampilan.

Tabel 11. Hasil klasterisasi pencari kerja yang mendominasi pendidikan DIPLOMA I/II/III/AKADEMI

No.	Cluster	Nama Kabupaten/Kota	Pendidikan Tertinggi	Jumlah Pencari Kerja	Tahun
1.	Cluster_5	Medan	DIPLOMA	302	2023
2.	Cluster_5	Medan	I/II/III/AKADEMI	714	2023
3.	Cluster_5	Medan	SD	20	2023
4.	Cluster_5	Medan	PROFESI	341	2023
5.	Cluster_5	Medan	DIPLOMA	1062	2023
....
55.	Cluster_5	Medan	I/II/III/AKADEMI	2	2023
56.	Cluster_5	Medan	TIDAK/BELUM TAMAT SD	154	2023

Tabel 11 menunjukkan hasil klasterisasi pencari kerja yang memiliki pendidikan DIPLOMA I/II/III/AKADEMI sebagai yang dominan. Data ini mencakup nomor cluster, nama Kabupaten/Kota (Medan), jenis pendidikan tertinggi (DIPLOMA I/II/III/AKADEMI), jumlah pencari kerja, dan tahun pengumpulan data (2023). Dari tabel tersebut, terlihat bahwa dalam klasterisasi ini, jumlah pencari kerja yang memiliki latar belakang pendidikan DIPLOMA I/II/III/AKADEMI mendominasi dalam Cluster_5 di wilayah Medan. Terdapat variasi jumlah yang signifikan dalam jumlah pencari kerja dengan pendidikan tersebut di dalam kelompok ini. Selain itu, terdapat juga entri yang menunjukkan adanya individu dengan pendidikan SD dan PROFESI, menunjukkan keberagaman latar belakang pendidikan di antara pencari kerja dalam cluster ini. Hal ini mencerminkan kompleksitas pasar kerja di wilayah tersebut, di mana terdapat permintaan akan tenaga kerja dengan berbagai tingkat pendidikan dan keterampilan.

4. KESIMPULAN

Berdasarkan Hasil yang diperoleh dari Implementasi Algoritma K-Means pada data pencari kerja di wilayah Medan berdasarkan tingkat pendidikan menghasilkan 6 cluster (k6) dengan nilai DBI 0.185 diantaranya sebagai berikut:



Cluster_0 terdiri dari 5 anggota, Cluster_1 terdiri dari 49 anggota, Cluster_2 terdiri dari 20 anggota, Cluster_3 terdiri dari 52 anggota, Cluster_4 terdiri dari 34 anggota dan Cluster_5 terdiri dari 56 anggota. Cluster_5 merupakan anggota terbanyak yakni berjumlah 56 anggota, masing-masing anggota terdiri dari : DIPLOMA I/II/III/AKADEMI dengan jumlah pencari kerja terbanyak berjumlah 1062 orang, PROFESI dengan jumlah pencari kerja terbanyak berjumlah 20 orang, SD dengan jumlah pencari kerja terbanyak berjumlah 714 orang dan TIDAK/BELUM TAMAT SD dengan jumlah pencari kerja terbanyak berjumlah 2 orang. Sehingga dapat disimpulkan pencari kerja yang mendominasi yakni pencari kerja dengan tingkat pendidikan DIPLOMA I/II/III/AKADEMI yang berjumlah 1062 orang. Berdasarkan pengelompokan data pencari kerja yang telah dilakukan dalam menentukan k optimum yang terbaik menghasilkan cluster (k6) dengan nilai DBI 0.185, sehingga semakin kecil nilai Davies Bouldin Index (DBI) maka kemiripan data dalam satu kelompok semakin berdekatan.

REFERENCES

- [1] Aditya, A., Jovian, I., & Sari, B. N. "Implementasi K-Means Clustering Ujian Nasional Sekolah Menengah Pertama di Indonesia Tahun 2018/2019". *Jurnal Media Informatika Budidarma*, 4(1), 51, 2020, <https://doi.org/10.30865/mib.v4i1.1784>
- [2] Harahap, T. R., & Nawawi, Z. M. "Pelayanan Permasalahan Dan Penempatan Tenaga Kerja Pada Dinas Tenaga Kerja (Disnaker) Kota Medan. Balance Jurnal Akuntansi Dan ...", 1(1), 96–109, 2022 <https://jurnal.risetilmiah.ac.id/index.php/jam/article/view/2%0Ahttps://jurnal.risetilmiah.ac.id/index.php/jam/article/download/2/2>
- [3] Ningsih, W., & Abdullah, F. "Analisis Perbedaan Pencari Kerja dan Lowongan Kerja Sebelum dan Pada Saat Pandemi Covid-19 di Kota Malang". *Journal of Regional Economics Indonesia*, 2(1), 42–56, 2021, <https://doi.org/10.26905/jrei.v2i1.6181>.
- [4] Nurcahyo, N. "Perlindungan hukum tenaga kerja berdasarkan peraturan perundang-undangan di Indonesia". *Jurnal Cakrawala Hukum*, 12(1), 69–78, 2021, <https://doi.org/10.26905/idjch.v12i1.5781>
- [5] Pa, P., Pardede, A. M. H., & Rahmadani, S. "Pengelompokan Data Pencari Kerja Terdaftar Berdasarkan Umur Dan Pendidikan Menggunakan Metode K-means Clustering Di Dinas Tenaga Kerja dan Perindustrian Perdagangan Kota Binjai". *Jurnal Informatika Kaputama (JIK)*, 6(3), 2022.
- [6] Utami, farathika putri. "Pengaruh Indeks Pembangunan Manusia (IPM), Kemiskinan dan Pengangguran Terhadap Pertumbuhan Ekonomi di Provinsi Aceh". *Jurnal Samudra Ekonomika*, 4(2), 101–113, 2020, <https://ejournalunsam.id/index.php/jse/article/view/2303>.
- [7] Wijaya, H. D., & Dwiasnati, S. "Implementasi Data Mining dengan Algoritma Naïve Bayes pada Penjualan Obat". *Jurnal Informatika*, 7(1), 1–7, 2020. <https://doi.org/10.31311/ji.v7i1.6203>
- [8] Dina Sunia, Kurniabudi, P. A. J. "Penerapan Data Mining untuk Clustering Data Penduduk Miskin Menggunakan Algoritma K-Means". *Jurnal Ilmiah Mahasiswa Teknik Informatika*, Vol 1 No 2, 121–134, 2019.
- [9] Halim, J. "Penerapan Data Mining Untuk Mengukur Tingkat Kepuasan Siswa Terhadap Pelayanan Di Bimbingan Belajar Al-Misbah Dengan Menggunakan Metode K-Means". Vol. 16, No 1, 2018.
- [10] R. R. Putra, N. A. Putri, and C. Wadisman, "Village Fund Allocation Information System for Community Empowerment in Klambir Lima Kebun Village," *J. Appl. ...*, vol. 3, no. 2, pp. 98–104, 2022, [Online]. Available: <https://journal.yrpiiku.com/index.php/jaets/article/view/681%0Ahttps://journal.yrpiiku.com/index.php/jaets/article/download/681/467>
- [11] Hermawati, F. A. *Data Mining*. Yogyakarta: Penerbit Andi, 2018.
- [12] Kusri dan taufiq, Emha. *Algoritma Data Mining*. CV. Andi Offset, Jogjakarta, 2019.
- [13] Larose, Daniel T. *Discovering Knowledge in Data: An Introduction to Data Mining*. John Wiley & Sons, Inc. 2018.
- [14] Sallaby, A. F., & Suryana, E. "Penerapan Data Mining untuk Menentukan Jumlah Pencari Kerja Terdaftar Berdasarkan Umur dan Pendidikan Menggunakan KMeans Clustering (Studi Kasus di Dinas Tenaga Kerja Dan Transmigrasi Provinsi Bengkulu)". *Journal of Technopreneurship and Information System (JTIS)*, 1(1), 2018a. <https://doi.org/10.36085/jtis.v1i2.28>
- [15] Gustientiedina, Gustientiedina, M. Hasmil Adiya, and Yenny Desnelita. "Penerapan Algoritma K-Means Untuk Clustering Data Obat-Obatan." *Jurnal Nasional Teknologi Dan Sistem Informasi*, vol 5. no 1 pp 17-24, 2019.
- [16] F. Kurnia, J. Kurniawan, and I. Fahmi, "Klasifikasi Keluarga Miskin Menggunakan Metode K-Nearest Neighbor Berbasis Euclidean Distance," in *Seminar Nasional Teknologi Informasi, Komunikasi dan Industri (SNTIKI)*, vol. 11, pp. 230–239, 2019.
- [17] A. M. Argina, "Penerapan Metode Klasifikasi K-Nearest Neighbor pada Dataset Penderita Penyakit Diabetes," *Indones. J. Data Sci.*, vol. 1, no. 2, pp. 29–33, 2020.
- [18] I. W. Supriana and L. G. Astuti, "Implementasi K-Nearest Neighbor Pada Penentuan Keluarga Miskin Bagi Dinas Sosial Kabupaten Tabanan," *J. Teknol. Inf. dan Komput*, vol. 5, no. 1, pp. 120–129, 2019
- [19] H. B. Suhartini, "Klasifikasi Algoritma K-Nearest Neighbor Berbasis Particle Swarm Optimization Untuk Kelayakan Bantuan Rehabilitasi Rumah Tidak Layak Huni Pada Desa Lenek Duren Kecamatan Aikmel Kabupaten Lombok Timur Suhartini1, Hariman," vol, vol. 2, pp. 79–85, 2019.
- [20] Sulistiyawati, Ari, and Eko Supriyanto. "Implementasi Algoritma K-means Clustering dalam Penentuan Siswa Kelas Unggulan." *Jurnal Tekno Kompak* vol 15.no 2 pp 25-36, 2021.
- [21] C. Rizal, Supriyandi, M. Amin. "Perancangan Aplikasi Pengelolaan Keuangan Desa Melalui E-Village Budgeting," *Bull. Comput. Sci. Res.*, vol. 3, no. 1, pp. 7–13, 2022, doi: 10.47065/bulletincsr.v3i1.181.