

Data Mining Klasifikasi Penduduk Penerima BST Menerapkan Metode K-Means

Rudha Raihan^{*}, Mesran

Fakultas Ilmu Komputer dan Teknologi Informasi, Program Studi Sistem Informasi, Universitas Budi Darma, Medan, Indonesia

Email: ^{1,*}rudhar25@gmail.com, ²mesran.skom.m.kom@gmail.com

Email Penulis Korespondensi: rudhar25@gmail.com

Abstrak—Program Bantuan Sosial Tunai (BST) merupakan salah satu upaya pemerintah dalam memberikan bantuan kepada kelompok masyarakat yang membutuhkan. Di Kecamatan Medan Polonia, penyaluran BST masih menghadapi tantangan dalam mendaftarkan dan mengelompokkan penduduk penerima secara efisien. Oleh karena itu, penelitian ini bertujuan untuk mengklasifikasi penduduk penerima BST menggunakan metode K-means clustering berdasarkan pendidikan terakhir, pengeluaran perbulan, jumlah tanggungan dan pekerjaan. Metode penelitian ini melibatkan pengumpulan data penduduk penerima BST dari kantor camat dan menggunakan aplikasi Rapid Miner untuk melakukan analisis clustering. Hasil pengelompokan data diidentifikasi menjadi tiga cluster dengan tingkat prioritas yang berbeda. Cluster 1 terdiri dari penduduk dengan tingkat prioritas tertinggi, sedangkan Cluster 2 dan Cluster 3 memiliki tingkat prioritas yang lebih rendah. Sehingga nama yang masuk kedalam Cluster 1 yang terdiri dari Ahmad Surya, Joko Santoso, Eko Setiawan, Dewi Lestari, Agus Wijaya, Jaya Pratama, Rudi Hermawan dan Andi Wijaya dianggap sebagai penerima BST dengan prioritas tertinggi.

Kata Kunci: Klasifikasi; Penduduk Penerima BST; Metode K-means Clustering; Rapid Miner

Abstract—The Cash Social Assistance (BST) program is one of the government's efforts to provide assistance to community groups in need. In Medan Polonia District, BST distribution still faces challenges in efficiently registering and categorizing the recipient population. Therefore, this study aims to classify the BST recipient population using the K-means clustering method based on last education, monthly expenses, number of dependents and employment. This research method involves collecting data on the BST recipient population from the sub-district office and using the Rapid Miner application to carry out clustering analysis. The results of data grouping were identified into three clusters with different priority levels. Cluster 1 consists of residents with the highest priority level, while Cluster 2 and Cluster 3 have a lower priority level. So the names included in Cluster 1 consisting of Ahmad Surya, Joko Santoso, Eko Setiawan, Dewi Lestari, Agus Wijaya, Jaya Pratama, Rudi Hermawan and Andi Wijaya are considered as BST recipients with the highest priority.

Keywords: Classification; BST Recipient Population; K-Means Clustering Method; Rapid Miner

1. PENDAHULUAN

Perkembangan peradaban manusia diiringi dengan populasi penduduk yang kurang mampu. Keberlangsungan kegiatan usaha yang dilakukan oleh instansi pemerintahan khususnya dalam hal pengolahan data untuk masyarakat kurang mampu yang memberikan dukungan terhadap keputusan-keputusan yang berlangsung. Bantuan Sosial Tunai (BST) merupakan program pemerintah yang bertujuan untuk membantu masyarakat yang kurang mampu dalam memenuhi kebutuhan dasar. Program ini telah dilaksanakan sejak tahun 2005 dan terus mengalami perubahan dalam pelaksanaannya[1].

Dalam pelaksanaan program ini, terdapat tantangan dalam menentukan penduduk yang berhak menerima bantuan tersebut. Dalam konteks ini, klasifikasi penduduk penerima BST menjadi penting untuk dilakukan. Klasifikasi ini bertujuan untuk mengidentifikasi kelompok penduduk yang memenuhi kriteria dan syarat yang telah ditetapkan oleh pemerintah.

Klasifikasi penduduk penerima BST berdasarkan faktor ekonomi dilakukan untuk memastikan bahwa bantuan yang diberikan tepat sasaran dan sesuai dengan kebutuhan masyarakat. Faktor ekonomi yang menjadi pertimbangan dalam klasifikasi ini meliputi pendapatan, aset, dan pengeluaran. Namun, pelaksanaan klasifikasi penduduk penerima BST berdasarkan faktor ekonomi tidaklah mudah. Terdapat beberapa tantangan yang dihadapi, seperti sulitnya mendapatkan data yang akurat dan valid mengenai faktor ekonomi penduduk, serta adanya perbedaan persepsi dan pemahaman mengenai kriteria penerima BST. Masalah di dalam klasifikasi penerima BST dapat diselesaikan dengan data mining diantaranya menggunakan metode K-Means Clustering[2].

Data mining adalah suatu proses analisis yang dilakukan untuk menggali informasi yang tersembunyi dan berharga dari kumpulan data yang besar. Dalam konteks klasifikasi penduduk penerima BST, data mining digunakan untuk mengidentifikasi pola atau aturan yang dapat membedakan antara penduduk penerima dan non-penerima BST[3]. Dengan menerapkan data mining dan metode K-means clustering, dapat dihasilkan pemetaan yang jelas dan terstruktur mengenai penduduk penerima BST, sehingga memudahkan dalam pengambilan keputusan terkait penyaluran bantuan sosial dengan lebih efektif dan efisien.

Faktor yang berkontribusi terhadap tingkat kemiskinan adalah kurangnya lapangan kerja yang memadai atau pendapatan yang tidak mencukupi untuk memenuhi kebutuhan hidup, seperti tanggungan keluarga dan anak-anak. Selain itu, hasil dari survei menunjukkan bahwa salah satu masalah dalam mengatasi kemiskinan adalah ketidaksesuaian antara bantuan yang diberikan oleh pemerintah dengan kebutuhan yang sebenarnya. Salah satu tantangan yang sering dihadapi oleh Pemerintah adalah distribusi bantuan sosial yang tidak merata dan tidak tepat sasaran.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Dalam penyusunan Laporan Kerja Praktek ini, penulis mengumpulkan data yang dipergunakan untuk membuat program dengan metode-metode sebagai berikut:

1. **Identifikasi Masalah**
Pada tahap ini, mengidentifikasi permasalahan yang muncul ditempuh dengan cara melakukan survei secara langsung di kecamatan medan polonia. Survei tersebut dilakukan untuk mengetahui permasalahan pada proses penerimaan BST.
2. **Studi Lapangan**
Wawancara akan dilakukan dengan pegawai kantor camat yang terlibat dalam pengolahan data masyarakat serta menanyakan beberapa hal dengan secara lisan maupun tertulis untuk mendengarkan dan memahami informasi dianggap perlu diketahui. Observasi juga dilakukan untuk mengamati proses pendataan dan penerimaan BST
3. **Studi Kepustakaan**
Studi pustaka dilakukan untuk mendapatkan referensi mengenai teori-teori tentang klasifikasi, metode K-Means yang mendukung penelitian. Langkah ini dilakukan dengan cara mengoperasikan aplikasi atau software rapid miner yang terkait dengan penelitian.
4. **Analisis Data**
Dalam penelitian ini, penggunaan metode riset pasar untuk analisis data memiliki pendekatan kualitatif. Metode ini melibatkan analisis tren, analisis similaritas, dan analisis kontradiksi guna memeriksa validitas hipotesis yang telah diajukan pada tahap sebelumnya. Jika hipotesis tersebut tidak terbukti valid, maka perlu dilakukan pengumpulan ulang data dan analisis data untuk menguji kembali hipotesis tersebut. Dengan demikian, analisis data bertujuan untuk memastikan keabsahan hipotesis dalam penelitian ini melalui penggunaan metode riset pasar yang bersifat kualitatif.
5. **Pengujian**
Bagaimana menganalisis penerapan metode K-Means menggunakan aplikasi rapid miner dan mengimplementasikan untuk memastikan fungsionalitas dan keandalannya.
6. **Penulisan Laporan**
Menyusun laporan kerja praktek yang mencakup pendahuluan, latar belakang, tujuan, metodologi, hasil penelitian, pembahasan, kesimpulan, dan saran. Kemudian, Menyajikan temuan penelitian dan rekomendasi untuk perbaikan atau peningkatan proses klasifikasi penduduk penerima BST.

2.2 Data Mining

Data mining atau disebut juga *Knowlgede Discovery in Database* (KDD) adalah serangkaian proses yang digunakan untuk menemukan informasi yang tersembunyi atau tidak terlihat secara manual dari sekumpulan data. Tujuannya adalah untuk mengungkapkan pengetahuan yang bermanfaat yang sebelumnya tidak diketahui. Istilah "mining" sendiri mengacu pada proses memperoleh barang berharga (*minerals*) dari jumlah material dasar yang besar (*earth*). Oleh karena itu, teknologi data mining merupakan proses yang melibatkan berbagai bidang ilmu seperti kecerdasan buatan (*artificial intelligence*), pembelajaran mesin (*machine learning*), statistika, dan basis data. Dalam literatur data mining, beberapa teknik yang umumnya dibahas meliputi *clustering*, *classification*, *association rule mining*, jaringan saraf tiruan (*neural network*), dan lain sebagainya[4][5][6]. Adapun tahap-tahap data mining yaitu [7][8][9]:

1. **Data selection** : Pada tahap ini berisi pemilihan tujuan penggalian data, mengidentifikasi data yang akan digali, memilih atribut input dan target.
2. **Data transformation** : Pada tahap ini dilakukan tranformasi yang meliputi organisasi data yang telah dipilih pada tahapan sebelumnya, melakukan konversi dari satu tipe data ke lainnya (misalnya dari simbolik ke numerik), mendefinisikan atribut-atribut, menurunkan dimensi data, menghilangkan noise (termasuk outlier), normalisasi, juga menentukan strategi untuk menangani data yang hilang.
3. **Data mining step per step**. Pada tahap ini data yang telah ditransformasi pada tahapan sebelumnya digali menggunakan satu atau lebih teknik untuk mendapatkan pola-pola yang menarik perhatian dan menjadi pengetahuan baru.
4. **Result interpretation and validation**. Untuk memahami arti dari pengetahuan baru yang telah tersintesa pada tahapan sebelumnya dan juga validitas hasil maka dilakukan uji aplikasi data mining menggunakan metode estimasi.
5. **Incorporation of the discovered knowledge**. Pada tahap ini berisi penyajian hasil-hasil dari teknik data mining kepada para pengambil kebijakan perusahaan atau pemerintahan yang akan dilihat adanya pertentangan dengan keputusan sebelumnya atau apakah pengetahuan baru tersebut dapat diaplikasikan dan akan memberikan nilai tambah bagi perusahaan atau kebijakan pemerintahan.

2.3 Penerima BST

Program Bantuan Sosial Tunai (BST) adalah program pemerintah yang memberikan bantuan sosial tunai kepada keluarga miskin di Indonesia. Penerima BST adalah keluarga miskin yang telah terdaftar dan memenuhi syarat yang ditetapkan oleh pemerintah. Penerima BST menerima bantuan sosial tunai secara berkala untuk memenuhi kebutuhan dasar mereka, seperti makanan, pakaian, dan pendidikan[10].

Terdapat beberapa faktor yang mempengaruhi penerima BST, di antaranya adalah tingkat kemiskinan, tingkat pendidikan, dan tingkat kesehatan. Keluarga yang hidup dalam kemiskinan cenderung memiliki akses yang terbatas terhadap sumber daya dan layanan publik, sehingga membutuhkan bantuan sosial untuk memenuhi kebutuhan dasar mereka. Tingkat pendidikan dan kesehatan juga mempengaruhi kemampuan keluarga untuk memenuhi kebutuhan dasar mereka, sehingga pemerintah perlu memperhatikan faktor-faktor tersebut dalam menentukan penerima BST.

2.4 Metode K-Means Clustering

K-Means Clustering adalah suatu metode penganalisaan data atau metode Data Mining yang melakukan proses pengelompokan (Clustering) data kedalam beberapa kelompok berdasarkan jarak, kriteria, kondisi dan karakteristik di berbagai partisi. Data dalam satu kelompok harus memiliki jarak terpendek, kriteria, kondisi atau yang hampir sama antara satu dengan lainnya[11][12][13].

K-Means Clustering memiliki objective yaitu meminimalisasi object function yang telah di atur pada proses clusterisasi. Dengan cara minimalisasi variasi antar 1 cluster dengan maksimalisasi variasi dengan data di cluster lainnya. Algoritma K-Means dapat mengelompokkan objek yang memiliki kemiripan. Tujuan dari Clustering Algoritma (K-Means) memiliki tujuan untuk meminimalisasikan fungsi objective yang telah di set dalam proses clustering. Tujuan tersebut dilakukan dengan cara meminimalikan variasi data yang ada didalam cluster dan memaksimalkan variasi data yang ada di cluster lainnya[14][15][16]. K-means clustering merupakan metode algoritma dasar, yang diterapkan sebagai berikut[17][18]:

1. Menyiapkan sampling data
2. Menentukan jumlah cluster.
3. Menentukan nilai centroid/titik pusat.
4. Menghitung jarak masing-masing centroid.
5. Mengelompokkan data berdasarkan jarak terpendek

Distance space dapat diimplementasikan untuk menghitung jarak data dan centroid[19]. Untuk menentukan jarak masing-masing centroid berdasarkan jarak terpendek, kriteria, kondisi atau karakteristik yang sama atau hampir sama, menggunakan model euclidean yang dirumuskan sebagai berikut:

$$d_{ij} = \sqrt{(X_{1i} - X_{1j})^2 + (X_{2i} - X_{2j})^2 + \dots + (X_{ki} - X_{kj})^2}$$

Keterangan :

d_{ij} = jarak dari data ke i ke pusat cluster j

X_{ki} = data dari ke – i pada attribute data ke – k

X_{kj} = data dari ke – j pada attribute data ke – k

3. HASIL DAN PEMBAHASAN

Dalam proses pendataan penduduk penerima BST, dilakukan pengklasifikasian menggunakan metode K-means clustering. Pengklasifikasian ini bertujuan untuk memilah penduduk menjadi kelompok-kelompok berdasarkan atribut-atribut yang relevan. Dalam proses ini, penduduk yang berkualitas baik maupun buruk dapat diidentifikasi dan dipisahkan. Proses pengklasifikasian ini memiliki potensi untuk meningkatkan penjualan, karena selain memilah penduduk berdasarkan kualitasnya, juga memungkinkan penambahan variasi produk yang dijual. Metode yang digunakan dalam pengklasifikasian ini adalah K-Means Clustering, dan prosesnya dilakukan dengan bantuan aplikasi RapidMiner dan penghitungan manual menggunakan Microsoft Excel.

3.1 Penerapan Metode K-Means Clustering

Pada tahap ini akan dilakukan proses utama yaitu pengelompokan data penduduk yang diambil dari database, menggunakan metode K-Means Clustering. Dari data penduduk yang diperoleh, di ambil 20 data penduduk untuk dijadikan sampel penerapan algoritma K-Means Clustering. Percobaan dilakukan dengan menggunakan parameter-parameter berikut:

Jumlah cluster : 3

Jumlah data : 20

Status ekonomi : Rendah

Penelitian ini data yang digunakan adalah pengelompokkan penduduk tidak mampu di Kecamatan Medan polonia. Kategori untuk penduduk tidak mampu yang ada didalam data seperti Pendidikan, Pekerjaan, Tanggungan, Pengeluaran.

Tabel 1. Transformasi Data

Kode Kriteria	Nama Kriteria	Bobot
Pengeluaran	<1 jt	1
	1 jt s.d 1.5 jt	2
	1.5 jt s/d 2 jt	3
	2 jt s/d 3 jt	4
Pendidikan	Tidak Bersekolah	1
	SD	2
	SMP	3
	SMA	4
	S1	5
Tanggungan	2 Orang	1
	3 Orang	2
	4 Orang	3
	5 Orang	4
	6 orang	5
Pekerjaan	Tidak Bekerja	1
	Buruh	2
	Pedagang	3
	Bertani	4
	Karyawan Swasta	5

Tabel 2. Rating kecocokan

Nama	Pengeluaran	Pendidikan	Tanggungan	Pekerjaan
Ahmad surya	3	4	3	3
Siti indah	3	3	4	2
Budi pratama	2	2	3	4
Rina fitriani	2	2	4	5
Joko santoso	2	2	1	2
Maya putri	3	1	2	3
Eko setiawan	4	3	3	4
Dewi lestari	3	4	5	3
Hadi nugroho	4	2	5	2
Rini kartika	3	2	3	1
Agus wijaya	3	3	2	5
Ana puspita	3	2	4	2
Jaya pranata	2	4	4	3
Tini rahayu	3	2	4	2
Dedi santoso	3	3	4	1
Yuni setyawati	3	2	4	2
Rudi hermawan	3	2	5	3
Nita safitri	2	2	5	2
Andi wijaya	4	3	4	3
Rini pratiwi	3	2	3	3

Iterasi ke-1

1. Penentuan Nilai *Centroid* Awal

Menentukan nilai centroid awal/titik pusat awal, dengan 3 record data sesuai dengan jumlah attribute kriteria. Nilai centroid awal/titik pusat awal diperoleh dari data penduduk.

Tabel 3. Nilai centroid iterasi ke-1

Agus wijaya	C1	3	3	2	5
Joko Santoso	C2	2	2	1	2
Maya putri	C3	3	1	2	3

2. Perhitungan Jarak

Dari 20 data yang dijadikan sebagai sampel yang telah dipilih. Lalu dilakukan jarak dari sisa sampel data dengan pusat cluster.

Tabel 3. Normalisasi sampel data

Kode Nama	Pengeluaran	Pendidikan	Tanggungan	Pekerjaan
-----------	-------------	------------	------------	-----------

M1	3	4	3	3
M2	3	3	4	2
M3	2	2	3	4
M4	2	2	4	5
M5	2	2	1	2
M6	3	1	2	3
M7	4	3	3	4
M8	3	4	5	3
M9	4	2	5	2
M10	3	2	3	1
M11	3	3	2	5
M12	3	2	4	2
M13	2	4	4	3
M14	3	2	4	2
M15	3	3	4	1
M16	3	2	4	2
M17	3	2	5	3
M18	2	2	5	2
M19	4	3	4	3
M20	3	2	3	3

Menghitung jarak dengan menggunakan model euclidean.

Cluster C1:

$$d(x_1, c_1) = \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2 + (c_1 - c_{1c})^2 + (d_1 - c_{1d})^2} = \sqrt{(3-3)^2 + (4-3)^2 + (3-2)^2 + (3-5)^2} = 2,449$$

$$d(x_2, c_1) = \sqrt{(a_2 - c_{1a})^2 + (b_2 - c_{1b})^2 + (c_2 - c_{1c})^2 + (d_2 - c_{1d})^2} = \sqrt{(3-3)^2 + (3-3)^2 + (4-2)^2 + (2-5)^2} = 3,606$$

$$d(x_3, c_1) = \sqrt{(a_3 - c_{1a})^2 + (b_3 - c_{1b})^2 + (c_3 - c_{1c})^2 + (d_3 - c_{1d})^2} = \sqrt{(2-3)^2 + (2-3)^2 + (3-2)^2 + (4-5)^2} = 2$$

$$d(x_4, c_1) = \sqrt{(a_4 - c_{1a})^2 + (b_4 - c_{1b})^2 + (c_4 - c_{1c})^2 + (d_4 - c_{1d})^2} = \sqrt{(2-3)^2 + (2-3)^2 + (4-2)^2 + (5-5)^2} = 2,449$$

$$d(x_5, c_1) = \sqrt{(a_5 - c_{1a})^2 + (b_5 - c_{1b})^2 + (c_5 - c_{1c})^2 + (d_5 - c_{1d})^2} = \sqrt{(2-3)^2 + (2-3)^2 + (1-2)^2 + (2-5)^2} = 3,464$$

$$d(x_6, c_1) = \sqrt{(a_6 - c_{1a})^2 + (b_6 - c_{1b})^2 + (c_6 - c_{1c})^2 + (d_6 - c_{1d})^2} = \sqrt{(3-3)^2 + (1-3)^2 + (2-2)^2 + (3-5)^2} = 2,828$$

$$d(x_7, c_1) = \sqrt{(a_7 - c_{1a})^2 + (b_7 - c_{1b})^2 + (c_7 - c_{1c})^2 + (d_7 - c_{1d})^2} = \sqrt{(4-3)^2 + (3-3)^2 + (3-2)^2 + (4-5)^2} = 1,732$$

$$d(x_8, c_1) = \sqrt{(a_8 - c_{1a})^2 + (b_8 - c_{1b})^2 + (c_8 - c_{1c})^2 + (d_8 - c_{1d})^2} = \sqrt{(3-3)^2 + (4-3)^2 + (5-2)^2 + (3-5)^2} = 3,742$$

$$d(x_9, c_1) = \sqrt{(a_9 - c_{1a})^2 + (b_9 - c_{1b})^2 + (c_9 - c_{1c})^2 + (d_9 - c_{1d})^2} = \sqrt{(4-3)^2 + (2-3)^2 + (5-2)^2 + (2-5)^2} = 4,472$$

$$d(x_{10}, c_1) = \sqrt{(a_{10} - c_{1a})^2 + (b_{10} - c_{1b})^2 + (c_{10} - c_{1c})^2 + (d_{10} - c_{1d})^2} = \sqrt{(3-3)^2 + (2-3)^2 + (3-2)^2 + (1-5)^2} = 4,243$$

$$d(x_{11}, c_1) = \sqrt{(a_{11} - c_{1a})^2 + (b_{11} - c_{1b})^2 + (c_{11} - c_{1c})^2 + (d_{11} - c_{1d})^2} = \sqrt{(3-3)^2 + (3-3)^2 + (2-2)^2 + (5-5)^2} = 0$$

$$d(x_{12}, c_1) = \sqrt{(a_{12} - c_{1a})^2 + (b_{12} - c_{1b})^2 + (c_{12} - c_{1c})^2 + (d_{12} - c_{1d})^2} = \sqrt{(3-3)^2 + (2-3)^2 + (4-2)^2 + (2-5)^2} = 3,742$$

$$d(x_{13}, c_1) = \sqrt{(a_{13} - c_{1a})^2 + (b_{13} - c_{1b})^2 + (c_{13} - c_{1c})^2 + (d_{13} - c_{1d})^2} = \sqrt{(2-3)^2 + (4-3)^2 + (4-2)^2 + (3-5)^2} = 3,162$$

$$d(x_{14}, c_1) = \sqrt{(a_{14} - c_{1a})^2 + (b_{14} - c_{1b})^2 + (c_{14} - c_{1c})^2 + (d_{14} - c_{1d})^2} = \sqrt{(3-3)^2 + (2-3)^2 + (4-2)^2 + (2-5)^2} = 3,742$$

$$d(x_{15}, c_1) = \sqrt{(a_{15} - c_{1a})^2 + (b_{15} - c_{1b})^2 + (c_{15} - c_{1c})^2 + (d_{15} - c_{1d})^2} = \sqrt{(3-3)^2 + (3-3)^2 + (4-2)^2 + (1-5)^2} = 4,472$$

$$d(x_{16}, c_1) = \sqrt{(a_{16} - c_{1a})^2 + (b_{16} - c_{1b})^2 + (c_{16} - c_{1c})^2 + (d_{16} - c_{1d})^2} = \sqrt{(3-3)^2 + (2-3)^2 + (4-2)^2 + (2-5)^2} = 3,472$$

$$d(x_{17}, c_1) = \sqrt{(a_{17} - c_{1a})^2 + (b_{17} - c_{1b})^2 + (c_{17} - c_{1c})^2 + (d_{17} - c_{1d})^2} = \sqrt{(3-3)^2 + (2-3)^2 + (5-2)^2 + (3-5)^2} = 3,742$$

$$d(x_{18}, c_1) = \sqrt{(a_{18} - c_{1a})^2 + (b_{18} - c_{1b})^2 + (c_{18} - c_{1c})^2 + (d_{18} - c_{1d})^2} = \sqrt{(2-3)^2 + (2-3)^2 + (5-2)^2 + (2-5)^2} = 4,472$$

$$d(x_{19}, c_1) = \sqrt{(a_{19} - c_{1a})^2 + (b_{19} - c_{1b})^2 + (c_{19} - c_{1c})^2 + (d_{19} - c_{1d})^2} = \sqrt{(4-3)^2 + (3-3)^2 + (4-2)^2 + (3-5)^2} = 3$$

$$d(x_{20}, c_1) = \sqrt{(a_{20} - c_{1a})^2 + (b_{20} - c_{1b})^2 + (c_{20} - c_{1c})^2 + (d_{20} - c_{1d})^2} = \sqrt{(3-3)^2 + (2-3)^2 + (3-2)^2 + (3-5)^2} = 2,449$$

Cluster C2:

$$d(x_1, c_2) = \sqrt{(a_1 - c_{2a})^2 + (b_1 - c_{2b})^2 + (c_1 - c_{2c})^2 + (d_1 - c_{2d})^2} = \sqrt{(3-2)^2 + (4-2)^2 + (3-1)^2 + (3-2)^2} = 3,162$$

$$d(x_2, c_2) = \sqrt{(a_2 - c_{2a})^2 + (b_2 - c_{2b})^2 + (c_2 - c_{2c})^2 + (d_2 - c_{2d})^2} = \sqrt{(3-2)^2 + (3-2)^2 + (4-1)^2 + (2-2)^2} = 3,317$$

$$d(x_3, c_2) = \sqrt{(a_3 - c_{2a})^2 + (b_3 - c_{2b})^2 + (c_3 - c_{2c})^2 + (d_3 - c_{2d})^2} = \sqrt{(2-2)^2 + (2-2)^2 + (3-1)^2 + (4-2)^2} = 2,828$$

$$d(x_4, c_2) = \sqrt{(a_4 - c_{2a})^2 + (b_4 - c_{2b})^2 + (c_4 - c_{2c})^2 + (d_4 - c_{2d})^2} = \sqrt{(2-2)^2 + (2-2)^2 + (4-1)^2 + (5-2)^2} = 4,243$$

$$d(x_5, c_2) = \sqrt{(a_5 - c_{2a})^2 + (b_5 - c_{2b})^2 + (c_5 - c_{2c})^2 + (d_5 - c_{2d})^2} = \sqrt{(2-2)^2 + (2-2)^2 + (1-1)^2 + (2-2)^2} = 0$$

$$d(x_6, c_2) = \sqrt{(a_6 - c_{2a})^2 + (b_6 - c_{2b})^2 + (c_6 - c_{2c})^2 + (d_6 - c_{2d})^2} = \sqrt{(3-2)^2 + (1-2)^2 + (2-1)^2 + (3-2)^2} = 2$$

$$d(x_7, c_2) = \sqrt{(a_7 - c_{2a})^2 + (b_7 - c_{2b})^2 + (c_7 - c_{2c})^2 + (d_7 - c_{2d})^2} = \sqrt{(4-2)^2 + (3-2)^2 + (3-1)^2 + (4-2)^2} = 3,606$$

$$d(x_8, c_2) = \sqrt{(a_8 - c_{2a})^2 + (b_8 - c_{2b})^2 + (c_8 - c_{2c})^2 + (d_8 - c_{2d})^2} = \sqrt{(3-2)^2 + (4-2)^2 + (5-1)^2 + (3-2)^2} = 4,69$$

$$d(x_9, c_2) = \sqrt{(a_9 - c_{2a})^2 + (b_9 - c_{2b})^2 + (c_9 - c_{2c})^2 + (d_9 - c_{2d})^2} = \sqrt{(4-2)^2 + (2-2)^2 + (5-1)^2 + (2-2)^2} = 3,464$$

$$d(x_{10}, c_2) = \sqrt{(a_{10} - c_{2a})^2 + (b_{10} - c_{2b})^2 + (c_{10} - c_{2c})^2 + (d_{10} - c_{2d})^2} = \sqrt{(3-2)^2 + (2-2)^2 + (3-1)^2 + (1-2)^2} = 2,449$$

$$d(x_{11}, c_2) = \sqrt{(a_{11} - c_{2a})^2 + (b_{11} - c_{2b})^2 + (c_{11} - c_{2c})^2 + (d_{11} - c_{2d})^2} = \sqrt{(3-2)^2 + (3-2)^2 + (2-1)^2 + (5-2)^2} = 3,464$$

$$d(x_{12}, c_2) = \sqrt{(a_{12} - c_{2a})^2 + (b_{12} - c_{2b})^2 + (c_{12} - c_{2c})^2 + (d_{12} - c_{2d})^2} = \sqrt{(3-2)^2 + (2-2)^2 + (4-1)^2 + (2-2)^2} = 3,162$$

$$d(x_{13}, c_2) = \sqrt{(a_{13} - c_{2a})^2 + (b_{13} - c_{2b})^2 + (c_{13} - c_{2c})^2 + (d_{13} - c_{2d})^2} = \sqrt{(2-2)^2 + (4-2)^2 + (4-1)^2 + (3-2)^2} = 3,742$$

$$d(x_{14}, c_2) = \sqrt{(a_{14} - c_{2a})^2 + (b_{14} - c_{2b})^2 + (c_{14} - c_{2c})^2 + (d_{14} - c_{2d})^2} = \sqrt{(3-2)^2 + (2-2)^2 + (4-1)^2 + (2-2)^2} = 3,162$$

$$d(x_{15}, c_2) = \sqrt{(a_{15} - c_{2a})^2 + (b_{15} - c_{2b})^2 + (c_{15} - c_{2c})^2 + (d_{15} - c_{2d})^2} = \sqrt{(3-2)^2 + (3-2)^2 + (4-1)^2 + (1-2)^2} = 3,464$$

$$d(x_{16}, c_2) = \sqrt{(a_{16} - c_{2a})^2 + (b_{16} - c_{2b})^2 + (c_{16} - c_{2c})^2 + (d_{16} - c_{2d})^2} = \sqrt{(3-2)^2 + (2-2)^2 + (4-1)^2 + (2-2)^2} = 3,162$$

$$d(x_{17}, c_2) = \sqrt{(a_{17} - c_{2a})^2 + (b_{17} - c_{2b})^2 + (c_{17} - c_{2c})^2 + (d_{17} - c_{2d})^2} = \sqrt{(3-2)^2 + (2-2)^2 + (5-1)^2 + (2-2)^2} = 4,243$$

$$d(x_{18}, c_2) = \sqrt{(a_{18} - c_{2a})^2 + (b_{18} - c_{2b})^2 + (c_{18} - c_{2c})^2 + (d_{17} - c_{2d})^2} = \sqrt{(2-2)^2 + (2-2)^2 + (5-1)^2 + (2-2)^2} = 4$$

$$d(x_{19}, c_2) = \sqrt{(a_{19} - c_{2a})^2 + (b_{19} - c_{2b})^2 + (c_{19} - c_{2c})^2 + (d_{19} - c_{2d})^2} = \sqrt{(4-2)^2 + (3-2)^2 + (4-1)^2 + (3-2)^2} = 3,873$$

$$d(x_{10}, c_2) = \sqrt{(a_{20} - c_{2a})^2 + (b_{20} - c_{2b})^2 + (c_{20} - c_{2c})^2 + (d_{20} - c_{2d})^2} = \sqrt{(3-2)^2 + (2-2)^2 + (3-1)^2 + (3-2)^2} = 2,449$$

Cluster C3:

$$d(x_1, c_3) = \sqrt{(a_1 - c_{3a})^2 + (b_1 - c_{3b})^2 + (c_1 - c_{3c})^2 + (d_1 - c_{3d})^2} = \sqrt{(3-3)^2 + (4-1)^2 + (3-2)^2 + (3-3)^2} = 3,162$$

$$d(x_2, c_3) = \sqrt{(a_2 - c_{3a})^2 + (b_2 - c_{3b})^2 + (c_2 - c_{3c})^2 + (d_2 - c_{3d})^2} = \sqrt{(3-3)^2 + (3-1)^2 + (4-2)^2 + (2-3)^2} = 3$$

$$d(x_3, c_3) = \sqrt{(a_3 - c_{3a})^2 + (b_3 - c_{3b})^2 + (c_3 - c_{3c})^2 + (d_3 - c_{3d})^2} = \sqrt{(2-3)^2 + (2-1)^2 + (3-2)^2 + (4-3)^2} = 2$$

$$d(x_4, c_3) = \sqrt{(a_4 - c_{3a})^2 + (b_4 - c_{3b})^2 + (c_4 - c_{3c})^2 + (d_4 - c_{3d})^2} = \sqrt{(2-3)^2 + (2-1)^2 + (4-2)^2 + (5-3)^2} = 3,162$$

$$d(x_5, c_3) = \sqrt{(a_5 - c_{3a})^2 + (b_5 - c_{3b})^2 + (c_5 - c_{3c})^2 + (d_5 - c_{3d})^2} = \sqrt{(2-3)^2 + (2-1)^2 + (1-2)^2 + (2-3)^2} = 2$$

$$d(x_6, c_3) = \sqrt{(a_6 - c_{3a})^2 + (b_6 - c_{3b})^2 + (c_6 - c_{3c})^2 + (d_6 - c_{3d})^2} = \sqrt{(3-3)^2 + (1-1)^2 + (2-2)^2 + (3-3)^2} = 0$$

$$d(x_7, c_3) = \sqrt{(a_7 - c_{3a})^2 + (b_7 - c_{3b})^2 + (c_7 - c_{3c})^2 + (d_7 - c_{3d})^2} = \sqrt{(4-3)^2 + (3-1)^2 + (3-2)^2 + (4-3)^2} = 2,646$$

$$d(x_8, c_3) = \sqrt{(a_8 - c_{3a})^2 + (b_8 - c_{3b})^2 + (c_8 - c_{3c})^2 + (d_8 - c_{3d})^2} = \sqrt{(3-3)^2 + (4-1)^2 + (5-2)^2 + (3-3)^2} = 4,243$$

$$d(x_9, c_3) = \sqrt{(a_9 - c_{3a})^2 + (b_9 - c_{3b})^2 + (c_9 - c_{3c})^2 + (d_9 - c_{3d})^2} = \sqrt{(4-3)^2 + (2-1)^2 + (5-2)^2 + (2-3)^2} = 3,464$$

$$d(x_{10}, c_3) = \sqrt{(a_{10} - c_{3a})^2 + (b_{10} - c_{3b})^2 + (c_{10} - c_{3c})^2 + (d_{10} - c_{3d})^2} = \sqrt{(3-3)^2 + (2-1)^2 + (3-2)^2 + (1-3)^2} = 2,449$$

$$d(x_{11}, c_3) = \sqrt{(a_{11} - c_{3a})^2 + (b_{11} - c_{3b})^2 + (c_{11} - c_{3c})^2 + (d_{11} - c_{3d})^2} = \sqrt{(3-3)^2 + (3-1)^2 + (2-2)^2 + (5-3)^2} = 2,828$$

$$d(x_{12}, c_3) = \sqrt{(a_{12} - c_{3a})^2 + (b_{12} - c_{3b})^2 + (c_{12} - c_{3c})^2 + (d_{12} - c_{3d})^2} = \sqrt{(3-3)^2 + (2-1)^2 + (4-2)^2 + (2-3)^2} = 2,449$$

$$d(x_{13}, c_3) = \sqrt{(a_{13} - c_{3a})^2 + (b_{13} - c_{3b})^2 + (c_{13} - c_{3c})^2 + (d_{13} - c_{3d})^2} = \sqrt{(2-3)^2 + (4-1)^2 + (4-2)^2 + (3-3)^2} = 3,742$$

$$d(x_{14}, c_3) = \sqrt{(a_{14} - c_{3a})^2 + (b_{14} - c_{3b})^2 + (c_{14} - c_{3c})^2 + (d_{14} - c_{3d})^2} = \sqrt{(3-3)^2 + (2-1)^2 + (4-2)^2 + (2-3)^2} = 2,449$$

$$d(x_{15}, c_3) = \sqrt{(a_{15} - c_{3a})^2 + (b_{15} - c_{3b})^2 + (c_{15} - c_{3c})^2 + (d_{15} - c_{3d})^2} = \sqrt{(3-3)^2 + (3-1)^2 + (4-2)^2 + (1-3)^2} = 3,464$$

$$d(x_{16}, c_3) = \sqrt{(a_{16} - c_{3a})^2 + (b_{16} - c_{3b})^2 + (c_{16} - c_{3c})^2 + (d_{16} - c_{3d})^2} = \sqrt{(3-3)^2 + (2-1)^2 + (4-2)^2 + (2-3)^2} = 2,449$$

$$d(x_{17}, c_3) = \sqrt{(a_{17} - c_{3a})^2 + (b_{17} - c_{3b})^2 + (c_{17} - c_{3c})^2 + (d_{17} - c_{3d})^2} = \sqrt{(3-3)^2 + (2-1)^2 + (5-2)^2 + (3-3)^2} = 3,162$$

$$d(x_{18}, c_3) = \sqrt{(a_{18} - c_{3a})^2 + (b_{18} - c_{3b})^2 + (c_{18} - c_{3c})^2 + (d_{18} - c_{3d})^2} = \sqrt{(2-3)^2 + (2-1)^2 + (5-2)^2 + (2-3)^2} = 3,464$$

$$d(x_{19}, c_3) = \sqrt{(a_{19} - c_{3a})^2 + (b_{19} - c_{3b})^2 + (c_{19} - c_{3c})^2 + (d_{19} - c_{3d})^2} = \sqrt{(4-3)^2 + (3-1)^2 + (4-2)^2 + (3-3)^2} = 3$$

$$d(x_{20}, c_3) = \sqrt{(a_{20} - c_{3a})^2 + (b_{20} - c_{3b})^2 + (c_{20} - c_{3c})^2 + (d_{20} - c_{3d})^2} = \sqrt{(3-3)^2 + (2-1)^2 + (3-2)^2 + (3-3)^2} = 1,414$$

Tabel 4. Rekapitulasi Hasil Perhitungan Iterasi Ke-1

	C1	C2	C3
M1	2,449	3,162	3,162
M2	3,606	3,317	3
M3	2	2,828	2
M4	2,449	4,243	3,162
M5	3,464	0	2
M6	2,828	2	0
M7	1,732	3,606	2,646
M8	3,742	4,69	4,243
M9	4,472	4,472	3,464
M10	4,243	2,449	2,449
M11	0	3,464	2,828
M12	3,742	3,162	2,449
M13	3,162	3,742	3,742
M14	3,742	3,162	2,449
M15	4,472	3,464	3,464
M16	3,742	3,162	2,449
M17	3,742	4,243	3,162
M18	4,472	4	3,464
M19	3	3,873	3
M20	2,449	2,449	1,414

Berdasarkan tabel 4, terlihat bahwa anggota C1 sebanyak 7 yang terdiri dari M1, M3, M4, M7, M8, M11, M13 dan M19, Anggota C2 hanya ada 3 yaitu M5, M10 dan M15, Anggota C3 terdiri dari M2, M6, M9, M12, M14, M16, M17, M18 dan M20. Menghitung titik cluster baru dengan cara menghitung nilai rata-rata dari data yang ada pada centroid yang sama:

$$C_k = \left(\frac{1}{n_k}\right) \sum d_1$$

Dimana n_k adalah jumlah dokumen dalam cluster k dan d_1 adalah dokumen dalam cluster k. Sehingga didapatkan titik centroid yang baru di tabel berikut:

Tabel 5. Nilai Centroid Iterasi Ke-2

C1	2,875	3,5	3,5	3,75
C2	2,667	2,333	2,667	1,333333

C3	3	2	4	2,333333
----	---	---	---	----------

Hitung Euclidean dari semua data ketitik pusat yang baru (C1,C2,C3) seperti yang telah dilakukan pada iterasi 1 setelah hasil perhitungan kita dapatkan, kemudian bandingkan dengan hasil tersebut. Jika posisi cluster pada iterasi 2 sama dengan iterasi 1, maka proses dihentikan, namun jika tidak maka proses dilanjutkan ke iterasi ke 3.

Tabel 6. Rekapitulasi Hasil Perhitungan Iterasi Ke-2

	C1	C2	C3
M1	1,436	3,18	2,333
M2	2,136	2,261	1,054
M3	1,887	1,054	2,186
M4	2,25	1,667	2,848
M5	3,544	4,876	4,372
M6	3,172	1,563	2,333
M7	1,25	2,261	2,404
M8	2,016	3,93	2,333
M9	2,562	3,232	1,764
M10	3,326	0,882	1,667
M11	2,25	3,232	3,887
M12	3,25	1,944	1,054
M13	2,016	4,773	3,232
M14	2,562	1,563	0,333
M15	3,01	2,261	1,667
M16	2,562	1,563	0,333
M17	2,016	3,93	2,333
M18	2,562	3,018	1,764
M19	1,436	3,667	2,333
M20	2,016	0,882	1,202

Berdasarkan tabel 6, terlihat bahwa Anggota C1 terdiri dari (M1,M5,M7,M8,M11,M13,M17,M19), Anggota C2 terdiri dari (M3,M4,M6,M10,M20) dan Anggota C3 terdiri dari (M2,M9,M12,M14,M15,M16,M18). Dikarenakan hasil cluster pada iterasi ke 2 tidak sama dengan posisi iterasi 1, maka proses dilanjutkan ke iterasi selanjutnya. Hal ini menunjukkan bahwa dalam proses K-means clustering, iterasi berulang dilakukan untuk mencapai konvergensi atau stabilitas dalam pengelompokan data. Setiap iterasi melibatkan perhitungan ulang centroid dan pembaruan penempatan data ke dalam cluster yang sesuai. Jika hasil clustering pada iterasi tertentu tidak sama dengan hasil pada iterasi sebelumnya, itu menunjukkan bahwa belum tercapai konvergensi dan perlu dilanjutkan ke iterasi berikutnya untuk mencapai pengelompokan yang lebih optimal. Setelah dilakukan proses pengelompokan, proses berhenti pada iterasi ke-4 dimana hasil pengelompokan iterasi ke-3 dan ke-4 sama. Yaitu seperti tabel 7 berikut.

Tabel 7. Nilai Centroid Iterasi Ke-4

C1	3	4,125	3,375	3,25
C2	2,5	1,75	3	3,75
C3	3	2,375	4,125	1,75

Tabel 8. Rekapitulasi Hasil Perhitungan Iterasi Ke-4

	C1	C2	C3
M1	0,468	2,424	2,339
M2	1,794	2,424	0,685
M3	2,494	0,612	2,733
M4	2,995	1,696	3,423
M5	2,995	4,228	4,209
M6	3,423	1,541	2,823
M7	1,723	1,969	2,778
M8	1,649	3,142	2,229
M9	2,543	3,298	1,49
M10	3,117	2,806	1,403
M11	2,229	2,806	4,209
M12	3,423	2,208	1,403
M13	1,49	3,518	3,077
M14	2,543	2,092	0,468
M15	2,592	3,221	0,984

	C1	C2	C3
M16	2,543	2,092	0,468
M17	1,649	3,142	2,229
M18	2,543	2,979	1,49
M19	1,212	2,979	2,284
M20	2,172	0,935	1,723

Berdasarkan tabel 8, terlihat bahwa Anggota C1 terdiri dari (M1,M5,M7,M8,M11,M13,M17,M19), Anggota C2 terdiri dari (M3,M4,M6,M20) dan Anggota C3 terdiri dari (M2,M9,M10,M12,M14,M15,M16,M18). Dalam hasil akhir penduduk penerima BST, terdapat pengelompokan menjadi tiga cluster. Cluster 1 terdiri dari Ahmad Surya, Joko Santoso, Eko Setiawan, Dewi Lestari, Agus Wijaya, Jaya Pratama, Rudi Hermawan dan Andi Wijaya. Cluster ini memiliki tingkat prioritas yang tinggi sebagai penerima BST, yang berarti mereka memenuhi kriteria yang lebih sesuai dengan program tersebut. Cluster 2 hanya terdiri dari Budi Pratama, Rina Fitriani, Maya Putri dan Rina Pratiwi yang memiliki tingkat prioritas yang sedikit rendah dibandingkan dengan Cluster 1. Sedangkan Cluster 3 terdiri dari Siti Indah, Hadi Nugroho, Rini Kartika, Ana Puspita, Tini Rahayu, Dedi Santoso, Yuni Setyawati dan Nita Safitri yang juga memiliki tingkat prioritas yang lebih rendah. Hal ini menunjukkan bahwa dalam pengelompokan data menggunakan metode K-means clustering, penduduk dalam Cluster 1 dianggap sebagai penerima BST dengan prioritas tertinggi, diikuti oleh penduduk dalam Cluster 2 dan Cluster 3. Selanjutnya digunakan aplikasi Rapid Miner untuk menghasilkan tes uji, gambar 1 di bawah adalah hasil uji algoritma Rapid Miner.

Row No.	id	cluster	pengeluaran	pendidikan	tanggungan	pekerjaan
1	1	cluster_0	3	4	3	3
2	2	cluster_2	3	3	4	2
3	3	cluster_1	2	2	3	4
4	4	cluster_1	2	2	4	5
5	5	cluster_0	2	2	1	2
6	6	cluster_1	3	1	2	3
7	7	cluster_0	4	3	3	4
8	8	cluster_0	3	4	5	3
9	9	cluster_2	4	2	5	2
10	10	cluster_2	3	2	3	1
11	11	cluster_0	3	3	2	5
12	12	cluster_0	3	2	4	2
13	13	cluster_0	2	4	4	3
14	14	cluster_2	3	2	4	2
15	15	cluster_2	3	3	4	1
16	16	cluster_0	3	2	4	2
17	17	cluster_2	3	2	5	3
18	18	cluster_0	2	2	5	2
19	19	cluster_1	4	3	4	3

Gambar 1. Hasil tes Rapid Miner

4. KESIMPULAN

Klasifikasi penduduk penerima BST berdasarkan faktor ekonomi bertujuan untuk memastikan bantuan tepat sasaran. Faktor ekonomi seperti pendapatan, aset, dan pengeluaran digunakan dalam klasifikasi ini. Namun, ada beberapa tantangan, termasuk kesulitan mendapatkan data yang akurat dan perbedaan pemahaman kriteria penerima BST. Data mining, khususnya metode K-Means Clustering, digunakan untuk mengatasi masalah ini. Hasilnya, penduduk dibagi menjadi tiga cluster, dengan Cluster 1 memiliki prioritas tertinggi sebagai penerima BST, diikuti oleh Cluster 2 dan Cluster 3. Ini mengindikasikan bahwa metode ini membantu memprioritaskan bantuan kepada mereka yang paling membutuhkan.

REFERENCES

- [1] E. Redy Susanto and A. Savitri Puspaningrum, "Rancang Bangun Rekomendasi Penerima Bantuan Sosial Berdasarkan Data Kesejahteraan Rakyat," *Tekno Kompak*, vol. 15, no. 1, pp. 1–12, 2019.
- [2] D. N. Alfiansyah, V. R. S. Nastiti, and N. Hayatin, "Penerapan Metode K-Means pada Data Penduduk Miskin Per Kecamatan Kabupaten Blitar," *J. Repos.*, vol. 4, no. 1, pp. 49–58, 2022, doi: 10.22219/repositor.v4i1.1416.

- [3] A. P. Windarto, “Penerapan Datamining Pada Ekspor Buah-Buahan Menurut Negara Tujuan Menggunakan K-Means Clustering Method,” *Techno.Com*, vol. 16, no. 4, pp. 348–357, 2017, doi: 10.33633/tc.v16i4.1447.
- [4] N. Rofiqo, A. P. Windarto, and D. Hartama, “Penerapan Clustering Pada Penduduk Yang Mempunyai Keluhan Kesehatan Dengan Datamining K-Means,” *KOMIK (Konferensi Nas. Teknol. Inf. dan Komputer)*, vol. 2, no. 1, pp. 216–223, 2018, doi: 10.30865/komik.v2i1.929.
- [5] S. Widaningsih, “Perbandingan Metode Data Mining Untuk Prediksi Nilai Dan Waktu Kelulusan Mahasiswa Prodi Teknik Informatika Dengan Algoritma C4,5, Naïve Bayes, Knn Dan Svm,” *J. Tekno Insentif*, vol. 13, no. 1, pp. 16–25, 2019, doi: 10.36787/jti.v13i1.78.
- [6] H. Maulidiya and A. Jananto, “Asosiasi Data Mining Menggunakan Algoritma Apriori dan FP-Growth sebagai Dasar Pertimbangan Penentuan Paket Sembako,” *Proceeding SENDIU 2020*, vol. 6, pp. 36–42, 2020.
- [7] A. S. L. T. H. Hafizah, “Data Mining Estimasi Biaya Produksi Ikan Kembung Rebus Dengan Regresi Linier Berganda,” *J. Sist. Inf. Triguna Dharma (JURSI TGD)*, no. Vol 1, No 6 (2022): EDISI NOVEMBER 2022, pp. 888–897, 2022, [Online]. Available: <https://ojs.trigunadharma.ac.id/index.php/jsi/article/view/5732/1938>
- [8] Y. L. Nainel, E. Buulolo, and I. Lubis, “Penerapan Data Mining Untuk Estimasi Penjualan Obat Berdasarkan Pengaruh Brand Image Dengan Algoritma Expectation Maximization (Studi Kasus: PT. Pyridam Farma Tbk),” *JURIKOM (Jurnal Ris. Komputer)*, vol. 7, no. 2, p. 214, 2020, doi: 10.30865/jurikom.v7i2.2097.
- [9] M. Azhari, Z. Situmorang, and R. Rosnelly, “Perbandingan Akurasi, Recall, dan Presisi Klasifikasi pada Algoritma C4.5, Random Forest, SVM dan Naive Bayes,” *J. Media Inform. Budidarma*, vol. 5, no. 2, p. 640, 2021, doi: 10.30865/mib.v5i2.2937.
- [10] N. Salsabila, “Klasifikasi Barang Menggunakan Metode Clustering K-Means Dalam Penentuan Prediksi Stok Barang,” *Cent. Libr. Maulana Malik Ibrahim State Islam. Univ. Malang*, p. 89, 2018.
- [11] F. Harahap, “Perbandingan Algoritma K Means dan K Medoids Untuk Clustering Kelas Siswa Tunagrahita,” *TIN Terap. Inform. Nusantara*, vol. 2, no. 4, pp. 191–197, 2021.
- [12] M. A. Rofiq, A. Qoiriah, S. Kom, and M. Kom, “Pengelompokan Kategori Buku Berdasarkan Judul Menggunakan Algoritma Agglomerative Hierarchical Clustering Dan K-Medoids,” *J. Informatics Comput. Sci.*, vol. 2, no. 03, pp. 220–227, 2021.
- [13] B. Harli Trimulya Suandi As and L. Zahrotun, “PENERAPAN DATA MINING DALAM MENGELOMPOKKAN DATA RIWAYAT AKADEMIK SEBELUM KULIAH DAN DATA KELULUSAN MAHASISWA MENGGUNAKAN METODE AGGLOMERATIVE HIERARCHICAL CLUSTERING (Implementation Of Data Mining In Grouping Academic History Data Before Students And Stud,” *J. Teknol. Informasi, Komput. dan Apl.*, vol. 3, no. 1, pp. 62–71, 2021, [Online]. Available: <http://jtika.if.unram.ac.id/index.php/JTIKA/>
- [14] P. Metode, K. U. Clustering, M. Berdasarkan, and N. Akademik, “PROGRAM STUDI TEKNOLOGI INFORMASI Fakultas Teknik Universitas Muhammadiyah Yogyakarta,” pp. 1–21, 2015.
- [15] M. A. K-means, “1 , 2 , 3 1,” vol. 1, no. 2, pp. 161–166, 2021.
- [16] W. Purba, W. Siawin, and . H., “Implementasi Data Mining Untuk Pengelompokan Dan Prediksi Karyawan Yang Berpotensi Phk Dengan Algoritma K-Means Clustering,” *J. Sist. Inf. dan Ilmu Komput. Prima(JUSIKOM PRIMA)*, vol. 2, no. 2, pp. 85–90, 2019, doi: 10.34012/jusikom.v2i2.429.
- [17] . F., F. T. Kesuma, and S. P. Tamba, “Penerapan Data Mining Untuk Menentukan Penjualan Sparepart Toyota Dengan Metode K-Means Clustering,” *J. Sist. Inf. dan Ilmu Komput. Prima(JUSIKOM PRIMA)*, vol. 2, no. 2, pp. 67–72, 2020, doi: 10.34012/jusikom.v2i2.376.
- [18] S. A. Rahmah, “KLASTERISASI POLA PENJUALAN PESTISIDA MENGGUNAKAN METODE K-MEANS CLUSTERING (STUDI KASUS DI TOKO JUANDA TANI KECAMATAN HUTABAYU RAJA),” vol. 1, no. 1, pp. 1–5, 2020.
- [19] E. Rouza, Basorudin, and Efrida, “Identifikasi dan Klasifikasi UMKM di Kabupaten Rokan Hulu Menggunakan Metode K-Means,” *J. Ilm. Univ. Pengaraian*, vol. 7, no. 01, pp. 32–40, 2021.