

Penerapan Metode K-Means Clustering Untuk Mengelompokkan Data Kelayakan Penerima Bantuan Renovasi Rumah

Gina Sonia*, Raissa Amanda Putri

Fakultas Sains dan Teknologi, Ilmu Komputer, Universitas Islam Negeri Sumatra Utara, Medan, Indonesia

Email: ^{1,*}soniadomo28@gmail.com, ²raissa.ap@uinsu.ac.id

Email Penulis Korespondensi: soniadomo28@gmail.com

Submitted: 15/09/2023; Accepted: 25/09/2023; Published: 27/09/2023

Abstrak—Pemerintah memberikan salah satu bantuan yaitu renovasi rumah didaerah Desa Kuala Bangka di kecamatan kualuh hilir, Kabupaten Labuhan Batu Utara dari tahun 2015 sampai sekarang. Namun dengan bantuan tersebut Desa Kuala Bangka terkadang mendapatkan masalah dalam penentuan kelayakan penerimaan bantuan dari pemerintah, dengan permasalahan pada penentuan kelayakan penerimaan bantuan dari pemerintah maka penulis akan melakukan penelitian dengan mengelompokkan atau clustering kelayakan penerima bantuan renovasi rumah dengan menerapkan Algoritma K-Means. Proses Algoritma K-Means Clustering ini nantinya dapat mengelompokkan setiap data kedalam cluster sehingga data yang mempunyai karakteristik yang sama akan dikelompokkan dalam sebuah cluster yang samadan sebaliknya jika data yang mempunyai karakteristik yang berbeda maka dikelompokkan kedalam sebuah cluster yang lain, tujuan dari pengelompokkan akan menentukan cluster 0 dan cluster 1 penerima bantuan renovasi rumah, cluster 0 merupakan layak dan cluster 1 merupakan tidak layak. Attribute yang digunakan dalam penelitian ini adalah jumlah anggota keluarga, pekerjaan, kondisi rumah, dan pendapatan. Hasil yang didapatkan dari 170 data adalah cluster 0 sebanyak 91 dan cluster 1 sebanyak 79 data, dari hasil tersebut maka 91 orang layak dalam penerima bantuan renovasi rumah dan 79 tidak layak dalam penrima bantuan renovasi rumah.

Kata Kunci: Data Mining; Algoritma K-Means Clustering; Penerima Bantuan

Abstract—The government provided one form of assistance, namely the renovation of houses in the Kuala Bangka Village area in Kualuh Hilir District, North Labuhan Batu Regency from 2015 until now. However, with this assistance, Kuala Bangka Village sometimes has problems in determining the feasibility of receiving assistance from the government, therefore the author will conduct research by categorizing the eligibility of recipients of house renovation assistance by applying the K-Means algorithm. The K-Means Clustering algorithm is an algorithm that can classify data accurately according to the previous problem. The grouping aims to determine cluster 0 and cluster 1 as recipients of house renovation assistance, cluster 0 is feasible and cluster 1 is not feasible. The attributes used in this research are number of family members, employment, housing conditions, and income. The results obtained from 170 data were cluster 0 with 91 data and cluster 1 with 79 data. From these results, 91 people were eligible to receive home renovation assistance and 79 were not eligible to receive home renovation assistance.

Keywords: Data Mining; K-Means Clustering Algorithm; Aid Recipients

1. PENDAHULUAN

Kemiskinan merupakan hal yang sering terjadi di negara berkembang dan merupakan salah satu dari masalah yang cukup sulit untuk dipecahkan, salah satunya terjadi di Indonesia. Pemerintah telah melakukan beberapa upaya untuk mengentaskan kemiskinan, seperti program bantuan renovasi rumah, bantuan sosial, bantuan langsung tunai (BLT), program keluarga harapan (PKH), dll. Berdasarkan hasil survei pemerintah warga tidak mendapatkan bantuan yang mereka butuhkan sehingga distribusi tunjangan yang tidak merata dan tidak konsisten.[1]

Desa Kuala Bangka merupakan salah satu desa di Kecamatan Kualuh Hilir, Kabupaten Labuhan Batu Utara, banyak penduduk Desa Kuala Bangka yang berprofesi sebagai petani, pedagang, buruh tani, PNS, Honorer, bekerja sebagai pekerja keras ada juga bekerja diluar kota bahkan di luar negeri seperti desa pada umumnya namun Desa Kuala Bangka lebih besar berprofesi sebagai buruh, petani, dan pedagang. Oleh sebab itu Desa Kuala Bangka banyak mendapat bantuan dari program pemerintah berupa renovasi rumah bagi masyarakat yang kurang mampu. Namun dengan bantuan tersebut banyak penerima bantuan yang tidak tepat sasaran dengan adanya program bantuan renovasi rumah, sehingga dampak yang ditimbulkan mengakibatkan bantuan yang tidak sesuai dengan Masyarakat dikarenakan kurangnya penyelidikan yang lebih dalam, dengan adanya permasalahan ini penulis meneliti lebih lanjut permasalahan tersebut, penelitian ini dilakukan selama lebih kurang 2 bulan dengan hasil penelitian yang didapatkan, dalam penelitian pada desa kuala bangka didapatkan penentuan kelayakan penerima bantuan tidak ditentukan berdasarkan range rata – rata pendapatan, pekerjaan, maupun jumlah keluarga, tapi dilakukan hanya pada Masyarakat yang memiliki surat tidak mampu atau hanya mendaftarkan secara langsung bahwa dirinya tidak mampu. Maka dengan permasalahan tersebut penulis menentukan kelayakan penerima bantuan dengan attribute pendapatan, jumlah anggota keluarga, pekerjaan, dan kelakayakan rumah, dengan menerapkan Algoritma K-Means Clustering metode yang digunakan dengan menentukan kelayakan berdasarkan attribute yang digunakan, kelayakan akan dibagi menjadi 2 yaitu cluster 0 atau layak dan cluster 1 tidak layak. [2]. [3]

Algoritma K-Means Clustering banyak dilakukan oleh para peneliti sebelumnya untuk mengelompokkan data. Diantara peneliti sebelumnya yaitu dengan judul Penerapan Data Mining Dalam Mengelompokkan Calon Penerima Beasiswa Dengan Menggunakan Algoritma K-Means dari penelitian tersebut diperoleh hasil bahwasanya Algoritms K-Means Clustering dapat diterapkan dalam pengelolaan kasus data penerimaan beasiswa dengan adanya bantuan Rapid Miner maka proses perhitungannya lebih mudah dalam membantu mendapatkan hasil di penelitian ini.[2]–[4]

Penelitian ini didukung dengan penelitian terkait pada tahun 2019 yang dilakukan oleh Dina Sunia dkk dengan judul Penerapan Data Mining Untuk Clustering Data Penduduk Miskin Menggunakan Algoritma K-Means, pada penelitian ini juga membahas tentang kasus kemiskinan yang tidak tepat sasaran, dalam penelitian ini menentukan 5 cluster dengan hasil 13 penduduk masuk sebagai anggota klaster 1, 153 penduduk berada di klaster 2, 129 penduduk di klaster 3, 138 penduduk berada di klaster 4, dan 79 penduduk berada di klaster 5, dengan hasil tersebut penelitian sukses dilakukan dalam menentukan clustering data penduduk miskin.[5]

Penelitian lainnya pada tahun 2020 yang dilakukan oleh Nur Afriani Manihuruk dkk, dengan judul Penerapan Data Mining Dalam Mengelompokkan Calon Penerima Beasiswa Dengan Menggunakan Algoritma K-Means, penelitian ini bertujuan untuk mencari pengelompokan data yang ada pada siswa yang berhak menerima beasiswa, penelitian ini dilakukan pengujian menggunakan rapid miner dengan 3 cluster yaitu cluster tingkat tinggi, cluster tingkat sedang dan cluster tingkat rendah, dan hasil dari pengujian rapid miner diperoleh hasil cluster 1 dengan hasil 73 data, cluster 2 dengan hasil 30 data, dan cluster 3 dengan hasil 25 data.[3], [6], [7]

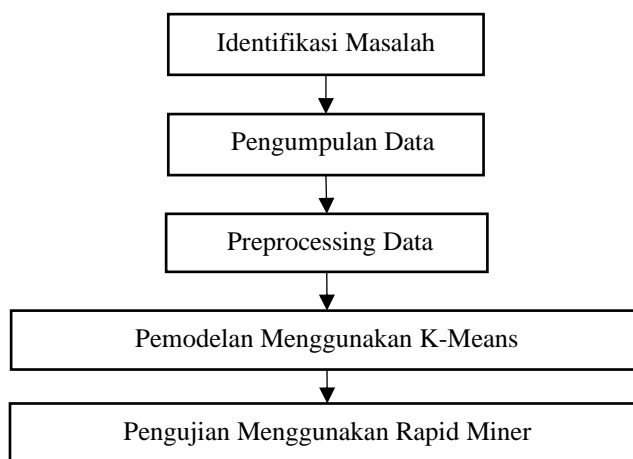
Penelitian lainnya yang mendukung dengan penelitian terkait dilakukan pada tahun 2022 oleh Aviv Fitria Yulia dan Handoyo Widi Nugroho, dengan judul Implementasi Algoritma K-Means Classifier Sebagai Pendukung Keputusan Penerima Dana Bantuan Siswa Miskin (Studi Kasus : SMKN Sukoharjo), pada penelitian ini bertujuan dalam mengambil keputusan penerima bantuan dana siswa yang kurang mampu, attribute yang digunakan dalam penerima bantuan adalah nilai rata-rata rapor, penghasilan orang tua siswa, tanggungan orang tua siswa, dan jarak tempuh ke sekolah, dalam penelitian ini cluster di bagi menjadi 3, yaitu cluster 0 layak menerima bantuan siswa miskin, cluster 1 dapat di pertimbangkan menerima bantuan siswa miskin dan cluster 2 tidak layak menerima bantuan siswa miskin, maka hasil yang didapatkan dari pengujian manual dan pengujian rapid miner pada penelitian ini didapatkan hasil yang sama, yaitu cluster 0 sebanyak 393 siswa, cluster 1 sebanyak 278 siswa dan cluster 2 sebanyak 373 dari dataset yang di uji sebanyak 1044.[8]

Penelitian berikutnya yang mendukung penelitian terkait dilakukan pada tahun 2023 oleh Febriansyah dan Siti Muntari, dengan judul Penerapan Algoritma K-Means untuk Klasterisasi Penduduk Miskin pada Kota Pagar Alam, memiliki tujuan dalam mengelompokan penduduk miskin dengan 3 cluster dan pengujian dilakukan pada rapid miner maka hasil dari proses clustering data dalam menentukan penerima PKH melalui aplikasi RapidMiner sama dengan hasil yang diterapkan di sistem yang telah dibangun dengan jumlah cluster yang terdiri dari tiga cluster. Cluster dimulai dari cluster 0, cluster 1, dan cluster 2. Data yang berada pada cluster 0 berjumlah 156 data, cluster 1 berjumlah 82 data, dan cluster_2 berjumlah 233 data.[9]

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Pada penelitian ini terdapat metodologi penelitian yang digunakan secara garis besar terdiri dari 5 langkah penyelesaian. Berikut ini adalah langkah-langkah penelitian pada Gambar 1 dibawah ini:



Gambar 1. Tahapan Penelitian

Pada Gambar 1 diatas merupakan gambar kerangka penelitian yang akan di gunakan pada penelitian ini. Berikut penjelasan tentang tahapan penelitian

a. Mengidentifikasi Masalah

Langkah Pertama yang dilakukan pada penelitian ini adalah mengidentifikasi masalah peneliti berupaya untuk mengamati dan memahami permasalahan dengan adanya permasalahan bantuan renovasi rumah yang tidak tepat sasaran.

b. Study Literatur

Pada bagian ini penulis melakukan serangkaian kegiatan yang berkenaan metode pengumpulan daftar pustaka, membaca dan mencatat, untuk menyelesaikan persoalan dengan menelusuri sumber-sumber tulisan yang pernah dibuat sebelumnya. serta mengolah bahan penelitian atau mencari referensi teori yang relevan dengan kasus atau permasalahan yang berhubungan dengan penelitian ini.[3], [4], [10]

c. Pengumpulan Data

Untuk memperoleh data, peneliti melakukan 2 cara, yaitu melalui observasi dan wawancara. Dari segi Observasi, peneliti melakukan pengamatan terhadap lokasi penelitian penulis sehingga memperoleh hasil yang akurat, Sedangkan dari segi wawancara, peneliti secara langsung mewawancarai Sekretaris Kepala Desa yaitu bapak Ilham Tanjung.[1], [2], [11], [12]

d. Preprocessing Data

Setelah data berhasil didapatkan kemudian dilakukan processing dari data primer hingga mendapatkan data baru. Data yang diambil secara primer (data mentah) terkadang memiliki inkonsistensi, noise, informasi yang kurang lengkap dan nilai yang hilang. Data preprocessing menerangkan tipe-tipe proses yang melaksanakan data mentah untuk mempersiapkan proses prosedur yang lainnya. Tujuannya preprocessing dalam data mining adalah mentransformasi data ke suatu format yang prosesnya lebih mudah dan efektif untuk kebutuhan pemakai. Setelah data primer (mentah) di preprocessing, maka selanjutnya masuk ketahapan pengolahan data.[1], [13]-[15]

e. Pemodelan Menggunakan K-Means

Pada tahap ini merupakan tahap dalam melakukan penerapan menggunakan algoritma k-means, tahap ini akan melakukan perhitungan secara manual dalam menerapkan algoritma k-means.[7], [16]

f. Pengujian Menggunakan Rapid Miner

Setelah pemodelan selesai kemudian dilakukan pengujian kedalam rapid miner, dilakukan pengujian dalam rapid miner untuk mendapatkan hasil yang sesuai dengan pemodelan.[6], [7]

2.2 Algoritma K-Means Clustering

Algoritma K-Means adalah Metode clustering non hierarchical berbasis jarak yang membagi data kedalam cluster dan algoritma ini bekerja pada atribut numerik. Algoritma K-Means termasuk dalam partitioning clustering yang memisahkan data ke k daerah bagian yang terpisah. Algoritma K Means sangat terkenal karena kemudahan dan kemampuannya untuk mengklaster data besar dan outlier dengan sangat cepat. [1], [2]

Algoritma K-Means pada awalnya mengambil sebagian dari banyaknya komponen dari populasi untuk dijadikan pusat cluster awal. Pada step ini pusat cluster dipilih secara acak dari sekumpulan populasi data. Berikutnya K-Means menguji masing-masing komponen didalam populasi data dan menandai komponen tersebut kesalah satu pusat cluster yang telah didefinisikan tergantung dari jarak minimum antar komponen dengan tiap-tiap pusat cluster. Posisi pusat cluster akan dihitung kembali sampai semua komponen data digolongkan kedalam tiap-tiap cluster dan terakhir akan terbentuk posisi cluster baru[1], [4], [6], [8], [9], [17]

Algoritma K-Means pada dasarnya melakukan 2 proses yakni proses pendeteksian lokasi pusat cluster dan proses pencarian anggota dari tiap-tiap cluster. Proses algoritma K-Means: [1], [4], [10]

- Tentukan k sebagai jumlah cluster yang dibentuk
- Tentukan k centroid (titik pusat cluster) awal secara random
- Hitung jarak setiap objek ke masing-masing centroid dari masing-masing cluster. Untuk menghitung jarak antara objek dari tiap centroid dapat menggunakan Euclidian Distance

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

Keterangan:

d = determinan (Euclidian Distance)

x = titik pusat centroid

y = data

n = jumlah ke i

i = data ke

- Alokasikan masing-masing objek ke dalam centroid yang paling dekat. Untuk melakukan pengalokasian objek kedalam masing-masing cluster pada saat iterasi secara umum dapat dilakukan dengan cara hard k-means dimana secara tegas setiap objek dinyatakan sebagai anggota cluster dengan mengukur jarak kedekatan sifatnya terhadap titik pusat cluster tersebut.
- Lakukan iterasi, kemudian tentukan posisi centroid baru dengan menggunakan persamaan 1
- Ulangi langkah 3 jika posisi centroid baru tidak sama.[1], [4], [18], [19]

3. HASIL DAN PEMBAHASAN

3.1 Identifikasi Masalah

Tahap awal dalam sebuah penelitian adalah mencari permasalahan yang akan diangkat dan diteliti, Berdasarkan Analisa yang dilakukan penulis, penulis menemukan masalah yang akan dijadikan sebagai rumusan dalam tujuan utama untuk menyelesaikan penelitian ini, fokus utama pada rumusan masalah ini adalah menentukan layak dan

tidak layak penerima bantuan yang akan dijadikan sebagai 2 cluster yang akan menggunakan algoritma k-means sebagai penerapan dalam menentukan layak dan tidak layak penerima bantuan.[8]

3.2 Processing Data

Preprocessing adalah tahapan mengubah data mentah menjadi format baru yang lebih efisien dan efektif agar dapat melakukan tahap proses selanjutnya. Pada penelitian ini preprocessing ada beberapa tahapan yaitu tahapan pembersihan data, pemberian bobot data, dan normalisasi data.[8]

a. Pembersihan data

Pembersihan data adalah proses identifikasi dan memperbaiki kesalahan dan ketidakakuratan dalam pengumpulan data. Pembersihan data dapat dilakukan secara manual ataupun otomatis. Pada penelitian ini penulis melakukan pembersihan secara manual. Dimulai dengan pemilihan column atau attribute yang digunakan, berikut column awal sebelum digunakan.

Tabel 1. Column Awal

Nama
Nama Keluarga
J/K
Tempat Lahir
Tgl Lahir
Nik
Agama
Jenis Pekerjaan
Status Perkawinan
Pendidikan
Dusun
Kualitas rumah
Nama Ayah
Nama Ibu
No Kk
Pendapatan

Pada tabel 1 merupakan column awal. Dapat dilihat pada tabel 1 diatas terdapat 15 column, disini penulis hanya mengambil column sesuai dengan attribute yang digunakan, attribute yang digunakan adalah pekerjaan, jumlah anggota keluarga, kualitas rumah, dan pendapatan, berikut data sementara sebelum menghitung jumlah anggota dalam keluarga. Pada tabel 2 merupakan tabel data awal

Tabel 2. Data Awal

No	Keluarga	Kepala Keluarga	Pekerjaan	Kualitas Rumah	Pendapatan
1	J. Sri Hardono Heltina Br Hombing Briyan Harnata Sinaga Christian Harsavin Sinaga Puspita Nathaniela Br Sinaga Jordhy Harditya Sinaga	J. Sri Hardono	Buruh	Rumah Tidak Layak	1700000
2	Robinson Sinaga Setialam Br Lumban Batu Marihhot Candro P. Sinaga Morina Wati Sinaga Monaldi Sinaga Median Sinaga Mini Ceria Sinaga	Robinson Sinaga	Buruh	Rumah Tidak Layak	1900000
3	Balut Parulian Sihombing Romauli Br Simbolon Dani Sihombing Holong Sihombing	Balut Parulian Sihombing	Petani/Pekebun	Rumah Layak	3600000
4	Koslen Naibaho Rena Renata Sinaga Iusmider Naibaho	Koslen Naibaho	Pegawai	Rumah Layak	2700000

5	Nocince Br. Naibaho Wilinton Naibaho David Pardosi Lenni Marlina Br Sitanggang Della Puspita Br Pardosi Surya Wandani Pardosi Kevin Robintang Pardosi Mexsel Ananta Pardosi Venti Sudarti Br Sinaga Christian Rumapea	David Pardosi	Buruh	Rumah Tidak Layak	1700000
6	Iskandar Juliana Anjar Partawi Sekar Kemuning Sri Utami Muhammad Samudra	Iskandar	Buruh	Rumah Layak	1600000
7	Sanga Braly Erianto Sihombing May Santi Pardede Kayla Br. Sihombing	Sanga Braly Erianto Sihombing	Buruh	Rumah Layak	1600000
8	Amin Rais Sibarani Rama Yanti Hutabarat	Amin Rais Sibarani	Buruh	Rumah Layak	1600000
9	Harapan Togatorop Mesta Limbong	Harapan Togatorop	Pedagang	Rumah Layak	3400000
10	Zebpri Sinaga Ester Dian Pratiwi Br. Girsang	Zebpri Sinaga	Buruh	Rumah Layak	1900000
...
170	Sugiarto Yuspina Simanjuntak Zevannya Siboro Rusmini Fajar Styawan Satria Ramadhani	Sugiarto	Petani/Pekebun	Rumah Tidak Layak	2500000

Dapat dilihat pada tabel 2 diatas bahwa attribute jumlah anggota keluarga belum ditentukan, dalam menentukan attribute jumlah anggota keluarga didapatkan dari jumlah keluarga salah satu contoh dalam menentukan jumlah anggota keluarga dapat dilihat pada baris ke 1 keluarga dari J. Sri Hardono ada 6 jumlah anggota keluarga, dalam menghitung anggota keluarga dimulai dari kepala keluarga, maka hasil untuk jumlah anggota keluarga dapat dilihat dibawah ini. Pada tabel 3 merupakan tabel hasil pembersihan data

Tabel 3. Hasil Pembersihan Data

No	Nama Kepala Keluarga	Pekerjaan	Jumlah Anggota Keluarga	Kualitas Rumah	Pendapatan
1	J. Sri Hardono	Buruh	6	Rumah Tidak Layak	1700000
2	Robinson Sinaga	Buruh	7	Rumah Tidak Layak	1900000
3	Balut Parulian Sihombing	Petani/Pekebun	4	Rumah Layak	3600000
4	Koslen Naibaho	Pegawai	5	Rumah Layak	2700000
5	David Pardosi	Buruh	8	Rumah Tidak Layak	1700000
6	Iskandar	Buruh	5	Rumah Layak	1600000
7	Sanga Braly Erianto Sihombing	Buruh	3	Rumah Layak	1600000
8	Amin Rais Sibarani	Buruh	2	Rumah Layak	1600000
9	Harapan Togatorop	Pedagang	2	Rumah Layak	3400000
10	Zebpri Sinaga	Buruh	2	Rumah Layak	1900000
11	Rusmawarni Sigalingging	Petani/Pekebun	2	Rumah Layak	3800000
12	Desmaria Br Siallagan	Buruh	2	Rumah Layak	1800000
13	Januari Ambarita	Buruh	8	Rumah Tidak Layak	1500000
14	Sarmin	Petani/Pekebun	6	Rumah Tidak Layak	2500000
15	Marungkil Butar Butar	Petani/Pekebun	4	Rumah Tidak Layak	2200000
16	Mhd Sumisnan	Buruh	3	Rumah Tidak Layak	1500000



17	Marpos Niroha Sihombing	Pedagang	5	Rumah Layak	3200000
18	Toni Mangasi Butar Butar	Buruh	8	Rumah Tidak Layak	1600000
19	Sudirman Sianturi	Petani/Pekebun	5	Rumah Layak	3200000
20	Mula Jantes Panjaitan	Buruh	9	Rumah Tidak Layak	1800000
...
170	Sugiarto	Petani/Pekebun	6	Rumah Tidak Layak	2500000

b. Pemberian bobot

Pemberian bobot merupakan tahapan merubah data yang berupa teks kedalam data label agar meningkatkan keakuratan dalam clustering menggunakan k-means. Pada bagian ini dapat membantu algoritma k-means untuk menyesuaikan dengan lebih baik terhadap karakteristik data dan menghasilkan hasil yang lebih relevan dengan attribute yang digunakan yaitu pekerjaan dan kualitas rumah, Berikut ini merupakan variabel penilaian kriteria.[17]

Tabel 4. Bobot Nama Pekerjaan

Atribut	Kriteria	Bobot
Nama Pekerjaan	Buruh	4
	Petani/Pekebun	3
	Pedagang	2
	Pegawai	1

Pada Tabel 4 dibawah merupakan bobot nama pekerjaan

Tabel 5. Bobot Kualitas Rumah

Atribut	Kriteria	Bobot
Kualitas Rumah	Rumah Tidak Layak	2
	Rumah Layak	1

Pada Tabel 5 dibawah merupakan bobot kualitas rumah .Sehingga hasil tabel setelah pemberian bobot menjadi seperti dibawah ini:

Tabel 6. Hasil Pembobotan

No	Nama Kepala Keluarga	Bobot Pekerjaan	Jumlah Anggota Keluarga	Bobot Kualitas Rumah	Pendapatan
1	J. Sri Hardono	4	6	2	1700000
2	Robinson Sinaga	4	7	2	1900000
3	Balut Parulian Sihombing	3	4	1	3600000
4	Koslen Naibaho	1	5	1	2700000
5	David Pardosi	4	8	2	1700000
6	Iskandar	4	5	1	1600000
7	Sanga Braly Erianto Sihombing	4	3	1	1600000
8	Amin Rais Sibarani	4	2	1	1600000
9	Harapan Togatorop	2	2	1	3400000
10	Zebpri Sinaga	4	2	1	1900000
11	Rusmawarni Sigalingging	3	2	1	3800000
12	Desmaria Br Siallagan	4	2	1	1800000
13	Januari Ambarita	4	8	2	1500000
14	Sarmin	3	6	2	2500000
15	Marungkil Butar Butar	3	4	2	2200000
16	Mhd Sumisnan	4	3	2	1500000
17	Marpos Niroha Sihombing	2	5	1	3200000
18	Toni Mangasi Butar Butar	4	8	2	1600000
19	Sudirman Sianturi	3	5	1	3200000
20	Mula Jantes Panjaitan	4	9	2	1800000
...
170	Sugiarto	3	6	2	2500000

Pada tabel 6 merupakan tabel hasil pembobotan

c. Normalisasi data

Normalisasi data adalah proses mengatur atau memanipulasi struktur data agar menjadi lebih konsisten dan mempunyai relasi yang baik. Pada bagian ini besaran angka antara atribut pendapatan dan jumlah anggota keluarga cukup jauh dimana atribut dengan jumlah data yang besar dapat mendominasi hasil cluster dan membuat hasil cluster



data menjadi tidak valid. Jika memang mempunyai satuan yang berbeda secara signifikan, standarisasi data menggunakan rumus Z-Score atau didalam rapid miner dikenal Z-Transformation dapat menghasilkan keseimbangan nilai perbandingan antar data saat sebelum dan sesudah proses.[7], [16], [20]

Rumus yang digunakan z-score sebagai berikut:

$$z_i = \frac{x_i - \bar{x}}{s} \tag{2}$$

Keterangan:

z_i = Nilai *z-score* ke i

x_i = Nilai datum ke i

\bar{x} = Nilai *average*

s = Nilai *standart deviasi*

Dalam menentukan average atau rumus rata – rata menggunakan rumus seperti berikut:

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} \tag{3}$$

Keterangan:

\bar{x} = Nilai *average*

n = Banyaknya jumlah data

x_i = Nilai atau objek ke i

Untuk menentukan nilai standart deviasi memiliki rumus seperti berikut ini:

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}} \tag{4}$$

Keterangan:

s = Nilai *standart deviasi*

\bar{x} = Nilai *average*

x_i = Nilai atau objek ke i

Dalam menghitung z-score terlebih dahulu menentukan nilai average kemudian menentukan nilai standart deviasi kemudian menentukan nilai z-score, berikut table hasil nilai average dari setiap attribute yang digunakan: Pada Tabel 7 dibawah ini merupakan hasil nilai avarage

Tabel 7. Hasil Nilai Avarage

No	Bobot Pekerjaan	Jumlah Anggota Keluarga	Bobot Kualitas Rumah	Pendapatan
1	4	6	2	1700000
2	4	7	2	1900000
3	3	4	1	3600000
4	1	5	1	2700000
5	4	8	2	1700000
6	4	5	1	1600000
7	4	3	1	1600000
8	4	2	1	1600000
9	2	2	1	3400000
10	4	2	1	1900000
11	3	2	1	3800000
12	4	2	1	1800000
13	4	8	2	1500000
14	3	6	2	2500000
15	3	4	2	2200000
16	4	3	2	1500000
17	2	5	1	3200000
18	4	8	2	1600000
19	3	5	1	3200000
20	4	9	2	1800000
...
170	3	6	2	2500000
\bar{x}	3,158823529	4,441176471	1,305882353	2.377.647

Adapun proses menghitung rata – rata dapat dilihat dibawah ini:

$$\bar{x} = \frac{4+4+3+1+4+4+4+4+2+4+3+4+4+3+3+4+2+4+3+4+\dots+3}{170}$$



$$\bar{x} = \frac{537}{170} = 3,158823529$$

Dalam menghitung rata - rata pada attribute lain juga sama dengan menjumlahkan semua data lalu membagi total data, sehingga didapatkan nilai rata – rata. Setelah didapatkan nilai rata – rata dari masing – masing attribute kemudian akan dihitung nilai setiap objek. Pada Tabel 8 dibawah ini merupakan nilai bobot pekerjaan deviasi x_i

Tabel 8. Nilai Bobot Pekerjaan Deviasi x_i

Bobot Pekerjaan	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$
4	0,841176471	0,707577855
4	0,841176471	0,707577855
3	-0,158823529	0,025224913
1	-2,158823529	4,660519031
4	0,841176471	0,707577855
4	0,841176471	0,707577855
4	0,841176471	0,707577855
4	0,841176471	0,707577855
2	-1,158823529	1,342871972
4	0,841176471	0,707577855
3	0,841176471	0,707577855
4	-0,158823529	0,025224913
4	0,841176471	0,707577855
3	0,841176471	0,707577855
3	-0,158823529	0,025224913
4	-0,158823529	0,025224913
2	0,841176471	0,707577855
4	-1,158823529	1,342871971
3	0,841176471	0,707577855
4	-0,158823529	0,025224913
...
3	-0,158823529	0,025224913
	$\sum_{i=1}^{170}(x_i - \bar{x})^2$	170,7117647

Untuk menghitung proses nilai objek deviasi dapat dilihat dibawah ini:

$$(x_1 - \bar{x}) = (4 - 3,158823529) = 0,841176471$$

$$(x_1 - \bar{x})^2 = (0,841176471)^2 = 0,707577855$$

Maka untuk hasil tabel yang lainnya dapat dilihat dibawah ini. Pada Tabel 9 dibawah ini merupakan nilai jumlah anggota keluarga deviasi x_i

Tabel 9. Nilai Jumlah Anggota Keluarga Deviasi x_i

Jumlah Anggota Keluarga	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$
6	1,558823529	2,429930796
7	2,558823529	6,547577855
4	-0,441176471	0,194636678
5	0,558823529	0,312283737
8	3,558823529	12,66522491
5	0,558823529	0,312283737
3	-1,441176471	2,076989619
2	-2,441176471	5,959342561
2	-2,441176471	5,959342561
2	-2,441176471	5,959342561
2	-2,441176471	5,959342563
2	-2,441176471	5,959342563
8	-2,441176471	5,959342563
6	3,558823529	12,66522491
4	1,558823529	2,429930795
3	-0,441176471	0,194636679
5	-1,441176471	2,076989621
8	0,558823529	0,312283737
5	3,558823529	12,66522491
...
6	1,558823529	2,429930796



$$\sum_{i=1}^{170} (x_i - \bar{x})^2 \quad 709,9117647$$

Pada Tabel 10 dibawah ini merupakan nilai bobot kualitas rumah pendapatan deviasi x_i

Tabel 10. Nilai Bobot Kualitas Rumah Pendapatan Deviasi x_i

Bobot Kualitas Rumah	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$
2	0,694117647	0,481799308
2	0,694117647	0,481799308
1	-0,305882353	0,093564014
1	-0,305882353	0,093564014
2	0,694117647	0,481799308
1	-0,305882353	0,093564014
1	-0,305882353	0,093564014
1	-0,305882353	0,093564014
1	-0,305882353	0,093564014
1	-0,305882353	0,093564014
1	-0,305882353	0,093564014
1	-0,305882353	0,093564014
2	-0,305882353	0,093564014
2	0,694117647	0,481799308
2	0,694117647	0,481799308
2	0,694117647	0,481799308
1	0,694117647	0,481799308
2	-0,305882353	0,093564014
1	0,694117647	0,481799308
2	-0,305882353	0,093564014
...
2	0,694117647	0,481799308
	$\sum_{i=1}^{170} (x_i - \bar{x})^2$	36,09411765

Pada Tabel 11 dibawah ini merupakan nilai pendapatan deviasi x_i

Tabel 11. Nilai Pendapatan Deviasi x_i

Pendapatan	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$
1700000	-677647	4,59205
1900000	-477647	2,28147
3600000	1222353	1,49415
2700000	322353	1,03911
1700000	-677647	4,59205
1600000	-777647	6,04735
1600000	-777647	6,04735
1600000	-777647	6,04735
3400000	1022353	1,04521
1900000	-477647	2,28147
3800000	1422353	2,02309
1800000	-577647	3,33676
1500000	-877647	7,70264
2500000	122353	1497025
2200000	-177647	3155845
1500000	-877647	7,70264
3200000	822353	6,76264
1600000	-777647	6,04735
3200000	822353	6,76264
1800000	-577647	3,33676
...
2500000	122352,9412	14970242215
	$\sum_{i=1}^{170} (x_i - \bar{x})^2$	1,07695

Setelah semua nilai objek deviasi didapatkan kemudian menghitung standart deviasi dengan rumus pada persamaan 4 yang dijelaskan sebelumnya

Standart deviasi bobot pekerjaan



$$s = \sqrt{\frac{170,7117647}{170-1}} = \sqrt{1,010128785} = 1,005051633$$

Standart deviasi jumlah anggota keluarga

$$s = \sqrt{\frac{709,9117647}{170-1}} = \sqrt{4,20066133} = 2,049551495$$

Standart deviasi bobot kualitas rumah

$$s = \sqrt{\frac{36,09411765}{170-1}} = \sqrt{0,213574661} = 0,462141386$$

Standart deviasi pendapatan

$$s = \sqrt{\frac{1,07695}{170-1}} = \sqrt{6,37249} = 798,278$$

Setelah standart deviasi dan average didapatkan kemudian dapat melanjutkan menghitung z-score, berikut salah satu proses dalam menghitung z-score:

$$Z(\text{bobot pekerjaan}) = \frac{4 - 3,158823529}{1,005051633} = 0,836948514$$

$$Z(\text{jumlah keluarga}) = \frac{6 - 4,441176471}{2,049551495} = 0,760568121$$

$$Z(\text{kualitas rumah}) = \frac{2 - 1,305882353}{0,462141386} = 1,501959504$$

$$Z(\text{pendapatan}) = \frac{1700000 - 2.377.647}{798,278} = -0,848885316$$

Berikut tabel normalisasi setelah dilakukan perhitungan z-score:

Tabel 12. Hasil Normalisasi

No	Bobot Pekerja	Jumlah Anggota Keluarga	Bobot Kualitas Rumah	Pendapatan
1	0,836948514	0,760568121	1,501959504	-0,848885316
2	0,836948514	1,248479746	1,501959504	-0,598346247
3	-0,158025244	-0,215255129	-0,66188046	1,531235839
4	-2,147972759	0,272656496	-0,66188046	0,403810029
5	0,836948514	1,736391371	1,501959504	-0,848885316
6	0,836948514	0,272656496	-0,66188046	-0,97415485
7	0,836948514	-0,703166753	-0,66188046	-0,97415485
8	0,836948514	-1,191078378	-0,66188046	-0,97415485
9	-1,152999001	-1,191078378	-0,66188046	1,28069677
10	0,836948514	-1,191078378	-0,66188046	-0,598346247
11	-0,158025244	-1,191078378	-0,66188046	1,781774908
12	0,836948514	-1,191078378	-0,66188046	-0,723615781
13	0,836948514	1,736391371	1,501959504	-1,099424385
14	-0,158025244	0,760568121	1,501959504	0,15327096
15	-0,158025244	-0,215255129	1,501959504	-0,222537644
16	0,836948514	-0,703166753	1,501959504	-1,099424385
17	-1,152999001	0,272656496	-0,66188046	1,030157701
18	0,836948514	1,736391371	1,501959504	-0,97415485
19	-0,158025244	0,272656496	-0,66188046	1,030157701
20	0,836948514	2,224302996	1,501959504	-0,723615781
...
170	-0,158025244	0,760568121	1,501959504	0,15327096

3.3 Algoritma K-Means Clustering

Setelah didapatkan hasil dari normalisasi kemudian data bisa dilakukan clustering menggunakan algoritma k-means, Langkah pertama dalam dilakukan pada tahap k means adalah menentukan titik pusat centroid, penentuan titik pusat dilakukan dengan data random dan tidak memiliki kesamaan antara data, berikut titik pusat centroid:[5], [6]. Pada Tabel 13 dibawah ini merupakan titik pusat centroid



Tabel 13. Titik Pusat Centroid

Cluster	Bobot Pekerjaan	Jumlah Anggota Keluarga	Bobot Kualitas Rumah	Pendapatan
C0	0,836948514	1,248479746	1,501959504	-0,598346247
C1	-1,152999001	-1,191078378	-0,66188046	1,28069677

Setelah didapatkan titik pusat centroid kemudian adalah menghitung jarak masing - masing antara data, proses menghitung jarak menggunakan data yang sudah dinormalisasi pada tabel 12 hasil normalisasi, berikut cara menghitung jarak masing – masing data:

$$D(1 - 0) = \sqrt{(0,836 - 0,836)^2 + (0,760 - 1,248)^2 + (1,501 - 1,501)^2 + ((-0,848) - 0,598)^2} = 0,54847751$$

$$D(1 - 1) = \sqrt{(0,836 - 1,152)^2 + (0,760 - 1,191)^2 + (1,501 - 0,661)^2 + ((-0,848) - 1,280)^2} = 4,121424319$$

Dalam proses diatas diammbil dari salah satu data yang pertama, maka untuk hasil jarak dari masing – masing data dapat dilihat pada table dibawah ini. Pada tabel 14 dibawah ini merupakan hasil dari Iterasi ke 1

Tabel 14. Iterasi ke 1

Nama Kepala Keluarga	Bobot Pekerjaan	Jumlah Anggota Keluarga	Bobot Kualitas Rumah	Pendapatan	C0	C1	Cluster
J. Sri Hardono	0,8369	0,7606	1,50196	-0,8489	0,548	4,1214	0
Robinson Sinaga	0,8369	1,2485	1,50196	-0,5983	0	4,2573	0
Balut Parulian Sihombing	-0,158	-0,215	-0,6619	1,5312	3,514	1,416	1
Koslen Naibaho	-2,148	0,2727	-0,6619	0,4038	3,943	1,9752	1
David Pardosi	0,8369	1,7364	1,50196	-0,8489	0,548	4,6634	0
Iskandar	0,8369	0,2727	-0,6619	-0,9742	2,403	3,3447	0
Sanga Braly Erianto Sihombing	0,8369	-0,703	-0,6619	-0,9742	2,938	3,0467	0
Amin Rais Sibarani	0,8369	-1,191	-0,6619	-0,9742	3,283	3,0074	1
Harapan Togatorop	-1,153	-1,191	-0,6619	1,2807	4,257	0	1
Zebpri Sinaga	0,8369	-1,191	-0,6619	-0,5983	3,261	2,7369	1
...
Sugiarto	-0,158	0,7606	1,50196	0,1533	1,339	3,2791	0

Setelah menghitung jarak masing – masing data maka didapatkan iterasi ke 1 yang diapat dilihat pada tabel 14 diatas, maka hasil cluster dapat dilihat tabel dibawah ini. Pada tabel 15 dibawah ini merupakan hasil cluster iterasi ke 1

Tabel 15. Hasil Cluster Iterasi Ke 1

Cluster	Total
0	83
1	87

Iterasi akan terus berlanjut sampai hasil cluster tidak berubah, sebelum melanjut ke iterasi selanjutnya maka harus menentukan titik centroid baru, proses mendapatkan titik pusat centroid baru dapat dilihat dibawah ini:

$$bobot\ pekerjaan_{(0)} = \frac{0,8369+0,8369+0,8369+0,8369+0,8369+(-0,158)+(-0,158)+\dots+(-0,158)}{83}$$

$$bobot\ pekerjaan_{(0)} = 0,48930708$$

$$bobot\ pekerjaan_{(1)} = \frac{(-0,158)+(-2,148)+0,8369+(-1,158)+0,8369+(-1,158)+0,8369+\dots\pm 0,76}{87}$$

$$bobot\ pekerjaan_{(1)} = -0,466810203$$

$$anggota\ keluarga_{(0)} = \frac{0,7606+1,2485+1,7364+0,2727+(-0,703)+1,7364+0,7606+\dots+0,7606}{83}$$

$$anggota\ keluarga_{(0)} = 0,560700708$$

$$anggota\ keluarga_{(1)} = \frac{(-0,21)+0,27+(-1,191)+(-1,19)+(-1,19)+(-1,19)+(-1,19)+\dots+(-0,215)}{87}$$

$$anggota\ keluarga_{(1)} = -0,534921366$$

$$kualitas\ rumah_{(0)} = \frac{1,501+1,501+1,501+(-0,661)+(-0,661)+1,501+1,501+\dots+1,501}{83}$$

$$kualitas\ rumah_{(0)} = 0,511285786$$



$$kualitas\ rumah_{(1)} = \frac{(-0,661)+(-0,661)+(-0,661)+(-0,661)+(-0,661)+(-0,661)+\dots+1,501}{87}$$

$$kualitas\ rumah_{(1)} = -0,487778394$$

$$pendapatan_{(0)} = \frac{(-0,848)+(-0,598)+(-0,848)+(-0,974)+(-0,974)+(-1,09)+0,153+\dots+0,153}{83}$$

$$pendapatan_{(0)} = -0,468548898$$

$$pendapatan_{(1)} = \frac{1,531+0,403+(-0,974)+1,280+(-0,598)+1,781+(-0,723)+1,030+\dots+0,153}{87}$$

$$pendapatan_{(1)} = 0,44700642$$

Untuk tabel *centroid* baru dari Iterasi sebelumnya dapat dilihat dibawah ini. Pada tabel 16 dibawah ini merupakan titik pusat centroid baru iterasi ke 1

Tabel 16. Titik Pusat Centroid Baru Iterasi Ke 1

Cluster	Bobot Pekerja	Jumlah Anggota Keluarga	Bobot Kualitas Rumah	Pendapatan
C0	0,48930708	0,560700708	0,511285786	-0,468548898
C1	-0,466810203	-0,534921366	-0,487778394	0,44700642

Untuk menghitung Iterasi selanjutnya sama seperti sebelumnya untuk mendapatkan Iterasi ke 13, seperti dijelaskan sebelumnya Iterasi akan berhenti sampai hasil cluster tidak berubah, pada penelitian ini proses Iterasi berhenti sampai Iterasi ke 13, berikut hasil Iterasi ke 13:

Tabel 17. Iterasi Ke 13

Nama Kepala Keluarga	Bobot Pekerja	Jumlah Anggota Keluarga	Bobot Kualitas Rumah	Pendapatan	C0	C1	Cluster
J. Sri Hardono	0,8369	0,7606	1,50196	-0,8489	1,685	2,9966	0
Robinson Sinaga	0,8369	1,2485	1,50196	-0,5983	1,966	3,0207	0
Balut Parulian Sihombing	-0,158	-0,215	-0,6619	1,5312	2,612	1,1732	1
Koslen Naibaho	-2,148	0,2727	-0,6619	0,4038	3,242	1,5382	1
David Pardosi	0,8369	1,7364	1,50196	-0,8489	2,299	3,3768	0
Iskandar	0,8369	0,2727	-0,6619	-0,9742	0,744	2,6658	0
Sanga Braly Erianto Sihombing	0,8369	-0,703	-0,6619	-0,9742	0,98	2,746	0
Amin Rais Sibarani	0,8369	-1,191	-0,6619	-0,9742	1,371	2,9106	0
Harapan Togatorop	-1,153	-1,191	-0,6619	1,2807	3,148	1,4354	1
Zebpri Sinaga	0,8369	-1,191	-0,6619	-0,5983	1,377	2,6798	0
...
Sugiarto	-0,158	0,7606	1,50196	0,1533	2,149	1,9936	1

Berikut hasil cluster dari Iterasi ke 13. Pada tabel 18 dibawah ini merupakan hasil cluster iterasi ke 13

Tabel 18. Hasil Cluster Iterasi Ke 13

Cluster	Total
0	91
1	79

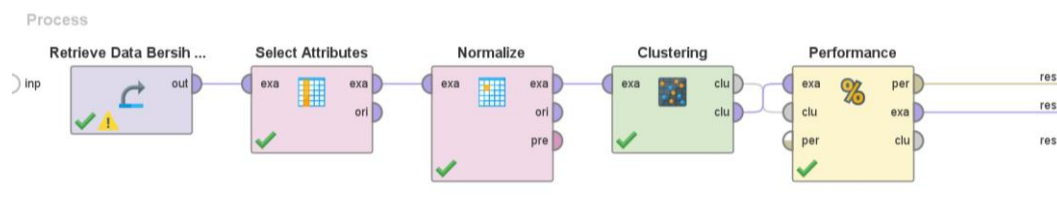
Dan untuk hasil centroid terakhir dapat dilihat dibawah ini. Pada tabel 19 dibawah merupakan hasil titik centroid terakhir

Tabel 19. Titik Centroid Terakhir

Cluster	Bobot Pekerja	Jumlah Anggota Keluarga	Bobot Kualitas Rumah	Pendapatan
C0	0,771345849	-0,006150146	0,003916452	-0,807587667
C1	-0,888512306	0,007084346	-0,004511357	0,930259211

3.4 Pengujian Rapid Miner

Langkah terakhir adalah pengujian system di rapid miner, pengujian ini bertujuan untuk melihat hasil yang dikeluarkan sama atau tidak dengan yang dilakukan sebelumnya, berikut pengujiannya:[13]



Gambar 2. Proses Rapid Miner

Pada gambar 2 diatas merupakan proses rapid miner. Dalam proses rapid miner data sudah melalu processing data yang sudah dijelaskan sebelumnya, kemudian pilih attribute yang digunakan lalu masuk ketahap normalisasi kemudian tahap terakhir clustering menggunakan k-means, maka hasil cluster yang didapatkan dapat dilihat gambar bawah ini:

Cluster Model

```
Cluster 0: 91 items
Cluster 1: 79 items
Total number of items: 170
```

Gambar 3. Hasil Cluster Rapid Miner

Pada gambar 3 merupakan hasil cluster rapid miner. Dapat dilihat pada gambar diatas hasil dari C0 sebanyak 91 dan C1 sebanyak 79, yang dimana hasil ini sama yang didapatkan pada pemodelan sebelumnya dapat dilihat pada tabel 18, kemudian melihat hasil akhir centroid pada rapid miner:

Attribute	cluster_0	cluster_1
BOBOT PEKERJAAN	0.771	-0.889
JUMLAH ANGGOTA KELUARGA	-0.006	0.007
BOBOT KUALITAS RUMAH	0.004	-0.005
PENDAPATAN	-0.808	0.930

Gambar 4. Hasil Centroid Pada Rapid Miner

Pada gambar 4 merupakan hasil centroid pada rapid miner

4. KESIMPULAN

Dari hasil pemodelan dan pengujian yang dilakukan maka didapatkan hasil dari pemodelan dan pengujian yang sama maka hasil dari 170 data yang digunakan dan didapatkan jumlah hasil cluster yaitu cluster 0 atau layak sebanyak 91 kepala keluarga layak mendapatkan bantuan renovasi rumah, dan cluster 1 atau tidak layak ada sebanyak 79 kepala keluarga tidak layak mendapatkan bantuan, hasil yang didapatkan berdasarkan dari attribute pekerjaan, jumlah anggota keluarga, kualitas rumah, dan pendapatan. Maka hasil dari identifikasi masalah sudah didapatkan dan pengelompokan menggunakan algoritma k-means berhasil dilakukan.

REFERENCES

- [1] N. Komang Sri Julyantari Et Al., "Implementasi K-Means Untuk Pengelompokan Status Gizi Balita (Studi Kasus Banjar Titih)" *Jurnal Janitra Informatika Dan Sistem Informasi*, Vol. 1, No. 2, Pp. 92-101, 2021
- [2] L. G. Rady Putra And A. Anggrawan, "Pengelompokan Penerima Bantuan Sosial Masyarakat Dengan Metode K-Means," *Matrik : Jurnal Manajemen, Teknik Informatika Dan Rekayasa Komputer*, Vol. 21, No. 1, Pp. 205-214, Nov. 2021
- [3] N. Afriani Manihuruk, M. Zarlis, E. Irawan, And H. Satria Tambunan, "Penerapan Data Mining Dalam Mengelompokkan Calon Penerima Beasiswa Dengan Menggunakan Algoritma K-Means," Vol. 4, No. 1, 2020
- [4] R. Rosmini, A. Fadlil, And S. Sunardi, "Implementasi Metode K-Means Dalam Pemetaan Kelompok Mahasiswa Melalui Data Aktivitas Kuliah," *It Journal Research And Development*, Vol. 3, No. 1, Pp. 22-31, Aug. 2018
- [5] W. Utomo, "Penerapan Algoritma K-Means Untuk Pengelompokkan Penyakit Kronis Pada Warga Lansia (Studi Kasus Pada: Posyandu Lansia Rw 07)," *Jurnal Media Informatika Budidarma*, Vol. 4, Pp. 1153-1161, 2020
- [6] M. Murdiaty, A. Angela, And C. Sylvia, "Pengelompokan Data Bencana Alam Berdasarkan Wilayah, Waktu, Jumlah Korban Dan Kerusakan Fasilitas Dengan Algoritma K-Means," *Jurnal Media Informatika Budidarma*, Vol. 4, No. 3, P. 744, Jul. 2020
- [7] R. Askia Kurniawan, M. Siddik Hasibuan, R. Syahrul Ramadhan, "Penerapan Algoritma K-Means Untuk Clustering Tempat Makan Di Batubara," *Cosie (Journal Of Computer Science And Informatics Engineering)*, Vol. 01, No. 1, 10-18, Jan 2022



- [8] A. Fitria Yulia And H. Widi Nugroho, “Implementasi Algoritma K-Means Classifier Sebagai Pendukung Keputusan Penerima Dana Bantuan Siswa Miskin (Studi Kasus : Smkn Sukoharjo),” Seminar Nasional Hasil Penelitian Dan Pengabdian Masyarakat, Pp, 48-58, Aug. 2022
- [9] F. Febriansyah And S. Muntari, “Penerapan Algoritma K-Means Untuk Klasterisasi Penduduk Miskin Pada Kota Pagar Alam,” Jiska (Jurnal Informatika Sunan Kalijaga), Vol. 8, No. 1, Pp. 66 – 77, Jan. 2023
- [10] H. Haviluddin, S. J. Patandianan, G. M. Putra, N. Puspitasari, And H. S. Pakpahan, “Implementasi Metode K-Means Untuk Pengelompokan Rekomendasi Tugas Akhir,” Informatika Mulawarman : Jurnal Ilmiah Ilmu Komputer, Vol. 16, No. 1, P. 13, Mar. 2021
- [11] P. Alkhairi And A. P. Windarto, “Penerapan K-Means Cluster Pada Daerah Potensi Pertanian Karet Produktif Di Sumatera Utara,” Seminar Nasional Teknologi Komputer & Sains (Sainteks), Pp, 762-767, Jan. 2019
- [12] L. G. Rady Putra And A. Anggrawan, “Pengelompokan Penerima Bantuan Sosial Masyarakat Dengan Metode K-Means,” Matrik : Jurnal Manajemen, Teknik Informatika Dan Rekayasa Komputer, Vol. 21, No. 1, Pp. 205–214, Nov. 2021
- [13] H. Haviluddin, S. J. Patandianan, G. M. Putra, N. Puspitasari, And H. S. Pakpahan, “Implementasi Metode K-Means Untuk Pengelompokan Rekomendasi Tugas Akhir,” Informatika Mulawarman : Jurnal Ilmiah Ilmu Komputer, Vol. 16, No. 1, P. 13, Mar. 2021
- [14] R. A. Putri, N. I. Maghfirani, G. R. Setyawan, A. A. Rayhan, And N. A. Rakhmawati, “Analisis Pengelompokan Peraturan Kementerian Dengan Menggunakan K-Means Clustering,” Jurnal Sisfokom (Sistem Informasi Dan Komputer), Vol. 9, No. 2, Pp. 152–157, May 2020
- [15] S. Amaliyah And S. Rianti Agustini, “Penerapan Data Mining Untuk Menentukan Kelompok Prioritas Penerima Bantuan Pkh Menggunakan Metode Clustering K-Means Pada Desa Kuala Dendang,” Jurnal Informatika dan Rekayasa Komputer, Vol. 3, No. 1, Pp. 453-459, Apr 2023.
- [16] N. Komang Sri Julyantari Et Al., “Implementasi K-Means Untuk Pengelompokan Status Gizi Balita (Studi Kasus Banjar Titih),” Jurnal Janitra Informatika Dan Sistem Informasi, Vol. 1, No. 2, Pp. 92–101, 2021
- [17] U. S, “Penerapan Data Mining Dengan Mengimplementasikan Algoritma K-Means Dalam Proses Clustering Untuk Pengelompokan Mahasiswa Calon Penerima Beasiswa Kip,” Building Of Informatics, Technology And Science (Bits), Vol. 5, No. 1, Jun. 2023
- [18] Y. F. S. Y. Damanik, S. Sumarno, I. Gunawan, D. Hartama, And I. O. Kirana, “Penerapan Data Mining Untuk Pengelompokan Penyebaran Covid-19 Di Sumatera Utara Menggunakan Algoritma K-Means,” Jurnal Ilmu Komputer Dan Informatika, Vol. 1, No. 2, Nov. 2021
- [19] K. K. Ningrum, J. Maulindar, And A. Farida, “Penerapan Algoritma K-Means Untuk Pengelompokan Penilaian Akhir Semester Di Sdn Kadokan 01 Sukoharjo,” Infotech Journal, Vol. 9, No. 1, Pp. 190–197, Mei 2023
- [20] R. Rosmini, A. Fadlil, And S. Sunardi, “Implementasi Metode K-Means Dalam Pemetaan Kelompok Mahasiswa Melalui Data Aktivitas Kuliah,” It Journal Research And Development, Vol. 3, No. 1, Pp. 22–31, Aug. 2018