

IMPELEMENTASI DATA MINING UNTUK CLUSTERING PENDUDUK MISKIN DI DESA SETROHADI MENGGUNAKAN ALGORITMA PILLAR K-MEANS

Rafly Aditya Perdana, Putri Aisyiyah Rakhma Devi

Teknik Informatika, Universitas Muhammadiyah Gresik

Jl. Sumatera No. 101, Gn. Malang, Randuagung, Kebomas, Gresik, Jawa Timur, 61162

Raflyadityap27@gmail.com

ABSTRAK

Kemiskinan adalah keadaan dimana terjadi ketidakmampuan untuk memenuhi kebutuhan dasar seperti makanan, pakaian, tempat berlindung, pendidikan, dan Kesehatan. Masalah kemiskinan masih menjadi perhatian penting yang harus ditindaklanjuti oleh pemerintah. Menurut salah satu masyarakat di Desa Setrohadi, terdapat masalah penyaluran bantuan yang kurang merata dan tepat sasaran. Hal tersebut disebabkan karena kurangnya ketelitian serta rumitnya pengolahan data dalam menentukan penduduk miskin yang menjadi prioritas utama penerima bantuan serta belum tersedianya sistem yang dapat digunakan untuk mengetahui penduduk miskin di Desa Setrohadi berdasarkan kriteria yang sudah ditentukan, dari permasalahan tersebut dapat dibuat Sistem Pengelompokan Penduduk Miskin Berbasis *Web* Menggunakan Metode *Pillar K-Means*. Dimana sistem yang digunakan dapat memberikan informasi mengenai pemetaan hasil klasterisasi data yang proses *clustering*nya menggunakan metode *Pillar K-Means*, yang nantinya dapat membantu pihak desa dalam mengelompokkan penduduk yang tergolong miskin dan data yang diperoleh lebih akurat. Hasil pengelompokan data penduduk miskin di Desa Setrohadi menggunakan Algoritma *Pillar K-Means* dengan 2 *cluster*, *cluster* 1 berjumlah 127 data kepala keluarga dan *cluster* 2 berjumlah 96 data kepala keluarga dengan nilai evaluasi performa menggunakan DBI didapatkan hasil 0,9911.

Kata kunci : *Pillar Algorithm, K-Means Algorithm, Data Mining, Clustering, Poor Residents.*

1. PENDAHULUAN

Permasalahan yang mengkhawatirkan dan sering terjadi di dunia, khususnya di Indonesia adalah kemiskinan. Banyaknya warga yang kurang mampu masih menjadi permasalahan yang sulit diatasi oleh pemerintah. Masalah kemiskinan masih menjadi perhatian penting yang harus ditindaklanjuti oleh pemerintah [1]. Kemiskinan adalah keadaan dimana terjadi ketidakmampuan untuk memenuhi kebutuhan dasar seperti makanan, pakaian, tempat berlindung, pendidikan, dan Kesehatan [2]. Salah satu permasalahan kemiskinan yang belum terselesaikan adalah hasil survey yang diberikan pemerintah tidak sesuai dengan kebutuhan penduduk [3].

Menurut ibu Mufliah salah satu penduduk di Desa Setrohadi, terdapat masalah penyaluran bantuan yang kurang merata dan tepat sasaran. Hal tersebut disebabkan karena kurangnya ketelitian serta rumitnya pengolahan data dalam menentukan penduduk miskin yang menjadi prioritas utama penerima bantuan serta belum tersedianya sistem yang dapat digunakan untuk mengetahui penduduk miskin di Desa Setrohadi berdasarkan kriteria yang sudah ditentukan. Salah satu aspek penting yang mendukung strategi penanggulangan kemiskinan adalah dengan tersedianya data kemiskinan yang akurat dan tepat sasaran [4]. Oleh karena itu, pihak kelurahan harus lebih teliti dalam pemilihan warga desa yang akan diberi bantuan sehingga warga yang akan diberi bantuan benar-benar layak menerima dan berasal dari keluarga yang kurang mampu.

Analisis data tersebut bisa dilakukan dengan berbagai macam cara, salah satunya yaitu

menggunakan teknik *Data Mining*. *Data Mining* merupakan proses pengumpulan dan analisis data historis untuk menemukan keteraturan, pola, dan hubungan dalam *database* yang besar [5]. Tujuan utama *data mining* adalah untuk menemukan pengetahuan dari data atau informasi yang kita miliki. *Data Mining* memiliki berbagai metode atau bahkan model yang dapat dipakai. Salah satu metode *data mining* ialah teknik *clustering*. *Clustering* merupakan proses pengelompokan data yang kesamaan datanya dalam suatu kelompok lebih besar daripada kesamaan data dengan data kelompok lain [6]. Data-data yang mempunyai kesamaan kriteria akan dikumpulkan kedalam kelompok atau *cluster* yang dibedakan [7]. Penerapan *Data Mining Clustering* di Desa Setrohadi menggunakan metode *K-Means* bertujuan untuk mengolah data penduduk miskin agar menghasilkan informasi baru yang dapat digunakan pihak kelurahan dalam mengoptimalkan penerima bantuan di desa setrohadi.

Metode *K-Means* adalah metode pengelompokan yang melakukan partisi set data ke dalam sejumlah *k-cluster* yang sudah ditetapkan di awal [8]. Metode *K-Means* mudah beradaptasi cepat dan umum dalam penggunaannya. Dalam penyelesaiannya, Metode *K-Means* akan mengelompokkan data-data ke dalam dataset ke suatu *cluster* berdasarkan jarak terdekat [9]. Namun, pemilihan *centroid* yang dilakukan secara acak diawal perhitungan menjadi kelemahan dari proses tersebut. Oleh karena itu, digunakan metode Algoritma *Pillar* yang merupakan improvisasi dari *K-Means* untuk mengatasi masalah tersebut dengan memilih pusat cluster awal secara cermat. Tujuan

utama dari algoritma ini adalah untuk mengoptimalkan nilai *centroid* pada Algoritma *K-Means* dan meningkatkan kinerja proses *clustering* [10].

Berdasarkan beberapa penelitian sebelumnya yang telah menggunakan Algoritma *K-Means* telah berhasil untuk menyelesaikan permasalahan *Clustering*. Sehingga pada penelitian ini digunakanlah metode *clustering* dengan menggunakan Algoritma *Pillar* yang merupakan improvisasi metode *K-Means* dengan harapan dapat memberikan hasil kinerja yang lebih baik untuk performa *clustering*. Ada beberapa kriteria yang harus digunakan dalam penentuan yaitu usia, penghasilan, jumlah keluarga, jumlah tanggungan anak sekolah, status rumah, jenis lantai dan jenis dinding. Penelitian ini akan menghasilkan *output clustering* data penduduk miskin di Desa Setrohadi, Kecamatan Dudusampayan, Kabupaten Gresik.

2. TINJAUAN PUSTAKA

2.1. Data Mining

Data Mining merupakan suatu proses pengumpulan dan analisis data untuk menemukan keteraturan, pola, maupun hubungan pada kumpulan data yang luas. Beberapa jenis pendekatan yang digunakan dalam data mining diantaranya prediksi, perbandingan, klasifikasi, *cluster*, dan perkiraan [11]. Data Mining juga merupakan proses statistik, matematika, kecerdasan buatan, dan *machine learning* untuk ekstraksi dan identifikasi suatu informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar [12].

Terdapat tiga elemen penting dalam data mining diantaranya adalah 1) Data Mining merupakan prosedur mekanis pada data yang ada, 2) Data yang hendak diproses merupakan data yang luas 3) Poin dari data mining yaitu untuk mendapatkan pola yang dapat memberikan informasi yang bermanfaat [13]. Data Mining memiliki manfaat untuk mengelola data mentah, dimensi data yang tinggi dan data campuran yang sifatnya berbeda menjadi kumpulan informasi yang dapat dijadikan sebagai penunjang keputusan secara efektif (Sari et al., 2020).

2.2. Clustering

Clustering merupakan suatu metode untuk pengelompokan data yang berkarakteristik sama antara data satu dengan yang lainnya. *Clustering* membagi data dalam satu himpunan ke dalam beberapa kelompok yang kesamaan datanya dalam suatu kelompok lebih besar daripada kesamaan data tersebut dengan data pada kelompok lain dan data yang mempunyai kesamaan kriteria akan dikelompokkan ke dalam data yang berbeda [14]. Partisi data dilakukan dengan mencari nilai jarak terdekat antara data dengan nilai *centroid* yang telah ditetapkan baik secara acak maupun dengan *Initial Set of Centroids* untuk menentukan nilai *centroid* dengan objek yang berurutan. *Clustering* berpotensi untuk mengetahui struktur pada data yang bisa digunakan lebih lanjut

pada berbagai macam aplikasi seperti klasifikasi, pengolahan gambar dan sosialisasi pola [15]. *Clustering* memiliki sifat tidak punya data latih dimana karakteristik tiap *cluster* tidak ditentukan sebelumnya melainkan berdasarkan kemiripan atribut-atribut dari suatu kelompok [16].

Tujuan *clustering* data dibedakan menjadi dua yaitu pengelompokan untuk pemahaman struktur alami data dan pengelompokan untuk penggunaan data [17]. Pengelompokan untuk pemahaman struktur alami data bertujuan untuk proses awal kemudian dilanjutkan dengan pekerjaan inti seperti peringkasan atau *summarization* (rata-rata standar deviasi) sedangkan pengelompokan untuk penggunaan data bertujuan untuk mencari kelompok yang paling *representatif* terhadap data dan memberikan abstraksi dari setiap objek data.

2.3. Algoritma K-Means

K-Means merupakan salah satu algoritma pengelompokan data yang mempartisi data ke dalam dua atau lebih kelompok. Algoritma ini pertama kali diusulkan oleh MacQueen (1996) dan dikembangkan oleh Hartigan dan Wong tahun 1975. *K-Means* mempunyai kemampuan dalam mengelompokan data dengan jumlah yang cukup besar dan waktu komputasi yang cepat dan efisien [18]. Algoritma *K-Means* memiliki ketelitian yang cukup tinggi terhadap ukuran objek [19].

Algoritma *K-Means* dapat merubah data yang ingin di *cluster* menuju beberapa titik yang akan digunakan sebagai acuan dalam pengklasteran data. Data yang di klasterisasi harus numerik [20]. Algoritma *K-Means* hanya mengambil sebagian dari banyaknya dari komponen yang didapatkan kemudian dijadikan pusat *cluster* awal, pada penentuan pusat *cluster* ini dipilih secara acak dari populasi data.

K-Means ialah metode klasterisasi yang paling banyak digunakan diberbagai bidang [21]. Hal tersebut dikarenakan metode *K-Means* bersifat sederhana dan memiliki kemampuan untuk melakukan *cluster* data yang berjumlah besar dalam waktu yang singkat.

Adapun langkah-langkah metode *K-Means Clustering* adalah sebagai berikut [22]:

- 1) Menentukan jumlah *cluster*
- 2) Melakukan inisialisasi cara yang sering digunakan adalah cara random atau acak.
- 3) Mengalokasikan semua data ke *cluster* terdekat. Dalam tahap ini perlu dihitung jarak tiap data ke tiap pusat *cluster* dengan menggunakan teori jarak *Euclidean* atau *Manhattan*.
- 4) Kembali ke step 3, apabila masih ada data yang berpindah *cluster* atau apabila perubahan nilai *centroid*.

2.4. Algoritma Pillar

Perbedaan antara Algoritma *K-Means* dan Algoritma *Pillar* terletak pada pemilihan *centroid* awal. Pada Algoritma *K-Means* pemilihan *centroid*

dilakukan secara acak sebanyak nilai K yang ditentukan, namun pada Algoritma *Pillar* penempatan *centroid* awal ditentukan di dalam ruang fitur dengan menempatkan inisialisasi masing-masing *centroid* awal yang memiliki akumulasi jarak terjauh [23]. Algoritma ini terinspirasi pada penempatan pilar di setiap sudut bangunan yang jaraknya jauh sehingga masa bangunan terkonsentrasi pada setiap pilar. Langkah-langkah metode Algoritma *Pillar* [24].

1. Menentukan k=jumlah kluster.
2. Menghitung rata-rata semua data =Max semua data.
3. Hitung jarak data dengan pusat kluster $D[n] = dis(X, m)$ seperti pada persamaan 1.

$$dis(X, m) = | \sum_{j=1}^N X_j - mj | \quad (1)$$

4. Menghitung $nmin$ dan $nbdis = a n/k$ semua data dibagi cluster
Ket : $Nmin$ (Alpha = 0,1)
 $Nbdis$ (Beta = 0,9)
5. Pengecekan keterpenuhan *range* data dengan hasil jarak $Nmin$ dan $Nbdis = IF(AND(jika\ memenuhi\ 'v',\ jika\ tidak\ memenuhi\ 'false'))$
Ket : *False* = tidak memenuhi *range*
V = memenuhi *range*
6. Pengurutan kandidat berdasarkan nilai jarak terkecil hingga terbesar
7. Pemilihan data yang memenuhi rata-rata dari jarak (langkah ke 6) dengan jarak *centroid* seperti contoh :
 - Jika jumlah *centroid* yang terpilih adalah 4 maka dilakukan pemilihan jumlah data dan yang terpilih ialah data no 1 dan 3.

2.5. Jarak Euclidean

Jarak Euclidean merupakan metode perhitungan jarak yang digunakan untuk mengukur jarak dari 2 (dua) buah titik dalam *Euclidean space* (meliputi bidang *Euclidean* dua dimensi, tiga dimensi, atau bahkan lebih) [25]. Adapun rumus dalam pengukuran jarak *Euclidean* seperti pada persamaan 2.

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

Keterangan :

- d : Jarak antara x dan y
- x : Data pusat *cluster*
- y : Data pada atribut
- i : Setiap data
- n : Jumlah data,
- x_i : Data pada pusat *cluster* ke i
- y_i : Data pada setiap data ke i

2.6. Davies Bouldin Index DBI

Davies-Bouldin Index (DBI) merupakan pengukuran yang digunakan sebagai bahan evaluasi dan pertimbangan dari hasil *clustering*. Dimana dalam perhitungan DBI ada beberapa tahap, pertama dengan

menghitung nilai SSW, kedua menghitung nilai SSB, ketiga menghitung nilai Ratio, keempat menghitung nilai DBI.

Langkah pertama Perhitungan nilai SSW dilakukan dengan menghitung jarak setiap data pada *centroid* lalu menghitung rata-ratanya menggunakan persamaan 3.

$$SSW_i = \frac{1}{m_i} \sum_{j=i}^{c_i} d(X_j, C_j) \quad (3)$$

Dimana :

- m_i = jumlah data dalam *cluster* ke-i
- c_i = *centroid cluster* ke-i
- $d(X_j, C_j)$ = jarak setiap data ke *centroid* i yang dihitung menggunakan *Euclidean Distance*.

Langkah kedua setelah diketahui nilai SSW maka selanjutnya dilakukan perhitungan SSB untuk menghitung jarak antar *centroid* menggunakan persamaan 4.

$$SSB_{ij} = d(X_i, X_j) \quad (4)$$

Dimana :

- $d(X_i, X_j)$ = jarak antara data ke-i dengan data ke-j di *cluster* lain.

Langkah ketiga setelah menghitung nilai SSW dan nilai SSB, selanjutnya mencari nilai rasio antar *cluster* dengan menggunakan persamaan 5.

$$R_{ij, \dots, n} = \frac{SSW_i + SSW_j + \dots + SSW_n}{SSB_i + \dots + SSB_{ni, nj}} \quad (5)$$

Dimana :

- SSW_i = *Sum of Square Within-Cluster* pada *Centroid* i
- SSB_i = *Sum of Square Between Cluster* data ke-i dengan j pada *cluster* yang berbeda

Langkah keempat setelah mengetahui rasio antar *cluster* kemudian dihitunglah nilai DBI dengan menggunakan persamaan 6.

$$DBI = \frac{1}{k} \sum_{i=1}^k \max_{i \neq j} (R_{i, j, \dots, k}) \quad (6)$$

Dari perhitungan *Davies Bouldin Index* (DBI) dapat disimpulkan bahwa jika semakin kecil nilai *Davies Bouldin Index* (DBI) yang diperoleh (non negatif ≥ 0) maka *cluster* tersebut semakin baik.

2.7. Black Box Testing

Black Box Testing adalah salah satu teknik pengujian perangkat lunak yang dilakukan tanpa mengetahui secara detail cara kerja kode program di dalamnya. Teknik ini melihat sistem atau program sebagai sebuah kotak hitam (*black box*) di mana input diberikan dan *output* yang dihasilkan dievaluasi, tanpa memperhatikan bagaimana alur program di dalamnya [26].

Tujuan dari *black box testing* adalah untuk memastikan bahwa sistem atau program berfungsi dengan benar sesuai dengan persyaratan fungsional dan non-fungsional yang telah ditetapkan, serta mengidentifikasi kesalahan atau kegagalan dalam program [27].

2.8. Kemiskinan

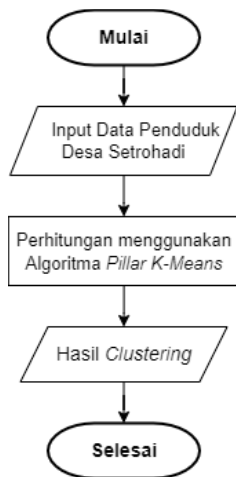
Kemiskinan merupakan permasalahan yang kompleks yang tidak hanya terjadi di negara berkembang seperti Indonesia, tetapi juga terjadi di negara maju [28]. Kemiskinan merupakan suatu situasi dimana sebuah rumah tangga kesulitan untuk memenuhi kebutuhan, dimana lingkungannya kurang memberikan peluang untuk meningkatkan kesejahteraan [29]. Kemiskinan disebabkan oleh dimensi alam, dimana alam tidak memberikan dukungan dengan kesuburan tanah sehingga menyebabkan masyarakat tidak mampu menggunakan lingkungan untuk usaha [30]. Mampu didefinisikan memiliki harta yang dapat digunakan untuk memenuhi kebutuhan sehari-hari. Sedangkan kaya adalah suatu kondisi dimana jumlah aset/penghasilan lebih besar daripada hutang/pengeluaran [31].

3. METODE PENELITIAN

3.1. Flowchart Sistem Clustering Penduduk Miskin

Pada tahap ini akan dijelaskan mengenai tahapan-tahapan dari Sistem Clustering Penduduk Miskin :

1. Admin memasukkan data penduduk ke dalam sistem.
2. Data yang telah dimasukkan kemudian diolah menggunakan Algoritma Pillar K-Means.
3. Menampilkan hasil clustering/pengelompokan data pada sistem.



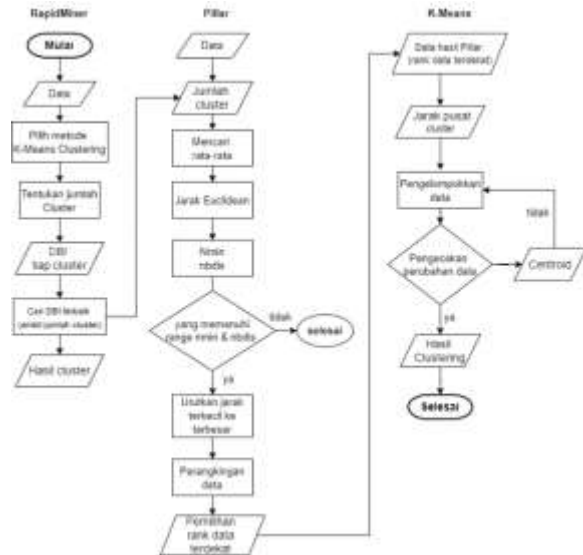
Gambar 1. Flowchart sistem clustering penduduk miskin

3.2. Flowchart Sistem Clustering Penduduk Miskin dengan Algoritma Pillar K-Means

Pada tahap ini akan dijelaskan mengenai tahapan-tahapan dari Sistem Pengelompokan Penduduk Miskin dengan Algoritma Pillar K-Means :

1. Evaluasi cluster menggunakan software RapidMiner.
2. Perhitungan titik awal centroid menggunakan perhitungan Algoritma Pillar.

3. Perhitungan jarak data menggunakan Algoritma K-Means.
4. Pengelompokan data berdasarkan cluster-nya sesuai hasil perhitungan dari semua tahapan.
5. Hasil pengklasteran data penduduk.



Gambar 2. Flowchart sistem dengan algoritma Pillar K-Means

4. HASIL DAN PEMBAHASAN

4.1. Implementasi



Gambar 3. Tampilan Halaman Login

Admin harus memasukkan Email dan Password untuk masuk ke dalam sistem. Jika Email dan Password yang dimasukkan benar bisa login kedalam sistem apabila salah sistem akan merefresh ulang halaman untuk pemasukan Email dan Password ulang.



Gambar 4. Tampilan Halaman Dashboard

Halaman dashboard yang berisi beberapa panduan dalam penggunaan sistem clustering pada

gambar 4 data penduduk miskin dengan Algoritma Pillar K-Means.



Gambar 5. Tampilan Halaman Data Penduduk

Halaman data penduduk untuk menambahkan data masyarakat kedalam sistem pada gambar 5. Input data masyarakat dapat dilakukan dengan mengisi form pengisian yang sudah disediakan oleh sistem, lalu sistem akan menampilkan data yang telah di inputkan. Terdapat *tools* untuk *edit* dan *delete* pada setiap data penduduk. Terdapat *tools* tambah data penduduk untuk menambahkan data penduduk baru, dapat dilihat pada gambar 6 berikut.



Gambar 6. Tampilan Halaman Tambah Data Penduduk

dilakukan proses *clustering* data. Sistem akan menampilkan hasil perhitungan *clustering*, dapat dilihat pada gambar 8 berikut.



Gambar 9. Tampilan Halaman Laporan Clustering

Halaman laporan hasil klasterisasi data penduduk desa setrohadi yang telah dihitung menggunakan Algoritma Pillar K-Means pada gambar 9. Dalam halaman tersebut ditampilkan hasil klasterisasi masyarakat berdasarkan Nama Kepala Keluarga dan penggolongan clusternya, dan ada *tools* untuk print file (pdf) data penduduk berdasarkan hasil perhitungan. Dapat dilihat pada gambar 10 berikut.



Gambar 10. Tampilan Halaman Simpan File (PDF)



Gambar 7. Tampilan Halaman Proses Clustering



Gambar 11. Tampilan Halaman Admin

Halaman admin yang berisi tambah admin baru, nama, *email*, *role* (hak akses), dan *tools* untuk edit dan hapus data admin pada gambar 11. Terdapat *tools* tambah admin untuk menambahkan data admin baru, dapat dilihat pada gambar 12 berikut.



Gambar 8. Tampilan Halaman Hasil Proses Clustering



Gambar 12. Tampilan Halaman Tambah Admin

Halaman proses *clustering*, untuk memulai proses klasterisasi data menggunakan Algoritma Pillar K-Means pada gambar 7, dimana kita harus menekan tombol Lakukan Proses Clustering data dan sistem akan memulai proses *clustering* data di sistem. Setelah

4.2. Pengujian Sistem

Tabel 1. Pengujian *black box*

Test Case	Input Data	Expected Output	Status
Form Login	Verifikasi Username dan Password	Sistem akan otomatis mengosongkan pengisian <i>username</i> dan <i>password</i> apabila pengisiannya salah.	Success
		Apabila admin memasukkan <i>username</i> dan <i>password</i> dengan benar maka akan masuk ke dalam sistem.	Success
Form Dashboard	-	Sistem menampilkan Panduan penggunaan aplikasi.	Success
Form Data Penduduk	Tombol tambah Data Penduduk	Sistem menampilkan form pengisian data penduduk.	Success
	Tools Edit Data	Sistem akan melakukan pengeditan data penduduk yang dipilih.	Success
	Tools Hapus Data	Sistem akan melakukan penghapusan data penduduk yang dipilih.	Success
Form Proses Clustering	Tombol Input Clustering data penduduk	Sistem memulai proses <i>clustering</i> data menggunakan Algoritma <i>Pillar K-Means</i> .	Success
Form Laporan Clustering	Tombol Print (pdf)	Sistem menampilkan hasil akhir dari proses <i>clustering</i> . Dimana terdapat tabel mulai dari Iterasi pertama hingga Iterasi terakhir dan terdapat pengelompokan data dari masing masing <i>cluster</i> , serta terdapat Tools untuk cetak file (pdf).	Success
Form Data Admin	Tombol tambah admin	Sistem menampilkan form pengisian data admin yang meliputi nama, <i>email</i> dan <i>password</i> .	Success
	Tools Edit akun admin	Sistem akan melakukan pengeditan akun admin yang dipilih.	Success
	Tools Delete akun admin	Sistem akan melakukan penghapusan data yang dipilih.	Success

4.3. Hasil Perhitungan

Tabel 2. Hasil akhir perhitungan clustering

Hasil Perhitungan	Hasil Centroid	A	B	C	D	E	F	G
Hasil Clustering Algoritma <i>Pillar</i>	C1	2,0104	2,3646	2,3229	2,8438	1,9479	2,4479	1,6250
Hasil Clustering Algoritma <i>K-Means</i>	C1	2,6142	4,8898	2,1260	2,6614	2,0079	2,9921	2,0079
Hasil Pengisian Davies Bouldin Index (DBI) <i>Pillar</i>	C2	2,0104	2,3646	2,3229	2,8438	1,9479	2,4479	1,6250
Hasil Pengisian Davies Bouldin Index (DBI) <i>K-Means</i>	C1	2,6142	4,8898	2,1260	2,6614	2,0079	2,9921	2,0079
Hasil Pengisian Davies Bouldin Index (DBI) <i>Pillar</i>	C2	2,0104	2,3646	2,3229	2,8438	1,9479	2,4479	1,6250
Hasil Pengisian Davies Bouldin Index (DBI) <i>K-Means</i>	C1	2,6142	4,7244	2,1260	2,6614	2,0079	2,9134	1,9606
Hasil Pengisian Davies Bouldin Index (DBI) <i>K-Means</i>	C2	2,0833	2,3750	2,3229	2,8438	1,9167	2,3646	1,5625

5. KESIMPULAN

Dengan menggunakan Algoritma *Pillar K-Means*, dua kelompok data penduduk miskin Desa Setrohadi, cluster 1 memiliki 127 data kepala keluarga dan cluster 2 memiliki 96 data kepala keluarga. Hasil implementasi penentuan titik *centroid* menggunakan Algoritma *Pillar K-Means* menunjukkan bahwa dua kelompok adalah yang terbaik. Nilai DBI Algoritma *Pillar* 0,911 dan nilai Algoritma *K-Means* 1,212 menunjukkan bahwa hasilnya cukup baik karena mendekati angka 0 (tidak negatif lebih dari 0). Sistem *clustering* data penduduk miskin yang dibangun dengan metode *Waterfall* memenuhi kebutuhan pengguna dan mudah digunakan.

Penelitian ini tentunya tidak terlepas dari kekurangan. Oleh karena itu, peneliti memberikan saran untuk penelitian selanjutnya agar lebih baik. Pada proses *clustering* data, penentuan titik awal atau

centroid sangat berpengaruh terhadap hasil dari *clustering*, peneliti sendiri menggunakan Algoritma *Pillar* untuk mengatasi masalah pemilihan pusat *cluster* awal agar hasil menjadi optimal. Namun untuk menilai apakah hasil tersebut telah optimal atau belum maka diperlukan metode lain sebagai pembanding.

DAFTAR PUSTAKA

- [1] S. Nurhafiza, H. Rahma, Sani Nahampun Muhammad, and I. A. A. Rahmadyan, "Penerapan K-Means Clustering Untuk Mengelompokkan Tingkat Kemiskinan di Provinsi Riau," *SENTIMAS Semin. Nas. Penelit. dan Pengabd. Masy.*, vol. 1, no. 1, pp. 59–65, 2022.
- [2] D. Sunia, Kurniabudi, and P. Alam Jusia, "Penerapan Data Mining Untuk Clustering Data Penduduk Miskin Menggunakan Algoritma K-Means," *J. Ilm. Mhs. Tek. Inform.*, vol. 1, no. 2, pp. 121–134, 2019.
- [3] 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.
- [4] S. Sari and J. N. Utamajaya, "Sistem Pendukung Keputusan Penerima Bantuan Langsung Tunai Dana Desa Menggunakan Metode Algoritma K-Means Clustering," *J. JUPITER*, vol. 14, no. 1, pp. 150–160, 2022.
- [5] F. Andrianti and R. Firmansyah, "Penerapan Clustering Data Kurang Mampu Di Desa Situmekar Menggunakan Algoritma K-Means," *eProsiding Tek. Inform.*, vol. 1, no. 1, pp. 88–95, 2020, [Online]. Available:

- <http://eprosiding.ars.ac.id/index.php/pti>
- [6] 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.
- [7] A. N. Nursia, W. Ramdhan, and W. M. Kifti, "Analisis Kelayakan Penerima Bantuan Covid-19 Menggunakan Metode K-Means," *Build. Informatics, Technol. Sci.*, vol. 3, no. 4, pp. 574–583, 2022, doi: 10.47065/bits.v3i4.1399.
- [8] Yuni Radana Sembiring, Saifullah, and Riki Winanjaya, "Implementasi Data Mining Dalam Mengelompokkan Jumlah Penduduk Miskin Berdasarkan Provinsi Menggunakan Algoritma," *KESATRIA J. Penerapan Sist. Inf. (Komputer Manajemen) Vol. 2, No. 2*, vol. 2, no. 2, pp. 125–132, 2021.
- [9] K. Aprilia and F. Sembiring, "Analisis Garis Kemiskinan Makanan Menggunakan Metode Algoritma K-Means Clustering," *SISMATIK (Seminar Nas. Sist. Inf. dan Manaj. Inform.)*, pp. 1–10, 2021.
- [10] A. L. Ramdani and H. B. Firmansyah, "Clustering Application for UKT Determination Using Pillar K-Means Clustering Algorithm and Flask Web Framework," *Indones. J. Artif. Intell. Data Min.*, vol. 1, no. 2, pp. 53–59, 2018, doi: 10.24014/ijaidm.v1i2.5126.
- [11] A. Ikhwan and N. Aslami, "Implementasi Data Mining untuk Manajemen Bantuan Sosial Menggunakan Algoritma K-Means," *J. Teknol. Inf.*, vol. 4, no. 2, pp. 208–217, 2020, doi: 10.36294/jurti.v4i2.2103.
- [12] P. Parjito and P. Permata, "Penerapan Data Mining Untuk Clustering Data Penduduk Miskin Menggunakan Algoritma Hard C-Means," *AINET J. Inform.*, vol. 3, no. 1, pp. 31–37, 2021.
- [13] M. A. P. Sari and U. Chotijah, "Pengelompokan Anggota Divisi Himpunan Mahasiswa Jurusan Pada Universitas 'Xyz' Dengan Metode K-Means Clustering," *Antivirus J. Ilm. Tek. Inform.*, vol. 16, no. 1, pp. 52–62, 2022, doi: 10.35457/antivirus.v16i1.2139.
- [14] K. Serliani, Y. A. Mustofa, and I. Surya Kumala, "Clustering Petani Penerima Pupuk Berdasarkan Luas Lahan Menggunakan Algoritma K-Means," *J. Nas. cosPhi*, vol. 4, no. 2, pp. 27–30, 2020.
- [15] S. Rahmayani, S. Sumarno, and Z. A. Siregar, "Analysis of K-Means Algorithm for Clustering of Covid-19 Social Assistance Recipients," *JOMLAI J. Mach. Learn. Artif. Intell.*, vol. 1, no. 1, pp. 77–84, 2021, doi: 10.55123/jomlai.v1i1.166.
- [16] E. Fammaldo and L. Hakim, "Penerapan Algoritma K-Means Clustering Untuk Pengelompokan Tingkat Kesejahteraan Keluarga Untuk Program Kartu Indonesia Pintar," *J. Ilm. Teknol. Infomasi Terap.*, vol. 5, no. 1, pp. 23–31, 2018, doi: 10.33197/jitter.vol5.iss1.2018.249.
- [17] R. Hidayat, "Clustering Menggunakan Algoritma K-Means untuk Mengelompokkan Wilayah Rawan Kejahatan di Wilayah Kabupaten Solok," *J. Ilm. Mhs. Manajemen, Bisnis dan Akunt.*, vol. 4, no. 5, pp. 646–654, 2022, doi: 10.32639/jimmba.v4i5.169.
- [18] D. Missa, S. Achmadi, and A. Mahmudi, "Penerapan Metode Clustering Dengan Algoritma K-Means Pada Pengelompokan Data Penghasilan Orang Tua Siswa," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 5, no. 1, pp. 125–133, 2021, doi: 10.36040/jati.v5i1.3275.
- [19] W. Febrianti and Mustopa Husein Lubis, "Identifikasi Calon Penerima Bantuan Satu Keluarga Satu Sarjana (SKSS) Menggunakan Algoritma K-Means," *J. Inform. Ekon. Bisnis*, vol. 4, no. 2, pp. 53–58, 2022, doi: 10.37034/infv.v4i2.117.
- [20] S. Suhartini and R. Yuliani, "Penerapan Data Mining untuk Mengcluster Data Penduduk Miskin Menggunakan Algoritma K-Means di Dusun Bagik Endep Sukamulia Timur," *Infotek J. Inform. dan Teknol.*, vol. 4, no. 1, pp. 39–50, 2021, doi: 10.29408/jit.v4i1.2986.
- [21] D. A. D. I. L. Muhammad Syamsul Arifin, "Penerapan Metode K-Means Untuk Cluster Calon Penerima Kartu Jombang Sehat (Kjs) Berbasis Website," *J. Ilm. Inov. Teknol. Inf.*, vol. 6, no. 2, pp. 1–10, 2022, [Online]. Available: <http://ejournal.unhasy.ac.id/index.php/inovate/article/view/3170>
- [22] I. Nasution, A. P. Windarto, and M. Fauzan, "Penerapan Algoritma K-Means Dalam Pengelompokan Data Penduduk Miskin Menurut Provinsi," *Build. Informatics, Technol. Sci.*, vol. 2, no. 2, pp. 76–83, 2020, doi: 10.47065/bits.v2i2.492.
- [23] A. Primandana, S. Adinugroho, and C. Dewi, "Optimasi Penentuan Centroid pada Algoritme K-Means Menggunakan Algoritme Pillar (Studi Kasus: Penyandang Masalah Kesejahteraan Sosial di Provinsi ...)," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 3, no. 11, pp. 10678–10683, 2019, [Online]. Available: <http://j-ptiik.ub.ac.id/index.php/j-ptiik/article/download/6748/3264>
- [24] K. A. Saputra and I. N. W. Wijaya, "Penerapan Algoritma Pillar Untuk Inisialisasi Titik Pusat K-Means Klaster Dinamis," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 7, no. 6, pp. 1213–1220, 2020, doi: 10.25126/jtiik.202072538.
- [25] M. Nishom, "Perbandingan Akurasi Euclidean Distance, Minkowski Distance, dan Manhattan Distance pada Algoritma K-Means Clustering berbasis Chi-Square," *J. Inform. J. Pengemb. IT*, vol. 4, no. 1, pp. 20–24, 2019, doi: 10.30591/jpit.v4i1.1253.
- [26] T. Hidayat and M. Muttaqin, "Pengujian sistem informasi pendaftaran dan pembayaran wisuda

- online menggunakan black box testing dengan metode equivalence partitioning dan boundary value analysis,” *J. Tek. Inform. UNIS*, vol. 6, no. 1, pp. 2252–5351, 2018, [Online]. Available: www.ccssenet.org/cis
- [27] W. Yahya Dwi and A. Muna Wardah, “Pengujian Blackbox Sistem Informasi Penilaian Kinerja Karyawan Pt Inka (Persero) Berbasis Equivalence Partitions Blackbox Testing of Pt Inka (Persero) Employee Performance Assessment Information System Based on Equivalence Partitions,” *J. Digit. Teknol. Inf.*, vol. 4, no. 1, pp. 22–26, 2021, [Online]. Available: <http://jurnal.um-palembang.ac.id/index.php/digital>
- [28] R. J. Kasim, S. Bahri, and S. Amir, “Implementasi Metode K-Means Untuk Clustering Data Penduduk Miskin Dengan Systematic Random Sampling,” *Pros. SISFOTEK*, vol. 5, no. 1, pp. 95–101, 2021.
- [29] E. Adawiyah, “Kemiskina dan Faktor-Faktor Penyebabnya,” *Khidm. Sos. J. Soc. Work Soc. Serv.*, vol. 1, no. 1, pp. 43–50, 2020.
- [30] K. K. Oki, M. Yanti Akoit, and H. Bubu, “Analisis Faktor-Faktor Yang Mempengaruhi Kemiskinan Di Desa Takarai Kabupaten Malaka Analysis of Factors Affecting Poverty In Takarai Village, Malaka Regency,” *EkoPem J. Ekon. Pembang.*, vol. 5, no. 3, pp. 46–54, 2020, [Online]. Available: <http://jurnal.unimor.ac.id/JEP>
- [31] H. A. Yunus, “Konsep Hidup Kaya dan Berkah,” *Jurnal MADINASIKA Manaj. dan Kegur.*, vol. 1, no. 1, pp. 1–9, 2019, [Online]. Available: <https://ejournal.unma.ac.id/index.php/madinasika>