

Pengelompokan Data Status Pertanahan *Letter C* menggunakan Algoritma *Partitioning Around Medoids*

Putri Merly Deleo Karina^{#1}, Indyah Hartami Santi^{#2}, Wahyu Dwi Puspitasari^{#3}

[#]Program Studi Teknik Informatika Universitas Islam Balitar

Jl. Majapahit No.2- 4, Sananwetan, Kec. Sananwetan, Kota Blitar, 66137, Indonesia

¹putrimerly262@gmail.com

²indyahartamisanti@gmail.com

³pushpitasari23@gmail.com

Abstrak

Letter C merupakan pengarsipan data pertanahan yang dikelola oleh desa. Pengelolaan data status pertanahan *Letter C* di Desa Pandanarum masih dilakukan secara manual yaitu dengan menuliskan *history* jual beli kedalam buku *Letter C*. Pada Desa Pandanarum terdapat 3960 data *Letter C* sehingga dapat mempengaruhi dalam proses pelayanan kepada warga. Pada penelitian ini menggunakan data sampel menggunakan teori *Slovin* sebesar 363 data dan diambil dari buku ketiga *Letter C*. Penelitian ditujukan untuk pengelompokan status pertanahan buku *Letter C*. Pengelompokan dilakukan dengan menggunakan algoritma *Partitioning Around Medoids* dengan 6 atribut yaitu persil, kelas desa, luas tanah, luas beli, luas jual, dan keterangan. Data Status Pertanahan *Letter C* dikelompokkan kedalam 3 *cluster* yaitu sudah terjual, belum terjual, dan habis terjual. Dari hasil pengelompokan diperoleh 6 data pada *cluster* sudah terjual, 344 data pada *cluster* belum terjual, dan 4 data pada *cluster* habis terjual dengan nilai *Davies Bouldin Index* sebesar 0,7804. Dengan pengelompokan *Partitioning Around Medoids* ini diharapkan dapat membantu petugas desa dalam melakukan pelayanan dan dapat menjaga kondisi dari buku *Letter C*.

Kata kunci: *Partitioning Around Medoids*, *Letter C*, Status Pertanahan

Clustering Land Status Data in *Letter C* Using the *Partitioning Around Medoids* Algorithm

Abstract

The *Letter C* serves as a land data archive managed by the village. The management of the land status data in *Letter C* in Pandanarum Village is still done manually by recording the history of sales and purchases in the *Letter C* books. In Pandanarum Village, there are 3960 *Letter C* data entries, which can affect the service process for residents. This study uses a sample size determined by *Slovin's* formula, amounting to 363 data entries taken from the third *Letter C* book. The study aims to classify the land status in the *Letter C* books. Classification is done using the *Partitioning Around Medoids* algorithm with 6 attributes: plot number, village class, land area, purchase area, sale area, and description. The land status data in *Letter C* is classified into 3 *clusters*: sold, unsold, and fully sold. From the classification results, there are 6 data points in the sold *cluster*, 344 data points in the unsold *cluster*, and 4 data points in the fully sold *cluster*, with a *Davies-Bouldin Index* value of 0.7804. This *Partitioning Around Medoids* classification is expected to assist village officials in providing services and maintaining the condition of the *Letter C* books.

Keywords: *Partitioning Around Medoids*, *Letter C*, Land Status

I. PENDAHULUAN

Desa Pandanarum merupakan salah satu dari sebelas kantor Desa yang berada di kecamatan sutojayan. Desa Pandanarum memiliki luas wilayah sebesar 369 hektar dengan 8300 penduduk[1]. Dengan luas wilayah dan padatnya penduduk maka diperlukan pelayanan yang baik kepada masyarakat khususnya terkait pelayanan pertanahan desa. Pelayanan ini mencakup tentang pengarsipan data *Letter C* dan PBB[2]. *Letter C* sendiri yakni dokumen pengarsipan data tanah dan digunakan untuk bukti atas kepemilikan tanah[3]. Pada desa Pandanarum terdapat pengarsipan terkait data *Letter C* yang masih dilakukan secara manual. Pengarsipan manual ini mengakibatkan kurang optimalnya dalam proses pelayanan kepada masyarakat dan dapat menimbulkan

kerusakan pada buku *Letter C* dikarenakan sering membolak-balikan halaman buku pada saat melakukan pelayanan. Hal ini diketahui dari pengumpulan data dengan teknik wawancara pada 1 November 2023. Gambar 1 merupakan dokumentasi dari kondisi buku *Letter C* pada Kelurahan Pandanarum.



Gambar 1. Kondisi Buku *Letter C*

Pada desa Pandanarum terdapat tiga buku *Letter C* dimana buku-buku tersebut saling berkelanjutan. Total data buku *Letter C* pada desa Pandanarum yaitu 3960 data nomor registrasi dengan 28 blok, 157 persil, 4 kelas desa daratan dan 3 kelas desa sawah. Dari hasil dari wawancara pada 1 November 2023 disarankan oleh pihak desa bahwa dalam penelitian ini menggunakan data dari buku ketiga *Letter C*. Hal ini dikarenakan pada buku ketiga *Letter C* terkait *history* kepemilikan tanah sudah jelas dan tidak ada penumpukan data.

Dari permasalahan pada pengarsipan data *Letter C* diatas, dapat diatasi dengan penggabungan antara pengarsipan data manual dengan teknologi informasi. Digitalisasi *Letter C* ini dapat menjaga kondisi dan data dari buku *Letter C*. Dalam digitalisasi ini digunakan teknik data mining dengan metode *clustering*. *Clustering* berfungsi untuk identifikasi kelompok yang sama pada sebuah dataset[4]. *Clustering* yang paling umum digunakan yaitu K-Means dan *Partitioning Around Medoids*. Dalam implementasinya algoritma *Partitioning Around Medoids* menghasilkan hasil pengelompokan yang lebih baik dibandingkan dengan algoritma *clustering* lainnya[5]. Hal ini dikarenakan algoritma *Partitioning Around Medoids* menghasilkan hasil yang stabil dan cepat dalam *konvergensi* data [6].

Dalam penerapan algoritma *Partitioning Around Medoids* digunakan evaluasi *Davies Bouldin Index* dalam memilih *cluster* terbaik. Dengan menggunakan evaluasi tersebut diharapkan jumlah *cluster* yang dibentuk dapat lebih optimal [7]. Selain itu, DBI juga dapat digunakan untuk mengatasi kelemahan algoritma dalam penentuan pusat *cluster* dengan mengukur validitas *cluster* terbaik[8]. Penelitian ini menerapkan jenis penelitian deskriptif dengan tahapan rumusan masalah, informasi yang diperlukan, prosedur pengumpulan data, pengolahan data, dan kesimpulan [9]. Pada tahap pengolahan data dilakukan dengan menerapkan metode *Cross-Industry Standart Process for Data Mining*.

Penelitian ini melanjutkan dari penelitian [1] yang melakukan penelitian terkait perancangan model bisnis dalam proses pengumpulan data serta pencarian data *Letter - C* pertanahan dengan hasil penelitian yaitu mengidentifikasi dan analisis kebutuhan sistem, merancang system. Penelitian ini juga melanjutkan dari penelitian [3] yang melakukan penelitian terkait pengelolaan dan pengamanan data arsip dengan basis web agar efektif dan efisien dengan hasil penelitian yaitu Sebuah sistem aplikasi pengelolaan dan pengamanan data *Letter C* berbasis web. Kebaruan dari dua penelitian diatas yaitu pada penelitian ini dilakukan pengelompokan status pertanahan dengan algoritma *Partitioning Around Medoids* dari buku 3 *Letter C* di Desa Pandanarum.

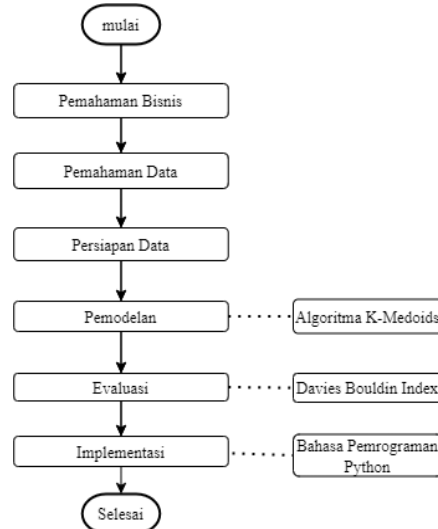
Selanjutnya terdapat beberapa penelitian yang sudah dilakukan terkait algoritma *Partitioning Around Medoids* yaitu pengelompokan daerah rawan banjir menggunakan algoritma *K-Medoids* [10] dengan menghasilkan penentuan daerah yang rawan akan banjir dengan menggunakan 7 atribut dan dikelompokkan kedalam 3 *cluster*. Selain itu terdapat penelitian pengelompokan ketahanan pangan [11] dengan menghasilkan

pengelompokan daerah yang hasil tanaman pangan tinggi di Sumatra dengan melakukan pengelompokan kedalam 3 *cluster* dan menggunakan metode evaluasi *Silhouette Coefficient*. Pembeda dari dua penelitian terkait algoritma *Partitioning Around Medoids* tersebut yaitu pada penelitian ini melakukan pemodelan algoritma *Partitioning Around Medoids* untuk pengelompokan data status pertanahan.

Dari identifikasi permasalahan diatas dapat dilakukan penerapan algoritma *Partitioning Around Medoids* dalam mengelompokkan status pertanahan. Dalam melakukan pengelompokan ini digunakan 6 atribut yaitu persil, kelas desa, luas tanah, luas jual, luas beli, dan keterangan. Dari 6 atribut tersebut dilakukan pengelompokan data ke 3 *cluster* yaitu habis terjual, sudah terjual, dan belum terjual. Hasil dari penelitian ini dapat membantu staff Desa Pandanarum dalam melayani masyarakat dan dapat menjaga kondisi dari buku *Letter C*.

II. METODOLOGI

Metodologi penelitian ini menerapkan metode *Cross-Industry Standart Process for Data Mining*. Metode ini merupakan metodologi data mining yang digunakan untuk membantu dalam menjalankan proyek dari fase perencanaan sampai implementasi[12]. Metode *CRISP-DM* ini juga dapat diartikan sebagai metodologi pengembangan proyek data dengan menetapkan tugas dan abstraksi serta untuk memberi fasilitas dalam implementasi dengan melalui tindakan pengambilan keputusan[13]. Metode ini digunakan sebagai kerangka panduan untuk memantau kesesuaian antara tugas dan keterampilan [14]. Metode *CRISP-DM* sudah teruji baik karena telah menerapkan pendekatan terstruktur dalam perencanaan proyek[15]. Gambar 2 merupakan flowchart dari metode *CRISP-DM*.



Gambar 2. Flowchart Metode *Cross-Industry Standart Process for Data Mining*

Berikut merupakan penjabaran dari gambar 2:

A. Pemahaman Bisnis

Pada fase pertama dilakukan pemahaman masalah yang terjadi pada penelitian. Pemahaman Bisnis pada penelitian ini yaitu pendataan *Letter C* masih dilakukan secara manual terutama pada status pertanahan. Penulisan secara

manual dapat menyulitkan petugas dalam melakukan pelayanan kepada masyarakat.

B. Pemahaman Data

Fase kedua yaitu dilakukan pengumpulan data yang berkaitan dengan penelitian. Pada proses pengumpulan data dilakukan 2 cara yaitu observasi dan wawancara. Pada proses pengumpulan data dengan observasi dilakukan pengamatan terkait Buku *Letter C*. Gambar 3 merupakan lembar *checklist* dari hasil observasi.

LEMBAR OBSERVASI			
Hari / Tanggal		: 1 November 2023	
Lokasi Observasi		: Kantor Desa Pandanarum Kecamatan Sutojayan	
Peneliti		: Putri Merly Deleo Karina	
NO	ASPEK YANG DIAMATI	YA	TIDAK
1	Buku <i>Letter C</i> tersedia di Kantor Desa Pandanarum	✓	
2	Struktur buku sesuai dengan pedoman yang berlaku	✓	
3	Buku <i>Letter C</i> tersimpan dengan baik	✓	
4	Kondisi buku pertama <i>Letter C</i> dalam keadaan baik		✓
5	Kondisi buku kedua <i>Letter C</i> dalam keadaan baik	✓	
6	Kondisi buku ketiga <i>Letter C</i> dalam keadaan baik	✓	
7	Penyimpanan buku <i>Letter C</i> terorganisir dengan rapi		✓
8	Kelengkapan data pemilik tanah pada setiap halaman buku <i>Letter C</i>		✓
9	Kelengkapan data pada setiap halaman		✓
10	Kejelasan penulisan pada buku pertama <i>Letter C</i>		✓
11	Kejelasan penulisan pada buku kedua <i>Letter C</i>	✓	
12	Kejelasan penulisan pada buku ketiga <i>Letter C</i>	✓	

Gambar 3. Lembar Checklist Hasil Observasi

Pada proses pengumpulan data dengan wawancara membahas terkait buku besar *Letter C*. Gambar 4 merupakan instrumen pengumpulan data wawancara.

Pertanyaan	Jawaban
Apakah kendala dalam melakukan pelayanan terkait data pertanahan pada Kelurahan Pandanarum?	Kendala dalam pelayanan data pertanahan yaitu proses pencarian serta hubungan status kepemilikan tanah.
Dalam proses pencarian apakah sering terjadi kerancuan data?	Benar. Karena pada buku Registrasi C hanya berisikan identitas nama, sedangkan nama warga bisa saja ganda.
Pada Kelurahan ini terdapat 2 jenis buku <i>Letter C</i> , apakah perbedaannya?	Buku kedua merupakan lanjutan dari buku pertama, pada buku pertama berisikan data dari nomor 1 hingga 3608 dan buku kedua berisikan data dari nomor 3609 hingga 3924.
Bagaimana penulisan dalam buku Registrasi C apabila terdapat transaksi jual beli tanah?	Staff desa hanya menuliskan tanah dijual ke nomor registrasi berapa dengan luas berapa pada kolom sebab perubahan pada buku Registrasi C
Apa kendala dalam status kepemilikan tanah?	Sulitnya menghubungkan antara pemilik tanah sebelumnya dan sesudahnya.

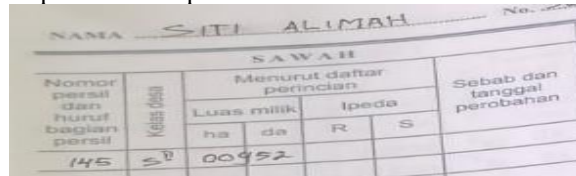
Gambar 4. Instrumen Pengumpulan Data Wawancara

Dari hasil pengumpulan data dapat dilakukan analisis data dengan memilah data yang akan digunakan. Pada penelitian ini menggunakan 2 jenis data yaitu data primer didapatkan dari wawancara kepada perangkat desa dan data sekunder didapatkan dari buku *Letter C*.

C. Persiapan Data

Pada persiapan data dilakukan perhitungan sampel data menggunakan teori slovin dari data populasi sebesar 3960 data *Letter C*. Teori slovin digunakan pada penelitian ini karena untuk pengambilan sampel dengan hasil yang dapat digeneralisasikan[16]. Data sampel yang didapatkan dari perhitungan sampel dengan *error tolerance* 5% yaitu sebesar 363 data *Letter C*. Pengambilan data dilakukan menggunakan Teknik Random Sampling dari buku ketiga *Letter C*. Proses berikutnya akan dilakukan cleaning data dengan tujuan untuk menghilangkan data yang kosong dan data yang tidak sesuai. Data kosong yang dimaksud yaitu pada halaman buku *Letter C* hanya terdapat nomor registrasi dan nama pemilik saja. Sedangkan data yang tidak sesuai yang dimaksud yaitu apabila terdapat nomor

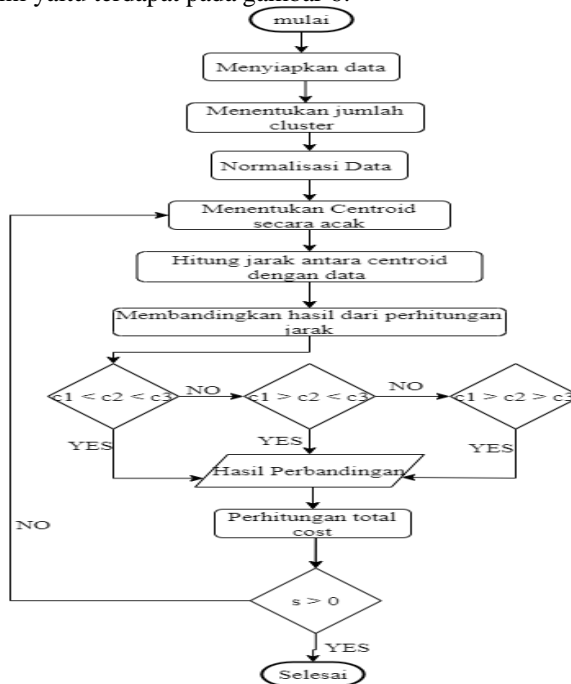
persil lebih dari 157 dan kelas desa selain kelas desa pada *Letter C* desa Pandanarum. Dari proses cleaning data didapatkan 9 data yang kosong dan data yang tidak sesuai. Sehingga data awal yang digunakan pada perhitungan *Partitioning Around Medoids* sebesar 354 data *Letter C*. Atribut yang digunakan dalam pengelompokan data status pertanahan yaitu nomor persil, kelas desa, luas tanah, luas jual, luas beli, dan keterangan. Hal ini diperoleh dari atribut yang terdapat dalam buku *Letter C*. Gambar 5 merupakan atribut pada buku *Letter C*.



Gambar 5. Atribut Buku *Letter C*

D. Pemodelan

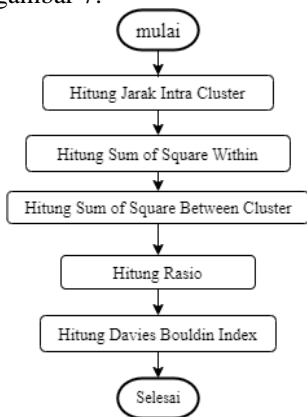
Pada fase pemodelan dilakukan penerapan data mining dimana pada penelitian ini melibatkan teknik *clustering* dengan penerapan algoritma *Partitioning Around Medoids*. Algoritma *Partitioning Around Medoids* merupakan algoritma *clustering* pada data mining [17]. Peneliti menggunakan algoritma *Partitioning Around Medoids* karena memiliki kelebihan pada proses perhitungan data dimana dengan semakin bertambahnya skala maka kinerja algoritma ini akan menjadi lebih kuat dan baik [18]. Selain itu, algoritma ini sudah diuji oleh penelitian [5] dengan hasil bahwa algoritma *Partitioning Around Medoids* menghasilkan hasil pengelompokan yang lebih baik. Perhitungan algoritma pada penelitian ini menggunakan 6 atribut dan dilakukan pengelompokan data kedalam 3 *cluster*. Langkah – langkah dari algoritma ini yaitu terdapat pada gambar 6:



Gambar 6. Flowchart Algoritma *Partitioning Around Medoids*

E. Evaluasi

Fase kelima yaitu evaluasi yaitu menguji algoritma yang digunakan. Pada penelitian ini digunakan pengujian algoritma *Partitioning Around Medoids* menggunakan *Davies Bouldin Index*. Tujuan dari evaluasi untuk mengetahui kualitas dari *cluster* yang telah digunakan. Langkah – langkah dari *Davies Bouldin Index* yaitu terdapat pada gambar 7:



Gambar 7. Flowchart Evaluasi *Davies Bouldin Index*

F. Implementasi

Fase yang terakhir yaitu implementasi dimana pada fase ini dilakukan implementasi dari perhitungan algoritma. Pada penelitian ini dilakukan implementasi algoritma *Partitioning Around Medoids* menggunakan Bahasa pemrograman *Python*.

Pada penelitian ini menerapkan metode *CRISP-DM* sampai langkah evaluasi. Untuk langkah implementasi dilakukan untuk bahan peneitian selanjutnya.

III. HASIL DAN PEMBAHASAN

Hasil dan pembahasan dari penelitian ini yaitu pemodelan algoritma *Partitioning Around Medoids* untuk pengelompokan data status pertanahan dan evaluasi menggunakan *Davies Bouldin Index*.

3.1 Pemodelan Algoritma *Partitioning Around Medoids*

A. Persiapan Data

Pada persiapan data dilakukan pemilihan atribut yang akan digunakan. Pada penelitian ini dilakukan pemisahan atribut dari kolom sebab perubahan pada buku *Letter C*. Kolom sebab perubahan dipisah menjadi 3 atribut yaitu luas jual, luas beli, dan keterangan. Sehingga atribut yang digunakan pada pengelompokan ini yaitu persil, kelas desa, luas tanah, luas jual, luas beli, dan keterangan. Pada atribut kelas desa dan keterangan dilakukan transformasi terlebih dahulu sebagai penunjang proses berikutnya yaitu imputasi data. Transformasi data pada sribut kelas desa terdapat pada table 1.

Tabel I
Transformasi Atribut Kelas Desa

Kelas Desa	Transformasi Kelas Desa
d-I	1
d-II	2
d-III	3
d-IV	4

s-I	5
s-II	6
s-III	7

Sedangkan transformasi pada atribut keterangan terdapat pada table 2.

Tabel II
Transformasi Atribut Keterangan

Keterangan	Transformasi Keterangan
JUAL	1
BELI	2

Pada tabel persiapan data bagian akhir ditambahkan dengan baris rata-rata. Hal ini digunakan sebagai penunjang dari proses berikutnya yaitu imputasi data. Tabel III merupakan data awal *Letter C*.

Tabel III
Data Awal *Letter C*

No. Reg	Persil	Kls. Ds.	I. Kls. Ds.	Luas TN	Luas JL	Luas BL	Ket	I. Ket.
3925	62	s-II	6	1114	450	-	JUAL	1
3841	2	d-II	2	740	-	-	-	-
3743	58	d-IV	4	742	-	-	-	-
3672	42	d-II	2	1547	1505	-	JUAL	1
...
3671	105	d-I	1	213	213	-	JUAL	1
3669	74	d-II	2	2048	667	-	JUAL	1
RATA-RATA					566,1	544,5		1,92

Keterangan dari tabel II yaitu, No Reg sebagai nomor registrasi, Kls Ds sebagai kelas desa, I Kls Ds sebagai transformasi kelas desa, Luas TN sebagai luas tanah, Luas JL sebagai luas jual, Luas BL sebagai luas beli, Ket sebagai keterangan, dan I Ket sebagai transformasi keterangan.

Dari tabel III yaitu tabel persiapan data dilanjutkan untuk proses imputasi data yang digunakan untuk mengisikan *missing values*. *Missing value* yaitu nilai yang tidak ada atau kosong pada dataset. *Missing value* dapat digantikan menggunakan salah satu metode imputasi data dengan nilai rata-rata[19]. Pada penelitian ini *missing value* terdapat pada atribut luas jual, luas beli, dan keterangan. Pada tabel imputasi nilai sitambahkan 2 baris pada bagian akhir yaitu baris nilai *minimum* dan nilai *maximum*. Hal ini digunakan sebagai penunjang proses berikutnya yaitu normalisasi data. Tabel IV merupakan imputasi data pada *Letter C*.

Tabel IV
Imputasi Data *Letter C*

No Reg	Persil	Kelas Desa	Luas Tanah	Luas Jual	Luas Beli	Ket
3925	62	6	1114	450	544,5	1
3841	2	2	740	566,1	544,5	1,92
3743	58	4	742	566,1	544,5	1,92
3672	42	2	1547	1505	544,5	1
...
3671	105	1	213	213	544,5	1
3669	74	2	2048	667	544,5	1
MAX	155	7	3097	1505	3097	2
MIN	1	1	42	175	49	1

B. Menentukan Jumlah Cluster

Pada penentuan jumlah *cluster* disesuaikan dengan buku *Letter C* dimana pada kolom sebab perubahan. Pada

kolom ini apabila tanah sudah terjual akan dituliskan nomor registrasi pembeli, apabila tanah hanya sebagian terjual dituliskan nomor registrasi pembeli dan luas tanah yang dijual, dan apabila tanah belum terjual dituliskan nomor registrasi penjual atau pada kolom sebab perubahan tidak tertulis apapun. Maka dari itu dilakukan pengelompokan data kedalam 3 cluster yaitu sudah terjual, belum terjual, dan habis terjual.

C. Normalisasi Data

Normalisasi data ditujukan untuk mengatasi rentang data yang terlalu jauh[20]. Selain itu normalisasi data juga digunakan untuk mempercepat perhitungan algoritma karena merubah skala nilai data dari 0 hingga 1[21]. *Min max normalization* dapat memberikan tingkat keakuratan data yang baik daripada normalisasi data dengan metode yang lain[22]. Pada normalisasi digunakan persamaan *Min max normalization* [21]. Berikut merupakan persamaan dari *min max normalization*:

$$X = \frac{x_i - \text{Min}(x)}{\text{Max}(x) - \text{Min}(x)} \tag{1}$$

Normalisasi pada penelitian ini dihitung menggunakan persamaan 1 dimana nilai *minimum* dan nilai *maximum* dapat dilihat pada tabel IV. Hasil dari normalisasi keseluruhan dapat dilihat pada tabel V.

Tabel V
Normalisasi Data Letter C

No Reg	Persil	Kelas Desa	Luas Tanah	Luas Jual	Luas Beli	Ket
3925	0,396	0,833	0,351	0,207	0,163	0,000
3841	0,006	0,167	0,228	0,294	0,163	0,922
3743	0,370	0,500	0,229	0,294	0,163	0,922
3672	0,266	0,167	0,493	1,000	0,163	0,000
...
3671	0,675	0,000	0,056	0,029	0,163	0,000
3669	0,474	0,167	0,657	0,370	0,163	0,000

D. Menentukan Centroid Acak Iterasi I

Langkah berikutnya yaitu menentukan titik pusat atau *centroid* untuk perhitungan pada *iterasi I* dengan mengambil pada data secara acak dengan jumlah sesuai dengan *cluster* yang digunakan. Pemilihan *centroid* secara acak pada algoritma *Partitioning Around Medoids* menjadi suatu pembeda dengan algoritma *clustering* yang lain [23]. Pada penelitian ini menggunakan 3 *cluster* sehingga memilih *centroid* secara acak sebanyak 3 data. *Centroid* yang digunakan yaitu data ke 15, data ke 45, dan data ke 335. Tabel VI merupakan 3 *centroid* acak pada *iterasi I*.

Tabel VI
Centroid Acak Iterasi I

No	Persil	Kelas Desa	Luas Tanah	Luas Jual	Luas Beli	Ket
15	0,396	0,833	0,351	0,207	0,163	0,000
45	0,779	0,500	0,078	0,294	0,076	1,000
335	0,675	0,000	0,044	0,000	0,163	0,000

E. Hasil Perhitungan Iterasi I

Selanjutnya dilakukan perhitungan jarak menggunakan persamaan *Euclidean Distance*, berikut merupakan persamaan dari *Euclidean Distance*:

$$d(i,j) = \sqrt{(x_{i1}-x_{j1})^2 + \dots + (x_{ik}-x_{jk})^2} \tag{2}$$

Euclidean Distance merupakan persamaan untuk menghitung jarak antara dua titik yaitu data normalisasi dengan titik *centroid*[24]. Dalam perhitungan jarak pada penelitian ini menggunakan data dari tabel V dengan *centroid* yang ada pada tabel VI. Setelah mendapatkan hasil perhitungan jarak, dilakukan perbandingan jarak dengan mencari nilai minimum antar *cluster*. Selanjutnya menghitung total *cost* dari *iterasi I* yaitu 199,1800. hasil akhir perhitungan *iterasi I* terdapat pada table VII.

Tabel VII
Hasil Akhir Perhitungan Iterasi I

No Reg	C1	C2	C3	Jarak Min	Ket Cluster
3925	0,000	1,1607	0,953	0,000	C1
3841	1,2119	0,8628	1,2023	0,8628	C2
3743	0,9920	0,4513	1,1463	0,4513	C2
3672	1,0539	1,4324	1,1819	1,0539	C1
...
3671	0,9440	1,1573	0,0312	0,0312	C3
3669	0,7554	1,2458	0,7623	0,7554	C1
TOTAL COST				199,1800	

Dari *iterasi I* didapatkan 5 data pada *cluster I*, 344 data pada *cluster II*, dan 5 data pada *cluster III*. Setelah perhitungan *iterasi I* selesai dilanjutkan pada perhitungan *iterasi II*. Hal ini dikarenakan masih terdapat suatu total *cost* yaitu pada *iterasi I*, sedangkan syarat penentuan berhenti atau tidak pada *iterasi* dihitung dengan selisih antar total *cost*. Sehingga pada penelitian ini dilanjutkan untuk perhitungan *iterasi II*.

F. Menentukan Centroid Acak Iterasi II

Perhitungan *Iterasi II* dilakukan dengan langkah-langkah yang sama pada perhitungan *iterasi I* yaitu langkah pertama menentukan *centroid* secara acak. *Centroid* acak *iterasi II* menggunakan data ke 302, data ke 153, dan data ke 335.

G. Hasil Perhitungan Iterasi II

Hasil perhitungan *iterasi II* diperoleh dengan cara yang sama seperti *iterasi I* yaitu dilakukan perhitungan jarak menggunakan persamaan *Euclidean Distance*, membandingkan *cluster* minimum, dan menghitung total *cost*. Table VIII merupakan hasil perhitungan akhir *iterasi II*.

Tabel VIII
Hasil Akhir Perhitungan Iterasi II

No Reg	C1	C2	C3	Jarak Min	Ket Cluster
3925	0,7058	1,1618	0,9537	0,7058	C1
3841	0,9471	0,2359	1,2023	0,2359	C2
3743	1,0087	0,3592	1,1463	0,3592	C2
3672	0,8571	1,1984	1,1819	0,8571	C1
...
3671	0,5943	1,0758	0,0312	0,0312	C3
3669	0,5126	1,0590	0,7623	0,5126	C1
TOTAL COST				166,3929	

Dari *iterasi II* didapatkan 6 data pada *cluster I*, 344 data pada *cluster II*, dan 4 data pada *cluster III*. Selanjutnya dilakukan perhitungan *cost* antara *iterasi II* dengan *iterasi I* dengan cara menghitung selisih antara kedua *cost iterasi*. Total *cost iterasi I* dan *iterasi II*

diperoleh hasil -32,7871. Dikarenakan hasil perhitungan *cost* masih bernilai *negative* maka dilanjutkan ke *iterasi* II.

H. Menentukan Centroid Acak Iterasi III

Perhitungan *Iterasi* III dilakukan dengan langkah-langkah yang sama pada perhitungan *iterasi* I dan *Iterasi* II yaitu langkah pertama menentukan *centroid* secara acak. *Centroid* acak *iterasi* III menggunakan data ke 298, data ke 9, dan data ke 197.

I. Hasil Perhitungan Iterasi III

Hasil perhitungan *iterasi* III diperoleh dengan cara yang sama seperti *iterasi* I dan *Iterasi* II yaitu dilakukan perhitungan jarak menggunakan persamaan *Euclidean Distance*, membandingkan *cluster* minimum, dan menghitung total *cost*. Table IX merupakan hasil perhitungan akhir *iterasi* III.

Tabel IX
Hasil Akhir Perhitungan Iterasi III

No Reg	C1	C2	C3	Jarak Min	Ket Cluster
3925	0,7554	1,2195	0,9440	0,7554	C1
3841	1,1214	0,2986	1,1938	0,2986	C2
3743	1,0772	0,3788	1,1373	0,3788	C2
3672	0,6834	1,2479	1,1530	0,6834	C1
...
3671	0,7387	1,1525	0,0000	0,0000	C3
3669	0,0000	1,0977	0,7387	0,0000	C1
TOTAL COST				179,3771	

Perhitungan total *cost* antara selisih *iterasi* III dengan *iterasi* II diperoleh hasil 12,9842 sehingga perhitungan *iterasi* dihentikan dikarenakan hasil selisih *cost* bernilai positif dan hasil *cluster* tidak berubah. Table X merupakan hasil akhir dari perhitungan algoritma *Partitioning Around Medoids*.

Tabel X
Hasil Akhir Perhitungan Algoritma Partitioning Around Medoids

No Reg	Jarak Min	Ket Cluster
3925	0,9537	C3
3841	0,2872	C2
3743	0,3702	C2
3672	0,0000	C1
...
3671	0,0312	C3
3669	0,6834	C1

3.2 Evaluasi Davies Bouldin Index

Setelah selesai dilakukan pemodelan algoritma *Partitioning Around Medoids* dilakukan evaluasi menggunakan *Davies Bouldin Index*. Dataset yang digunakan pada evaluasi ini yaitu table V dan *centroid* acak pada *iterasi* III. Langkah-langkah dari evaluasi *Davies Bouldin Index* terdapat pada *flowchart* gambar 4 dengan penjelasan sebagai berikut:

A. Hitung Jarak Intra Cluster

Jarak intra *cluster* dihitung sesuai hasil *cluster* sesuai dengan hasil akhir algoritma *Partitioning Around Medoids*. Pada perhitungan jarak intra *cluster* digunakan persamaan *Euclidean Distance* persamaan 2. Pada *cluster* I terdapat 6 data yaitu data ke 12, 15, 295, 298, 302, dan data ke 304. Pada *cluster* II terdapat 344 data yaitu data ke

1, 2, 3, hingga data ke 354. Pada *cluster* III terdapat 4 data yaitu data ke 245, 297, 299, dan data ke 335. Dari masing-masing data tersebut dilakukan perhitungan terhadap masing-masing *centroid*.

B. Hitung Sum of Square Within

Sum of Square Within atau biasa disebut dengan SSW dihitung dari nilai rata-rata setiap *cluster*. Rata-rata ini diperoleh dari total keseluruhan jarak intra *cluster* disetiap *cluster* dibagi dengan jumlah data disetiap *cluster*. Pada penelitian ini terdapat 3 *Sum of Square Within* yaitu *Sum of Square Within cluster* I, *Sum of Square Within cluster* II, dan *Sum of Square Within cluster* 3. Table XI merupakan hasil dari *Sum of Square Within*.

Tabel XI
Hasil dari Sum of Square Within

	TOTAL DATA	JUMLAH DATA	NILAI SSW
Cluster I	1,855	6	0,309
Cluster II	175,745	344	0,511
Cluster III	1,119	4	0,280

C. Hitung Sum of Square Between Cluster

Sum of Square Between atau biasa disebut dengan SSB dihitung menggunakan persamaan 2 yaitu persamaan *Euclidean Distance* dimana jarak yang dihitung merupakan antar *centroid*. Pada penelitian ini terdapat 3 SSB yaitu SSB *centroid* 1 dengan *centroid* II, SSB *centroid* 1 dengan *centroid* 3, dan SSB *centroid* II dengan *centroid* III. Table XII merupakan hasil dari *Sum of Square Between*.

Tabel XII
Hasil Sum of Square Between

	Centroid I	Centroid II	Centroid III
Centroid I	0	1,098	0,739
Centroid II	1,098	0	1,152
Centroid III	0,739	1,152	0

D. Hitung Rasio

Perhitungan rasio diperoleh dari penjumlahan 2 SSW dibagi dengan SSB. Sehingga diperoleh hasil rasio *cluster* I dengan *cluster* 2 yaitu 0,7470, rasio *cluster* II dengan *cluster* III yaitu 0,7972, dan rasio *cluster* II dengan *cluster* III yaitu 0,6860.

E. Hitung Davies Bouldin Index

Langkah terakhir pada evaluasi *Davies Bouldin Index* yaitu menghitung DBI itu sendiri. Perhitungan ini diperoleh dari rata-rata nilai maximum rasio. Table XIII merupakan hasil akhir dari evaluasi *Davies Bouldin Index*.

Tabel XIII
Hasil Akhir dari Evaluasi Davies Bouldin Index

No	1	2	3	Rasio Max	DBI
1	0	1,098	0,739	0,7972	0,7804
2	1,098	0	1,152	0,7470	
3	0,739	1,152	0	0,7972	

Berdasarkan hasil evaluasi *Davies Bouldin Index* 3 *cluster* pada table XIII yaitu 0,7804. Dengan langkah-langkah evaluasi *Davies Bouldin Index* yang sama

dilakukan evaluasi terhadap 2 *cluster* dan memperoleh hasil 0,8023. Menurut [25] hasil evaluasi DBI terkecil dan mendekati 0 bernilai *positif* merupakan *cluster* yang digunakan baik. Sehingga pengelompokan data status pertanahan menggunakan 3 *cluster* termasuk *cluster* terbaik.

Hasil dari pengelompokan data status pertanahan pada penelitian ini yaitu 6 data termasuk kedalam *cluster* 1 sudah terjual, 344 data termasuk kedalam *cluster* 2 belum terjual, dan 4 data termasuk kedalam *cluster* 3 habis terjual. Dari hasil tersebut pihak desa dapat mengetahui pengelompokan dari masing-masing status pertanahan tanpa membuka arsip buku *Letter C*. hal ini dapat membantu staff desa dalam melakukan pelayanan dan dapat menjaga kondisi dari buku Registrasi C.

IV. KESIMPULAN

Berdasarkan penelitian yang telah dilakukan, disimpulkan bahwa jumlah *cluster* untuk pengelompokan data status pertanahan yaitu 3 *cluster* diantaranya sudah terjual, belum terjual, dan habis terjual merupakan *cluster* terbaik dengan nilai DBI 0,7804. Hasil pengelompokan data status pertanahan yaitu 6 data termasuk *cluster* I dengan kategori sudah terjual, 344 data termasuk *cluster* II dengan kategori belum terjual, dan 4 data termasuk *cluster* III dengan kategori habis terjual. Hasil dari penelitian ini dapat dilanjutkan ke penelitian berikutnya terkait implementasi menggunakan Bahasa pemrograman *Python* sehingga dapat merealisasikan hasil pengelompokan data status pertanahan pada penelitian ini. Dari hasil penelitian diharapkan dapat menjadi masukan kepada pihak kantor desa dalam melayani masyarakat terkait administrasi *Letter C*. Dengan hasil penelitian ini dapat meminimalisir dalam menggunakan buku *Letter C* sehingga dapat menjaga kondisi buku *Letter C*.

DAFTAR PUSTAKA

- [1] I. H. Santi, F. Febrinita, dan W. D. Puspitasari, "Engineering Design Business Process Modelling *Letter C* Land Data Archiving System with Software Requirement Specifications Approach," *J. Intell. Decis. Support Syst.*, vol. 6, no. 4, hal. 231–240, 2023, doi: <https://doi.org/10.35335/idss.v6i4.187>.
- [2] A. Sahrina, Purwanto, F. Rosyida, D. Prasetyono, I. Sulistio, dan C. E. Pamungkas, "Digitalisasi Dokumen Pertanahan Desa Berbasis Webgis Di Desa Sukorejo Kecamatan Gondanglegi Kabupaten Malang," *MARTABE J. Pengabd. Masy.*, vol. 6, no. 12, 2023, doi: [10.31604/jpm.v6i12.4299-4306](https://doi.org/10.31604/jpm.v6i12.4299-4306).
- [3] E. Setiawan, I. H. Santi, dan S. N. Budiman, "Sistem Pengelolaan Dan Pengamanan Arsip Data *Letter C* Desa (Studi Kasus : Kantor Desa Gondang)," *J. Mhs. Tek. Inform.*, vol. 6, no. 2, 2022, doi: <https://doi.org/10.36040/jati.v6i2.5602>.
- [4] N. Sureja, B. Chawda, dan A. Vasant, "An improved K-medoids *clustering* approach based on the crow search algorithm," *J. Comput. Math. Data Sci.*, vol. 3, no. April, hal. 100034, 2022, doi: [10.1016/j.jcmds.2022.100034](https://doi.org/10.1016/j.jcmds.2022.100034).
- [5] I. M. Karo Karo, S. Dewi, Mardiana, F. Ramadhani, dan P. Harliana, "K-Means and K-Medoids Algorithm Comparison for *Clustering* Forest Fire Location in Indonesia," *J. Ecotipe (Electronic, Control. Telecommun. Information, Power Eng.*, vol. 10, no. 1, hal. 86–94, 2023, doi: [10.33019/jurnalecotipe.v10i1.3896](https://doi.org/10.33019/jurnalecotipe.v10i1.3896).
- [6] W. Y. E, A. Jian, L. Yan, dan W. HongGang, "Optimization of K-medoids Algorithm for Initial *Clustering* Center," *J. Phys. Conf. Ser.*, vol. 1487, no. 1, 2020, doi: [10.1088/1742-6596/1487/1/012011](https://doi.org/10.1088/1742-6596/1487/1/012011).
- [7] Y. Sopyan, A. D. Lesmana, dan C. Juliane, "Analisis Algoritma K-Means dan *Davies Bouldin Index* dalam Mencari *Cluster* Terbaik Kasus Perceraian di Kabupaten Kuningan," *Build. Informatics, Technol. Sci.*, vol. 4, no. 3, hal. 1464–1470, 2022, doi: [10.47065/bits.v4i3.2697](https://doi.org/10.47065/bits.v4i3.2697).
- [8] M. Mughnyanti, S. Efendi, dan M. Zarlis, "Analysis of determining *centroid clustering* x-means algorithm with *davies-bouldin* index evaluation," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 725, no. 1, 2020, doi: [10.1088/1757-899X/725/1/012128](https://doi.org/10.1088/1757-899X/725/1/012128).
- [9] Salim, *Penelitian Pendidikan: Metode, Pendekatan, dan Jenis*. Kencana, 2019. [Daring]. Tersedia pada: https://www.google.co.id/books/edition/Penelitian_Pendidikan_Metode_Pendekatan/2fq1DwAAQB_AJ?hl=en&gbpv=0
- [10] D. Kurmiati, M. Z. Fauzi, A. Falegas, dan Indria, "Klasterisasi Daerah Rawan Gempa Bumi di Indonesia Menggunakan Algoritma K-Medoids," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 1, hal. 47–57, 2021, doi: <https://doi.org/10.57152/malcom.v1i1.61>.
- [11] N. P. Dharshinni dan C. Fandi, "Penerapan Metode K-Medoids *Clustering* Untuk Mengelompokkan Ketahanan Pangan," *J. MEDIA Inform. BUDIDARMA*, vol. 6, no. 4, hal. 2301, Okt 2022, doi: [10.30865/mib.v6i4.4939](https://doi.org/10.30865/mib.v6i4.4939).
- [12] F. Sulianta, *Basic Data Mining from A to Z*. Feri Sulianta, 2023. [Daring]. Tersedia pada: https://www.google.co.id/books/edition/Basic_Data_Mining_from_A_to_Z/JcLhEAAAQBAJ?hl=en&gbpv=0
- [13] C. E. D. Vanegas, J. C. G. Mejía, F. A. V. Agudelo, dan D. E. S. Duran, "A Representation Based on Essence for the *CRISP-DM* Methodology," *Comput. y Sist.*, vol. 27, no. 3, hal. 675–689, 2023, doi: [10.13053/CyS-27-3-3446](https://doi.org/10.13053/CyS-27-3-3446).
- [14] J. Bokrantz, M. Subramaniam, dan A. Skoogh, "Realising the promises of artificial intelligence in manufacturing by enhancing *CRISP-DM*," *Prod. Plan. Control*, vol. 0, no. 0, hal. 1–21, 2023, doi: [10.1080/09537287.2023.2234882](https://doi.org/10.1080/09537287.2023.2234882).
- [15] B. A. Kuncoro, *Pengenalan Prinsip Data Science untuk Pemula*.

- https://www.google.co.id/books/edition/Pengenalan_Prinsip_Data_Science_untuk_Pe/O8kQEAAAQBAJ?hl=en&gbpv=0, 2020. [Daring]. Tersedia pada:
- https://www.google.co.id/books/edition/Pengenalan_Prinsip_Data_Science_untuk_Pe/O8kQEAAAQBAJ?hl=en&gbpv=0
- [16] Rifkhan, *Pedoman Metodologi Penelitian Data Panel Dan Kuesioner*. Penerbit Arab, 2023. [Daring]. Tersedia pada: https://www.google.co.id/books/edition/PEDOMAN_METODOLOGI_PENELITIAN_DATA_PANEL/UN2vEAAAQBAJ?hl=en&gbpv=1&dq=PEDOMAN+METODOLOGI+PENELITIAN+DATA+PANEL+DAN+KUESIONER&pg=PA39&printsec=frontcover
- [17] Mesran dkk., *Sistem Pendukung Keputusan & Data Mining: Metode dan Penerapannya Dalam Pengambilan Keputusan*. Green Press, 2020. Diakses: 14 Desember 2023. [Daring]. Tersedia pada: https://www.google.co.id/books/edition/Sistem_Pendukung_Keputusan_Data_Mining_M/eDXgDwAAQBAJ?hl=en&gbpv=0
- [18] Z. Zhang, T. Wu, X. Sun, dan J. Yu, "MPDP k-medoids: Multiple partition differential privacy preserving k-medoids clustering for data publishing in the Internet of Medical Things," *Int. J. Distrib. Sens. Networks*, vol. 17, no. 10, 2021, doi: 10.1177/15501477211042543.
- [19] J. Han, M. Kamber, dan J. Pei, *Data Mining: Concepts and Techniques*. 2011. doi: 10.1016/C2009-0-61819-5.
- [20] N. H. Harani dan F. A. Nugraha, *Segmentasi Pelanggan Menggunakan Python*. Kreatif, 2020. [Daring]. Tersedia pada: https://www.google.co.id/books/edition/Segmentasi_Pelanggan_Menggunakan_Python/Bqr7DwAAQBAJ?hl=en&gbpv=1&dq=Segmentasi+Pelanggan+Menggunakan+Python&pg=PA148&printsec=frontcover
- [21] I. Werdiningsih, D. C. R. Novitasari, dan D. Z. Haq, *Pengelolaan Data Mining dengan Pemrograman Matlab*. 2022. [Daring]. Tersedia pada: https://www.google.co.id/books/edition/Pengelolaan_Data_Mining_dengan_Pemrogram/CgOdEAAQBAJ?hl=en&gbpv=1&dq=Pengelolaan+Data+Mining+dengan+Pemrograman+Matlab&printsec=frontcover
- [22] T. B. Sasongko, "Komparasi dan Analisis Kinerja Model Algoritma SVM dan PSO-SVM (Studi Kasus Klasifikasi Jalur Minat SMA)," *J. Tek. Inform. dan Sist. Inf.*, vol. 2, no. 2, 2016, doi: <https://doi.org/10.28932/jutisi.v2i2.627>.
- [23] D. D. Aulia dan Nurahman, "Comparison Performance of K-Medoids and K-Means Algorithms In Clustering Community Education Levels," *J. Nas. Pendidik. Tek. Inform.*, vol. 12, no. 2, hal. 273–282, 2023, doi: 10.23887/janapati.v12i2.59789.
- [24] D. Jollyta, M. Siddik, H. Mawengkang, dan S. Efendi, *Teknik Evaluasi Cluster Solusi Menggunakan Python Dan Rapidminer*. Deepublish, 2021. [Daring]. Tersedia pada: https://www.google.co.id/books/edition/Teknik_Evaluasi_Cluster_Solusi_Menggunakan/3rcgEAAAQBAJ?hl=en&gbpv=0
- [25] Mustika dkk., *Data Mining Dan Aplikasinya*. Penerbit Widina, 2021. [Daring]. Tersedia pada: https://www.google.co.id/books/edition/DATA_MINING_DAN_APLIKASINYA/53FXEAAAQBAJ?hl=en&gbpv=0