



Perbandingan Jarak Euclidean dan Manhattan pada Pemetaan Potensial Tanaman Padi Menggunakan K-Means

Khoirul Rizky Asrofi, Risqi Darma Rusdyan, Vion Age Trichahyo*

Fakultas Ilmu Eksakta, Program Ilmu komputer, Universitas Nahdlatul Ulama Blitar, Blitar

Jl. Masjid No.22, Kauman, Kec. Kepanjenkidul, Kota Blitar, Jawa Timur, Indonesia

Email: ¹rizkykhoirul@gmail.com, ²risqidarma@gmail.com, ^{3,*}vionage@gmail.com

Email Penulis Korespondensi: vionage@gmail.com

Submitted: 30/12/2024; Accepted: 31/01/2025; Published: 31/01/2025

Abstrak—Penelitian ini bertujuan untuk menganalisis potensi produksi padi di Jawa Timur dengan menggunakan 2 metode yaitu Euclidean Distance dan Manhattan Distance. Dengan menerapkan kedua metode ini, Dinas Pertanian dapat memperoleh pemahaman yang lebih mendalam mengenai potensi produksi padi di setiap kabupaten atau kota. Hasil analisis menunjukkan bahwa K-Means dengan Euclidean Distance dan Manhattan Distance dapat digunakan untuk mengidentifikasi wilayah dengan potensi padi rendah, sedang, dan tinggi. Perbandingan antara kedua metode menggunakan Silhouette Coefficient menunjukkan bahwa Euclidean Distance memiliki nilai yang lebih unggul, yaitu 0,61, dibandingkan dengan Manhattan Distance yang mencapai 0,58. Temuan ini memberikan solusi praktis bagi Dinas Pertanian dalam upaya meningkatkan ketahanan pangan dan produksi padi di Jawa Timur, serta memberikan kontribusi akademis dalam bidang analisis data pertanian.

Kata Kunci: Euclidean distance; Manhattan distance; Padi; K-means

Abstract—This research aims to analyze the potential of rice production in East Java by using 2 methods, namely Euclidean Distance and Manhattan Distance. By applying these two methods, the Department of Agriculture can gain a deeper understanding of the potential of rice production in each district or city. The analysis results show that K-Means with Euclidean Distance and Manhattan Distance can be used to identify areas with low, medium, and high rice potential. Comparison between the two methods using Silhouette Coefficient shows that Euclidean Distance has a superior value of 0.61, compared to Manhattan Distance which reaches 0.58. The findings provide practical solutions for the Department of Agriculture in an effort to improve food security and rice production in East Java, as well as making an academic contribution in the field of agricultural data analysis

Keywords: Euclidean Distance; Manhattan Distance; Rice; K-Means

1. PENDAHULUAN

Perkembangan Teknologi Informasi dan Komunikasi yang sudah sangat pesat telah digunakan dalam berbagai bidang salah satunya ialah bidang pertanian. Sektor pertanian merupakan salah satu sektor utama sebagai penghasil pangan[1]. Sektor pertanian adalah salah satu dari banyak sektor yang mendorong pertumbuhan ekonomi Indonesia[2]. Tanaman pangan memiliki banyak jenis, salah satunya adalah tanaman padi. Padi adalah sumber bahan pangan yang menghasilkan beras. Dengan jumlah total sekitar 59,2 juta ton beras, Indonesia menempati peringkat ketiga di dunia dalam produksi beras, menurut FAO (FAO)[3]. Secara tidak langsung, potensi beras di Jawa Timur diukur dengan jarak Euclidean dan Manhattan. Secara keseluruhan, beras memainkan peran penting dalam sistem ketahanan pangan global. Beras tidak hanya memberikan manfaat ekonomi dan sosial, tetapi juga merupakan sumber kalori utama. Oleh karena itu, untuk memastikan pasokan pangan yang cukup dan merata di masa depan, strategi ketahanan pangan yang lebih luas harus mencakup mempertahankan produksi beras, meningkatkan hasil panen, dan melindungi petani padi. Selain menjadi komoditas strategis yang memainkan peran penting dalam perekonomian dan ketahanan pangan nasional, beras merupakan prioritas utama dalam revitalisasi pertanian ke depan[4]. Pemilihan antara jarak Euclidean dan Manhattan dalam penelitian perbandingan produksi padi di Jawa Timur harus mempertimbangkan tujuan analisis dan karakteristik data. Sementara jarak Manhattan lebih relevan untuk mempertimbangkan faktor-faktor praktis dalam distribusi dan aksesibilitas produksi padi, jarak Euclidean mungkin lebih tepat untuk analisis potensial hasil[5].

Meskipun beras merupakan produk penting untuk memenuhi kebutuhan pangan di provinsi Jawa Timur, tidak dapat dipastikan bahwa beras merupakan produk unggulan di setiap kota atau kabupaten. Akibatnya, untuk memenuhi kebutuhan beras di Jawa Timur, perlu dilakukan upaya untuk meningkatkan produksi beras di setiap kota atau kabupaten. Data tentang hasil panen atau produksi beras di setiap kota atau kabupaten dapat digunakan sebagai acuan untuk menentukan potensi lahan pertanian padi di daerah tersebut. Algoritma K-Means mengelompokkan data berdasarkan titik pusat klaster (Centeroid) yang paling dekat[6]. Untuk mengelompokkan data, K-Means berusaha untuk meminimalkan kesamaan data antar klaster dan memaksimalkan kesamaan data dalam klaster[7]. Penelitian ini bertujuan untuk memetakan daerah potensial padi di Jawa Timur dari tahun 2017-2023 dengan mengelompokkan potensi produksi padi dengan menggunakan perbandingan jarak euclidean dan manhattan menggunakan K-means. Dalam penelitian lainnya seperti Perbandingan Penggunaan Jarak Manhattan, Jarak Euclid, dan Jarak Minkowski dalam Klasifikasi Menggunakan Metode KNN pada Data Iris oleh Adi Setiawan dengan hasil diperoleh akurasi rata-rata 96,6667 % [8].

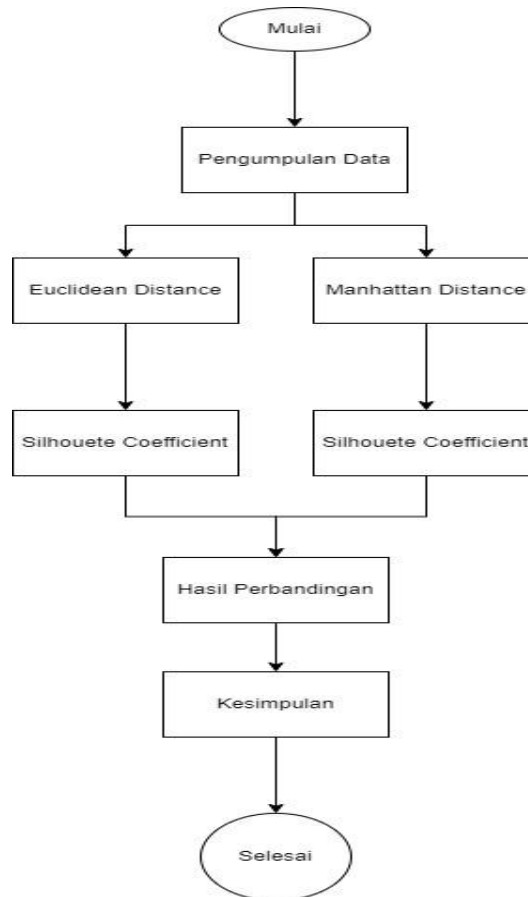
Dari latar belakang tersebut, untuk menentukan kota atau kabupaten mana saja yang memiliki potensial padi di Jawa Timur, perlu dilakukan pengelompokan potensi padi. Ini dapat dilakukan dengan menggunakan

perhitungan jarak Euclidean dan Manhattan pada algoritma K-Means untuk memetakan potensial padi. Ini dapat membantu Dinas Pertanian Jawa Timur membuat keputusan tentang kota atau kabupaten mana yang memiliki potensial padi rendah, sedang, atau tinggi untuk menentukan mana yang harus diprioritaskan untuk menanam padi. Dimungkinkan untuk meningkatkan kontribusi terhadap kebijakan pertanian dan strategi pengelolaan lahan di Jawa Timur dengan mengutamakan pengembangan kebijakan yang berkelanjutan, peningkatan produktivitas pertanian, dan pengelolaan sumber daya alam yang efisien. Selain itu, studi ini dapat menyelidiki dampak sosial dan ekonomi kebijakan-kebijakan tersebut terhadap kesejahteraan petani dan ketahanan pangan.

2. METODOLOGI PENELITIAN

2.1 Tahap Penelitian

Metode perbandingan jarak Euclidean dan Manhattan digunakan dalam penelitian ini. Diagram alir berikut menunjukkan tindakan yang akan dilakukan oleh para peneliti.



Gambar 1. Flowchart

Langkah awal yang dilakukan dalam penelitian ini adalah mencari data produksi padi yang ada di kota/kabupaten di Jawa Timur. Data diperoleh dari website BPS Jawa Timur dengan alamat <https://www.bpsjatim.co.id>. Data yang diperoleh yaitu hasil produksi padi mulai dari tahun 2017-2023 dengan jumlah 29 kabupaten dan 9 kota di Jawa Timur[9]. Selanjutnya, data yang dikumpulkan diproses dengan menggunakan metode jarak, seperti jarak Euclidean dan jarak Manhattan. Langkah ini bertujuan untuk menentukan seberapa dekat data antara titik-titik dalam ruang. Selain itu, hasil clustering dievaluasi dengan silhouette coefficient. Tujuan dari langkah ini adalah untuk menentukan metode jarak mana yang paling efektif untuk membentuk cluster. Selain itu, setelah perbandingan dilakukan, metode jarak mana yang paling efektif untuk menggambarkan potensi produksi padi di Jawa Timur. Data awal yang digunakan dalam penelitian ini terdapat dalam Tabel 1.

Tabel 1. Produksi padi

Tahun	2017	2018	2019	2020	2021	2022	2023
Kabupaten							
Pacitan	120 134.0	93 787	91,941.66	83,377.94	88,116.57	90,955.25	92 993
Ponorogo	432 485.0	377 367	322,205.50	396,816.11	404,665.04	370,435.11	392 994
Trenggalek	167 690.0	104 712	112,213.57	109,241.07	116,456.34	117,346.67	114 875



Tahun	2017	2018	2019	2020	2021	2022	2023
Kabupaten							
Tulungagung	295 498.0	219 252	196,430.58	223,382.97	237,916.96	221,637.05	235 502
Blitar	291 959.0	221 520	224,027.19	201,803.84	247,366.27	217,566.97	240 224
Kediri	267 043.0	220 453	222,837.82	218,624.90	198,222.01	172,474.88	183 534
Malang	421 214.0	284 583	281,072.26	283,096.75	273,358.61	283,895.29	279 366
Lumajang	437 054.0	342 752	283,894.22	294,865.05	295,075.52	303,468.63	308 646
Jember	910 979.0	745 410	616,858.41	602,794.88	615,697.87	613,237.38	616 726
Banyuwangi	772 429.0	532 815	445,253.86	462,126.36	513,490.07	462,584.81	454 768
Bondowoso	399 426.0	282 307	251,371.94	270,271.50	258,951.46	246,388.27	247 779
Situbondo	294 043.0	186 375	167,665.54	160,972.37	151,157.12	141,914.27	157 273
Probolinggo	244 924.0	179 915	193,774.50	196,516.22	190,180.14	187,277.08	201 044
Pasuruan	692 520.0	327 338	269,463.05	279,072.07	264,950.78	254,578.42	260 062
Sidoarjo	204 425.0	239 183	234,788.11	225,346.46	202,501.40	196,839.63	195 855
Mojokerto	333 511.0	309 535	339,755.88	316,759.39	297,042.32	287,251.32	302 891
Jombang	444 741.0	439 002	344,236.34	345,349.53	326,826.64	343,427.84	346 873
Nganjuk	503 707.0	404 586	399,845.68	444,108.03	429,311.01	387,897.28	418096
Madiun	513 384.0	425 023	419,292.44	453,540.93	461,798.12	419,977.93	437 593
Magetan	337 652.0	315 400	260,671.35	283,304.50	307,279.68	263,822.71	280 592
Ngawi	751 885.0	753 199	777,190.36	829,467.64	786,475.65	785,037.99	771 251
Bojonegoro	802 258.0	757 441	692,073.16	737,397.60	674,002.00	715,198.84	705 963
Taban	556 049.0	598 039	519,934.36	520,811.26	489,418.62	502,136.24	501 741
Lamongan	887 072.0	924 212	839,724.43	873,786.03	792,662.09	920,935.59	798 705
Gresik	445 099.0	362 252	367,717.66	409,985.27	379,666.19	410,323.14	417 429
Bangkalan	300 218.0	247 867	201,620.00	202,885.36	195,323.29	193,329.37	186 258
Sampang	189 258.0	208 665	156,219.52	187,323.59	195,600.69	172,558.93	167 191
Pamekasan	110 044.0	118 139	91,313.32	100,727.49	96,723.97	108,020.20	108 483
Sumenep	221 707.0	230 486	186,090.84	229,595.79	221,979.41	223,000.46	217 810
Kota/Municipality							
Kediri	8 449.0	16 659	7,032.99	8,784.91	9,534.86	10,449.52	8 272
Blitar	6 105.0	5 137	6,084.58	5,258.60	5,793.30	5,259.99	5 637
Malang	13 650.0	10 140	13,909.70	12,459.91	11,311.40	11,524.42	10 319
Probolinggo	10 160.0	5 940	7,165.20	7,667.56	8,924.08	8,067.16	7 502
Pasuruan	15 721.0	14 029	10,804.71	10,037.32	8,304.59	7,939.57	8 020
Mojokerto	4 975.0	4 903	3,565.88	4,650.47	4,414.93	4,324.18	3 293
Madiun	12 987.0	8 662	8,259.39	14,225.77	13,505.81	11,585.31	12 170
Surabaya	9 685.0	14 519	9,596.78	10,798.50	9,832.67	8,180.52	8 069
Batu	2 653.0	6 318	5,031.10	5,153.04	5,750.19	5,912.16	4 863
Jawa Timur	12 432	10 537	9,580,934	10,022,387	9,789,587.6	9,686,760.3	9 710
	793.0	922			7	8	661

2.2 Pre-processing data

Langkah ini dilakukan untuk memastikan data input yang akan diproses dalam perhitungan memiliki kualitas data yang baik, serta untuk menghindari kesalahan, nilai yang hilang, dan data yang tidak konsisten yang dapat menyebabkan masalah selama proses penelitian. Proses preprocessing data dalam penelitian ini termasuk reduction dan cleaning serta transformation.

2.2.1 Data Reduction

Data file yang dikumpulkan mengandung data atribut yang digunakan, tetapi tidak digunakan dalam penelitian. Terdapat tiga atribut yang dikumpulkan dari BPS Jatim: luas panen, produksi padi, dan produktivitas. Dalam penelitian ini, sejumlah 7 atribut digunakan, yang berasal dari produksi padi dari tahun 2017 hingga 2023.

2.2.2 Data cleaning

Data cleaning diterapkan pada beberapa data produksi padi tertentu, pembersihan data dilakukan secara manual dengan menggunakan Microsoft excel, dengan memilih semua data seperti pada tahun sebelumnya. Ini dilakukan karena kesalahan penggunaan spasi pada tabel data tahun 2018 dan 2023 yang dapat menyebabkan kesalahan saat sistem dijalankan.

2.2.3 Data transformation

Transformation adalah tahap membuat data yang akan diproses menjadi sesuai dengan model ataupun algoritma yang ingin digunakan dalam tahap pemrosesan data. Teknik bertujuan untuk mengubah atribut numerik menjadi atribut kategorikal, sehingga menjadi beberapa tingkatan kategori[10]. Data yang telah siap untuk proses data



mining dapat diproses melalui proses data transformation[11]. Pada penelitian ini transformation data digunakan pada atribut “2017 dan 2023”. Adapun data yang digunakan dalam pengelompokan dapat dilihat pada Tabel 2 [12].

Tabel 2. Data transformation

No.	Kriteria	Satuan
1.	Kota/kabupaten	Nama kota/kabupaten
2.	Produksi padi mulai tahun 2017-2023	Ton

2.2.4 Perhitungan manual

a. Euclidean distance

Euclidean distance adalah perhitungan jarak dari dua buah titik dalam euclidean space untuk mempelajari hubungan antara sudut dan jarak. Euclidean ini berkaitan dengan theorema pythagoras. Dalam matematika, jarak euclidean pythagoras adalah perhitungan yang digunakan untuk mengukur jarak antara dua titik dalam dimensi yang sama[13]. Nilai jarak antara data training dan data masukan baru (testing) dapat dihitung dengan menggunakan hasil perhitungan jarak euclidean[14]. Dalam berbagai aplikasi, teknik ini sering digunakan sebagai fungsi heuristik[15]. Perhitungan euclidean distance pada dua dimensi digunakan untuk dapat mengukur jarak pada koordinat lattitude dan longitude. Euclidean distance digunakan untuk mengukur tingkat kemiripan jarak antara data dengan rumus euclidean[16]. Berikut adalah rumus perhitungan Euclidean distance, yaitu:

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

Keterangan:

i : index dari atribut

n : jumlah data

x_i : atribut dari data ke-*i*, (*i* = 1, 2, 3, ..., ..., *n*)

y_i : atribut dari pusat Cluster ke-*i*, (*i* = 1, 2, 3, ..., ..., *n*)

Persamaan euclidean distance dapat digunakan pada penelitian perbandingan antara dua titik. Semakin kecil jarak euclidean semakin mirip kedua sifat tersebut[17]. Pada penelitian ini perhitungan euclidean distance digunakan sebagai salah satu metode perhitungan jarak yang digunakan.

b. Manhattan distance

Manhattan distance atau adapula yang menyebutnya City block distance biasa digunakan dalam melakukan perhitungan jarak antara dua titik. Manhattan distance mencerminkan jarak antar titik di jalan perkotaan dalam 1 blok[18]. Manhattan distance adalah jarak yang dihitung dari selisih absolute (mutlak) antara dua objek data.[19] Perhitungan manhattan menerapkan konsep dari perhitungan selisih asli antara nilai kedua titik. Perhitungan jarak manhattan adalah penerapan gagasan pencarian selisih murni antar data[20]. Rumus matematik dari manhattan distance, yaitu:

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

Keterangan:

d: jarak antara *x* dan *y*

x_i: atribut dari data ke-*i*, (1,2,3,.....,n)

y_i: atribut pusat Cluster ke-*i*, (1,2,3,.....,n)

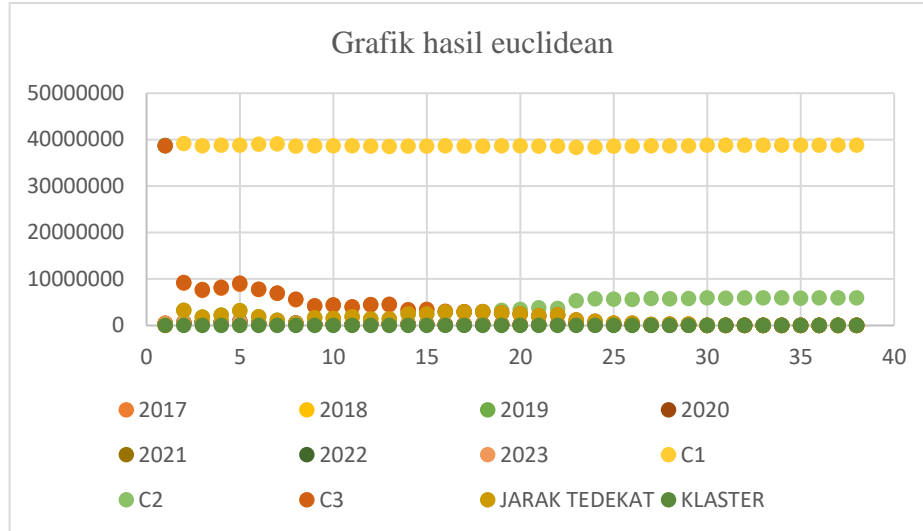
Perhitungan manhattan distance merupakan penjumlahan hasil absolut dari pengurangan antar dua titik. Perhitungan jarak manhattan digunakan untuk mengukur jarak antara dua titik dengan cara menghitung total jarak horizontal dan vertikal.

3. HASIL DAN PEMBAHASAN

Hasil analisis ini menunjukkan bahwa transformasi dan pra pemrosesan data sangat penting untuk pengelompokan. Sebelum perhitungan jarak euclidean dan manhattan memberikan gambaran yang berbeda tentang kemiripan wilayah, dan nilai skor silhouette yang jelas menunjukkan kualitas pengelompokan. Untuk meningkatkan ketahanan pangan dan produksi beras di Jawa Timur, pemahaman yang lebih baik tentang pola produksi beras dapat memungkinkan pengembangan kebijakan pertanian. Hasil perhitungan euclidean distance ada pada gambar 1, dan hasil dari manhattan ada pada Gambar 2.

3.1 Hasil Grafik Perhitungan Euclidean Distance

Perhitungan jarak euclidean distance menggunakan menghasilkan visualisasi grafik yang dapat dilihat pada Gambar 2.



Gambar 2. Hasil Perhitungan Metode Euclidean Distance

3.2 Hasil Pusat Cluster Akhir Metode Euclidean Distance

Perhitungan jarak pada algoritma k-means menghasilkan pusat cluster akhir, dimana proses perulangan iterasi berhenti ketika tidak ada data pada suatu cluster yang berpindah. Pusat cluster akhir metode euclidean distance dapat dilihat pada Tabel 5.

Tabel 3. Hasil pusat cluster akhir euclidean distance

Kota/kabupaten	C1	C2	C3
Ponorogo	783119.59	246104.9032	835143.993
Tulungagung	2165921.00	2522311.5	2991028.066
Blitar	2129779.70	2486514.062	2957129.836
Kediri	1913711.94	2238447.672	2705672.375
Malang	3342199.34	3780618.316	4254354.141
Lumajang	3486371.53	3939989.293	4418772.231
Jember	8175582.64	8721599.125	9217297.153
Banyuwangi	6794735.56	7312119.634	7795500.256
Bondowoso	3137950.58	3562482.101	4032615.06
Probolinggo	1735583.65	2019200.463	2480937.725
Pasuruan	6036756.37	6498207.285	6949723.963
Sidoarjo	1338947.95	1605097.692	2097321.799
Mojokerto	2460174.35	2901468.001	3401518.614
Jombang	3537898.27	4021233.953	4515623.733
Madiun	52177.53	228818.1563	1090060.161
Magetan	2515870.84	2942568.257	3432003.046
Ngawi	6589342.68	7181718.497	7716431.784
Bojonegoro	7087470.21	7658208.928	8175649.752
Tuban	4615595.93	5161418.598	5678984.501
Lamongan	7953502.17	8547615.741	9077050.914
Gresik	3534723.35	4027154.773	4529145.608
Bangkalan	2219393.60	2571327.791	3033857.795
Sumenep	1490908.12	1779947.173	2265151.23
Situbondo	1071309.62	103147.1393	455713.2188
Nganjuk	243875.39	135389.046	1040925.413
Sampang	1088167.87	157240.955	442267.6325
Pacitan	1328965.64	442950.8668	221626.2252
Trenggalek	1255401.57	360109.1056	289639.3589
Pamekasan	1306699.10	421762.9265	243504.8256
Kota Kediri	1555894.97	677372.0754	12999.95408
Kota Blitar	1536797.59	657117.8774	58219.49317
Kota Malang	1550759.29	671458.7005	17508.5022
Kota Probolinggo	1561488.89	682242.6636	7645.355608
Kota Pasuruan	1552426.73	673490.8165	16565.55191
Kota Mojokerto	1545335.44	665728.0807	46938.13166

Kota/kabupaten	C1	C2	C3
Kota Madiun	1552502.22	673150.0581	16544.481
Kota Surabaya	1555176.65	676428.2347	12945.78995
Kota Batu	1555809.95	676388.8142	23376.6384

3.3 Hasil Jumlah Data di Setiap Cluster

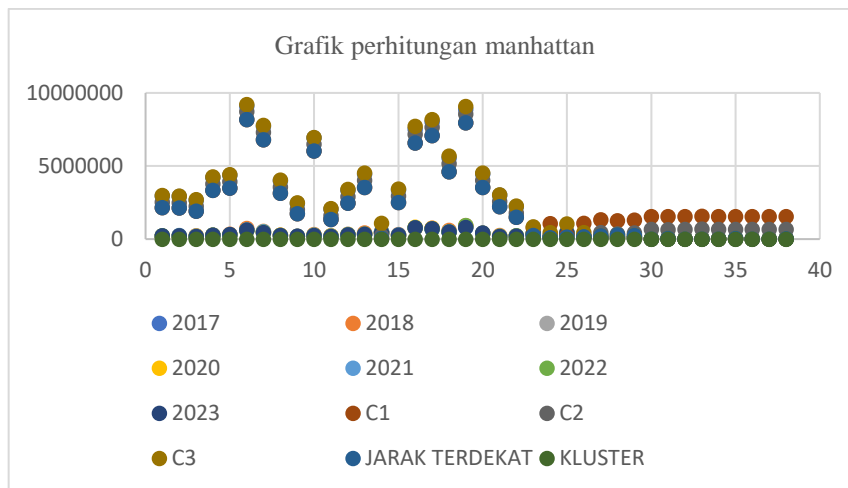
Hasil clustering menggunakan algoritma k-means dengan metode perhitungan jarak euclidean distance menghasilkan 3 cluster dan data yang terkelompok pada masing-masing cluster, hasil dapat dilihat pada Tabel 4.

Tabel 4. Hasil jumlah data setiap Centroid

Cluster	Jumlah data
C1	1
C2	16
C3	22

3.4 Hasil Grafik Perhitungan Metode Manhattan Distance

Perhitungan jarak manhattan distance menggunakan menghasilkan visualisasi grafik yang dapat dilihat pada Gambar 3.



Gambar 3. Hasil grafik manhattan distanc

3.5 Hasil Pusat Cluster Akhir Metode Manhattan Distance

Untuk melakukan perhitungan k-means manhattan distance, langkah pertama adalah menentukan nilai dataset yang akan digunakan. Data yang digunakan untuk produksi padi di Jawa Timur berasal dari tahun 2017–2023. Tahap kedua adalah menentukan nilai klaster atau k yang fungsinya digunakan untuk menghitung persamaan jarak pada metode k-means yaitu manhattan distance. Pada penelitian ini penulis menggunakan nilai k=3. Hasil clustering menggunakan algoritma k-means dengan metode perhitungan jarak manhattan distance menghasilkan 3 cluster dan data yang terkelompok pada masing-masing cluster, hasil dapat dilihat pada Tabel 5

Tabel 5. Hasil pusat cluster manhattan distance

Kabupaten/Regency	C1	C2	C3
Pacitan	-2478708.12	-843120	625624.9
Ponorogo	-1067747.78	590381.9	2661287
Trenggalek	-2392590.89	-757003	806854.2
Tulungagung	-1861121.98	-225534	1593939
Blitar	-1839196.27	-203609	1608787
Kediri	-1950641.93	-315054	1447509
Malang	-1635587.63	0	2070905
Lumajang	-1508098.12	159169.5	2230075
Jember	0	2615118	4686023
Banyuwangi	-801136.44	1536881	3607787
Bondowoso	-1742102.37	-106515	1920815
Situbondo	-2228431.24	-592844	1223720
Probolinggo	-1995962.6	-360375	1357950
Pasuruan	-1936801.22	241398.4	2312304

Kabupaten/Regency	C1	C2	C3
Sidoarjo	-1809656.94	-174069	1463258
Mojokerto	-1380021.63	255566	2151065
Jombang	-1198771.19	483870.4	2554776
Nganjuk	-919608.54	880965.1	2951871
Madiun	-795904.12	1024024	3094929
Magetan	-1526327.3	109260.3	2013042
Ngawi	1050991.1	3347921	5418826
Bojonegoro	580072.06	2977748	5048653
Tuban	-323714.06	1581544	3652449
Lamongan	1363207.6	3930511	6001417
Gresik	-997471.28	685886.4	2756792
Bangkalan	-1972680.52	-337093	1491821
Sampang	-2001444.81	-365857	1241136
Pamekasan	-2386382.56	-750795	697770.5
Sumenep	-1812490.04	-176902	1494989
Kota Kediri	-2847461.26	-1211874	33501.79
Kota Blitar	-2872680.07	-1237092	3594.98
Kota Malang	-2843731.11	-1208143	47633.94
Kota Probolinggo	-2864639.54	-1229052	19745.51
Kota Pasuruan	-2856331.35	-1220744	39175.7
Kota Mojokerto	-2879569.08	-1243981	-5554.03
Kota Madiun	-2844324.26	-1208737	45714.79
Kota Surabaya	-2848434.07	-1212846	35000.98
Kota Batu	-2869371.05	-1233783	0

Dari Tabel 5 hasil pusat kluster akhir metode manhattan distance menunjukkan bahwa pada masing-masing kabupaten/kota di Jawa Timur menempati cluster pada kolom cluster manhattan.

3.6 Hasil Jumlah Data di Setiap Cluster

Hasil clustering menggunakan algoritma k-means dengan metode perhitungan jarak manhattan distance menghasilkan 3 cluster dan data yang terkelompok pada masing-masing cluster, hasil dapat dilihat pada Tabel 6

Tabel 6. Hasil jumlah data setiap cluster

Cluster	Jumlah data
C1	38

3.7 Silhouette Coefficient

Metode ini menggabungkan teknik cohesion dan separation untuk evaluasi cluster [21]. Silhouette dari masing-masing cluster atau bahkan keseluruhan cluster dapat dihitung dari hasil kerja suatu algoritma clustering untuk memberikan informasi tentang kualitas hasil clustering pada proses clustering [22]. Berikut adalah hasil perbandingan metode k-means euclidean distance dan manhattan distance menggunakan metode silhouette coefficient yaitu:

Silhouette Score: 0.61

Gambar 4. Euclidean Distance

Silhouette Score: 0.58

Gambar 5. Manhattan Distance

Dari data Gambar 4 dan Gambar 5 menunjukkan bahwa metode k-means dengan euclidean distance memperoleh nilai silhouette coefficient 0,61 sedangkan manhattan distance 0,58. Dapat disimpulkan bahwa metode k-means dengan model euclidean distance itu baik untuk distributif nilai secara merata, dan k-means dengan model manhattan distance baik ketika mengelompokkan dengan distributif yang berbeda.

4. KESIMPULAN

Hasil penelitian menunjukkan bahwa metode yang digunakan dapat memengaruhi hasil analisis dan Keputusan yang dibuat. Kedua metode ini memungkinkan Dinas Pertanian Jawa Timur untuk mendapatkan pemahaman yang lebih mendalam tentang kemungkinan produksi padi di setiap kabupaten atau kota. K-Means dengan model



euclidean distance dan manhattan distance dapat digunakan dalam penentuan wilayah dengan potensial padi rendah, sedang, dan tinggi. Hasil dari perbandingan euclidean distance dan manhattan distance menggunakan silhouette coefficient menunjukkan bahwa euclidean distance lebih unggul 0,61 dibandingkan manhattan distance 0,58. Akibatnya, penelitian ini menawarkan solusi praktis untuk meningkatkan ketahanan pangan dan produksi padi di Jawa Timur selain manfaat akademis.

REFERENCES

- [1] H. Husdi and Y. Lasena, "Real Time Analisis Berbasis Internet Of Things Untuk Prediksi Iklim Lahan Pertanian," *J. Media Inform. Budidarma*, vol. 4, no. 3, pp. 834–840, 2020, [Online]. Available: <http://ejurnal.stmik-budidarma.ac.id/index.php/mib/article/view/2165>
- [2] D. H. Pujiningtyas and Y. Nangameka, "Mapping Potency of Excellent Comodity of Food Crops in Situbondo Regency," *J. Ilm. Agribios*, vol. 16, no. 1, pp. 43–54, 2018.
- [3] F. Marisa et al., "Digitasi Produktivitas Panen Padi Berbasis K-Means Clustering," *SMARTICS J.*, vol. 7, no. 1, pp. 21–26, 2021.
- [4] A. L. Zahra, S. Tiara, R. Ada, and A. F. Ardini, "Implementasi Clustering Algoritma K-Means Pada Produksi Beras di Provinsi Jawa Timur Tahun 2022," *J. Comput. Inf. Syst. Ampera*, vol. 5, no. 3, pp. 2775–2496, 2024
- [5] Y. A. Singgalen, "Analisis Sentimen Pengunjung Pulau Komodo dan Pulau Rinca di Website Tripadvisor Berbasis CRISP-DM," *J. Inf. Syst. Res.*, vol. 4, no. 2, pp. 614–625, 2023, doi: 10.47065/josh.v4i2.2999.
- [6] R. S. Auria, D. A. Puspitaningrum, and B. Widayanto, "Potensi Subsektor Tanaman Pangan Komoditas Padi di Kabupaten Klaten Provinsi Jawa Tengah," *AGRIFITIA J. Agribus. Plant.*, vol. 2, no. 1, pp. 14–21, 2022, doi: 10.55180/aft.v2i1.199.
- [7] R. M. Koretsky, "Implementasi Metode Data Mining K-Means Clustering Terhadap Data Pembayaran Transaksi Menggunakan Bahasa Pemrograman Python Pada Cv Digital Dimensi," *Raspberry Pi OS Syst. Adm. with Syst. Python*, vol. 8, no. 1, pp. 175–305, 2023, doi: 10.1201/b23421-3.
- [8] L. Maulida, "Penerapan Datamining Dalam Mengelompokkan Kunjungan Wisatawan Ke Objek Wisata Unggulan Di Prov. Dki Jakarta Dengan K-Means," *JISKA (Jurnal Inform. Sunan Kalijaga)*, vol. 2, no. 3, p. 167, 2018, doi: 10.14421/jiska.2018.23-06.
- [9] B. P. Statistik, "Badan Pusat Statistik Provinsi Jawa Timur." <https://jatim.bps.go.id/id> (accessed Jun. 14, 2024).
- [10] I. Ismawati, "Kajian Data Mining Profil Siswa Baru Dalam Penentuan Strategi Promosi Dengan Metode Two Step Clustering," *J. Ilm. Teknol. Infomasi Terap.*, vol. 5, no. 3, pp. 47–59, 2019, doi: 10.33197/jitter.vol5.iss3.2019.302.
- [11] Haris Kurniawan, Sarjon Defit, and Sumijan, "Data Mining Menggunakan Metode K-Means Clustering Untuk Menentukan Besaran Uang Kuliah Tunggal," *J. Appl. Comput. Sci. Technol.*, vol. 1, no. 2, pp. 80–89, 2020, doi: 10.52158/jacost.v1i2.102.
- [12] D. I. Kawasan, K. Makmur, D. Kadek, and P. Kumala, "Kajian Perkembangan Pola Ruang Fisik Spasial Permukiman Kumuh," pp. 2045–2058, 2024.
- [13] Y. Risha, "Sejarah Geometri : Euclid Hingga Konsep Geometri Modern," vol. 7, pp. 17009–17020, 2024.
- [14] A. Yudhana, S. Sunardi, and A. J. S. Hartanta, "Algoritma K-Nn Dengan Euclidean Distance Untuk Prediksi Hasil Penggergajian Kayu Sengon," *Transmisi*, vol. 22, no. 4, pp. 123–129, 2020, doi: 10.14710/transmisi.22.4.123-129.
- [15] E. Yulianto and A. Setiawan, "Optimasi Rute Sales Coverage Menggunakan Algoritma Cheapest Insertion Heuristic Dan Layanan Google Maps Api," *Intern. (Information Syst. Journal)*, vol. 1, no. 1, pp. 39–54, 2018, doi: 10.32627/internal.v1i1.326.
- [16] C. A. Awaliyah, A. Prasetyadi, and A. Junaidi, "Sistem Rekomendasi Desain Website Berdasarkan Tingkat Kemiripan Menggunakan Euclidean Distance," *J. Dinda Data Sci. Inf. Technol. Data Anal.*, vol. 2, no. 2, pp. 75–81, 2022, doi: 10.20895/dinda.v2i2.543.
- [17] S. Martha and W. Andani, "Komparasi Algoritma K-Nearest Neighbor Dengan Euclidean Distance Dan Manhattan Distance Untuk Klasifikasi Stunting Balita (Studi Kasus : Puskesmas Kelurahan Parit Mayor)," vol. 13, no. 2, pp. 285–292, 2024.
- [18] R. Alya Shafira, Yahfizham, and A. Muliani Harahap, "Menentukan Jarak Terpendek Dalam Pengiriman Barang Dengan Perbandingan Euclidean Distance Dan Manhattan Distance," *J. Sci. Soc. Res.*, vol. VI, no. 3, pp. 678–685, 2023.
- [19] A. C. Fauzan, S. S. Wibowo, and M. Ahmad, "Performa Klasifikasi Berbasis Jarak untuk Deteksi Covid-19 Varian Delta dan Omicron Menggunakan Citra CT-Scan Paru-Paru," *J. Syst. Comput. Eng.*, vol. 4, no. 1, pp. 62–81, 2023, doi: 10.47650/jsce.v4i1.714.
- [20] Y. Miftahuddin, S. Umaroh, and F. R. Karim, "Perbandingan Metode Perhitungan Jarak Euclidean, Haversine, Dan Manhattan Dalam Penentuan Posisi Karyawan," *J. Tekno Insentif*, vol. 14, no. 2, pp. 69–77, 2020, doi: 10.36787/jti.v14i2.270.
- [21] K. Pratama Simanjuntak and U. Khaira, "Hotspot Clustering in Jambi Province Using Agglomerative Hierarchical Clustering Algorithm," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 1, no. 1, pp. 7–16, 2021.
- [22] S. Paembonan and H. Abduh, "Penerapan Metode Silhouette Coefficient untuk Evaluasi Clustering Obat," *PENA Tek. J. Ilm. Ilmu-Ilmu Tek.*, vol. 6, no. 2, p. 48, 2021, doi: 10.51557/pt_jiit.v6i2.659.