

**Algoritma *Partitioning Around Medoids* (PAM) dalam Mengelompokkan Kecamatan di Kabupaten Jember Berdasarkan Produktivitas Tanaman Padi**  
***Partitioning Around Medoids (PAM) Algorithm in Classifying Districts in Jember Regency Based on Rice Productivity***

**Adinda Putri Rengganis<sup>1</sup>, Deni Arifianto<sup>2</sup>, Hardian Oktavianto<sup>3</sup>**

<sup>1</sup>Mahasiswa Fakultas Teknik, Universitas Muhammadiyah Jember  
e-mail: [rengganisadinda@gmail.com](mailto:rengganisadinda@gmail.com)<sup>1</sup>

<sup>2,3</sup>Dosen Fakultas Teknik, Universitas Muhammadiyah Jember

**Abstrak**

Indonesia adalah negara agraris yang sebagian besar penduduknya tinggal didaerah pedesaan dan bermata pencaharian sebagai petani. Jika mengingat peranan pertanian dalam sistem perekonomian negara, maka perlu dilakukan upaya untuk meningkatkan produksi pertanian agar menjadi penyanggah kebutuhan pokok. Salah satu sumber makanan pokok yang sebagian besar dikonsumsi masyarakat Indonesia adalah beras. Apabila persediaan beras di Kabupaten Jember tidak mencukupi dikarenakan sedang terjadi penurunan produksi, maka dapat menyebabkan kelangkaan beras, kelemahan pangan, kenaikan harga beras, dll. Dalam hal tersebut, diperlukan upaya untuk mengetahui tingkat produksi suatu daerah atau kecamatan guna mengatasi penurunan produksi. Kecamatan yang akan dijadikan prioritas dapat diketahui dengan melakukan pengelompokan, satu diantaranya adalah memakai algoritma *Partitioning Around Medoids* (PAM) dengan metode *Silhouette Coefficient* (SC) untuk menghitung *cluster* optimal dalam rangka menentukan *cluster* yang paling baik. Data yang digunakan adalah data pada 31 kecamatan di Kabupaten Jember periode tahun 2019-2021. Berdasarkan pengujian dari 2 sampai 10 *cluster*, dihasilkan 2 *cluster* sebagai *cluster* optimal yang mana nilai *Silhouette Coefficient* (SC) yaitu sebesar 0.4678087. Ada 15 anggota kecamatan di *cluster* 1, dan ada 16 anggota kecamatan di *cluster* 2.

**Kata kunci:** produksi, tanaman padi, *partitioning around medoids*, *clustering*, *sillhouette coefficient*.

**Abstract**

Indonesia is an agricultural country where most of the population lives in rural areas and earn a living as farmers. If you remember the role of agriculture in the country's economic system, it is necessary to make efforts to increase agricultural production so that it becomes a buffer for basic needs. One of the staple food sources consumed by the majority of Indonesian people is rice. If the supply of rice in Jember Regency is insufficient due to a decline in production, it can cause rice scarcity, food weakness, increase in rice prices, etc. In this case, efforts are needed to determine the production level of a region or sub-district in order to overcome production declines. Districts that will be prioritized can be identified by grouping, one of which is using the *Partitioning Around Medoids* (PAM) algorithm with the *Silhouette Coefficient* (SC) method to calculate optimal clusters in order to determine the best clusters. The data used is data on 31 sub-districts in Jember Regency for the 2019-2021 period. Based on testing from 2 to 10 clusters, 2 clusters are produced as optimal clusters where the *Sillhouette Coefficient* (SC) value is 0.4678087. There are 15 sub-district members in cluster 1, and there are 16 sub-district members in cluster 2.

**Keywords:** production, *partitioning around medoids*, *clustering*, *sillhouette coefficient*.

## 1. PENDAHULUAN

Indonesia adalah negara agraris yang sebagian besar penduduknya tinggal didaerah pedesaan dan bermata pencaharian sebagai petani. Diketahui 75% dari penduduk Indonesia tinggal didaerah pedesaan dan mayoritas sangat menggantungkan hidupnya pada sektor pertanian. Sehingga dapat dikatakan bahwa sektor pertanian memegang peran penting dalam ekonomi Indonesia (Claudia Olvi Rondonuwu, 2017). Mengingat pentingnya peranan pertanian dalam sistem perekonomian negara, maka perlu dilakukan upaya untuk meningkatkan produksi pertanian agar dapat menjadi penyangga kebutuhan utama sebagai kebutuhan pokok (Wildan Mukhtari, 2018).

Salah satu sumber makanan pokok yang sebagian besar dikonsumsi masyarakat Indonesia adalah beras. Beras sebagai sumber pangan pokok yang vital di Indonesia tidak dapat digantikan oleh bahan pangan lainnya (Kementan, 2016). Salah satu daerah penyangga pangan nasional di Indonesia adalah Kabupaten Jember. Sebagai salah satu lumbung padi di Provinsi Jawa Timur dengan produksi sebesar 978.373 ton selama tahun 2014 dengan produktivitas sebesar 59,55 Kw/Ha. Perkembangan produksi padi dari tahun 2012-2014 mengalami peningkatan diiringi dengan peningkatan produktivitas. Produksi padi meningkat dari 930.027 ton pada tahun 2013 menjadi 978.373 ton pada tahun 2014. Sedangkan produksi padi pada tahun 2018 sebesar 984.201 ton dengan produktivitas sebesar 59,88 Kw/Ha. Namun, perkembangan produksi padi dari tahun 2016-2018 cenderung mengalami penurunan. Produksi padi menurun dari 986.653 ton pada tahun 2016 menjadi 984.201 ton pada tahun 2018. Oleh karena itu, produktivitas tanaman pangan khususnya padi perlu terus ditingkatkan. Apabila persediaan beras di Kabupaten Jember tidak mencukupi dikarenakan sedang terjadi penurunan produksi, maka dapat menyebabkan kelangkaan beras, kelemahan pangan, kenaikan harga beras dan akibat buruk lainnya (Badan Pusat Statistik Kabupaten Jember, 2013-2018).

Sehubungan dengan penjelasan diatas, maka peningkatan efisiensi tanaman pangan khususnya padi menjadi penting, mengingat kebutuhan akan beras terus meningkat sejalan dengan perkembangan penduduk, peningkatan upah individu, dll (Badan Pusat Statistik Jawa Timur, 2016). Dengan demikian, untuk meningkatkan produksi beras maka Dinas Pertanian Kabupaten Jember berupaya untuk terus mengoptimalkan hasil pertanian padi agar semakin meningkat. Untuk meningkatkan produksi, hasil panen harus dikelompokkan berdasarkan luas panen, produktivitas, dan produksi padi di setiap kecamatan. Proses pengelompokan dapat menggunakan metode *clustering*. Tujuannya adalah untuk mengidentifikasi daerah dengan hasil produksi padi yang belum optimal. Daerah-daerah ini memerlukan perhatian dan penanganan yang efektif karena berkaitan dengan pengambilan kebijakan penyaluran bantuan yang lebih tepat sasaran yang dilakukan oleh Dinas Pemerintahan.

Mengacu pada Penelitian sebelumnya yang dilakukan oleh (Dharshinni & Fandi, 2022) berjudul “Penerapan algoritma *Partitioning Around Medoids* (PAM) untuk mengelompokkan ketahanan pangan”. Penelitian tersebut memanfaatkan informasi hasil panen tanaman pangan di Provinsi Sumatera Utara pada tahun 2018-2020 dengan menggunakan algoritma *Partitioning Around Medoids* (PAM) dengan metode *Davies Bouldin Index* (DBI) dan metode *Silhouette Coefficient* (SC), penelitian ini menghasilkan *Davies Bouldin Index* (DBI) senilai 0,062 dan *Silhouette Coefficient* (SC) senilai 0,8980 dengan jumlah *cluster* sebanyak 3 *cluster* sehingga *Cluster\_0* didominasi tanaman pangan jagung yang mengalami peningkatan produksi sebesar 5% dan kacang tanah sebesar 5%, *Cluster\_1* didominasi dengan penurunan jumlah hasil produksi kedelai sebesar 38%, dan *Cluster\_2* didominasi oleh penurunan hasil kacang hijau sebesar 33%.

## 2. STUDI PUSTAKA

### A. Luas Panen

Luas panen padi adalah salah satu faktor yang dapat mempengaruhi pertumbuhan produksi padi di Indonesia. Luas panen padi menurut jurnal yang diakui adalah luas lahan yang ditanami padi dan dipanen pada suatu periode tertentu. Berdasarkan sumber yang ditemukan, dapat disimpulkan bahwa luas panen padi dihitung berdasarkan pengamatan yang objektif menggunakan metodologi Kerangka Sampel Area (KSA) yang dikembangkan oleh BPPT dan BPS. Metodologi KSA telah mendapat pengakuan dari LIPI. Produksi padi diperoleh dari hasil perkalian antara luas panen (bersih) dengan produktivitas. Luas panen tanaman padi dilahan sawah harus dikoreksi dengan besaran konversi galengan (Badan Pusat Statistik, 2021).

### B. Produktivitas Padi

Produktivitas padi adalah kemampuan suatu tanah yang menghasilkan suatu tanaman padi dengan sistem pengolahan tertentu. Produksi padi dihitung dengan menghitung jumlah produksi padi dalam bentuk Gabah Kering Giling (GKG) per satuan luas lahan, atau kuintal per hektar. Produktivitas padi adalah nilai yang menunjukkan hasil produksi rata-rata per satuan luas per komoditas tanaman padi selama periode laporan satu tahun. Meningkatkan kualitas padi yang dihasilkan dapat dicapai dengan mengolah lahan yang ada melalui peningkatan produktivitas lahan. Penerapan berbagai sistem tanam dan jarak tanam dapat meningkatkan produktivitas padi. Sebagian petani menghadapi masalah dengan penggunaan pupuk ketika mereka ingin meningkatkan produktivitas pertanian mereka. (Sugiyono, 2021).

### C. Produksi

Produksi padi adalah jumlah hasil panen padi yang dihasilkan pada suatu periode tertentu. Produksi komoditi pertanian, lebih diartikan sebagai hubungan praktis antara beberapa faktor produksi yang sebenarnya, seperti tanah, benih/bibit, kompos, dll. Semua

faktor saling memenuhi antara satu sama lain dalam memperoleh hasil produksi yang lebih maksimal nantinya (Andani, 2018). Peningkatan produksi padi dapat dipengaruhi oleh faktor-faktor sebagai berikut : luas lahan garapan, jumlah tenaga kerja efektif, dosis pupuk, dosis pestisida, pengalaman petani dalam berusahatani, jarak lahan garapan dengan tempat tinggal petani, dan sistem pengairan atau irigasi (Diantoro K, 2018).

### D. Clustering

*Clustering* adalah sebuah proses pemisahan atau pemecahan informasi yang digunakan sebagai pengolah data yang tidak diketahui label maupun kelasnya. *Clustering* memiliki nama lain yang disebut *cluster analysis*. Pola kerja *clustering* dapat dilakukan dengan mengelompokkan sekumpulan data ke dalam beberapa kelas. Objek-objek yang ada dalam kelas tersebut memiliki kesamaan yang tinggi apabila dibandingkan dengan objek lain yang ada pada kelas tersebut, tetapi mempunyai kesamaan yang rendah apabila dibandingkan dengan objek yang ada pada *cluster* lain (Handoko, 2016).

### E. Normalisasi

Normalisasi data adalah suatu proses melakukan penskalaan nilai atribut dari sebuah data sehingga dapat diposisikan pada range tertentu (Nasution et al., 2019). Ada beberapa metode yang digunakan untuk normalisasi seperti normalisasi *min-max*, normalisasi *z-score*, *decimal scaling*, *Sigmoidal* dan lain-lain (Nasution et al., 2019). Penelitian ini dilakukan dengan tahap normalisasi *min-max*. Normalisasi *min-max* adalah salah satu metode normalisasi yang menghasilkan perbandingan nilai yang seimbang antara data sebelum diproses dan setelah diproses (Nasution et al., 2019).

Metode normalisasi *min-max* dapat dilakukan dengan rumus sebagai berikut (Permana & Nur Salisah, 2022):

$$x' = \frac{x_i - \min(x)}{\max(x) - \min(x)}$$

Keterangan:

$x_i$  = nilai tertentu yang akan dinormalisasi  
 $x'$  = nilai hasil normalisasi  
 $\min(x)$  = nilai minimal dari sebuah atribut  
 $\max(x)$  = nilai maksimal dari sebuah atribut

#### F. Partitioning Around Medoids (PAM)

*Partitioning Around Medoids* (PAM) merupakan algoritma *clustering* yang hampir serupa dengan *K-Means*. Perbedaan antara kedua perhitungan ini ialah perhitungan *Partitioning Around Medoids* (PAM) memakai objek sebagai perwakilan (*medoid*) sebagai *centroid* untuk setiap *cluster*, sedangkan *K-Means* memakai rata-rata sebagai *centroid* (Kaur, et al., 2014). Algoritma *Partitioning Around Medoids* (PAM) mempunyai keunggulan untuk mengatasi kekurangan pada algoritma *K-Means* yang sensitif terhadap permasalahan data dimana terdapat data yang memiliki perbedaan sangat besar dengan data-data lain yang ada dikumpulan data. Hal ini mengakibatkan terjadi persebaran data yang tidak seimbang. Keunggulan lainnya adalah terjadinya efek samping dari sistem pengelompokan tidak bergantung pada urutan masuk dataset (Furqon, et al., 2015).

Langkah-langkah algoritma *Partitioning Around Medoids* (PAM) (Abdurrahman et al., 2021):

1. Menginisialisasi *centroid* sebanyak  $k$  (jumlah *cluster*).
2. Mengalokasikan setiap data (objek) ke *cluster* terdekat memakai persamaan jarak *Euclidian Distance* dengan persamaan:
 
$$d(x, y) = \|x - y\|$$

$$= \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad ; 1, 2, 3, \dots, n \quad (1)$$
3. Memilih secara acak objek pada masing-masing *cluster* sebagai kandidat *medoid* baru.
4. Menghitung jarak setiap objek yang ada pada masing-masing *cluster* dengan kandidat *medoid* baru.
5. Menghitung total simpangan (S) dengan menghitung nilai total *distance* baru – total *distance* lama. Jika  $S < 0$ , maka tukar objek

dengan data *cluster* untuk membentuk sekumpulan  $k$  objek baru sebagai *medoid*.

6. Mengulangi langkah 3 sampai 5 hingga tidak terjadi perubahan *medoid*, sehingga didapatkan *cluster* beserta anggota *cluster* masing-masing.

#### G. Silhouette Coefficient (SC)

*Silhouette Coefficient* (SC) ialah metode evaluasi untuk menentukan mutu dari *cluster*, seberapa bagus objek ditempatkan disuatu *cluster* (Mario, Herry, & Nasution, 2016).

Berikut langkah-langkah dalam menghitung *Silhouette Coefficient* (SC) (Handoyo, Mangkudjaja, & Nasution, 2014):

1. Hitung rata-rata jarak dari suatu objek dengan semua objek lain yang berada dalam satu *cluster*.

$$a(i) = \frac{1}{|A|-1} \sum_{j \in A, j \neq i} d(i, j) \quad (2)$$

Keterangan:

$a(i)$  : rata-rata jarak objek  $i$  ke semua objek lain didalam 1 *cluster*.

$d(i, j)$  : jarak objek  $i$  dan objek  $j$ , dimana  $i \neq j$ .

$|A|$  : total objek dalam *cluster* A

2. Menghitung rata-rata jarak objek dengan semua objek yang berada pada *cluster* lain, dan diambil nilai terkecil.

$$d(i, C) = \frac{1}{|C|} \sum_{j \in C} d(i, j) \quad (3)$$

Kemudian mencari nilai terkecilnya dengan persamaan berikut:

$$b(i) = \min_{C \neq A} d(i, C) \quad (4)$$

Keterangan:

$d(i, C)$  : jarak antara objek  $i$  dengan *cluster* C.

$d(i, j)$  : jarak objek  $i$  dengan objek  $j$ , dimana  $j$  anggota *cluster* C

$b(i)$  : rata-rata objek dengan semua objek lain yang berada pada *cluster* C, dengan  $C \neq A$ .

3. Hitung nilai *silhouette coefficient* (SC) dengan persamaan berikut:

$$s(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))} \quad (5)$$

Nilai *Silhouette Coefficient* (SC) terdapat pada *range* nilai -1 sampai 1. Apabila hasil nilai *Silhouette Coefficient* (SC)

mendekati nilai 1, maka semakin baik pengelompokan data dalam satu *cluster*. Untuk sebaliknya jika *Silhouette Coefficient* (SC) mendekati nilai -1, maka semakin buruk pengelompokan data didalam satu *cluster* (Wira et al., 2019).

#### H. RStudio

RStudio adalah bagian dari IDE (*Integrated Development Environment*), yang tidak umum digunakan dalam bahasa pemrograman R. Beberapa opsi untuk menyunting bahasa R dirancang untuk membuatnya lebih mudah digunakan, tetapi RStudio adalah yang paling banyak digunakan oleh pengguna (Budiaji, 2019).

### 3. METODOLOGI PENELITIAN

#### A. Pengumpulan Data

Penelitian ini menggunakan data yang berasal dari publikasi buku di Badan Pusat Statistik (BPS) Kabupaten Jember. Data yang akan diolah pada penelitian ini menggunakan data Kecamatan di Kabupaten Jember pada tahun 2019-2021. Terdapat 31 Kecamatan di Kabupaten Jember. Dari data yang tersedia, peneliti menggunakan data tersebut untuk mengelompokkan daerah dengan produktivitas padi yang terjadi kenaikan dan penurunan pada Kecamatan yang ada di Kabupaten Jember.

**Tabel 1.** Data sampel Kecamatan di Kabupaten Jember

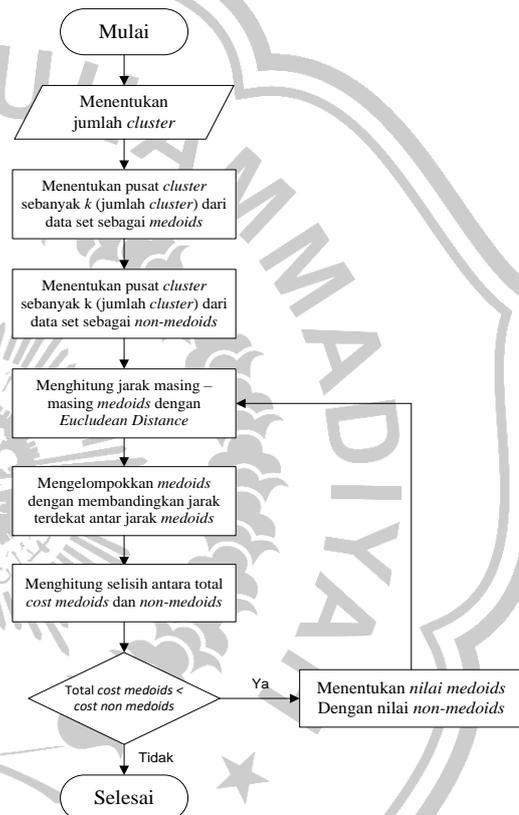
NO	KECAMATAN	LP (Ha)	PS (Kw/Ha)	PI (Ton)
1	Kencong	6065	6,90	41850
2	Gemukmas	7602	6,77	51444
3	Puger	6051	6,95	42055
4	Wuluhan	4995	6,85	34219
5	Ambulu	3700	6,86	25385
6	Tempurejo	3682	6,49	23897
...	...	...	...	...
31	Patrang	3404	6,06	20640

Sumber : Data Awal

Pada data tersebut mempunyai tiga atribut yakni Luas Panen (LP), Produktivitas (PS) dan Produksi Padi (PI).

#### B. Partitioning Around Medoids (PAM)

Diagram alur dari penerapan metode *clustering* menggunakan *Partitioning Around Medoids* (PAM) yang digunakan untuk menyelesaikan masalah pengelompokan produktivitas padi pada Kecamatan di Kabupaten Jember:



**Gambar 1.** Algoritma *Partitioning Around Medoids* (PAM)

#### C. Silhouette Coefficient (SC)

Pada proses pengelompokan Kecamatan menggunakan algoritma *Partitioning Around Medoids* (PAM) yang telah dilakukan pada rentang 2 *cluster* sampai 10 *cluster*. Selanjutnya akan dilakukan pengujian *cluster* optimal menggunakan *silhouette coefficient* (SC) untuk menentukan *cluster* optimal atau yang terbaik. Untuk menentukan *cluster* optimal menggunakan *silhouette coefficient* (SC)

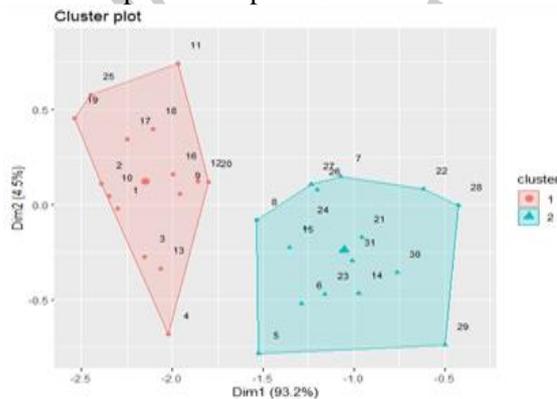
langkah pertama adalah menghitung rata-rata jarak dari suatu objek dengan semua objek lain yang berada dalam satu *cluster* menggunakan persamaan. Lalu Menghitung rata-rata jarak objek dengan semua objek yang berada pada *cluster* lain, dan diambil nilai terkecil menggunakan persamaan. Kemudian mencari nilai terkecilnya dengan persamaan. Proses selanjutnya menghitung nilai *sillhouette coefficient* (SC) menggunakan persamaan.

#### 4. HASIL DAN PEMBAHASAN

Pada tahap ini membahas tentang hasil yang telah di peroleh dari proses perhitungan yang telah dilakukan. Data yang telah didapat akan di *cluster* memakai algoritma *Partitioning Around Medoids* (PAM) yang kemudian diolah dengan memakai metode *Silhouette Coefficient* (SC). Data yang digunakan yakni data pada Kecamatan di Kabupaten Jember pada tahun 2019-2021 yang berjumlah 31 Kecamatan.

##### A. *Partitioning Around Medoids* (PAM) pada RStudio

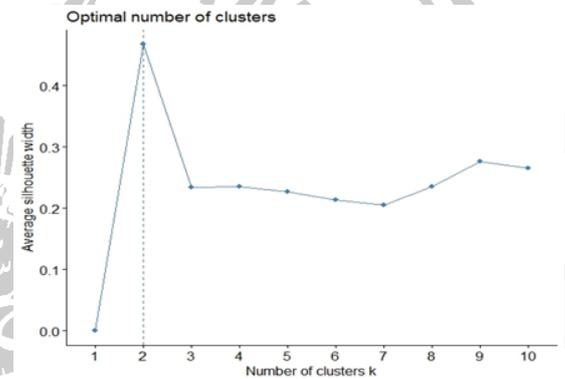
Data jumlah luas panen, produktivitas, dan produksi padi diolah menggunakan tools RStudio yang di *cluster* menggunakan algoritma *Partitioning Around Medoids* (PAM) dari 2 sampai dengan 10 *cluster*. Hasil dari perintah *Partitioning Around Medoids* (PAM) pada RStudio ditampilkan ke dalam plot dari setiap *cluster* yang terbentuk dan digambarkan dengan bentuk simbol serta warna yang berbeda-beda. Berikut contoh plot pada 2 *cluster* dapat dilihat pada Gambar 2.



Gambar 2. Plot di RStudio

Keterangan pada Gambar plot di RStudio: *Cluster* 1 ditandai dengan warna merah yang memiliki 15 anggota data (Kecamatan) sedangkan *Cluster* 2 ditandai dengan warna biru yang memiliki 16 anggota data (Kecamatan).

Dapat dilihat pada gambar grafik *silhouette coefficient* (SC) dibawah bahwa metode SC menggunakan pendekatan nilai rata-rata untuk mengetahui nilai k optimal dari *cluster* yang terbentuk. Semakin tinggi nilai rata-ratanya maka akan semakin baik. Grafik dibawah menunjukkan bahwa jumlah *cluster* optimal yang akan terbentuk adalah k=2 yang memiliki nilai rata-rata paling tinggi jika dibandingkan dengan jumlah k lainnya.



Gambar 2. Grafik *Cluster* Optimal pada RStudio

##### B. Penentuan *Cluster* Optimal

Menentukan *cluster* yang terbaik atau optimal dari proses perhitungan menggunakan algoritma *Partitioning Around Medoids* (PAM) sebelumnya, dilanjutkan dengan menerapkan metode pengujian *cluster* optimal yaitu *silhouette coefficient* (SC). Adapun hasil keluaran yang didapat yaitu berupa nilai seperti pada Tabel 2 berikut:

Tabel 2. Perhitungan *silhouette coefficient* (SC)

<i>Cluster</i>	Nilai <i>Silhouette Coefficient</i> (SC)
2	0.4678087
3	0.2343113

<i>Cluster</i>	<b>Nilai <i>Silhouette Coefficient</i> (SC)</b>
4	0.2358232
5	0.2268639
6	0.2141862
7	0.2056124
8	0.2348294
9	0.2766252
10	0.2654746

Sumber : Hasil Perhitungan *cluster* optimal

Berdasarkan hasil pengujian yang telah dilakukan, telah didapatkan *cluster* optimal untuk tingkat produksi di Kabupaten Jember ialah 2 *cluster*.

## 5. KESIMPULAN DAN SARAN

### A. Kesimpulan

Berdasarkan hasil yang didapat dari perhitungan pada penelitian yang telah dilakukan, maka implementasi atau penerapan *Partitoning Around Medoids* (PAM) pada data jumlah luas panen, produktivitas dan produksi padi di Kecamatan untuk mengelompokkan Kecamatan di Kabupaten Jember menghasilkan *cluster* optimal pada 2 *cluster* dengan nilai *Silhouette Coefficient* (SC) yaitu sebesar 0,4678087.

Diketahui bahwa nilai *koefisien partisi* 2 sampai 10 *cluster* memiliki nilai keanggotaan yang berbeda-beda, dan nilai yang lebih tinggi serta mendekati nilai 1 adalah *cluster* yang optimal jika dibandingkan dengan *cluster* lainnya. *Cluster* optimal dari penelitian ini terdapat pada 2 *cluster* yaitu *cluster* 1 dan *cluster* 2, untuk keanggotaan masing-masing *cluster* yaitu *cluster* 1 terdapat 15 anggota kecamatan sedangkan *cluster* 2 terdapat 16 anggota kecamatan.

### B. Saran

Penelitian ini dapat dikembangkan menggunakan metode *clustering* yang berbeda, untuk uji validasi *cluster* dapat menggunakan

teknik yang berbeda seperti metode *Dunn Index*, *Elbow*, *Davies Bouldin Index* (DBI), *Partition Coefficient Index* (PCI), dll. Dan dapat meningkatkan jumlah atribut dan variabel pada data yang digunakan.

## 6. REFERENSI

- A. Supriyatna, I. Carolina, W. Widiati, and C. Nuraeni, "Rice Productivity Analysis by Province Using K-Means Cluster Algorithm," in *IOP Conference Series: Materials Science and Engineering*, 2020, vol. 771, no. 1, doi: 10.1088/1757-899X/771/1/012025.
- Badan Pusat Statistik Jawa Timur. *Produksi Padi dan Palawija di Jawa Timur 2015/2016*. Jawa Timur: Badan Pusat Statistika Jawa Timur.
- Badan Pusat Statistik. 2020, 2021, 2022.. *Jember Publikasi Statistik Daerah Kabupaten Jember 2019*, Jember: Badan Pusat Statistika Kabupaten Jember.
- Dennis, A., Baskoro, D.A., Ambarwati L., Wicaksana I.W.S. 2012. *Belajar Data Mining dengan RapidMiner*. Jakarta: Gramdia Pustaka Utama.
- Karmila, Windarto, A.P., Fauzan, M. 2019. Analisis Metode K-Medoid Pada Kasus Obesitas Balita Menurut Provinsi Di Indonesia. *Prosiding Seminar Nasional Riset Information*.
- Kaur, N. U., & Singh, D. 2014. *K-Medoids Clustering Algorithm – A Review*. [pdf] *International Journal of Computer Application and Technology (IJCAT)*. ISSN. 2349-1841 Vol. 1, Issue 1. April 2014.
- Kementan RI, 2016. *Mengenai Beras sebagai sumber pangan*, Jakarta.
- Pramesti, D. F., Lahan, Tanzil Furqon, M., & Dewi, C. (2017). Implementasi Metode K-Medoids Clustering Untuk Pengelompokan Data. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 1(9), 723–732.

Retno, S. 2019. Peningkatan Akurasi Algoritma K-Means dengan *Clustering Purity* sebagai Titik Pusat *Cluster* Awal (*Centroid*).[pdf] repository.usu.ac.id.

Ridlo, M. R., Defiyanti, S., & Primajaya, A. (2017). Implementasi Algoritma K-Means Untuk Pemetaan Produktivitas Panen Padi Di Kabupaten Karawang.

Triyanto, W.A. 2015. Algoritma K-Medoids untuk Penentuan Strategi Pemasaran Produk. Kudus: Jurnal Simetris/Vol. 6 No. 1.

