

OPTIMASI K-MEANS DENGAN PARTICLE SWARM OPTIMIZATION PADA PENGELOMPOKAN DAERAH STUNTING

K-Means Optimization With Particle Swarm Optimization Algorithm to Determine Cluster of Stunting Area

Harliana*¹, Raden Mohamad Herdian Bhakti², Otong Saeful Bachri³, Fery Sofian Efendi⁴

¹Program Studi Ilmu Komputer, Fakultas Eksakta, Universitas Nahdlatul Ulama Blitar, Indonesia

^{2,3}Program Studi Teknik Informatika, Fakultas Teknik, Universitas Muhadi Setiabudi Brebes, Indonesia

⁴Program Studi Manajemen Informatika, Jurusan Teknologi Informasi, PSDKU Polinema Kediri, Indonesia

e-mail: *¹harliana.hifzhiya@gmail.com, ²herdian.bhakti@umus.ac.id,

³otongsaifulbahriumus@gmail.com, ⁴feri.sofian@gmail.com

Abstrak

Stunting suatu kondisi balita yang memiliki tinggi badan lebih pendek dari usia normalnya. Berdasarkan data PSG Dinas Kesehatan Kabupaten Kediri di bulan februari 2018, tingkat prevelensi stunting di seluruh kecamatan Kabupaten Kediri sekitar 19,79%. Melihat tingkat prevelensi yang tinggi tersebut, maka penelitian ini akan melakukan clustering terhadap 37 kecamatan yang ada di Kabupaten Kediri berdasarkan prosentasi tingkat stunting tertinggi sampai dengan terendah. Untuk mendapatkan hasil yang maksimal maka algoritma k-means yang digunakan akan dioptimasi dengan PSO. Berdasarkan hasil penelitian didapatkan bahwa rata-rata hasil nilai silhouette coefficient dan akurasi pada PSO k-means akan menghasilkan nilai yang lebih tinggi bila dibandingkan dengan k-means murni. Namun apabila dilihat dari waktu komputasinya maka k-means murni memiliki waktu yang lebih cepat bila dibandingkan dengan PSO k-means.

Kata kunci—stunting, optimasi, k-means, particle swarm intelligence

Abstract

Stunting is a condition of toddlers who have a height that is shorter than their average age. Based on PSG data from the Kediri District Health Office in February 2018, the stunting prevalence rate in all Kediri Districts was around 19.79%. This study will group 37 sub-districts in Kediri Regency based on the highest to lowest stunting rates to reduce the prevalence rate. The k-means algorithm used will be optimized with PSO to get maximum results. Based on the results of the study, it is known that the average silhouette coefficient and accuracy on the PSO k-means will produce a higher value when compared to the pure k-means. However, when viewed from the computational time, pure k-means have a faster time when compared to PSO k-means

Keywords—stunting, optimization, k-means, particle swarm intelligence

PENDAHULUAN

Stunting merupakan suatu kondisi balita dengan tinggi badan yang lebih dari minus dua standar deviasi median dari standar pertumbuhan anak normal[1]. Dengan kata lain seorang balita dikatakan stunting ketika memiliki masalah pada gizinya sehingga mengakibatkan tubuh yang pendek dan sangat pendek berdasarkan usianya[2]. Berdasarkan data Pantauan Status Gizi (PSG) Dinas Kesehatan Kabupaten Kediri, tingkat prevelensi stunting di Jawa Timur pada tahun 2017 sekitar 26,7% dengan Kabupaten Kediri sendiri memiliki tingkat prevelensi sebesar 19,79% per Februari 2018. Untuk mencegah tingkat prevelensi yang lebih tinggi, maka pada tahun 2019 Kabupaten Kediri menjadi salah satu kabupaten dengan lokus stunting Jawa Timur

Informasi Artikel:

Submitted: September 2021, **Accepted:** Oktober 2021, **Published:** November 2021

ISSN: 2685-4902 (media online), Website: <http://jurnal.umus.ac.id/index.php/intech>

oleh Kementerian Perencanaan Pembangunan Nasional dan Kepala Badan Pembangunan Nasional bersama 11 kota dan kabupaten lainnya yaitu Kabupaten Malang, Trenggalek, Probolinggo, Bondowoso, Nganjuk, Jember, Lamongan, Bangkalan, Sampang, Pamekasan, dan Sumenep.

Dari 37 kecamatan yang ada di Kabupaten Kediri, ternyata masih ada 16 kecamatan dengan prosentase stunting diatas 15% dimana 2 kecamatan tersebut adalah Kecamatan Kandangan dan Plemahan dengan tingkat prosentasi stunting sebesar 23,61% dan 20,01% atau setara dengan 323 dan 603 anak usia balita dari 1.368 dan 3.014 balita yang diperiksa. Untuk menekan angka stunting ini, Dinas Kesehatan Kabupaten Kediri telah mencanangkan program Gerakan Peduli Keluarga (Garpu) dengan 9 kegiatan inovasi yaitu sarasehan atau rembuk stunting, balita cerdas (ibu balita cerdas, sehat dan pintar), buku KIA lestari, lumbung pitutur paes manten, KPM (kader pembangunan manusia) di desa lokus, germas melalui pemanfaatan daun kelor, wirausahawan jamban, perubahan perilaku pada O-KPK (Organisasi Kader Penyuluh Kesehatan), dan kelas parenting yang akan mengedukasi keluarga yang memiliki anak PAUD. Melalui kegiatan sarasehan atau rembuk stunting yang telah dilakukan dari bulan Februari 2019 hingga Februari 2020, tingkat stunting yang terjadi di Desa Panjul Kecamatan Plosoklaten berhasil ditekan sekitar 1,38% atau setara mengalami penurunan dari 112 menjadi 81 balita yang mengalami stunting.

Berdasarkan hal tersebut, maka penelitian ini akan melakukan *clustering* terhadap desa-desa di Kabupaten Kediri untuk melihat prosentasi tingkat stunting dari yang tertinggi sampai dengan yang terendah. Hal ini bertujuan untuk membantu Dinas Kesehatan Kabupaten Kediri dalam menerapkan skala prioritas penanganan yang tepat berdasarkan 9 kegiatan inovasi Gerakan Peduli Keluarga yang telah dicanangkan. Hasil dari *clustering* ini juga diharapkan dapat membantu pemerintah untuk menekan prevelensi stunting balita di tingkat nasional. Algoritma *clustering* yang akan dipakai dalam penelitian ini adalah K-means yang akan dioptimasi dengan PSO (*Particle Swarm Optimization*). PSO akan digunakan untuk mengoptimalkan nilai dari jarak antar cluster sehingga dapat menghasilkan pusat cluster baru yang lebih optimal [3]. Selain dengan PSO, penentuan nilai centroid yang optimal juga pernah dilakukan dengan menggunakan algoritma genetic[4][5][6] dan algoritma pillar [7]. Sedangkan untuk optimasi terhadap jumlah cluster dapat dilakukan dengan melakukan algoritma optimasi rule of thumb [8], pendekatan elbow [9], dan index davies bouldin [10].

METODE PENELITIAN

Metode penelitian yang dilakukan dalam penelitian ini adalah penelitian eksperimen dengan alur seperti gambar 1. Penelitian akan dimulai dengan proses observasi dan wawancara untuk mendapatkan dataset stunting di Dinas Kesehatan Kabupaten Kediri. Observasi dilakukan untuk mengetahui penyebab stunting yang terjadi, sedangkan wawancara dilakukan untuk mengetahui alur dan makna dari program garpu yang dicanangkan. Setelah data tentang stunting tersebut didapatkan maka langkah selanjutnya akan dilakukan preprocessing data. Preprocessing dilakukan untuk menghilangkan data pencilan / data yang outlier dari dataset. Untuk mendapatkan jumlah cluster optimal akan dihitung dengan menggunakan metode elbow melalui perhitungan *Sum of Square Error (SSE)* di persamaan (1) [11]

$$SSE = \sum_{k=1}^k \sum_{x=i} |x_i - c_k|^2$$

persamaan (1)

Dengan:

- k : cluster ke-c
- x_i : jarak data obyek ke-i

c_k : pusat cluster ke-i

Adapun tahapan yang dilakukan untuk melakukan PSO K-Means adalah sebagai berikut[12] :

- a. Inisialisasi partikel yang berisi centroid setiap cluster dan kecepatan partikel melalui persamaan (2).

$$V_i^{t+1} = \omega \cdot v_i^t + c_1 \cdot r_1 \cdot (p_i - x_i) + c_2 \cdot r_2 \cdot (p_{g_i} - x_i) \quad \text{persamaan (2)}$$

Dengan :

- v_i : nilai kecepatan untuk partikel ke-i hingga ke-n.
- t : waktu iterasi.
- ω : nilai vector inertia.
- p_i : posisi terbaik untuk setiap artikel.
- p_{g_i} : posisi terbaik untuk semua artikel.
- c_1, c_2 : konstanta cognitive dan social.
- r_1, r_2 : bilangan yang dibangkitkan secara acak antara 0-1

- b. Tentukan partikel *best* dan global *best* awal.
 c. Mulai iterasi berdasarkan sejumlah partikel dan lakukan tahapan berikut ini untuk setiap data vector:
 i. Hitung dengan k-means.
 ii. Tentukan cluster setiap vector dengan proses K-Means clustering
 iii. Hitung nilai silhouette coefficient dengan persamaan (3).

$$s(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))} \quad \text{persamaan (3)}$$

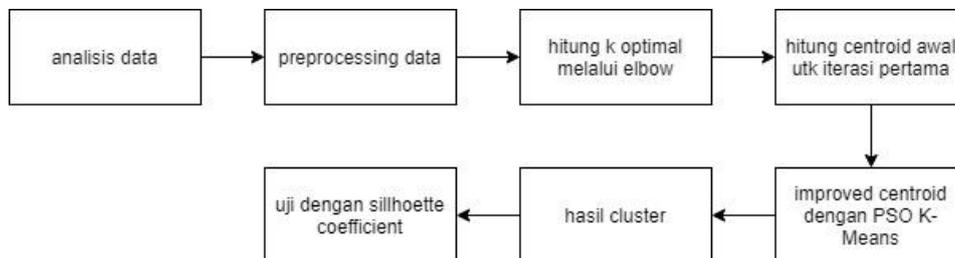
- iv. Pengulangan akan berhenti ketika artikel terakhir.
 v. Update centroid dengan menggunakan update kecepatan di persamaan (4) dan update partikel di persamaan (5).

$$\omega = (\omega_{\max} - \omega_{\min}) \frac{(\text{Iterasi} - t)}{\text{Iterasi}} + \omega_{\min} \quad \text{persamaan (4)}$$

$$x_i^{t+1} = x_i + V_i^{t+1} \quad \text{persamaan (5)}$$

- vi. Update partikel best dan global best
 d. Iterasi berhenti ketika maksimum iterasi tercapai.

Rangkuman mengenai tahapan penelitian yang dilakukan terdapat pada Gambar 1.



Gambar 1. Alur penelitian

HASIL DAN PEMBAHASAN

Penelitian ini akan menggunakan dataset dari 37 kecamatan dan 37 puskesmas di Kabupaten Kediri. Untuk variabel yang digunakan dalam penelitian ini adalah:

- a. Tinggi badan menurut umur, yang terbagi menjadi sangat pendek, pendek, normal, dan tinggi.
- b. Berat badan menurut umur, yang terbagi menjadi berat badan sangat kurang, berat badan kurang, berat badan normal, dan resiko berat badan lebih.
- c. Berat badan menurut tinggi badan yang terbagi menjadi gizi buruk (severely wasting, gizi buruk (wasting), gizi baik (normal), beresiko gizi lebih, gizi lebih, dan obesitas.
- d. Supas (proyeksi balita).
- e. Jumlah balita yang diperiksa.
- f. Jumlah balita stunting (sangat pendek dan pendek).
- g. Jumlah balita underweight (berat badan sangat kurang dan berat badan kurang).
- h. Jumlah balita wasting (gizi buruk dan gizi kurang).

Gambar 2 adalah rangkuman rekap stunting yang terjadi di Kabupaten Kediri yang dikelompokkan berdasarkan kecamatan.

NO	PUSKESMAS	TINGGI BADAN ATAU PANJANG BADAN MENURUT UMUR				BERAT BADAN MENURUT UMUR				BERAT BADAN MENURUT TINGGI BADAN					
		Sangat Pendek	Pendek	Normal	Tinggi	Berat Badan Sangat Kurang	Berat badan Kurang	Berat Badan Normal	Risiko Berat Badan Lebih	Gizi Buruk (Severely Wasting)	Gizi Buruk (Wasting)	Gizi Baik (Normal)	Berisiko Gizi Lebih	Gizi lebih	Obesitas
1	KANDANGAN	85	240	1024	21	34	134	1105	112	32	74	1041	123	59	39
2	PLEMAHAN	183	420	2378	33	74	271	2402	256	88	189	2157	312	148	103
3	PARE	117	264	1572	37	48	195	1535	202	33	63	1693	111	41	36
4	KRAS	37	234	1179	6	29	152	1160	116	6	118	1103	137	57	35
5	PURWOASRI	106	216	1377	51	52	165	1344	202	65	117	1201	196	75	93
6	GURAH	136	335	2070	19	39	233	2055	243	32	117	1883	235	87	106
7	KEPUNG	94	297	1711	31	44	250	1679	171	30	106	1771	128	55	43
8	KAYEN KIDUL	49	167	950	15	23	129	919	110	27	66	891	118	36	43
9	NGADILUWIH	64	228	1339	23	14	133	1360	168	0	98	1285	153	62	56
10	PAGU	96	278	1807	14	30	199	1735	230	35	92	1608	265	98	96
11	KUNJANG	59	319	1817	41	31	253	1762	190	73	161	1652	198	87	65
12	PARE	49	129	967	42	23	124	928	115	39	50	868	82	33	19
13	PLOSOKLATEN	44	295	1704	32	50	243	1617	168	27	88	1725	144	61	30
14	KANDAT	120	309	1918	282	53	193	1719	669	15	246	1771	381	188	128
15	GAMPENREJO	77	198	1406	33	45	180	1381	180	55	81	1351	121	54	49
16	PAPAR	67	326	2153	22	49	254	2059	214	35	164	2029	206	74	58
17	NGANCAR	71	228	1679	28	26	161	1656	163	38	90	1572	215	53	36
18	PLOSOKLATEN	63	233	1716	28	25	167	1648	203	25	110	1562	213	78	50
19	GURAH	54	242	1843	24	24	214	1731	208	23	112	1761	156	69	41
20	PURWOASRI	23	172	1264	3	19	131	1152	160	13	84	1087	156	70	52
21	PARE	21	139	1107	4	23	109	1010	125	19	89	913	143	50	49
22	KRAS	39	113	1082	12	21	112	1011	105	29	87	958	103	33	39
23	GAMPENREJO	26	183	1499	5	19	137	1395	160	14	97	1285	169	78	68
24	WATES	64	204	1959	19	46	216	1765	221	39	149	1752	185	76	45
25	PUNCU	50	214	1974	29	60	264	1804	175	41	82	1801	134	51	37
26	BADAS	49	331	3015	20	34	296	2889	196	39	96	3064	143	54	19
27	BANTYAKAN	29	217	1986	30	16	97	1130	132	29	45	1029	101	37	30
28	MOJO	47	141	1590	24	26	198	1475	113	28	130	1468	109	51	15
29	SEMEN	48	218	2254	30	38	258	2047	212	32	136	2055	200	64	57
30	NGADILUWIH	18	147	1392	37	12	123	1529	154	20	125	1272	96	48	32
31	KAYEN KIDUL	36	105	1212	11	19	93	1129	124	10	89	1108	109	38	30
32	KEPUNG	38	236	2249	60	22	214	2309	138	19	127	2440	54	23	10
33	RINGINREJO	57	161	1861	64	44	217	1714	169	110	219	1516	181	56	61
34	WATES	35	189	2024	16	19	153	1920	180	2	54	1907	232	38	31
35	TAROKAN	11	296	2889	61	48	223	2603	383	59	147	2511	300	154	84
36	MOJO	59	197	2831	20	41	274	2575	218	40	154	2536	254	68	55
37	NGASEM	87	130	2023	2	51	260	2654	277	48	133	2667	198	89	86
JUMLAH :		2304	8347	65821	1229	1271	7024	61601	7182	1269	4189	60485	6271	2491	1926

Gambar 2. Dataset stunting kecamatan

Pada tahapan preprocessing, peneliti memanfaatkan fitur *replace missing value* dari rapidminer. Hasilnya adalah tidak ada data yang missing, sehingga dataset yang akan diolah masih sama seperti dataset awal. Berdasarkan perhitungan dengan persamaan (3) maka jumlah cluster optimal yang akan digunakan adalah 2 cluster (k=2) dengan centroid awal di iterasi pertama adalah titik tengah dari dataset.

Sebelum melakukan perhitungan PSO maka akan dilakukan inisialisasi terlebih dahulu untuk mendapatkan nilai kecepatan masing-masing partikel seperti jumlah partikel = 13, jumlah iterasi = 10, cognitive = 1,5 dan nilai w. Selanjutnya akan dihitung nilai fitness berdasarkan persamaan (3), rangkuman mengenai nilai fitness yang dihasilkan terangkum pada Tabel 1.

Table 1. Nilai fitness yang dihasilkan

Partikel	Fitness
1	0,3210

Partikel	Fitness
2	0,1156
3	0,1423
4	0,2109
5	0,1190
...	...
28	0,1291

Nilai fitness ini selanjutnya akan digunakan untuk mencari nilai pbest dan gbest dari fitness terbaik untuk setiap artikel (p_i) dan fitness terbaik untuk semua artikel (pg_i). Selanjutnya akan dihitung perubahan kecepatan melalui persamaan (2). Nilai w (vector inertia) pada persamaan (2) didapatkan melalui persamaan (4) dengan nilai konstanta cognitive dan social adalah 1 dan nilai r adalah bilangan random antara 0 dan 1. Perhitungan untuk nilai w (vector inertia) berdasarkan persamaan (4) adalah:

$$\omega = (0,6 - 0,2) * \frac{(10 - 1)}{10} + 0,2$$

$$\omega = 0,56$$

Sehingga didapatkan nilai v_i adalah:

$$V_{2,1} = 0,56(0) + (1)(1).(25-30) + (1)(1).(25-30)$$

$$V_{2,1} = -10$$

berdasarkan hasil perhitungan diatas maka diketahui nilai kecepatan yang dihasilkan pada partikel 2 dimensi 1 adalah -10, perhitungan dan cara yang sama juga akan dilakukan untuk setiap partikel. Selanjutnya akan dilakukan perhitungan untuk perubahan posisi melalui persamaan (5). Dan proses PSO ini akan terus dilakukan sampai dengan maksimal iterasi dengan gbest position sebagai solusi yang diberikan. Tabel 2 adalah gbest partikel dari centroid yang dihasilkan.

Tabel 2. Representasi Partikel

Cluster	Global best position
1	-0,12845
2	-0,38032

Pengujian

Untuk mengetahui performa PSO terhadap K-Means maka akan dilakukan 3 jenis pengujian melalui 10 kali pengujian, yaitu:

- Membandingkan hasil pengujian antara K-Means murni dan PSO K-Means berdasarkan rata-rata nilai silhouette coefficient yang dihasilkan oleh keduanya.
- Membandingkan rata-rata nilai akurasi diantara keduanya.
- Membandingkan rata-rata waktu komputasi.

Rangkuman hasil ketiga hasil pengujian tersebut terangkum pada Tabel 3.

Table 3. Hasil pengujian k-means dan PSO k-means

Jenis Algoritma	Rata-rata nilai Silhouette coefficient	Rata-rata nilai akurasi	Rata-rata waktu komputasi
k-means murni	0,12337	0,6572	0,42
PSO k-means	0,30731	0,8732	1,85

Berdasarkan table 3 diketahui bahwa PSO k-means dapat membentuk model suatu cluster menjadi lebih baik bila dibandingkan dengan k-means murni, hal ini terlihat dari rata-rata nilai silhouette coefficient PSO k-means yang lebih tinggi bila dibandingkan dengan k-means murni. Untuk nilai rata-rata akurasi yang dihasilkan juga menunjukkan bahwa PSO k-means memiliki nilai akurasi yang lebih tinggi bila dibandingkan dengan k-means murni yaitu 87%, namun apabila dilihat dari waktu komputasi yang dibutuhkan maka k-means murni lebih cepat bila dibandingkan dengan PSO k-means.

KESIMPULAN

Berdasarkan hasil penelitian yang telah didapatkan, dapat diketahui bahwa PSO k-means mampu melakukan clustering yang lebih baik bila dibandingkan dengan k-means murni dalam mengelompokkan daerah stunting di Kabupaten Kediri. Hal ini terlihat dari nilai silhouette coefficient dan akurasi yang dihasilkan oleh PSO k-means lebih tinggi bila dibandingkan dengan k-means murni yaitu sekitar 0,30731 dan 87%. Namun waktu yang diperlukan oleh PSO k-means dalam melakukan eksekusi cenderung lebih lama yaitu sekitar 1,85 detik bila dibandingkan k-means murni yang hanya sekitar 0,42 detik.

DAFTAR PUSTAKA

- [1] A. D. Laksono and H. Megatsari, "Determinan Balita Stunting di Jawa Timur: Analisis Data Pemantauan Status Gizi 2017," *Amerta Nutr.*, vol. 4, no. 2, p. 109, 2020, doi: 10.20473/amnt.v4i2.2020.109-115.
- [2] C. Dewanti, V. Ratnasari, and T. Rumiati, "Pemodelan Faktor-faktor yang Memengaruhi Status Balita Stunting di Provinsi Jawa Timur Menggunakan Regresi Probit Biner," *J. Sains Dan Seni Its*, vol. 8, no. 2, 2019, [Online]. Available: http://ejurnal.its.ac.id/index.php/sains_seni/article/view/48519.
- [3] D. Werdiningsih, D. Rahmawati, I. Cholissodin, and N. Santoso, "Optimasi K-Means untuk Pengelompokan Data Kinerja Akademik Dosen menggunakan Particle Swarm Optimization," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 3, no. 4, pp. 4102–4110, 2019.
- [4] M. Mursalim, P. Purwanto, and M. A. Soeleman, "Penentuan Centroid Awal Pada Algoritma K-Means Dengan Dynamic Artificial Chromosomes Genetic Algorithm Untuk Tuberculosis Dataset," *Techno.Com*, vol. 20, no. 1, pp. 97–108, 2021, doi: 10.33633/tc.v20i1.4230.
- [5] T. Taslim, D. Toresa, D. Jollyta, D. Suryani, and E. Sabna, "Optimasi K-Means dengan Algoritma Genetika untuk Target Pemanfaat Air Bersih Provinsi Riau," *Indones. J. Comput. Sci.*, vol. 10, no. 1, 2021.
- [6] D. A. Kuntjoro, B. D. Setiawan, and R. S. Perdana, "Algoritme Genetika Untuk Optimasi K-Means Clustering Dalam Pengelompokan Data Tsunami," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 2, no. 10, pp. 3865–3872, 2018, [Online]. Available: <http://j-ptiik.ub.ac.id>.
- [7] 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 Jawa Timur)," *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>.
- [8] M. Nishom and M. Y. Fathoni, "Implementasi Pendekatan Rule-Of-Thumb untuk Optimasi Algoritma K-Means Clustering," *J. Inform. J. Pengemb. IT*, vol. 3, no. 2, pp. 237–241, 2018, doi: 10.30591/jpit.v3i2.909.
- [9] A. Winarta and W. J. Kurniawan, "Optimasi Cluster K-Means Menggunakan Metode

- Elbow Pada Data Pengguna Narkoba Dengan Pemrograman Python,” *JTIK (Jurnal Tek. Inform. ...*, vol. 5, no. 1, 2021, [Online]. Available: <https://jurnal.kaputama.ac.id/index.php/JTIK/article/view/466>.
- [10] E. Muningsih, I. Maryani, and V. R. Handayani, “Penerapan Metode K-Means dan Optimasi Jumlah Cluster dengan Index Davies Bouldin untuk Clustering Propinsi Berdasarkan Potensi Desa,” *Evolusi J. Sains dan Manaj.*, vol. 9, no. 1, pp. 95–100, 2021, doi: 10.31294/evolusi.v9i1.10428.
- [11] D. A. I. C. Dewi and D. A. K. Pramita, “Analisis Perbandingan Metode Elbow dan Silhouette pada Algoritma Clustering K-Medoids dalam Pengelompokan Produksi Kerajinan Bali,” *Matrix J. Manaj. Teknol. dan Inform.*, vol. 9, no. 3, pp. 102–109, 2019, doi: 10.31940/matrix.v9i3.1662.
- [12] I. Wahyuni, Y. A. Auliya, A. Rahmi, and W. F. Mahmudy, “Clustering nasabah bank berdasarkan tingkat likuiditas menggunakan hybrid particle swarm optimization,” vol. 10, no. 2, pp. 24–33, 2016.