Implementasi Algoritma *K-Means* dengan *Particle Swarm Optimization* Untuk Pengelompokan Tingkat Kesejahteraan Di Provinsi Jawa Timur

Implementation of K-Means Algorithm with Particle Swarm Optimization for Welfare Level Grouping in East Java Province

Ahmad Dedi Hartono[1], Deni Arifianto[2], Ilham Saifudin[3]
[1]Mahasiswa Fakultas Teknik, Universitas Muhammadiyah Jember
e-mail: deddyhartono18@gmail.com[1]
[2,3]Dosen Fakultas Teknik, Universitas Muhammadiyah Jember

Abstrak

Sejahtera diartikan sebagai kondisi masyarakat yang berkecukupan dan mampu memenuhi kebutuhannya. Jika membahas kesejahteraan, adanya perbedaan sumber daya, kondisi demografi serta pembangunan disetiap wilayahnya menyebabkan ketimpangan. Terjadinya ketimpangan pembagunan antar wilayah, akan membawa impilkasi terhadap tingkat kesejahteraan pada wilayah bersangkutan. Implikasi tersebut ditimbulkan dalam bentuk kecemburuan dan ketidakpuasan yang dapat mengganggu ketentraman masyarkat. Dalam hal tersebut, diperlukan upaya untuk mengetahui tingkat kesejahteraan suatu wilayah atau kabupaten guna mengatasi ketimpangan kesejahteraan. Kabupaten yang akan dijadikan prioritas dapat diketahui dengan melakukan pengelompokan, satu diantaranya adalah memakai algoritma *K-Means* dengan *Particle Swarm Optimization* untuk menghitung *cluster* optimal dalam rangka menentukan *cluster* yang paling baik. Data yang digunakan adalah data pada 38 kabupaten/kota di jawa timur periode tahun 2019-2020. Berdasarkan pengujian dari 2 sampai 10 *cluster*, dihasilkan 2 *cluster* sebagai *cluster* optimum yang mana nilai *sillhouette coeficient* yaitu sebesar 0,35057. Ada 27 anggota kabupaten/kota di *cluster* 1, sedangkan ada 11 anggota kabupaten/kota di *cluster* 2.

Kata kunci: kesejahteraan, *k-means, particle swarm optimization, clustering, sillhouette coefficient*.

Abstract

Prosperity is defined as the condition of the community that is sufficient and able to meet their needs. When discussing welfare, the differences in resources, demographic conditions and development in each region cause inequality. The occurrence of development inequality between regions will have implications for the level of welfare in the region concerned. These implications are generated in the form of jealousy and dissatisfaction which can disturb the peace of the community. In this case, efforts are needed to determine the level of welfare of a region or district in order to overcome welfare inequality. Districts that will be prioritized can be identified by grouping, one of which is using the K-Means algorithm with Particle Swarm Optimization to measure the optimum cluster in order to determine the best cluster. The data used is data from 38 districts or cities in East Java for the 2019-2020 period. Based on testing from 2 to 10 clusters, 2 clusters were produced as the optimum cluster in which the sillhouette coefficient value was 0.35057. There are 27 district/city members in cluster 1, and there are 11 in cluster 2.

Keywords*: welfare, k-means, particle swarm optimization, clustering, sillhouette coefficient*.

1. PENDAHULUAN

Salah satu faktor yang sering digunakan untuk menyatakan keberhasilan suatu daerah atau pemerintah dalam mengelola kesejahteraan dan kemakmuran adalah tingkat perekonomian. Situasi masyarakat yang cukup dan mampu memenuhi kebutuhan disebut kemakmuran, sedangkan kemakmuran merupakan komponen yang memungkinkan individu hidup damai dan terhindar dari kecemburuan sosial.

Jika membahas tentang kesejahteraan, ada perbedaan kandungan sumber daya, perbedaan kondisi demografi serta pembangunan setiap wilayah menyebabkan terjadinya ketimpangan. Tingkat kesejahteraan masyarakat yang bersangkutan dipengaruhi oleh adanya disparitas pembangunan (Simbolon, 2017). Dampak tersebut biasanya berupa ketidakpuasan dan kecemburuan publik namun juga dapat berlanjut dengan dampak politik dan perdamaian publik.

Berdasarkan pernyataan diatas, diperlukan sebuah upaya untuk mengetahui tingkat kesejahteraan suatu wilayah guna mengatasi ketimpangan kesejahteraan. Satu dari beberapa cara yang bisa diterapkan antara lain menggunakan pengelompokan kesejahteraan. Pengelompokan dapat dilakukan dengan analisis *cluster.*

Analisis *cluster* merupakan suatu metode untuk mengelompokkan variabel atau objek ke dalam beberapa kelompok. Tujuan dari analisis klaster adalah mengelompokkan *n* objek yang memiliki karakteristik yang sama satu sama lain berdasarkan *p* variabel (Soemartini & Supartini, 2017). Objek tersebut akan dikategorikan menjadi satu atau lebih klaster (grup) dengan objek dalam satu klaster berbagi karakteristik

Beberapa peneliti sebelumnya menggunakan metode *K-Means* untuk metode *clustering* data. Salah satunya adalah penelitian tingkat kesejahteraan masyarakat yang dilakukan oleh (Tenriawaru & Kunci, 2018) dengan studi kasus "Kabupaten atau Kota Provinsi Sulawessi Tenggara Berdasarkan Indikator Kesejahteraan Rakyat. Dari riset tersebut menciptakan data yang dikumpulkan menjadi 3 *cluster*. Namun pada penelitian tersebut tidak melakukan pengukuran *cluster* optimal dalam menentukan *cluster* yang paling baik. Penelitian lainnya yang dilakukan oleh (Rustam, Santoso, & Supriyanto, 2018) dengan studi kasus "Optimasi *K-Means Clustering* Untuk Identifikasi Daerah Endemik Penyakit Menular (Study Kasus : Kota Semarang) menggunakan Algoritma *Particle Swarm Optimation* untuk acuan penentuan *cluster* optimum.

Oleh karena itu pemerintah perlu mengetahui tentang kabupaten/kota di jawa timur yang mana perlu ditingkatkan tingkat kesejahteraannya. Penelitian ini menggunakan data dari badan pusat statistik pada tahun 2019-2020 pada kabupaten atau kota di jawa timur. Dengan melakukan pengelompokan kita dapat mengetahui wilayah mana saja yang perlu diutamakan dalam hal seperti pembangunan pembangunan sekolah dan sarana kesehatan, serta peningkatan yang lain. Dengan begitu penelitian ini diharapakan mampu menjadi salah satu alternatif penelitian untuk mengelompokkan data tingkat kesejahteraan masyarakat sehingga dapat terwujudnya kesejahteraan pada seluruh wilayah di jawa timur.

2. STUDI PUSTAKA

A. Data Mining

Penambangan data adalah salah satu dari beberapa teknik berbasis komputer yang dapat digunakan untuk mengumpulkan data, menganalisisnya, dan mengidentifikasi pola pada data yang belum ditemukan sebelumnya. *Database*, pembelajaran mesin, kecerdasan buatan, jaringan saraf, statistik, pengenalan pola, sistem berbasis pengetahuan, akuisisi pengetahuan, pencarian informasi, visualisasi data, dan komputasi kinerja tinggi merupakan beberapa bidang teknologi yang berkontribusi pada *output* data mining.

Data mining dan *Knowledge Discovery in Database* sesekali digunakan secara bergantian untuk menggambarkan cara paling umum untuk menggali informasi yang tersimpan dalam kumpulan yang sangat luas atau banyak. Pada hakikatnya

dua istilah tersebut mempunyai rancangan yang tidak sama, namun saling berkaitan. Bahkan satu dari tahapan *Knowledge Discovery* didalam seluruh prosesnya merupakan *data mining.*

B. Clustering

Clustering adalah siklus untuk mengelompokkan data ke dalam beberapa kelompok atau pertemuan dengan tujuan agar data dalam satu kelompok memiliki tingkat kemiripan yang paling tinggi dan data antar kelompok memiliki kesamaan yang kecil (Tan, 2006).

Clustering dapat mengambil bagian yang signifikan dalam kehidupan sehari-hari, karena tidak dapat dipisahkan dari seberapa banyak informasi yang menghasilkan data untuk mengatasi masalah kehidupan. Salah satu metode utama untuk mengelola data ialah dengan meng-klasifikasikan atau pengelompokan data ke dalam sekelompok klasifikasi atau *cluster*. *Clustering* dapat ditemui di sebagian aplikasi yang ada pada bermacam-macam bidang. Misalnya, pengumpulan informasi digunakan untuk memeriksa informasi faktual, misalnya, pengelompokan untuk pembelajaran AI, penggalian data, pemeriksaan gambar dan bioinformatika.

Berdasarkan penjabaran secara garis besar terdapat 2 jenis metode Pengelompokan atau *clustering*:

- *Hierarchical Clustering Method*, yang mana jumlah klaster tidak ditentukan terlebih dahulu. Metode yang masuk ke dalam golongan ini adalah *Single Linkage, Complete Linkage, Average Linkage*, dll.
- *Non Hierarchical Clustering Method*, yang mana jumlah k sudah ditentukan di awal. Metode yang masuk kedalam golongan ini adalah *K-Means, Fuzzy C-Means, Kohonen SOM, LVQ*, dll.

C. Normalisasi Data

Normalisasi adalah metode penskalaan atau teknik pemetaan data. Di mana kami dapat melacak rentang baru dari jangkauan yang ada. Normalisasi ini akan sangat bermanfaat untuk tujuan prediksi atau peramalan. Berikut merupakan contoh teknik normalisasi yang ada antara lain normalisasi min-max, normalisasi z-skor dan normalisasi penskalaan desimal (Patro & sahu, 2015). Dalam penelitian ini, teknik normalisasi data yang digunakan ialah normalisasi min-max.

Normalisasi min-max memberikan perubahan langsung pada rentang data asli tetapi tetap mempertahankan hubungan antara data asli. Rumus normalisasi min-max ditampilkan dalam kondisi berikut:

$$x_j = \frac{x_i - \min value\ of\ x}{\max value\ of\ x - \min value\ of\ x} \quad (1)$$

Keterangan

x_j = data setelah normalisasi

x_i = data ke-*i*

D. *K-Means*

Algoritma *KMeans* adalah salah satu algoritma partisi, karena *K-Means* bergantung pada penentuan jumlah kelompok yang mendasarinya dengan mendefinisikan nilai centroid awal (Madhulatha, 2012). Algoritma *K-Means* menggunakan interaksi yang berurutan untuk memperoleh sekumpulan data *cluster*. Membutuhkan jumlah *cluster* pertama yang diinisialisasikan sebagai *input* dan memebentuk centroid akhir menjadi *output*. Teknik *K-Means* akan memilih *k* pola sebagai titik pertama centroid secara acak.

Langkah-langkah menggunakan metode *K-Means* dapat dilakukan dengan cara sebagai berikut:

- Menentukan jumlah *cluster*.
- Menginisialisai *k* menjadi pusat *cluster* awal (menggunakan nilai acak).
- Hitung jarak terdekat setiap data pada pusat *cluster* menggunakan rumus *Euclidean Distance*.
- Menghitung ulang *centroid* dengan anggota kelompok *cluster* baru.
- Menghitung ulang jarak data menggunakan *centroid* baru.

Ulangi cara nomer 3 sampai dimana nilai *centroid* tidak berganti lagi.

E. *Particle Swarm Optimization*

Particle Swarm Optimization yang dikemukakan Dr. Eberhart dan Dr. Kennedy tahun 1995, ialah metode tambahan yang menyalin siklus pada kelompok burung (sekelompok burung) beserta ikan (sekolah ikan) saat hidup. Semenjak pertama diperkenalkan, metode PSO telah berkembang, dalam segi penerapannya maupun dalam pengembangan algoritmanya (Haupt, Haupt, & Wiley, 2004).

Menurut (Supianto & Mar'i, 2018) optimasi *clustering* pada PSO dihitung berdasarkan perpindahan posisi paertikel dan *update* kecepatan. Langkah ini dilakukan untuk mengupdate pusat *cluster*. Adapaun rumus perhitungan *update* kecepatan ialah sebagai berikut:

$$V_i\,(t) = WV_i + c_1\,R_1\left(X_{pi} - X_i\right) + c_2\,R_2\left(X_{gi} - X_i\right) \quad (2)$$

dimana:

R_1 dan R_2 = nilai *random* antara [0,1]
c_1 dan c_2 = *coeficient acceleration*
X_{pi} = posisi terbaik partikel ke-*i*
X_{gi} = nilai rata-rata semua solusi
W = konstanta inersia
X_i= posisi partikel ke-*i*

Standarlisasi PSO menggunkan nilai *inersia weight* (W_{max} ,W_{min}) antara =0,4 sampai 1,4, dan *coeficient acceleration* antara 1,5 sampai 2,0 (Singh, Singh, & Singh, 2012).

Setelah menghitung kecepatan partikel, maka selanjutnya menghitung perpindahan posisi partikel menggunakan rumus *update* posisi sebagai berikut:

$$X_i\,(t) = X_t + V_t \quad (3)$$

Dimana :
X_t = posisi partikel
V_t = kecepatan baru

Setelah pusat cluster di *update* makan akan diperoleh titik pusat *cluster* yang baru. Ulangi perhitungan pada persamaan 1 dan 2 sampai terjadi *stoping condition* yaitu ketika posisi partikel menuju ke satu nilai yang sama untuk setiap *cluster*.

F. *Sillhouette Coefficient*

Sillhouette Coefficient ialah metode evaluasi untuk menentukan mutu dari *cluster*, seberapa bagus objek di tempatkan di suatu *cluster* (Mario, Herry, & Nasution, 2016). Tahap-tahap menghitung *Silhouette Coefficient* ialah seperti berikut:

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

   $$a(i) = \frac{1}{[A]-1} \sum j \in_{A, j \neq i} d(i,j) \quad (4)$$

   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 niali terkecil.

   $$d(i,C) = \frac{1}{[C]} \sum j \in_C d(i,j) \quad (5)$$

   Kemudian mencari niali terkecilnya dengan persamaan berikut:

   $$b(i) = min_{C \neq A}\, d(i,C) \quad (6)$$

   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*≠ *A*.
3. Hitung nilai *silhouette coefficient* dengan persamaan berikut:

   $$s(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))} \quad (7)$$

3. METODOLOGI PENELITIAN

A. Pengumpulan Data

Penelitian ini menggunakan data yang berasal dari website resmi Badan Pusat Statistik (BPS) Provinsi Jawa Timur. Data yang akan diolah pada penelitian ini menggunakan data Kabupaten atau Kota di Provinsi Jawa Timur pada tahun 2019-2020. Terdiri dari 38 Kabupaten/Kota di Jawa Timur. Dari data yang tersedia, peneliti menggunakan data tersebut untuk mengelompokkan kesejahteraan yang terjadi pada Kabupaten atau Kota yang ada di Provinsi Jawa Timur. Pada data tersebut mempunyai tujuh atribut yakni tenaga kesehatan(TK), fasikitas

kesehatan(FK), sekolah(S), guru(G), persentase sumber penerangan listrik(PSPL), tingkat partisipasi angkatan kerja(PSPL), dan indeks pembangunan masyarakat(IPM). Atribut pada tabel fasilitas kesehatan (FK) meliputi RS umum daerah, RS bersalin daerah, poliklinik, dan puskesmas.

Tabel 1. Data Kabupaten atau Kota di Jawa timur.

Kab/Kot	TK	FK	S	...	IPM
Pacitan	1396	42	709	...	68,16
Ponorogo	989	64	1009	...	70,56
Trenggalek	1793	34	730	...	69,46
Tulungagung	1403	75	981	...	72,62
Blitar	1997	65	1139	...	70,57
Kediri	1674	103	1234	...	71,85
Malang	1144	171	2291		70,35
Lumajang	1870	51	1111		65,33
Jember	674	119	2321		66,69
...	...	...	...	...	...
Banyuwangi	1889	121	1536	...	70,6

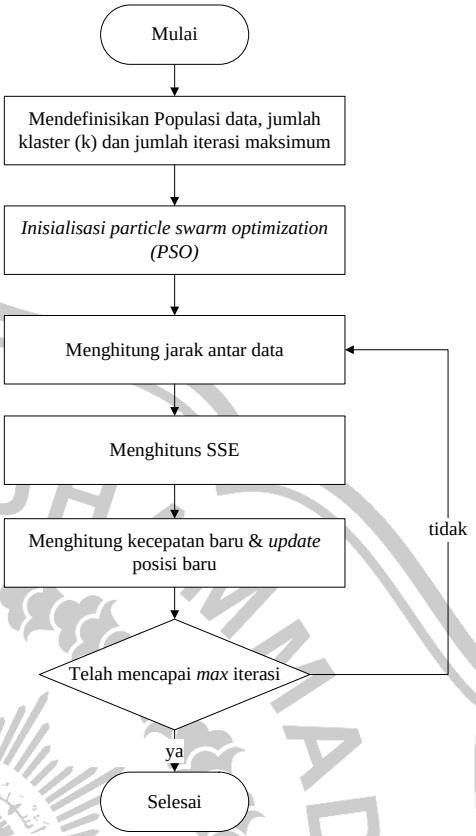
Sumber : Data Awal

B. Pra-proses Data

Data tersebut masih berupa data mentah yang nantinya masih akan diolah menggunakan cara normalisasi min-max. Setelah data mentah tersebut di normalisasi barulah data tersebut akan dilakukan perhitungan menggunakan metode *k-means* dengan *particle swarm optimization*. Pada bagian ini, normalisasi data dilakukan dengan cara normalisasi min-max.Normalisasi min-max memberikan transformasi linier pada data asli namun tetap mempertahankan hubungan antar data aslinya.

C. K-Means dengan *Particle Swarm Optimization*

Diagram alur dari penerapan metode clustering menggunakan *Kmeans* dengan *Particle Swarm Optimization* yang digunakan berdasarkan untuk menyelesaikan masalah pengelompokan kesejahteraan pada Kabupaten pada Jawa Timur:



Gambar 1 Algoritma *K-Means* dengan *Particle Swarm Optimization*

Sebelum masuk ketahapan *clustering*, terlebih dahulu melakukan normalisasi data menggunakan persamaan (1). Barulah masuk kelangkah *clustering* data menggunakan algoritma *K-Means* dengan *Particle Swarm Optimization* tahap pertama yaitu memasukkan data yang akan dilakukan *clustering*. Tahap selanjutnya yaitu melakukan inisialisasi jumlah *k* dan iterasi maksimal. Kemudian melakukan inisialisasi parameter pada *particle swarm optimization* seperti *r1, r2, c1, c2* dan lain-lain. Selanjutnya melakukan perhitungan jarak menggunakan rumus *euclidean distance*. Selanjutnya melakukan perhitungan *sse*. Lalu melakukan perhitungan kecepatan baru menggunakan persamaan (2), serta melakukan perhitungan posisi partikel baru menggunakan persamaan (3). Tahap selanjutnya melakukan perbaruan terhadap pusat *cluster* menggunakan kecepatan partikel baru dan posisi partikel baru, pada tahap ini iterasi +1. Proses selanjutnya melakukan pemeriksaan kondisi berhenti, jika iterasi sudah mencapai iterasi

maksimal atau memenuhi kondisi berhenti seperti yang sudah dijelaskan diatas.

D. Sillhouette Coefficient

Pada proses pengelompokan Kabupaten atau Kota menggunakan algoritma *K-Means* dengan *Particle Swarm Optimization* yang telah dilakukan pada rentang 2 *cluster* sampai 10 *cluster*. Selanjutnya akan dilakukan pengujian *cluster* optimum menggunakan *sillhouette coefficient* untuk menentukan *cluster* optimum atau yang terbaik. Untuk mentukan *cluster* optimum menggunakan *sillhouette coefficicent* langkah pertama adalah menghitung rata-rata jarak dari suatu objek dengan semua objek lain yang berada dalam satu *cluster* menggunakan persamaan (4). Lalu Menghitung rata-rata jarak objek dengan semua objek yang berada pada cluster lain, dan diambil niali terkecil menggunakan persamaan (5). Kemudian mencari niali terkecilnya dengan persamaan (6). Proses selanjutnya menghitung nilai *sillhouette coefficient* menggunakan persamaan (7).

4. HASIL DAN PEMBAHASAN

Pada tahap ini membahas tentang hasil yang telah di peroleh dari proses perhitungan yang telah di lakukan. Data yang telah di dapat akan di *cluster* memakai metode *K-Means* dengan *Particle Swarm Optimization* yang kemudian di olah untuk memperoleh *cluster* optimal atau *cluster* yang paling baik dengan memakai metode *Silhouette Coefficient.* Data yang digunakan yakni data pada Kabupaten/Kota di Provinsi Jawa Timur pada tahun 2019-2020 yang berjumlah 38 Kabupaen atau Kota.

A. Penentuan *Cluster* Optimum

Menentukan *cluster* yang terbaik atau optimum dari proses perhitungan menggunakan algoritma *k-means* dengan *particle swarm optimization* sebelumnya, dilanjutkan dengan menerapkan metode pengujian *cluster* optimum yaitu *sillhouette coefficient*. Adapun hasil keluaran yang didapat yaitu berupa nilai seperti pada **Tabel 2** berikut:

Tabel 2. Perhitungan *sillhouette coefficient*

Cluster	*Sillhouette Coefficient*
2	0,35057
3	0,30007
4	0,27811
5	0,25870
6	0,23866
7	0,22330
8	0,23027
9	0,22219
10	0,13680

Sumber : Hasil Perhitungan *cluster* optimum

Berdasaarkan hasil pengujian yang telah dilakukan, telah didaptakan *cluster* optimal untuk tingkat kesejahteraan di Jawa Timur ialah 2 *cluster*.

B. Penentuan Jumlah Partikel Optimum

Pengujian jumlah partikel dilakukan untuk mencari jumlah partikel terbaik yang mana memberikan nilai *Silhouette Coefficient* tertinggi. Proses pengujian partikel dilakukan terhadap masing-masing data. Adapun hasil keluaran yang didapat yaitu berupa hasil seperti pada **Tabel 3** berikut:

Tabel 3. Penentuan Jumlah Partikel

Partikel	*Sillhouette Coefficient*
10	0,38057
20	0,30380
30	0,29902
40	0,30950
50	0,29849
60	0,31642
70	0,29816
80	0,27896
90	0,28586
100	0,30371

Sumber: Hasil Perhitungan Jumlah Partikel Optimum

Untuk hasil pengujian yang telah dilakukan, telah didaptakan jumlah partikel optimal untuk tingkat kesejahteraan di Jawa Timur ialah 10 partikel.

C. Penentuan *Learning Factor* Optimum

Pengujian *Learning Factor* sama halnya dengan pengujian jumlah partikel yakni dilakukan untuk mencari nilai *learning factor* terbaik yang mana memberikan nilai *Silhouette Coefficient* tertinggi. Adapun hasil keluaran yang didapat yaitu berupa nilai seperti pada **Tabel 4** berikut:

Tabel 4 Penentuan *Learning Factor* Optimum

Learning Factor	*Sillhouette Coefficient*
[1,5;1,5]	0,30224
[1,5;1,8]	0,31791
[1,5;2,0]	0,24545
[1,8;2,0]	0,34893
[2,0;2,0]	0,33351

Sumber : Hasil Perhitungan *Learning Factor* Optimal

Pada pengujian yang telah dilakukan, telah didaptakan nilai *Learning Factor* optimal untuk tingkat kesejahteraan di Jawa Timur ialah [1,8;2,0].

D. Penentuan Bobot Inersia Optimum

Pengujian bobot inersia dilakukan untuk menemukan nilai bobot inersia yang paling baik yang mana memberikan nilai *Silhouette Coefficient* tertinggi. Adapun hasil keluaran yang didapat yaitu berupa nilai seperti pada **Tabel 5** berikut:

Tabel 5 Penentuan Bobot Inersia Optimum

Bobot Inersia	*Sillhouette Coefficient*
0,4	0,27983
0,5	0,37027
0,6	0,27535
0,7	0,27212
0,8	0,37655
0,9	0,33943
1	0,33064
1,1	0,36735
1,2	0,29195
1,3	0,24568
1,4	0,25843

Sumber : Hasil Perhitungan Bobot Inersia Optimum

Hasil pengujian yang telah dilakukan, telah didaptakan nilai bobot inersia yang optimal untuk tingkat kesejahteraan di Jawa Timur ialah 0,8.

E. Penentuan Jumlah Iterasi Optimum

Pengujian jumlah iterasi dilakukan untuk menemukan nialai *Silhouette Coefficient* tertinggi. Adapun hasil keluaran yang didapat yaitu berupa nilai seperti pada **Tabel 6** berikut:

Tabel 6 Penentuan Jumlah Iterasi Optimum

Jumlah Iterasi	*Sillhouette Coefficient*
10	0,31099
20	0,29563
30	0,34968
40	0,34687
50	0,28828
60	0,28330
70	0,29039
80	0,29660
90	0,27630
100	0,27782

Sumber : Hasil Perhitungan Jumlah Iterasi Optimum

Berdasarkan hasil pengujian yang telah dilakukan, telah didaptakan jumlah iterasi yang optimal untuk tingkat kesejahteraan di Jawa Timur ialah 30 iterasi.

5. KESIMPULAN DAN SARAN

A. Kesimpulan

Penerapan algoritma *K-Means* dengan *Particle Swarm Optimization* untuk menetukan tingkat kesejahteraan pada Provinsi Jawa Timur menghasilkan *cluster* 2 sebagai *cluster* yang optimum.

Penerapan algoritma *K-Means* dengan *Particle Swarm Optimization* untuk menetukan tingkat kesejahteraan pada Provinsi Jawa Timur juga menghasilkan nilai Parameter yang optimum, dimana nilai parameter tertinggi ialah 10 partikel, untuk nilai *Learning Factor* tertinggi ialah [1,8;2,0], dan untuk nilai Bobot Inersia tertinggi ialah 0,8, sedangkan untuk jumlah iterasi maksimal tertinggi ialah 30 iterasi.

Hasil pengelompokan tingkat kesejahteraan pada Provinsi Jawa Timur ialah 27 Kabupaten atau Kota untuk *cluster* 1 dan 11 Kabupaten atau Kota untuk *cluster* 2.

B. Saran

Memilih parameter yang tepat: Kedua algoritma ini memerlukan pengaturan parameter yang tepat untuk memberikan hasil yang optimal. Oleh karena itu, perlu melakukan eksperimen untuk menemukan

parameter yang cocok untuk masalah yang akan diselesaikan.

6. REFERENSI

BPS. (2015). Indikator Kesejahteraan Rakyat 2015. *Cerebrovascular Diseases*, 170–175. https://doi.org/10.1159/000490421

BPS. 2019. Kabupaten Jember Dalam Angka. Jember: Badan Pusat Statistik. URL : BPS. 2019. Provinsi Jawa Timur Dalam Angka. Jawa Timur: Badan Pusat Statistik. URL : https://jatim.bps.go.id/publication/2019/08/16/f668b9b7ca53a7998bc81453/provinsi-jawa-timur-dalam-angka-2019.html

BPS. 2019. Provinsi Jawa Timur Dalam Angka. Jawa Timur: Badan Pusat Statistik. URL : https://jatim.bps.go.id/publication/2019/08/16/f668b9b7ca53a7998bc81453/provinsi-jawa-timur-dalam-angka-2019.html

Haupt, R. L., Haupt, S. E., & Wiley, A. J. (2004). *Algorithms Second Edition*.

Kusrini dan Emha Taufiq. (2015). Proses Data Mining. *Data Mining*, 1–143.

Larose, D. T., & Larose, C. D. (2014). Discovering Knowledge in Data: An Introduction to Data Mining: Second Edition. In *Discovering Knowledge in Data: An Introduction to Data Mining: Second Edition* (Vol. 9780470908). https://doi.org/10.1002/9781118874059

Madhulatha, T. S. (2012). an Overview on Clustering Methods. *IOSR Journal of Engineering*, *02*(04), 719–725. https://doi.org/10.9790/3021-0204719725

Mario, A., Herry, S., & Nasution, H. (2016). Pemilihan Distance Measure Pada K-Means Clustering Untuk Pengelompokkan Member Di Alvaro Fitness. *Jurnal Sistem dan Teknologi Informasi*, *1*(1), 1–6.

Ndakularak, E., Setiawina, N. D., & I Ketut Djayasastra. (2014). *ANALISIS FAKTOR-FAKTOR YANG MEMPENGARUHI*. 140–153.

Patro, S. G. K., & sahu, K. K. (2015). Normalization: A Preprocessing Stage. *Iarjset*, 20–22. https://doi.org/10.17148/iarjset.2015.2305

Rustam, S., Santoso, H. A., & Supriyanto, C. (2018). Optimasi K-Means Clustering Untuk Identifikasi Daerah Endemik Penyakit Menular Dengan Algoritma Particle Swarm Optimization Di Kota Semarang. *ILKOM Jurnal Ilmiah*, *10*(3), 251–259. https://doi.org/10.33096/ilkom.v10i3.342.251-259

Sa'diyah, N. 2020. Implementasi *Algoritma Fuzzy C-Means* Dengan *Particle Swarm Optimization* (FCMPSO) Untuk Pengelompokan Proses Berpikir Siswa Dalam Proses Belajar. Malang: Universitas Brawijaya Malang

Shalabi, L. Al, Shaaban, Z., & Kasasbeh, B. (2006). Data Mining: A Preprocessing Engine. *Journal of Computer Science*, *2*(9), 735–739. https://doi.org/10.3844/jcssp.2006.735.739

Simbolon, T. R. (2017). *Analisa Keterkaitan Ketimpangan Pembangunan Antar Daerah Terhadap Pertumbuhan Ekonomi di Wilayah Sumatera*.

Singh, N., Singh, S., & Singh, S. B. (2012). *Hpso : a New Version of Particle Swarm Optimization Algorithm*. *3*(3), 123–134.

Soemartini, & Supartini, E. (2017). Analisis K-Means Cluster Untuk Pengelompokan Kabupaten / Kota Di Jawabarat Berdasarkan Indikator Masyarakat. *Prosiding*, (Knpmp Ii), 144–154.

Supianto, A. A., & Mar'i, F. (2018). Clustering Credit Card Holder Berdasarkan Pembayaran Tagihan Menggunakan

Improved K-Means Dengan Particle Swarm Clustering Credit Card Holder Based on Billing Payment Using Improved K-Means With Particle Swarm Optimization. *Jurnal Teknologi Informasi dan Ilmu Komputer(JTIIK)*, *5*(6), 737–744. https://doi.org/10.25126/jtiik.20185858

Tan, P.-N. (2006). Introduction to data mining. *Journal of School Psychology*, *19*(1), 51–56. https://doi.org/10.1016/0022-4405(81)90007-8

Tenriawaru, A., & Kunci, K. (2018). *Kabupaten/kota provinsi sulawesi tenggara berdasarkan indikator kesejahteraan rakyat*. *4*(2), 175–182.

Trisno, I. B. (2016). *Coba Python*.

Zulnaidi. (2007). Metodologi Penelitian. *USU Repository © 2007*, 1–20. Diambil dari http://repository.usu.ac.id/bitstream/handle/123456789/1728/07002017.pdf?sequence=1