

ANALISIS PENGELOMPOKAN FASILITAS KESEHATAN BPJS KETENAGAKERJAAN MENGGUNAKAN METODE KLUSTERISASI K-MEANS

¹⁾Tiza Novianti, ²⁾Jeni April, ³⁾Muhammad Faisal

^{1,2}Sistem Informasi, STMIK Profesional Makassar, ³Ilmu Komputer, STMIK Profesional Makassar
e-mail : ¹tizanoviantisucir11194028@gmail.com, ²jenniapriyani1194022@gmail.com, ³muhfaisal@stmikprofesional.ac.id

ABSTRACT

This research aims to be able to assist BPJS Kesehatan and other stakeholders in designing more targeted policies and interventions to improve the accessibility and quality of health services for the entire Indonesian population. Research on klusterisasi BPJS health facilities has a significant impact on improving the health service system in Indonesia, specifically the Makassar region. One of the non-hierarchical grouping procedures is to use the K-Means method. Based on grouping data in the last iteration, it is known that health facility participants are at a moderate level, when viewed the majority of data is gathered in cluster 2. It was concluded that the community has not maximally used assistance from the government in the form of health facility services. It is recommended in future studies to use datasets from primary sources in larger quantities, as well as use hybrid algorithms to obtain better results

Keywords: K-Means, BPJS, Faskes, clustering

ABSTRAK

Penelitian ini bertujuan untuk dapat membantu BPJS Kesehatan dan pemangku kepentingan lainnya dalam merancang kebijakan dan intervensi yang lebih tepat sasaran untuk meningkatkan aksesibilitas dan kualitas layanan kesehatan bagi seluruh penduduk Indonesia. Penelitian tentang klusterisasi fasilitas kesehatan BPJS memiliki dampak yang signifikan dalam meningkatkan sistem layanan kesehatan di Indonesia khusus wilayah makassar. Salah satu prosedur pengelompokkan pada non-hirarki adalah dengan menggunakan metode K-Means. Berdasarkan data pengelompokan pada iterasi terakhir diketahui bahwa peserta fasilitas kesehatan pada tingkat sedang jika dilihat mayoritas data berkumpul pada kluster 2. Disimpulkan bahwa masyarakat belum secara maksimal menggunakan bantuan dari pemerintah dalam bentuk layanan fasilitas kesehatan. disarankan pada penelitian selanjutnya menggunakan dataset dari sumber utama dalam jumlah yang lebih besar, serta menggunakan hybrid algoritma agar dapat diperoleh hasil yang lebih baik.

Kata Kunci: K-Means, BPJS, Faskes, klusterisasi

I. PENDAHULUAN

Program BPJS Kesehatan adalah program jaminan kesehatan nasional yang memberikan akses layanan kesehatan kepada seluruh penduduk Indonesia. Dalam pelaksanaannya, program ini melibatkan berbagai jenis fasilitas kesehatan, termasuk rumah sakit, klinik, puskesmas, dan praktik dokter. Namun, karakteristik dan kualitas pelayanan kesehatan yang disediakan oleh setiap fasilitas dapat bervariasi, sehingga perlu dilakukan penelitian untuk mengklasifikasikan fasilitas-fasilitas ini ke dalam kelompok-kelompok yang serupa berdasarkan karakteristik yang relevan. Terdapat lebih dari seratus fasilitas kesehatan, baik

dokter umum maupun puskesmas yang dapat melayani pasien pengguna BPJS Wilayah Makassar. Fasilitas kesehatan tersebut dapat dipilih oleh setiap peserta BPJS berdasarkan beberapa faktor, salah satunya adalah jenis pelayanan dari fasilitas kesehatan tersebut. Pada setiap layanan fasilitas kesehatan memiliki jumlah peserta yang beragam, antara lain terdapat fasilitas kesehatan yang memiliki banyak peserta terdaftar, namun juga terdapat fasilitas kesehatan yang memiliki sedikit peserta terdaftar karena kurang diminati peserta. Banyaknya peserta di suatu fasilitas kesehatan mengindikasikan adanya kualitas pelayanan yang baik di fasilitas kesehatan tersebut, sedangkan fasilitas kesehatan yang memiliki

sedikit peserta terdaftar mengindikasikan fasilitas kesehatan tersebut memiliki kualitas pelayanan yang kurang baik. adapun jenis layanan fasilitas kesehatan yang akan menjadi indikator adalah Pelayanan Kesehatan Primer(PKP), Pelayanan Kesehatan Spesialis(PKS). Pelayanan Kesehatan Primer menyediakan akses ke pelayanan kesehatan primer, seperti pemeriksaan kesehatan rutin, pengobatan ringan, konsultasi medis dasar, dan vaksinasi. Pelayanan Kesehatan Spesialistik menyediakan akses ke pelayanan kesehatan spesialistik, termasuk pemeriksaan dan pengobatan oleh dokter spesialis[1]. Metode K-Means bertujuan mengelompokkan data yang ada kedalam beberapa kelompok, dimana data dalam satu kelompok mempunyai karakteristik yang sama satu sama lainnya dan mempunyai karakteristik yang berbeda dengan data yang ada di dalam kelompok yang lain[2] disebut teknik data mining. Metode data mining dalam menangani data outlier adalah klusterisasi, dimana klusterisasi mendefinisikan outlier yang tidak ada di cluster[3], data mining menganalisa sejumlah besar kumpulan data observasi, menemukan suatu hubungan tidak terduga serta dapat merangkum data dengan cara yang baru sehingga dapat berguna dan dimengerti bagi pengguna[4], seperti data fasilitas kesehatan yang terdapat pada BPJS.

Penelitian ini dapat memberikan pemahaman yang lebih baik tentang variabilitas pelayanan kesehatan di wilayah makassar, melalui pengetahuan tentang bagaimana fasilitas kesehatan terorganisir dan beroperasi dalam setiap kelompok, serta dapat diketahui apakah terdapat kesenjangan aksesibilitas atau ketidaksesuaian dalam penyediaan layanan kesehatan di wilayah makassar, sehingga hasil penelitian ini dapat digunakan untuk mengarahkan upaya perbaikan dan pengembangan layanan fasilitas kesehatan terhadap peserta BPJS Ketenagakerjaan.

Penelitian terkait studi kasus pada BPJS menyatakan meskipun kualitas pelayanan yang diberikan pada pasien peserta BPJS dan non BPJS kesehatan pada salah satu pusat kesehatan masyarakat di kota Bandung sangat baik bagi peserta BPJS kesehatan dan non BPJS kesehatan, namun untuk kualitas pelayanan berupa kehandalan dan empati atau perhatian masih harus ditingkatkan[5]. Penelitian terdahulu dilakukan untuk menganalisis dan memvisualisasikan data BPJS pada platform *Twitter* dengan menggunakan metode K-Means klusterisasi untuk mengetahui topik yang diperbincangkan[6].

Penelitian tentang klusterisasi fasilitas kesehatan BPJS memiliki dampak yang signifikan dalam meningkatkan sistem layanan kesehatan di Indonesia khusus wilayah makassar. Melalui identifikasi pola fasilitas kesehatan yang serupa, penelitian ini dapat

membantu BPJS Kesehatan dan pemangku kepentingan lainnya dalam merancang kebijakan dan intervensi yang lebih tepat sasaran untuk meningkatkan aksesibilitas dan kualitas layanan kesehatan bagi seluruh penduduk Indonesia.

II. METODE PENELITIAN

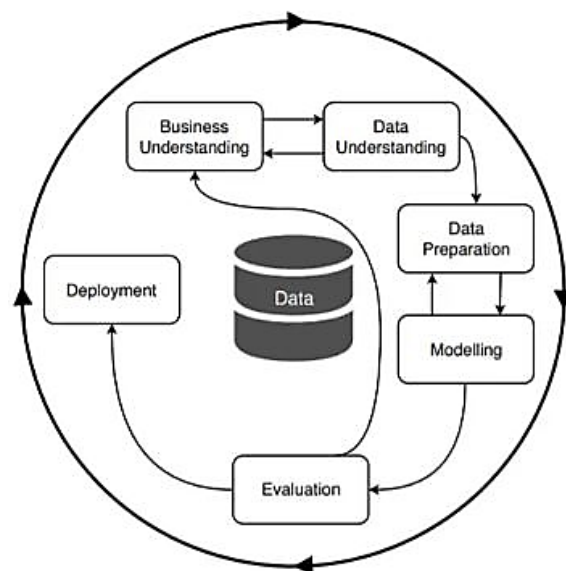
Pada bagian ini menjelaskan prosedur sistematis yang melibatkan langkah-langkah yang terorganisir dan terstruktur untuk memperoleh jawaban terhadap pertanyaan penelitian atau mencapai tujuan penelitian yang telah ditetapkan.

1. Data Mining

Data mining adalah mengekstraksi informasi akurat yang diinginkan dan bermakna dari kumpulan data melalui metode dan algoritma untuk membantu proses pengelompokan data[7]. Data mining dapat diterapkan untuk menemukan pola pengetahuan dari profil pasien dan data riwayat kesehatannya, dimana Pengetahuan yang diperoleh dapat digunakan analisis dan pengambilan keputusan, antara lain untuk memprediksi jenis penyakit, mengetahui pola penyebaran penyakit, dan melihat efektifitas pengobatan. Gambar 1. Menampilkan model data mining

GAMBAR 1.

Data Mining process model of data mining[8]



Pada gambar 1 ditampilkan tahapan pada proses data mining dimana tahapan dalam data mining ini dapat berulang dan iteratif, di mana setiap tahap dapat disesuaikan atau diperbaiki berdasarkan hasil evaluasi dan kebutuhan yang muncul selama proses.

III. HASIL DAN PEMBAHASAN

2. Metode K-Means

Salah satu prosedur pengelompokan pada non-hirarki adalah dengan menggunakan metode K-Means. K-Means merupakan metode pengelompokan yang paling terkenal dan banyak digunakan di berbagai bidang karena sederhana dan mudah diimplementasikan[9]. Metode K-Means merupakan metode pengelompokan yang bertujuan mengelompokkan obyek sedemikian sehingga jarak tiap-tiap obyek ke pusat kelompok di dalam satu kelompok adalah minimum.

Dasar algoritma K-Means adalah sebagai berikut :

- a) Diberikan nilai k sebagai jumlah *cluster* yang ingin dibentuk
- b) Bangkitkan k *centroid* (titik pusat klaster) awal secara *random*
- c) Hitung jarak setiap data ke masing-masing pusat *cluster* yaitu menggunakan *Euclidean Distance* . Jarak Euclidean adalah besarnya jarak suatu garis lurus yang menghubungkan antar objek diukur dengan rumus:

$$d_{(x,y)} = \sqrt{\sum_{j=1}^p (x_j - y_j)^2} \quad (1)$$

- d) Kelompokkan setiap data berdasarkan jarak terdekat antara data dengan pusatnya.
- e) Tentukan posisi pusat *cluster* baru (C_{kj}) dengan cara menghitung nilai rata-rata dari data-data yang ada pada pusat *cluster* yang sama

$$C_{kj} = \frac{x_{1j} + x_{2j} + \dots + x_{aj}}{a}, j = 1, 2, 3, \quad (2)$$

C_{kj} = pusat *cluster* ke-k pada variabel ke- j
 a = banyak data pada *cluster* ke-k

Dalam mendeskripsikan karakteristik masing-masing *cluster*, digunakan persamaan berikut :

$$X = \mu + Z. \sigma \quad (3)$$

Dimana:

- X = rata-rata sampel (variabel dalam klaster)
- μ = rata-rata populasi
- Z = nilai standarisasi
- σ = standar deviasi

Metode K-means merupakan algoritma tradisional yang terdiri dari sejumlah besar data berdimensi tinggi dalam memilih pusat pengelompokan awal dan ketidakpastian nilai K[10].

Penelitian ini menggunakan data sekunder peserta BPJS yang menggunakan layanan Fasilitas kesehatan Primer dan Spesialist pada tahun 2011-2022. Penerapan metode K-Means dijelaskan berdasarkan Tahapan yang telah ditentukan.

1. Penentuan jumlah kluster
 Jumlah kluster yang ditetapkan adalah 3 sebagai representasi jumlah pengguna banyak, sedang, kurang.
2. Penentuan Jumlah dataset
 Jumlah dataset yang ditetapkan sebanyak 12, yang mewakili data tahun 2011 sampai 2022. Dataset ditampilkan Pada tabel 1.

TABEL 1.
 Dataset Peserta BPJS

No	Tahun	PKP	PKS
1	2011	666	402
2	2012	311	182
3	2013	492	259
4	2014	505	289
5	2015	510	281
6	2016	655	464
7	2017	725	387
8	2018	162	964
9	2019	113	706
10	2020	613	329
11	2021	489	290
12	2022	591	311

3. Penetapan *attribute*
 Jumlah atribut yang digunakan sebanyak 2, terdiri dari Pelayanan Kesehatan Primer(PKP) dan Pelayanan Kesehatan Spesialist(PKS).
4. Penentuan pusat kluster
 Pada tahap awal ditentukan pusat kluster pada data ke-8 sebagai pusat cluster ke-1 terdiri dari nilai 162 dan 964. Data ke-7 sebagai pusat cluster ke-2 terdiri dari nilai 725 dan 387. Data ke-2 sebagai pusat cluster ke-3 terdiri dari 311 dan 182.
5. Perhitungan jarak pusat
 Proses perhitungan jarak pusat kluster menggunakan rumus 1. Hasil perhitungan ditampilkan pada tabel 2.

TABEL 2.
Perhitungan Jarak Pusat Kluster

Thn	PKP	PKS	C1	C2	C3	Jarak
2011	666	402	754.89	60.88	417.64	60.88
2012	311	182	796.07	461.98	0.00	0.00
2013	492	259	778.41	265.84	196.70	196.70
2014	505	289	757.15	240.84	221.55	221.55
2015	510	281	766.55	239.71	222.27	222.27
2016	655	464	702.17	104.06	444.81	104.06
2017	725	387	806.16	0.00	461.98	0.00
2018	162	964	0.00	806.16	796.07	0.00
2019	113	706	262.61	690.15	560.16	262.61
2020	613	329	778.86	126.13	335.88	126.13
2021	489	290	749.14	255.16	208.20	208.20
2022	591	311	781.31	154.05	308.29	154.05

Pada tabel 2. Ditampilkan nilai jarak terdekat dari pusat kluster, ditulis dalam contoh:

$$417.6 = \sqrt{((666 - 311)^2) + (402 - 182)^2}$$

Seterusnya dilakukan terhadap setiap dataset.

6. Pengelompokan data.

Pengempokan ditentukan berdasarkan jarak terdekat dari pusat kluster menggunakan rumus 2. Pada iterasi pertama diperoleh hasil pengelompokan seperti pada tabel 3.

TABEL 3.

Hasil Pengelompokan data Iterasi ke-1.

No.	C1	C2	C3
1		1	
2			1
3			1
4		1	
5			1
6		1	
7		1	
8	1		
9	1		
10		1	
11		1	
12		1	

Pada Tabel 2, ditampilkan cell yang berisi nilai 1 merupakan cluster terhadap dataset.

7. Penentuan Kluster Baru

Pada Iterasi kedua dan seterusnya dilakukan pembaruan pusat kluster agar setiap dataset dapat lebih terseleksi nilai kedekatannya dengan pusat kluster menggunakan rumus 3. Data pada pusat

cluster baru pada iterasi kedua ditampilkan pada Tabel 4.

TABEL 4.
Pusat kluster baru

Cluster baru ke-1	137.5	835
Cluster baru ke-2	606.29	353.14
Cluster baru ke-3	437.67	240.67

Nilai pusat kluster baru diperoleh berdasarkan total nilai dataset pada setiap kluater lalu dibagi jumlah dataset terpilih., ditulis pada contoh :

$$135 = (162 + 113) / 2.$$

$$835 = (964 + 706) / 2.$$

Maka diperoleh hasil perhitungan jarak terdekat dari pusat kluster pada tabel 5.

TABEL 5.

Hasil Pengelompokan data Iterasi ke-2

Thn	PKP	PKS	C1	C2	C3	Jarak
2011	666	402	683.23	77.15	279.58	77.15
2012	311	182	675.66	341.30	139.59	139.59
2013	492	259	676.35	148.07	57.34	57.34
2014	505	289	658.16	119.89	82.88	82.88
2015	510	281	667.59	120.31	82.82	82.82
2016	655	464	636.75	121.09	311.63	121.09
2017	725	387	738.82	123.45	322.45	123.45
2018	162	964	131.31	755.34	774.08	131.31
2019	113	706	131.31	606.50	567.40	131.31
2020	613	329	694.36	25.06	196.33	25.06
2021	489	290	648.52	133.20	71.20	71.20
2022	591	311	692.99	44.83	168.69	44.83

Berdasarkan data pada tabel 5, maka diperoleh hasil pengelompokan baru yang ditampilkan pada tabel 6.

TABEL 6.

Hasil Pengelompokan data Iterasi ke-2

No.	C1	C2	C3
1		1	
2			1
3			1
4			1
5			1
6		1	
7		1	
8	1		

9	1		
10		1	
11			1
12		1	

Proses iterasi dilakukan hingga hasil kluster tidak berubah. Pada penelitian ini proses iterasi hanya dilakukan sebanyak 3 kali, yang menghasilkan pengelompokan terbaik, dimana hasil ditampilkan pada tabel 7.

TABEL 7.

Hasil jarak pusat kluster pada Iterasi ke-3

Thn	PKP	PKS	C1	C2	C3	Jarak
2011	666	402	683.23	77.15	279.58	77.15
2012	311	182	675.66	341.30	139.59	139.59
2013	492	259	676.35	148.07	57.34	57.34
2014	505	289	658.16	119.89	82.88	82.88
2015	510	281	667.59	120.31	82.82	82.82
2016	655	464	636.75	121.09	311.63	121.09
2017	725	387	738.82	123.45	322.45	123.45
2018	162	964	131.31	755.34	774.08	131.31
2019	113	706	131.31	606.50	567.40	131.31
2020	613	329	694.36	25.06	196.33	25.06
2021	489	290	648.52	133.20	71.20	71.20
2022	591	311	692.99	44.83	168.69	44.83

Berdasarkan Perhitungan jarak terdekat dengan pusat kluster pada iterasi ke-3 maka diperoleh hasil pengelompokan yang ditampilkan pada Tabel 8.

TABEL 8.

Hasil Pengelompokan data Iterasi ke- 3

No.	C1	C2	C3
1		1	
2		1	
3		1	
4		1	
5		1	
6		1	
7		1	
8	1		
9			1
10		1	
11		1	
12		1	

Berdasarkan data pengelompokan pada iterasi terakhir diketahui bahwa peserta fasilitas kesehatan pada tingkat sedang jika dilihat mayoritas data berkumpul pada kluster 2.

IV. KESIMPULAN

Kesimpulan yang diperoleh pada penelitian ini adalah banyaknya peserta pengguna fasilitas kesehatan PKP dan PKS berada pada status sedang, sehingga dapat diketahui bahwa masyarakat belum secara maksimal menggunakan bantuan dari pemerintah dalam bentuk layanan fasilitas kesehatan..

V. SARAN

Saran-saran pada penelitian selanjutnya adalah menggunakan dataset dari sumber utama dalam jumlah yang lebih besar, serta menggunakan hybrid algoritma agar dapat diperoleh hasil yang lebih baik.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada STMIK Profesional yang telah memberikan dukungan selama peelitian ini dilaksanakan.

DAFTAR PUSTAKA

- [1] I. Thobibah, E. Ernawaty, and N. A. Damayanti, "Analisis Faktor Penentu Kepesertaan Bpjs," *J. Keperawatan Silampari*, vol. 4, no. 1, pp. 162–170, Dec. 2020, doi: 10.31539/jks.v4i1.1474.
- [2] U. N. Putra *et al.*, "Penerapan Data Mining K-Means Klusterisasi Untuk Mengelompokkan Berbagai Jenis Merk Smartphone," pp. 241–249, 2021.
- [3] W. P. Nurmawati, D. J. Ratnaningsih, S. Nisrina, A. Rahim, M. Malthuf, and W. Kusuma, "Clustrering of BPJS National Health Insurance Participant Using DBSCAN Algorithm," *J. Varian*, vol. 6, no. 1, pp. 25–34, Nov. 2022, doi: 10.30812/varian.v6i1.1886.
- [4] G. Gustientiedina, M. H. Adiya, and Y. Desnelita, "Penerapan Algoritma K-Means Untuk Klusterisasi Data Obat-Obatan," *J. Nas. Teknol. dan Sist. Inf.*, vol. 5, no. 1, pp. 17–24, Apr. 2019, doi: 10.25077/TEKNOSI.v5i1.2019.17-24.
- [5] R. Ekawati and N. Nurhalimah, "Kualitas Pelayanan Bagi Peserta BPJS Kesehatan dan Non BPJS Kesehatan," *Pros. FRIMA (Festival Ris. Ilm. Manaj. dan Akuntansi)*, no. 3, pp. 483–494, Feb. 2022, doi:

- 10.55916/frima.v0i3.317.
- [6] A. B. Saputra, P. W. Cahyo, M. Habibi, and A. Priadana, “Analysis and visualization of BPJS on twitter using K-means klusterisasi,” *Int. J. Heal. Sci. Technol.*, vol. 3, no. 3, pp. 109–117, 2022, doi: 10.31101/ijhst.v3i3.2466.
- [7] S. R. Durugkar, R. Raja, K. K. Nagwanshi, and S. Kumar, “Introduction to data mining,” *Data Min. Mach. Learn. Appl.*, pp. 1–19, 2022, doi: 10.1002/9781119792529.ch1.
- [8] F. Martinez-Plumed *et al.*, “CRISP-DM Twenty Years Later: From Data Mining Processes to Data Science Trajectories,” *IEEE Trans. Knowl. Data Eng.*, vol. 33, no. 8, pp. 3048–3061, Aug. 2021, doi: 10.1109/TKDE.2019.2962680.
- [9] X. Chu, J. Lei, X. Liu, and Z. Wang, “KMEANS Algorithm Klusterisasi for Massive AIS Data Based on the Spark Platform,” in *2020 5th International Conference on Control, Robotics and Cybernetics (CRC)*, Oct. 2020, pp. 36–39, doi: 10.1109/CRC51253.2020.9253451.
- [10] H. Zhao, “Research on Improvement and Parallelization of K-Means Klusterisasi Algorithm,” in *2021 IEEE 3rd International Conference on Frontiers Technology of Information and Computer (ICFTIC)*, Nov. 2021, pp. 57–61, doi: 10.1109/ICFTIC54370.2021.9647385.