

PENGELOMPOKAN WILAYAH MENURUT POTENSI FASILITAS KESEHATAN DAN KEJADIAN COVID-19 MENGGUNAKAN ALGORITMA FUZZY K-PROTOTYPES

Hermawan Prasetyo

Badan Pusat Statistik Provinsi Jawa Tengah
email: herpas@bps.go.id

Abstrak

COVID-19 telah masuk ke Indonesia dan menyebar ke seluruh wilayah Indonesia hingga saat ini. Penanganan COVID-19 tidak lepas dari peran fasilitas kesehatan dalam pengendalian COVID-19. Fasilitas kesehatan merupakan potensi suatu wilayah, yang dapat diperoleh datanya melalui PODES. Data terkait fasilitas kesehatan pada suatu wilayah dapat dianalisis dengan metode clustering. Metode ini dapat dilakukan pada tipe data numerik, kategorik, dan campuran. Pada penelitian ini dilakukan pengelompokan wilayah desa di Kota Semarang menjadi dua kluster dengan algoritma fuzzy k-prototypes. Dari hasil pengelompokan dapat dilihat bahwa kluster 2 memiliki rata-rata jumlah fasilitas kesehatan, persentase kejadian COVID-19, dan rata-rata jumlah penderita COVID-19 yang lebih tinggi dibandingkan kluster 1. Jumlah fasilitas kesehatan juga berbanding lurus dengan jumlah kejadian dan penderita COVID-19. Akan tetapi, jumlah fasilitas kesehatan yang ada mengalami ketimpangan jika dibandingkan jumlah kejadian dan penderita COVID-19. Hal ini dapat dijadikan masukan bagi Pemerintah Kota Semarang dalam menangani COVID-19 pada wilayah yang mengalami ketimpangan jumlah fasilitas kesehatan.

Keywords: COVID-19, fasilitas kesehatan, clustering, potensi desa

1. PENDAHULUAN

Pada awal tahun 2020 *coronavirus* jenis baru (SARS-CoV-2) dengan penyakitnya yang disebut *Coronavirus disease 2019* (COVID-19)[1]. Virus ini berawal dari laporan dari Tiongkok kepada *World Health Organization* (WHO), yaitu terdapat 44 pasien pneumonia berat di Wuhan pada akhir tahun 2019[2]. COVID-19 telah masuk ke Indonesia pertama kali dilaporkan tanggal 2 Maret 2020[3]. Virus ini kemudian menyebar sampai ke seluruh wilayah Indonesia hingga mencapai 2.615.529 orang yang terkonfirmasi COVID-19 pada 14 Juli 2021[4].

Di tengah pandemi COVID-19, peran fasilitas kesehatan tingkat pertama sangat penting dalam melakukan prevalensi, deteksi, dan respon dalam pencegahan dan pengendalian COVID-19[5]. Pelayanan primer di bidang kesehatan seperti puskesmas, klinik pratama, dan dokter praktik mandiri sebagai ujung tombak pelayanan kesehatan harus bergerak membantu memotong

mata rantai penularan[6]. Selain itu, apotek juga berperan penting dalam masa pandemi. Apoteker dalam apotek memberikan layanan kefarmasian untuk mengurangi beban pasien di fasilitas kesehatan seperti rumah sakit dan praktik dokter umum[6].

Fasilitas kesehatan merupakan salah satu potensi wilayah, yang tersebar di seluruh wilayah Indonesia. Badan Pusat Statistik (BPS) mengumpulkan data potensi wilayah melalui Pendataan Potensi Desa (PODES)[7]. Data yang dikumpulkan dalam PODES termasuk data tentang kesehatan, antara lain fasilitas kesehatan dan kejadian wabah, termasuk COVID-19. Tipe data yang ada pada data PODES berupa data numerik dan kategorik.

Analisis potensi wilayah dapat dilakukan dengan *clustering*. *Clustering* atau pengelompokan merupakan teknik dalam *data mining* yang menghasilkan kelompok-kelompok objek dengan objek-objek di dalam kluster sangat mirip dan objek-objek di luar kluster tidak

mirip[8]. *Clustering* dapat dilakukan pada tipe data numerik, kategorik, dan campuran keduanya. Salah satu metode *clustering* untuk data beratribut campuran adalah algoritma *k-prototypes*[9]. Dalam perkembangannya, algoritma ini dikembangkan menjadi algoritma *fuzzy k-prototypes* (FKP)[10].

Penelitian ini bertujuan untuk pengelompokan potensi kesehatan wilayah dengan metode FKP *clustering*. Variabel yang digunakan untuk pengelompokan yaitu jumlah fasilitas kesehatan, adanya kejadian COVID-19, dan jumlah penderita COVID-19. Data diambil dari PODES tahun 2020 yang menyajikan data potensi wilayah hingga tingkat desa. Adapun wilayah dalam penelitian ini adalah wilayah desa di Kota Semarang. Hasil penelitian ini diharapkan dapat menjadi bahan masukan bagi Pemerintah Kota Semarang dalam melihat potensi kesehatan dan kejadian COVID-19 suatu wilayah sehingga mempermudah dalam penanganan COVID-19.

2. METODE PENELITIAN

Pada penelitian ini dilakukan pengelompokan data potensi fasilitas kesehatan wilayah dan kejadian COVID-19 dengan menggunakan algoritma FKP. Untuk mendapatkan hasil pengelompokan dilakukan beberapa tahap, yaitu penyiapan data, data *preprocessing*, pemrosesan data, dan evaluasi.

a. Penyiapan Data

Data potensi fasilitas kesehatan dapat diambil dari hasil Pendataan Potensi Desa (PODES)[7]. *Dataset* yang digunakan dalam penelitian ini adalah data potensi fasilitas kesehatan dan kejadian COVID-19 yang diperoleh dari hasil PODES tahun 2020 (PODES2020-DESA). Adapun wilayah yang dipilih adalah Kota Semarang. *Dataset* terdiri dari 177 *record*, tiap *record* mewakili satu desa.

b. Data Preprocessing

Tahap selanjutnya yang dilakukan adalah tahap *data preprocessing*. Pada tahap ini dilakukan pemilihan atribut yang digunakan untuk pengelompokan data. Atribut dipilih dari dataset PODES2020-DESA pada Blok VI Pendidikan dan Kesehatan. Atribut terpilih terdiri

dari atribut dengan tipe data numerik dan kategorik.

Tabel 1. Atribut Terpilih pada *Dataset* PODES2020-DESA

No	Rincian	Deskripsi	Tipe Data
1	R603a	Jumlah rumah sakit	Numerik
2	R603c	Jumlah puskesmas dengan rawat inap	Numerik
3	R603d	Jumlah puskesmas tanpa rawat inap	Numerik
4	R603e	Jumlah puskesmas pembantu	Numerik
5	R603f	Jumlah poliklinik/balai pengobatan	Numerik
6	R603g	Jumlah tempat praktik dokter	Numerik
7	R603l	Jumlah apotek	Numerik
8	R605hK2	Kejadian COVID-19	Kategorik
9	R605hK3	Jumlah penderita COVID-19	Numerik

Pada tahap ini juga dilakukan *data preprocessing* berupa agregasi nilai atribut. Tujuan agregasi adalah agar mendapatkan satu atribut baru dari beberapa atribut yang ada. Agregasi dilakukan pada atribut R603a, R603c, R603d, R603e, R603f, R603g, dan R603l menjadi atribut baru R603, yaitu jumlah fasilitas kesehatan. Adapun hasil agregasi dapat dilihat pada Tabel 2.

Tabel 2. Hasil Agregasi Atribut

No	Rincian	Deskripsi	Tipe Data
1	R603	Jumlah fasilitas kesehatan	Numerik
2	R605hK2	Kejadian COVID-19	Kategorik
3	R605hK3	Jumlah penderita COVID-19	Numerik

c. Pemrosesan Data

Dari *dataset* yang telah siap kemudian dilakukan pemrosesan *clustering* data menggunakan algoritma FKP[10]. Algoritma FKP merupakan pengembangan dari algoritma *k-prototypes* yang diusulkan oleh Huang [9]. Misalkan X adalah suatu himpunan objek X_1, X_2, \dots, X_n dengan atribut numerik A_1, A_2, \dots, A_p dan atribut kategorik A_1, A_2, \dots, A_q serta k adalah jumlah kluster. Algoritma FKP dilakukan dengan meminimalkan *objective function* $F(W, Z)$, yang didefinisikan pada Persamaan 1.

$$F(W, Z) = \sum_{l=1}^k \sum_{i=1}^n w_{li}^\alpha d(X_i, Z_l) \quad (1)$$

dengan $\alpha \geq 1$ adalah suatu parameter *fuzzy*, $W = \{w_{li}\}$ adalah suatu matriks keanggotaan berukuran $k \times n$, dan $Z = \{Z_1, Z_2, \dots, Z_k\}$ adalah suatu himpunan k prototipe dengan m dimensi.

Algoritma FKP didefinisikan sebagai[10]:

- (1) Pilih satu himpunan prototipe awal $Z^{(l)}$ dan satu pengontrol iterasi ϵ .
- (2) Tentukan $W^{(l)}$ yang meminimalkan $F(W, Z^{(l)})$. Set $i = 1$.
- (3) Tentukan $Z^{(i+1)}$ yang meminimalkan $F(W^{(i)}, Z)$. Jika $|F(W^{(i)}, Z^{(i+1)}) - F(W^{(i)}, Z^{(i)})| < \epsilon$, maka berhenti. Jika tidak, maka set $(W^*, Z^*) = (W^{(i)}, Z^{(i+1)})$.
- (4) Tentukan $W^{(i+1)}$ yang meminimalkan $F(W, Z^{(i+1)})$. Jika $|F(W^{(i+1)}, Z^{(i+1)}) - F(W^{(i)}, Z^{(i+1)})| < \epsilon$, maka berhenti. Jika tidak, maka set $(W^*, Z^*) = (W^{(i+1)}, Z^{(i+1)})$. Tambahkan i dengan 1 dan kembali ke langkah (3).

d. Evaluasi

Setelah dilakukan pemrosesan data dengan algoritma FKP, langkah selanjutnya adalah evaluasi hasil pengelompokan. Metode evaluasi dalam penelitian ini adalah dengan indeks *partition coefficient* (PC) dan indeks *partition entropy* (PE)[8].

Indeks PC dapat dihitung dengan Persamaan 2 sebagai berikut,

$$V_{PC} = \frac{1}{n} \sum_{l=1}^c \sum_{i=1}^n w_{li}^2$$

dengan $W = (w_{li})$ ($1 \leq l \leq c, 1 \leq i \leq n$) adalah matriks keanggotaan dari c partisi *fuzzy* dari *dataset* D dengan n objek. Nilai indeks PC yang

semakin mendekati 1 menandakan bahwa hasil pengelompokan semakin tepat.

Indeks PE dapat dihitung dengan Persamaan 3 sebagai berikut,

$$V_{PE} = -\frac{1}{n} \sum_{l=1}^c \sum_{i=1}^n w_{li} \log_a(w_{li})$$

dengan a adalah bilangan basis logaritma dan $W = (w_{li})$ adalah matriks keanggotaan dari c partisi *fuzzy*. Nilai indeks PE yang semakin mendekati 0 menandakan bahwa hasil pengelompokan semakin tepat.

3. HASIL DAN PEMBAHASAN

Clustering dilakukan pada data PODES 2020 pada beberapa jumlah kluster. Nilai koefisien *fuzziness* (α) yang dipilih sebesar 2, yaitu nilai yang paling banyak digunakan pada *fuzzy clustering*[11]. Data diolah menggunakan *software* MATLAB. Hasil indeks PC dan PE berdasarkan beberapa percobaan dapat dilihat pada Tabel 3.

Tabel 3. Nilai indeks PC dan PE

Jumlah kluster	Indeks PC	Indeks PE
2	0,6805	0,4886
3	0,5081	0,8409
4	0,4201	1,0821
5	0,3580	1,2788

Dari hasil indeks PC dan PE pada Tabel 3 dapat dilihat bahwa hasil *clustering* terbaik terjadi pada saat jumlah kluster sebanyak dua. Pada jumlah kluster dua, nilai indeks PC paling tinggi dibandingkan jumlah kluster yang lain. Demikian juga dengan nilai indeks PE pada jumlah kluster dua paling rendah di antara jumlah kluster lainnya.

Clustering data menjadi dua kluster menghasikan nilai *centroid* seperti pada Tabel 4.

Tabel 4. Nilai *centroid* hasil *clustering*

Variabel	Kluster	Kluster
	1	(2) 2
Jumlah fasilitas kesehatan	4,6533	9,0202
Kejadian COVID-19	Ada	Ada
Jumlah penderita COVID-19	2,9871	4,4027

Dari Tabel 4 dapat dilihat bahwa berdasarkan nilai *centroid*, terdapat perbedaan antara klaster 1 dan 2 pada variabel jumlah fasilitas kesehatan dan jumlah penderita COVID-19. Pada klaster 1, nilai *centroid* jumlah fasilitas kesehatan sebesar 4,6533 sedangkan pada klaster 2 sebesar 9,0202. Untuk variabel jumlah penderita COVID-19, nilai *centroid* klaster 1 sebesar 2,9871 dan klaster 2 sebesar 4,4027. Jika dilihat dari variabel kejadian COVID-19, kedua klaster memiliki nilai *centroid* yang sama yaitu ada kejadian COVID-19.

Selanjutnya, dari hasil pengelompokan dilakukan analisis deskriptif pada kedua klaster yang dapat dilihat pada Tabel 5.

Tabel 5. Statistik variabel tiap klaster

Statistik	Klaster 1	Klaster 2
Jumlah desa	109	68
Rata-rata jumlah fasilitas kesehatan	3,17	11,57
Persentase kejadian COVID-19	72,22%	97,01%
Rata-rata jumlah penderita COVID-19	2,31	5,54

Dari Tabel 5 dapat dilihat bahwa wilayah desa yang termasuk dalam klaster 1 sejumlah 109 desa memiliki rata-rata jumlah fasilitas kesehatan sebanyak 3,17. Persentase desa dengan kejadian COVID-19 pada klaster ini sebesar 72,22 persen dengan rata-rata jumlah penderita COVID-19 sebanyak 2,31 orang. Sementara itu, sebanyak 68 desa termasuk dalam klaster 2 dan memiliki rata-rata jumlah fasilitas kesehatan sebanyak 11,57. Pada klaster ini, persentase desa dengan kejadian COVID-19 sebanyak 97,01 persen dengan rata-rata jumlah penderita COVID-19 sebanyak 5,54 orang.

Berdasarkan analisis di atas bahwa rata-rata jumlah fasilitas kesehatan dan penderita COVID-19 pada desa klaster 1 lebih sedikit daripada klaster 2. Selain itu, persentase kejadian COVID-19 pada klaster 1 juga lebih kecil daripada desa-desa klaster 2. Akan tetapi, jika melihat perbandingan jumlah fasilitas kesehatan terhadap

persentase kejadian COVID-19 dan jumlah penderita COVID-19, jumlah fasilitas kesehatan mengalami ketimpangan. Jumlah fasilitas kesehatan pada klaster 2 lebih banyak 3,65 kali lipat dibandingkan dengan klaster 1. Sementara itu, perbandingan kejadian COVID-19 pada klaster 2 adalah 1,34 kali lebih banyak daripada klaster 1 dan perbandingan jumlah penderita COVID-19 pada klaster 2 sebanyak 2,39 kali lipat dibandingkan klaster 1.

4. KESIMPULAN

Dari hasil pengelompokan desa menurut potensi fasilitas kesehatan dan kejadian COVID-19 di Kota Semarang menjadi dua klaster dapat disimpulkan bahwa jumlah kejadian dan penderita COVID-19 berbanding lurus dengan jumlah fasilitas kesehatan yang ada. Akan tetapi, terdapat ketimpangan jumlah fasilitas kesehatan dibandingkan dengan jumlah kejadian dan penderita COVID-19. Hal ini dapat dijadikan masukan kepada Pemerintah Kota Semarang sebagai target dalam penanganan COVID-19 pada wilayah dengan kejadian dan penderita COVID-19 yang cukup tinggi tetapi memiliki jumlah fasilitas kesehatan yang kurang.

5. REFERENSI

- [1] H. A. Diah Handayani, Dwi Rendra Hadi, Fathiyah Isbaniah, Erlina Burhan, “Penyakit Virus Corona 2019,” *Jurnal Respiriologi Indonesia*, vol. 40, no. 2, pp. 119–129, 2020.
- [2] Yuliana, “Corona virus diseases (Covid-19): Sebuah tinjauan literatur,” *Wellness And Healthy Magazine*, vol. 2, no. 1, pp. 187–192, 2020, doi: 10.30604/well.95212020.
- [3] A. Susilo *et al.*, “Coronavirus Disease 2019: Tinjauan Literatur Terkini,” *Jurnal Penyakit Dalam Indonesia*, vol. 7, no. 1, p. 45, 2020, doi: 10.7454/jpdi.v7i1.415.
- [4] Komite Penanganan Covid-19 dan Pemulihan Ekonomi Nasional, “Peta Sebaran COVID-19,” 2021. <https://covid19.go.id/peta-sebaran-covid19>.
- [5] S. Pangoempia, E. Grace, and A. Adisti, “Analisis Pengaruh Pandemi Covid-19 Terhadap Pelayanan Kesehatan Di Puskesmas Ranotana Weru Dan Puskesmas Teling Atas Kota Manado,” *Jurnal KESMAS*, vol. 10, no. 1, pp. 40–49, 2021.

- [6] H. Hadiyanto, "Peran Dokter Di Layanan Primer Pada Era Pandemi Covid-19," *Jurnal Kedokteran dan Kesehatan: Publikasi Ilmiah Fakultas Kedokteran Universitas Sriwijaya*, vol. 7, no. 3, pp. 15–24, 2020, doi: 10.32539/jkk.v7i3.11572.
- [7] Badan Pusat Statistik Provinsi Jawa Tengah, *Statistik Potensi Desa Provinsi Jawa Tengah 2018*. Semarang: Badan Pusat Statistik Provinsi Jawa Tengah, 2018.
- [8] G. Gan, C. Ma, and J. Wu, *Data Clustering: Theory, Algorithms, and Applications, Second Edition*. 2020.
- [9] Z. Huang, "Clustering large data sets with mixed numeric and categorical values," *Proceedings of the 1st Pacific-Asia Conference on Knowledge Discovery and Data Mining,(PAKDD)*, pp. 21–34, 1997, [Online]. Available: http://reference.kfupm.edu.sa/content/c/1/clustering_large_data_sets_with_mixed_nu_362883.pdf.
- [10] N. Chen, A. Chen, and L. X. Zhou, "Fuzzy K-prototypes algorithm for clustering mixed numeric and categorical valued data," *Ruan Jian Xue Bao/Journal of Software*, vol. 12, no. 8, pp. 1107–1119, 2001.
- [11] K. Zhou and S. Yang, "Fuzzifier Selection in Fuzzy C-Means from Cluster Size Distribution Perspective," *Informatica*, vol. 30, no. 3, pp. 613–628, 2019, doi: 10.15388/informatica.2019.221.