

Pengelompokan Kabupaten/Kota di Jawa Timur Berdasarkan Indikator Kesehatan Masyarakat dengan Pendekatan Metode *Ensemble* ROCK

Yovita Liana Salsabila dan Vita Ratnasari

Departemen Statistika, Institut Teknologi Sepuluh Nopember (ITS)

e-mail: vita_ratna@statistika.its.ac.id

Abstrak—Kesehatan menjadi salah satu faktor yang sangat penting dalam menghasilkan sumber daya manusia yang berkualitas, karena kesehatan merupakan modal awal untuk meningkatkan taraf hidup. Indeks Pembangunan Kesehatan Masyarakat (IPKM) menjadi salah satu alat untuk memonitor keberhasilan pembangunan Kesehatan masyarakat melalui penentuan peringkat provinsi dan kabupaten/kota. Provinsi Jawa Timur menempati peringkat ke-6 secara nasional, namun masih adanya kesenjangan peringkat IPKM antar kabupaten/kota yang mengindikasikan terdapat ketidakmerataan kualitas kesehatan di Jawa Timur. Ketidakmerataan ini dapat diatasi dengan salah satu metode statistik yang dapat membantu dalam pengambilan kebijakan yang tepat dalam memotret profil kesehatan masyarakat dengan tipe data campuran yaitu metode ensemble ROCK. Melalui penelitian ini dihasilkan bahwa kabupaten/kota di Jawa Timur terbagi menjadi tiga kelompok kualitas kesehatan yaitu kualitas kesehatan yang baik, sedang, dan buruk. Adapun variabel yang membedakan ketiga kelompok tersebut adalah jumlah dokter umum, jumlah bidan, jumlah perawat, persentase RT terhadap akses air minum layak, persentase RT terhadap akses sanitasi layak, persentase daerah UCI, penolong persalinan, jaminan kesehatan, dan IPM.

Kata Kunci—*Agglomerative Hierarchical Clustering* (AHC), *Ensemble* ROCK, Jawa Timur, Kesehatan Masyarakat, ROCK.

I. PENDAHULUAN

KESEHATAN termasuk kedalam tujuan ketiga dari 17 Agenda pada Tujuan Pembangunan Berkelanjutan (TPB). Tujuan ketiga tersebut adalah menjamin kehidupan yang sehat serta mendorong kesejahteraan hidup untuk seluruh masyarakat di segala umur, mulai dari bayi hingga penduduk usia lanjut. Indonesia sebagai salah satu negara dari 193 negara yang mengesahkan Agenda Pembangunan Berkelanjutan tersebut sehingga, pemerintah Indonesia berkewajiban untuk mewujudkan pembangunan kesehatan untuk seluruh penduduk Indonesia tanpa terkecuali. Salah satu indikator yang mampu mengetahui capaian pembangunan kesehatan yang menggambarkan kondisi kesehatan masyarakat sampai tingkat kabupaten/kota adalah Indeks Pembangunan Kesehatan Masyarakat (IPKM) [1]. IPKM menjadi salah satu alat monitor keberhasilan pembangunan kesehatan masyarakat melalui penentuan peringkat provinsi dan kabupaten/kota. Jawa Timur menempati peringkat ke-6 nilai IPKM tahun 2018. Meskipun sudah termasuk kedalam 10 besar kondisi kesehatan masyarakat terbaik di Indonesia, masih adanya disparitas terhadap peringkat kabupaten/kota di Jawa Timur yang menandakan masih adanya kesenjangan kondisi kesehatan antar masyarakat ditiap kabupaten/kota. Disparitas peringkat tersebut menunjukkan bahwa kualitas kesehatan masyarakat

ditiap kabupaten/kota di Jawa Timur belum merata dalam artian kondisi kesehatan antar masyarakat di Jawa Timur masih ada yang perlu ditingkatkan di beberapa daerah [2].

Demi mengatasi kesenjangan kondisi kesehatan masyarakat di Jawa Timur, diperlukan pemetaan terhadap profil kabupaten/kota berdasarkan kondisi kesehatan masyarakatnya. Pemetaan kabupaten/kota dapat dilakukan melalui salah satu metode statistika yaitu analisis pengelompokan. Metode pengelompokan yang biasa digunakan untuk tipe data campuran adalah metode *Ensemble* ROCK [3]. Metode ini memiliki *robustness* yang baik sehingga dapat mengelompokkan data secara efektif dan menunjukkan sifat skalabilitas yang baik [4]. Adapun penelitian yang pernah dilakukan terhadap kondisi kesehatan dilakukan oleh Vienza dan Ratnasari (2014) tentang faktor yang mempengaruhi indeks kesehatan kabupaten dan kota di Jawa Timur menghasilkan bahwa faktor yang paling berpengaruh adalah persentase pertolong persalinan dan persentase imunisasi bayi [5]. Adapun analisis pengelompokan yang dilakukan oleh Khotimah, dkk tentang pengelompokan kelurahan di Kota Surabaya berdasarkan kriteria pembentukan kampung keluarga berencana. Penelitian ini menggunakan metode pengelompokan hirarki dan dihasilkan lima kelompok optimum [6]. Kemudian penelitian yang menggunakan metode *ensemble* ROCK pernah dilakukan oleh Ichsan (2018) menghasilkan 2 *final cluster* dibedakan menjadi kelompok yang memiliki kondisi pembangunan SDM dan ekonomi yang tinggi dan rendah [7]. Melalui hal-hal tersebut, maka akan dilakukan penelitian terhadap kesehatan masyarakat dengan pendekatan metode *ensemble* ROCK. Hasil dari penelitian ini diharapkan dapat membantu menetapkan kebijakan yang tepat dalam pembangunan kesehatan di Provinsi Jawa Timur.

II. TINJAUAN PUSTAKA

A. Statistika Deskriptif

Statistika deskriptif adalah metode-metode yang berkaitan dengan pengumpulan dan penyajian suatu gugus data sehingga memberikan informasi yang berguna. Statistika deskriptif dapat disajikan dalam bentuk tabel, gambar, maupun grafik yang dapat digunakan untuk menjelaskan karakteristik suatu data. Penyajian statistika deskriptif dapat dilakukan secara numerik melalui nilai tertentu ataupun secara visual [8].

B. Analisis Multivariat

Multivariat berasal dari kata multi yang artinya banyak dan variat yang berarti variabel dan antar variabel saling berhubungan. Melalui definisi tersebut maka dapat dikatakan

bahwa analisis multivariat merupakan analisis yang berkaitan dengan jumlah variabel yang banyak atau lebih dari satu yang dianalisis secara simultan pada masing-masing pengamatan dan terdapat hubungan antar variabelnya [9]. Salah satu metode yang digunakan dalam analisis multivariat adalah analisis pengelompokan.

C. Analisis Pengelompokan

Analisis kelompok (cluster analysis) adalah metode dalam analisis multivariat untuk mengelompokkan n pengamatan kedalam C kelompok ($C \leq n$) berdasarkan karakteristiknya. Ukuran kemiripan dan ketidakmiripan merupakan hal yang sangat mendasar dalam analisis pengelompokan [9].

1) Pengelompokan Data Numerik

Metode pengelompokan data numerik didasarkan pada ukuran ketidakmiripan atau jarak. Ukuran ketidakmiripan yang biasa digunakan adalah jarak Euclidean yang ditampilkan pada persamaan (1).

$$d_{ij} = \sqrt{(x_i - x_j)^T(x_i - x_j)}, i, j = 1, 2, \dots, \text{ dan } i \neq j \quad (1)$$

dengan,

$$x_i^T = [x_{i1}, x_{i2}, \dots, x_{im}], x_j^T = [x_{j1}, x_{j2}, \dots, x_{jm}]$$

Teknik pengelompokan yang digunakan dalam penelitian ini adalah metode hirarki atau disebut juga Agglomerative Hierarchical Clustering (AHC). Berikut merupakan beberapa metode dalam pengelompokan menggunakan AHC.

a) Single Linkage (Pautan Tunggal)

Single linkage merupakan pengelompokan yang didasarkan pada jarak terdekat atau kesamaan yang banyak dengan perumusan pada persamaan (2).

$$d_{(UV)W} = \min\{d_{UV}, d_{VW}\} \quad (2)$$

b) Complete Linkage (Pautan Lengkap)

Complete linkage merupakan pengelompokan yang didasarkan pada jarak terjauh atau kesamaan yang sedikit dengan perumusan pada persamaan (3).

$$d_{(UV)W} = \max\{d_{UV}, d_{VW}\} \quad (3)$$

c) Average Linkage (Pautan Rataan)

Average linkage adalah metode dimana kelompok dibentuk berdasarkan nilai rata-rata seluruh individu pada kelompok lain dengan perumusan pada persamaan (4).

$$d_{(UV)W} = \frac{1}{N_{UVNW}} \sum_q \sum_r d_{qr} \quad (4)$$

2) Pengelompokan Data Kategorik

Pengelompokan data kategorik dilakukan menggunakan ukuran kemiripan atau jarak untuk data kategorik yang dapat dilakukan pengelompokan menggunakan metode ROCK (Robust Clustering using linKs). Pengelompokan dengan metode ROCK dilakukan dengan melakukan langkah berikut ini [10].

a) Menghitung Similaritas

Ukuran kemiripan antara pasangan pengamatan ke-i dan ke-j dihitung dengan rumus yang ditampilkan pada persamaan (5).

$$sim(X_i, X_j) = \frac{|X_i \cap X_j|}{|X_i \cup X_j|}, i \neq j \quad (5)$$

dimana,

$i : 1, 2, \dots, n$ dan $j = 1, 2, \dots, n$

X_i : Himpunan pengamatan ke-i dengan $X_i = \{x_{1i}, x_{2i}, \dots, x_{m_{kategorij}}\}$

X_j : Himpunan pengamatan ke-j dengan $X_j = \{x_{1j}, x_{2j}, \dots, x_{m_{kategorij}}\}$

b) Menentukan Tetangga (Neighbors)

Pengamatan X_i dan X_j dinyatakan sebagai tetangga jika nilai $sim(X_i, X_j) \geq \theta$. Nilai threshold θ yang digunakan berkisar antara 0 sampai 1 dengan menyesuaikan data yang ada.

c) Menghitung Link

Cara untuk menghitung link untuk semua kemungkinan pasangan dari n objek dapat menggunakan matriks A. Matriks A merupakan matriks berukuran $n \times n$ yang bernilai 1 jika X_i dan X_j dinyatakan tetangga dan bernilai 0 jika X_i dan X_j dinyatakan bukan tetangga.

d) Menentukan Local Heap (Goodness Measure)

Local heap yaitu nilai goodness measure untuk setiap kelompok dengan kelompok lainnya jika link $\neq 0$. Goodness measure dapat dihitung dengan rumus pada persamaan (6) berikut ini.

$$g(C_i, C_j) = \frac{link(C_i, C_j)}{(n_i + n_j)^{1+2f(\theta)} - n_i^{1+2f(\theta)} - n_j^{1+2f(\theta)}} \quad (6)$$

n_i dan n_j adalah jumlah anggota dalam kelompok ke-i dan kelompok ke-j sedangkan $f(\theta) = \frac{1-\theta}{1+\theta}$.

e) Menentukan Global Heap

Global heap adalah nilai maksimum goodness measure antar kolom di baris ke-i. Mengulangi langkah hingga mendapatkan nilai maksimum di global heap dan local heap.

Selama ukuran data $> k$, dengan k adalah jumlah kelas yang ditentukan, maka dilakukan penggabungan kelompok yang memiliki nilai local heap terbesar dengan global heap terbesar menjadi satu kelompok. Kemudian menghapus kelompok yang digabungkan dari local heap dan memperbaharui nilai global heap dengan hasil penggabungan. Lakukan langkah ini hingga menemukan jumlah kelompok yang diharapkan atau kelompok akan dibangkitkan secara otomatis ketika tidak ada lagi link antar kelompoknya.

3) Pengelompokan Data Campuran

Langkah-langkah dalam algoritma pengelompokan data campuran CEBMDC (Cluster Ensemble Based Mixed Data Clustering) memiliki tahapan sebagai berikut [3]:

1. Membagi data menjadi dua subdata, yaitu murni numerik dan murni kategorik.
2. Melakukan pengelompokan objek yang memiliki variabel numerik dengan algoritma pengelompokan data numerik, serta melakukan pengelompokan objek yang memiliki variabel kategorik dengan algoritma pengelompokan data kategorik.
3. Menggabungkan (combining) hasil pengelompokan dari variabel numerik dan kategorik yang disebut proses ensemble.

4. Melakukan pengelompokan *ensemble* menggunakan algoritma pengelompokan data kategorik untuk mendapatkan kelompok akhir (*final cluster*).

D. Kinerja Hasil Pengelompokan

Kelompok yang baik adalah kelompok yang memiliki kehomogenan yang tinggi antar anggota kelompok dan keheterogenan yang tinggi antar kelompok [11].

1) Kinerja Hasil Pengelompokan Data Numerik

Indeks validitas untuk menentukan jumlah kelompok optimum pada pengelompokan AHC dapat dituliskan sebagai berikut:

Sum of Square Tootal (SST)

$$SST = \sum_{l=1}^{m_{numerik}} \sum_{i=1}^n (x_{il} - \bar{x}_l)^2 \tag{7}$$

Sum of Square Within Group (SSW)

$$SSW = \sum_{c=1}^C \sum_{l=1}^{m_{numerik}} \sum_{i=1}^n (x_{ilc} - \bar{x}_{lc})^2 \tag{8}$$

Sum of Square Between Group (SSB)

$$SSB = SST - SSW \tag{9}$$

$$R^2 = \frac{SSB}{SST} = \frac{SST - SSW}{SST} \tag{10}$$

Penentuan jumlah kelompok yang terbentuk dapat dilihat berdasarkan nilai maksimum dari Pseudo-F dan minimum ICD Rate.

$$Pseudo - F = \frac{\left(\frac{R^2}{c-1}\right)}{\left(\frac{1-R^2}{n-c}\right)} \tag{11}$$

$$ICDRate = 1 - \frac{SSB}{SST} = 1 - R^2 \tag{12}$$

2) Kinerja Hasil Pengelompokan Data Kategorik

Adapun indeks validitas untuk menentukan jumlah kelompok optimum pada data numerik ditampilkan pada Tabel 1.

Simpangan Baku dalam Kelompok (S_W)

$$S_W = [MSW]^{\frac{1}{2}} \tag{13}$$

Simpangan Baku antar Kelompok (S_B)

$$S_B = [MSB]^{\frac{1}{2}} \tag{14}$$

Kinerja suatu metode pengelompokan data kategorik semakin baik apabila nilai rasio S_W dan S_B semakin kecil. Artinya terdapat homogenitas maksimum dalam kelompok dan heterogenitas maksimum antar kelompok [12].

E. One-Way ANOVA

ANOVA (Analysis of Variance) digunakan untuk membandingkan rata-rata dari beberapa populasi yang diwakili oleh beberapa kelompok sampel secara bersama. Pengujian hipotesis yang digunakan dalam ANOVA adalah sebagai berikut:

Hipotesis

$$H_0: \mu_1 = \mu_2 = \dots = \mu_c$$

$$H_1: \text{minimal ada satu } \mu_i \text{ yang tak sama, } i = 1, 2, \dots, C$$

Statistik Uji

Persamaan (15) menampilkan formulasi untuk statistik uji ANOVA.

$$F = \frac{\frac{(SST-SSW)}{c-1}}{\frac{SSW}{n-c}} \tag{15}$$

Adapun SST dan SSW diperoleh melalui persamaan (7) dan (8)

Daerah Kritis

Tolak H_0 jika $F > F_{c-1, n-c; \frac{\alpha}{2}}$ atau $p\text{-value} < \alpha$ yang artinya minimal terdapat satu populasi yang memiliki nilai berbeda atau dapat dikatakan juga rata-rata seluruh populasi tidak sama. Dalam hal pengujian, α berarti probabilitas untuk menolak H_0 padahal H_0 benar dan p-value adalah probabilitas dengan asumsi H_0 benar dari nilai statistik uji yang dihasilkan.

F. Asumsi MANOVA

Asumsi MANOVA yang harus dipenuhi adalah dependensi variabel, data berdistribusi normal multivariat, dan matriks varians kovarians homogen. Pengujian dependensi variabel digunakan uji Bartlett berikut ini.

Adapun pengujian yang dilakukan untuk menguji asumsi MANOVA adalah uji Bartlett untuk dependensi variabel [13]. Kriteria pemenuhan asumsi kedua yaitu normal multivariat dilakukan secara visual yaitu jika plot antara jarak Mahalanobis dan Kuantil Chi-Square membentuk garis lurus. Sedangkan pemenuhan asumsi secara inferensia, dilakukan melalui pengujian korelasi Pearson. Data berdistribusi normal multivariat jika koefisien korelasi Pearson $r_{dq} > r_{n,\alpha}$, dengan $r_{n,\alpha}$ adalah nilai tabel korelasi Pearson, n adalah banyak pengamatan dan α adalah taraf signifikansi. Asumsi terakhir adalah homogenitas matriks varians kovarians populasi dapat dilakukan menggunakan Uji Box's [11].

G. One-Way MANOVA

MANOVA merupakan perluasan dari teknik univariat ANOVA yang melibatkan lebih dari satu variabel dependen [9]. Pengujian hipotesis pada MANOVA adalah sebagai berikut:

Hipotesis

$$H_0: \mu_1 = \mu_2 = \dots = \mu_c$$

$$H_1: \text{minimal ada satu } \mu_i \text{ yang tidak sama, } i = 1, 2, \dots, C$$

Statistik Uji

Statistik uji yang digunakan dalam pengambilan keputusan pada perbedaan antar kelompok adalah wilk's lamda. Rumusan untuk statistik uji wilk's lamda, didefinisikan pada persamaan (16).

$$A^* = \frac{\left| \sum_{c=1}^C \sum_{j=1}^{n_c} (x_{cj} - \bar{x}_c)(x_{cj} - \bar{x}_c)^T \right|}{\left| \sum_{c=1}^C n_c (\bar{x}_c - \bar{x})(\bar{x}_c - \bar{x})^T + \sum_{c=1}^C \sum_{j=1}^{n_c} (x_{cj} - \bar{x}_c)(x_{cj} - \bar{x}_c)^T \right|} \tag{16}$$

Daerah Kritis

Tolak H_0 jika distribusi sampling $F > F_{c-1, n-c; \frac{\alpha}{2}}$ atau $p\text{value} < \alpha$ yang artinya minimal terdapat satu populasi yang menghasilkan rata-rata yang berbeda.

H. Uji Kruskal-Wallis

Uji Kruskal-Wallis digunakan pada teknik nonparametrik untuk menguji hipotesis nol yang beberapa sampelnya diambil dari populasi yang sama atau identik. Hipotesis hingga daerah kritis dalam pengujian ini adalah sebagai berikut:

Tabel 1.

Indeks Validitas Kinerja Pengelompokan Data Kategorik	
<i>Sum Square</i>	<i>Mean Square</i>
$SST = \frac{n}{2} - \frac{1}{2n} \sum_{k=1}^K n_k^2$	$MST = \frac{SST}{(n-1)}$
$SSW = \sum_{c=1}^C \left(\frac{n_c}{2} - \frac{1}{2n_c} \sum_{k=1}^K n_{kc}^2 \right)$	$MSW = \frac{SSW}{(n-C)}$
$SSB = \frac{1}{2} \left(\sum_{c=1}^C \frac{1}{n_c} \sum_{k=1}^K n_{kc}^2 \right) - \frac{1}{2n} \sum_{k=1}^K n_k^2$	$MSB = \frac{SSB}{(C-1)}$

Tabel 2.

Variabel Penelitian	
Variabel	Keterangan
X_1	Angka Kematian Bayi (AKB)
X_2	Angka Kematian Ibu (AKI)
X_3	Persentase Balita Gizi Kurang
X_4	Persentase Pemberian ASI Eksklusif
X_5	Angka Kesakitan
X_6	Persentase Kematian Akibat Demam Berdarah Dengue (DBD)
X_7	Jumlah Dokter Umum per 100.000 penduduk
X_8	Jumlah Bidan per 100.000 penduduk
X_9	Jumlah Perawat per 100.000 penduduk
X_{10}	Jumlah Puskesmas per 100.000 penduduk
X_{11}	Jumlah Posyandu Aktif per 100.000 penduduk
X_{12}	Persentase Rumah Tangga yang Memiliki Akses Terhadap Sumber Air Minum Layak
X_{13}	Persentase Rumah Tangga yang Memiliki Akses Terhadap Sanitasi Layak
X_{14}	Persentase <i>Universal Child Immunization</i> (UCI)

Tabel 3.

Variabel Penelitian (Lanjutan)		
Variabel	Keterangan	Kategori
X_{15}	Penolong Persalinan	0: Semua ditolong tenaga medis baik dokter, bidan, dan tenaga kesehatan lain 1: Masih ada yang ditolong tenaga bukan medis, dukun atau yang lainnya
X_{16}	Angka Keberhasilan Pengobatan TBC	0: > 90% 1: 85% - 90% 2: 85%
X_{17}	Kepemilikan Jaminan Kesehatan Mayoritas	0: BPJS Kesejahteraan BPI dan non PBI 1: Jamkesda, Asuransi Swasta, dan Perusahaan/ Kantor
X_{18}	Indeks Pembangunan Manusia (IPM)	0: Tinggi (> 80) 1: Menengah Atas (66 - 79,9) 2: Menengah Bawah (50-65,9)

Hipotesis

H0: k populasi mempunyai fungsi sebaran yang identik

H1: Minimal ada satu fungsi sebaran dari k populasi yang berbeda

Statistik Uji

Dengan statistik uji didapatkan melalui persamaan (17).

$$H = \frac{12}{N(N+1)} \sum_{i=1}^k \frac{R_i^2}{n_i} - 3(N+1) \tag{17}$$

dimana,

N = jumlah pengamatan

Ri = jumlah rank sampel ke-i

Tabel 4.

Karakteristik Kesehatan Masyarakat di Jawa Timur Tahun 2019 Berdasarkan Indikator Kesehatan Masyarakat Menurut Variabel Numerik

Variabel	Rata-rata	Median	Varians	Min	Maks
AKB	5,962	5,589	7,543	0,604	12,456
AKI	0,912	0,8674	0,1306	0,000	1,9800
Persentase Balita Gizi Kurang	8,103	8,100	5,069	4,300	13,000
Persentase Pemberian ASI	78,03	78,80	91,19	46,90	96,40
Angka Kesakitan	36,97	38,14	47,31	18,64	49,49
Persentase Kematian akibat DBD	1,318	1,050	2,063	0,000	8,000
Jumlah Dokter Umum	26,58	17,50	455,12	7,00	86,00
Jumlah Bidan	71,18	63,00	732,48	36,00	149,00
Jumlah Perawat	161,1	120,5	13562,5	64,0	602,0
Jumlah Puskesmas	2,658	3,000	0,718	1,000	5,000
Jumlah Posyandu	121,71	122,50	353,56	75,00	153,00
Persentase RT Akses Air Minum Layak	44,10	48,36	304,20	5,60	78,64
Persentase RT Akses Sanitasi Layak	72,20	77,93	356,18	25,51	97,43

Tabel 5.

Karakteristik Kesehatan Masyarakat di Jawa Timur Tahun 2019 Berdasarkan Indikator Kesehatan Masyarakat Menurut Variabel Kategorik

Variabel	Modus	Jumlah	Persentase
Daerah UCI	0	25	66%
Penolong Persalinan	0	21	55%
Keberhasilan Pengobatan TBC	0	17	45%
Jaminan Kesehatan	0	32	84%
IPM	1	29	76%

Tabel 6.

Statistik Uji Signifikansi Perbedaan *Final Cluster* Secara Multivariat

Variabel	Uji	Statistik Uji
Numerik	<i>One-Way</i> MANOVA	2,147
Kategorik	Perluasan Uji Kruskal-Wallis	1328,190

Jika terdapat ties atau nilai yang sama, statistik uji perlu dikoreksi dengan faktor. Sehingga statistik uji Kruskal-Wallis terkoreksi menjadi:

$$H_c = \frac{H}{1 - \sum T / (N^3 - N)} \tag{18}$$

Daerah Kritis

Statistik uji H mengikuti distribusi Chi-Square sehingga kriteria pengujiannya adalah tolak H0 jika $H \geq \chi_{(k-1); \alpha}^2$ atau p-value < α [14].

I. Perluasan Uji Kruskal-Wallis untuk Data Multivariat

Perluasan Kurskal-Wallis untuk data multivariat mentransformasi data original kedalam peringkat sehingga tidak berdistribusi (bebas distribusi) [15]. Diberikan bahwa:

$$\bar{R}_{i.k} = \sum_{j=1}^{n_i} \frac{R_{ijk}}{n_i} \tag{19}$$

Adapun $E(\bar{R}_{i.k}) = m = (n+1)/2$ dan $\mathbf{U}_i = (\bar{R}_{i.1} - m, \dots, \bar{R}_{i.p} - m)'$ menunjukkan rata-rata peringkat untuk kelompok ke-i dikoreksi dengan m untuk setiap variat. Estimasi dari matriks kovarians dalam grup disajikan dalam persamaan (20).

$$\mathbf{V} = \frac{1}{n-1} \sum_{i=1}^g \sum_{j=1}^{n_i} (\mathbf{R}_{ij} - m\mathbf{1}_p)(\mathbf{R}_{ij} - m\mathbf{1}_p)' \tag{20}$$

Tabel 7.
Hasil Pengujian One-Way ANOVA Variabel Numerik

Variabel	Statistik Uji (F_{hitung})	P-Value
AKB	1,274	0,292
AKI	1,353	0,271
Persentase Balita Gizi Kurang	0,024	0,967
Persentase Pemberian ASI	1,651	0,206
Angka Kesakitan	0,235	0,792
Persentase Kematian akibat DBD	0,097	0,908
Jumlah Dokter Umum	7,702	0,002
Jumlah Bidan	3,620	0,037
Jumlah Perawat	6,549	0,004
Jumlah Puskesmas	0,159	0,854
Jumlah Posyandu	0,353	0,705
Persentase RT Akses Air Minum Layak	12,371	0,000
Persentase RT Akses Sanitasi Layak	10,764	0,000

Tabel 8.
Hasil Uji Kruskal-Wallis Variabel Kategorik

Variabel	Statistik Uji H	P-Value
Persentase Cakupan Daerah UCI	16,689	0,001
Penolong Persalinan	15,405	0,000
Keberhasilan Pengobatan TBC	2,002	0,367
Jaminan Kesehatan	9,539	0,008
IPM	18,978	0,001

Di bawah hipotesis nol yang mana tidak ada perbedaan rata-rata grup untuk p variabel $E(\mathbf{U}_i) = \mathbf{0}_p$ dan statistik uji MKW (Multivariate Kruskal-Wallis) dinyatakan sebagai:

$$W^2 = \sum_{i=1}^g n_i \mathbf{U}_i' \mathbf{V}^{-1} \mathbf{U}_i \quad (21)$$

Statistik uji W^2 mengikuti distribusi Chi-Square dengan derajat bebas $p(g-1)$ sehingga kriteria pengujianya adalah tolak H_0 jika $W^2 \geq \chi_{p(k-1);\alpha}^2$ atau $p\text{-value} < \alpha$ [16].

J. Kondisi Kesehatan Masyarakat Provinsi Jawa Timur

Pada tahun 2019 juga masih terdapat 0,2 persen wanita yang pernah kawin usia 15-49 tahun yang penolong proses kelahiran anak hidup yang terakhir adalah bukan oleh tenaga medis. Hal tersebut mengindikasikan bahwa belum seluruhnya wanita yang melahirkan di Jawa Timur sadar akan pentingnya keselamatan dalam proses persalinan karena jika tidak di tolong oleh tenaga medis profesional akan meningkatkan risiko kematian pada bayi yang dilahirkan atau ibu yang melahirkan atau bahkan keduanya. Kondisi kesehatan masyarakat yang menjadi perhatian Pemerintah Provinsi Jawa Timur adalah status kesehatan ibu, bayi, dan anak balita yang masih rendah disebabkan ditandai dengan angka kematian ibu yang mengalami kenaikan selama 3 tahun yaitu pada 2015 – 2017 dan disebabkan karena belum semua ibu bersalin melahirkan di fasilitas kesehatan dan masih ada persalinan yang ditolong oleh dukun yang tercatat sebanyak 2.780 kasus pada tahun 2017 serta gizi buruk dan stunting yang masih terjadi pada balita yang mengalami peningkatan pada tahun 2018 [17].

III. METODOLOGI PENELITIAN

A. Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh melalui publikasi Badan Pusat

Tabel 9.
Rangkuman Perbandingan Karakteristik Masing-Masing Kelompok Kualitas Kesehatan Masyarakat di Jawa Timur

Variabel	Kelompok Kualitas Kesehatan		
	Buruk (Kelompok 1)	Sedang (Kelompok 3)	Baik (Kelompok 2)
Jumlah Dokter Umum	Sedikit	Sedang	Banyak
Jumlah Bidan	Sedikit	Sedang	Banyak
Jumlah Perawat	Sedang	Sedikit	Banyak
Persentase RT Akses Air	Sedang	Tinggi	Rendah
Minum Layak			
Persentase RT Akses Sanitasi Layak	Rendah	Sedang	Tinggi
Cakupan Daerah UCI	1 kab/kota diatas 90%	8 kab/kota diatas 90%	Seluruh kab/kota diatas 90%
Penolong Persalinan	Seluruh kab/kota ditolong bukan medis	53% kab/kota ditolong bukan medis	14% kab/kota ditolong bukan medis
Jaminan Kesehatan	BPJS	BPJS	BPJS dan NonBPJS
IPM	Menengah bawah	Menengah atas	Tinggi

Statistik (BPS) dan publikasi Dinas Kesehatan Provinsi Jawa Timur tahun 2019. Adapun yang menjadi acuan adalah Program Prioritas Pemerintah Provinsi Jawa Timur di bidang kesehatan dan Indikator Kinerja Upaya Kesehatan dari Renstra Dinas Kesehatan Provinsi Jawa Timur.

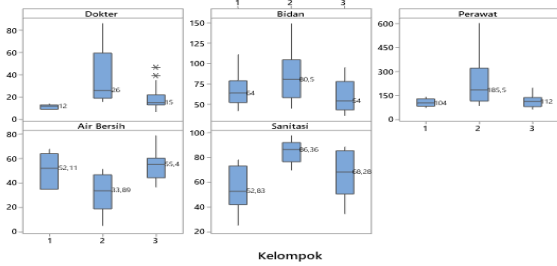
B. Variabel Penelitian

Variabel yang digunakan dalam penelitian ini terdiri atas 18 variabel yang terdiri atas 13 variabel numerik dan 5 variabel kategorik. Daftar variabel penelitian disajikan pada Tabel 2 dan Tabel 3.

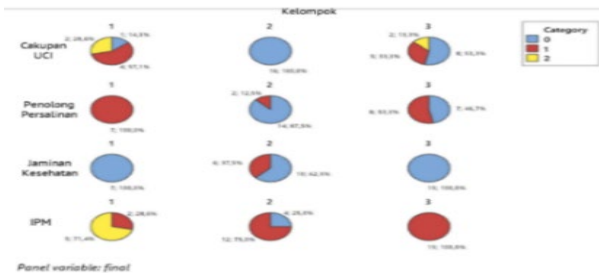
C. Langkah Analisis

Langkah analisis yang ditempuh dalam penelitian ini dirinci menurut tujuan penelitian. Untuk mencapai tujuan penelitian pertama yaitu mendeskripsikan karakteristik kabupaten/kota di Jawa Timur berdasarkan indikator kesehatan masyarakat sesuai dengan pokok bahasan pada bagian II.A, dilakukan sebagai berikut:

1. Menrumuskan masalah dan mengumpulkan data sekunder
2. Mendeskripsikan karakteristik data.
3. Mengelompokkan variabel berskala data numerik menggunakan *Agglomerative Hierarchical Clustering*.
4. Menentukan kelompok optimum melalui nilai *Pseudo-F* terbesar dan nilai *ICD Rate* terkecil.
5. Mengelompokkan data berkala kategori menggunakan metode ROCK dengan nilai *threshold* θ antara 0 dan 1.
6. Menentukan nilai rasio S_W dan S_B terkecil.
7. Menggabungkan hasil pengelompokan data numerik dan data kategorik.
8. Mengelompokkan *cluster* gabungan menggunakan metode ROCK menggunakan *threshold* θ antara 0 dan 1. Metode inilah yang disebut dengan metode *ensemble ROCK*.
9. Menentukan rasio S_W dan S_B terkecil yang kemudian kelompok tersebut disebut *final cluster*.
10. Melakukan uji signifikansi perbedaan *final cluster* menggunakan uji perluasan Kruskal-Wallis.



Gambar 1. Perbandingan karakteristik *final cluster* berdasarkan variabel numerik.



Gambar 2. Perbandingan karakteristik *final cluster* berdasarkan variabel kategorik.



Gambar 3. Peta persebaran kualitas kesehatan masyarakat di Jawa Timur.

11. Melakukan uji asumsi MANOVA untuk data numerik.
12. Melakukan uji signifikansi perbedaan *final cluster* secara multivariat pada masing-masing tipe data.
13. Melakukan pengujian signifikansi perbedaan *final cluster* secara univariat pada masing-masing tipe data.
14. Mendeskripsikan dan membandingkan karakteristik berdasarkan kelompok masing-masing.
15. Menarik kesimpulan dan saran.

IV. HASIL DAN PEMBAHASAN

A. Karakteristik Kabupaten/Kota di Jawa Timur Berdasarkan Indikator Kesehatan Masyarakat

Karakteristik indikator kesehatan sesuai dengan variabel numerik disajikan pada Tabel 4. Informasi yang diperoleh melalui Tabel 3 adalah bahwa nilai varians pada hampir seluruh variabel penelitian nilainya lebih besar dibandingkan rata-ratanya. Besarnya nilai varians tersebut mengindikasikan bahwa terdapat ketidakmerataan atau disparitas kondisi kesehatan masyarakat di Jawa Timur berdasarkan variabel numerik, dapat dikatakan pula bahwa kualitas kesehatan masyarakat di Jawa Timur sangat beragam, ada yang tinggi dan ada juga yang rendah. Adapun karakteristik indikator Kesehatan masyarakat di Jawa Timur sesuai dengan variabel kategorik disajikan pada Tabel 5.

Melalui Tabel 5 diketahui bahwa mayoritas masyarakat di Jawa Timur persentase cakupan daerah UCI sudah diatas 90% yang menandakan mayoritas balita di Jawa Timur sudah mendapatkan imunisasi lengkap. Kesadaran akan penolong persalinan dengan tenaga medis perlu di tingkatkan lagi karena hampir setengah kabupaten/kota di Jawa Timur proses bersalinnya masih ditolong bukan oleh tenaga medis. Keberhasilan pengobatan TBC dikatakan sudah cukup baik karena hampir sebagian kabupaten/kota keberhasilan pengobatan TBC diatas 90%. Kesadaran masyarakat akan pentingnya asuransi kesehatan dari pemerintah berupa BPJS kesehatan juga sudah baik karena mayoritas masyarakat sudah memiliki BPJS kesehatan. Kemudian untuk IPM di Jawa Timur mayoritas kabupaten/kota-nya masuk kedalam kategori menengah atas dengan skor antara 66 – 79,9 yang sudah termasuk baik.

B. Pengelompokan Kabupaten/Kota di Jawa Timur Berdasarkan Indikator Kesehatan Masyarakat

Pengelompokan kabupaten/kota menggunakan metode ensemble ROCK dimulai dengan pengelompokan variabel numerik menggunakan AHC, kemudian pengelompokan variabel kategorik menggunakan metode ROCK, dan terakhir pengelompokan kelompok gabungan menggunakan metode ROCK yang disebut dengan ensemble ROCK.

Adapun pengelompokan variabel numerik menggunakan AHC menghasilkan 3 kelompok optimum yang dihasilkan melalui pautan average linkage dengan nilai Pseudo-F sebesar 32,0475 dan ICD Rate 0,3531. Ketiga kelompok tersebut disebut sebagai kelompok numerik. Kemudian untuk pengelompokan variabel kategorik menggunakan algoritma ROCK dengan nilai *tetha* 0,07; 0,08; 0,10; 0,13; 0,15; 0,18; 0,22; 0,24; 0,29; dan 0,31 menghasilkan 3 kelompok optimum dengan nilai rasio SW dan SB terkecil sebesar 0,0869 yang dihasilkan oleh *tetha* sebesar 0,29, kemudian kelompok tersebut disebut dengan kelompok kategorik.

Melalui kelompok numerik dan kelompok kategorik, dilakukan pengelompokan kembali menggunakan metode ROCK dengan nilai *tetha* yang sama dengan pengelompokan data kategorik menghasilkan *final cluster* sebanyak 3 kelompok. *Final cluster* tersebut dihasilkan oleh *tetha* 0,18 yang memberikan nilai rasio *S_w* dan *S_B* terkecil yaitu 0,0292. *Final cluster* ini mengalokasikan 38 kabupaten/kota di Jawa Timur sebanyak 7 kabupaten/kota ke kelompok 1, 16 kabupaten/kota ke kelompok 2, dan 15 kabupaten/kota ke kelompok 3. Kemudian *final cluster* inilah yang menjadi acuan dalam mengidentifikasi variabel apa saja yang menjadi pembeda kualitas kesehatan masyarakat di Jawa Timur.

C. Uji Signifikansi Perbedaan Final Cluster Terhadap Variabel Pembentuk Final Cluster

Pengujian pertama yang dilakukan adalah uji signifikansi perbedaan *final cluster* terhadap dua variabel kategorikal yaitu kelompok numerik dan kelompok kategorik untuk mengetahui apakah *final cluster* benar memberikan perbedaan yang signifikan antar ketiga kelompok yang telah terbentuk. Pengujian ini menghasilkan nilai statistik uji W^2 sesuai dengan uji perluasan Kruskal-Wallis sebesar 97,222 yang bernilai lebih besar dari $\chi^2_{2,0,05}(5,991)$, sehingga keputusan yang diambil adalah Tolak H_0 yang menyatakan bahwa terdapat perbedaan yang signifikan antar ketiga

kelompok final cluster terhadap kelompok numerik dan kelompok kategorik sehingga dapat dilanjutkan ke pengujian signifikansi perbedaan final cluster pada masing-masing tipe data penelitian.

D. Uji Signifikansi Perbedaan Final Cluster Secara Multivariat

Pengujian secara multivariat ini dipisahkan antara variabel numerik dan variabel kategorik. Untuk variabel numerik dilakukan pengujian asumsi MANOVA terlebih dahulu. Asumsi pertama yang harus dipenuhi adalah dependensi variabel, dihasilkan statistik uji χ^2_{hitung} melalui uji Bartlett dihasilkan sebesar 221,229 yang mana jika dibandingkan dengan nilai $\chi^2_{78;0,05}$ (58,653) menghasilkan keputusan Tolak H_0 yang menandakan 13 variabel numerik saling berhubungan.

Pengujian asumsi kedua yaitu normal multivariat menghasilkan plot antaran jarak Mahalanobis dan kuantil Chi-Square membentuk garis lurus dan menghasilkan nilai korelasi Pearson sebesar 0,996 yang jika dibandingkan dengan nilai $r_{38;0,05}$ (0,3203) menghasilkan keputusan Tolak H_0 yang menandakan bahwa 13 variabel numerik berdistribusi normal multivariat, kemudian untuk pengujian asumsi terakhir yaitu matriks varians kovarians menggunakan uji Box's M menghasilkan statistik uji F 1,073 dan p-value 0,303 yang menandakan bahwa matriks varians kovarians dari tigabelas variabel numerik adalah homogen. Kedua asumsi sudah terpenuhi maka dapat dilanjutkan pada tahap pengujian signifikansi perbedaan final cluster secara multivariat untuk masing-masing tipe data. Tabel 6 menampilkan hasil pengujian signifikansi tersebut.

Sesuai dengan Tabel 6, statistik uji untuk variabel numerik menghasilkan nilai 2,147 dan menghasilkan p-value sebesar 0,012. Jika dibandingkan dengan taraf signifikansi 0,05, maka keputusan yang diambil untuk uji signifikansi perbedaan final cluster untuk variabel numerik adalah Tolak H_0 . Kemudian untuk variabel kategorik, statistik uji sesuai dengan Tabel 5 bernilai 1328,19 dan jika dibandingkan dengan nilai tabel $\chi^2_{10;0,05}$ (18,307) menghasilkan keputusan Tolak H_0 juga. Sehingga dapat dikatakan bahwa dengan keyakinan 95% setiap kelompok kabupaten/kota yang terbentuk melalui final cluster berpengaruh signifikan terhadap indikator kesehatan masyarakat yang diwakili oleh tiga belas variabel numerik dan lima variabel kategorik secara multivariat atau dapat dikatakan pula minimal terdapat satu variabel numerik dan satu variabel kategorik yang memberikan perbedaan yang signifikan dari ketiga kelompok kabupaten/kota yang terbentuk.

E. Uji Signifikansi Perbedaan Final Cluster Secara Univariat

Pengujian ini dilakukan untuk mengetahui variabel mana saja yang memberikan perbedaan yang signifikan diantara ketiga kelompok final cluster yang terbentuk. Untuk variabel numerik dilakukan menggunakan One-Way ANOVA dan dihasilkan keluaran yang ditampilkan pada Tabel 7.

Sesuai Tabel 6 dengan menggunakan taraf signifikansi 5%, dihasilkan bahwa variabel numerik yang memberikan perbedaan yang signifikan antara ketiga kelompok final cluster adalah variabel yang ditandai dengan cetak tebal yaitu persentase pemberian ASI, jumlah tenaga kesehatan, dan

persentase RT atas akses terhadap air minum dan sanitasi yang layak. Variabel tersebut signifikan karena menghasilkan nilai Fhitung yang lebih besar dari $F_{2;35;0,05}$ (3,267) dan nilai p-value yang lebih kecil dari 0,05. Adapun hasil pengujian univariat variabel kategorik disajikan pada Tabel 8.

Tabel 7 memberikan informasi bahwa jika digunakan taraf signifikansi 5%, variabel kategorik yang memberikan perbedaan diantara ketiga kelompok final cluster adalah persentase cakupan daerah UCI, penolong persalinan, jaminan kesehatan dan IPM karena variabel tersebut menghasilkan statistik uji H lebih besar dari nilai $\chi^2_{2;0,05}$ (5,991) dan menghasilkan p-value yang kurang dari 0,05.

F. Karakteristik Perbedaan Final Cluster

Pendeskripsian perbedaan karakteristik dilakukan pada variabel yang telah dinyatakan sebagai variabel pembeda pada masing-masing tipe data penelitian. Perbedaan karakteristik kelompok yang terbentuk pada variabel numerik ditampilkan pada Gambar 1.

Melalui Gambar 1, dapat dilihat bahwa secara keseluruhan kelompok 2 menghasilkan nilai tengah pada box plot yang cenderung paling tinggi diantara kelompok yang lain, sedangkan kelompok 1 menghasilkan nilai tengah pada box plot yang cenderung paling rendah. Sesuai dengan hal tersebut maka dapat dikatakan bahwa, kelompok kabupaten/kota di Jawa Timur yang memiliki kondisi kesehatan masyarakat yang paling baik adalah kelompok 2, disusul oleh kelompok 3, kemudian diakhiri oleh kelompok 1.

Adapun karakteristik pembeda untuk variabel kategorik ditampilkan pada Gambar 2. Sesuai dengan Gambar 2, diketahui bahwa kategori 0 yang berwarna biru merupakan kategori terbaik dan kategori 2 yang berwarna kuning adalah kategori paling buruk pada setiap variabel. Melalui kategori tersebut dapat dikatakan bahwa secara keseluruhan, kelompok yang paling banyak menghasilkan persentase terbesar untuk kategori berwarna biru adalah kelompok 2 dan sedangkan kelompok yang menghasilkan persentase terkecil untuk kategori berwarna biru adalah kelompok 1. Maka dari itu, dapat dikatakan bahwa kelompok kabupaten/kota di Jawa Timur yang menghasilkan kualitas kesehatan masyarakat terbaik ke terburuk secara berturut-turut berdasarkan variabel kategorik adalah kelompok 2, kelompok 3, dan kelompok 1.

Setelah dilakukan perbandingan perbedaan karakteristik di masing-masing kelompok, dapat diketahui bahwa kualitas Kesehatan masyarakat di Jawa Timur dikategorikan menjadi kelompok dengan kualitas Kesehatan yang baik, sedang, dan buruk. Adapun peta persebaran kabupaten/kota sesuai dengan kualitas kesehatannya disajikan pada Gambar 3.

Demi mempermudah dalam mengidentifikasi perbedaan karakteristik kualitas kesehatan masyarakat pada masing-masing kelompok yang terbentuk, maka dilakukan perbandingan karakteristik masing-masing kelompok kesehatan masyarakat di Jawa Timur yang dirangkum pada Tabel 9.

Pendeskripsian perbedaan karakteristik antar kelompok yang terbentuk sesuai Tabel 9 dapat dicontohkan bahwa kelompok kesehatan masyarakat dengan kualitas yang buruk memiliki jumlah tenaga kesehatan (dokter umum, bidan, dan perawat) per 100.000 penduduk yang sedikit, persentase RT

terhadap akses sanitasi layak, persentase cakupan daerah UCI, dan kategori IPM dengan kualitas yang rendah, serta seluruh kabupaten/kotanya masih ada yang proses persalinannya ditolong bukan oleh tenaga medis, namun kelompok ini memiliki persentase RT terhadap akses air minum layak dengan kualitas yang sedang dan mayoritas masyarakatnya sudah memiliki asuransi kesehatan berupa BPJS kesehatan. Pendeskripsian yang sama juga dapat dilakukan untuk kelompok kesehatan masyarakat yang sedang dan baik sesuai dengan Tabel 9.

Adapun variabel yang menghasilkan nilai terendah atau menjadi indikator kesehatan terburuk pada kelompok kesehatan masyarakat yang buruk adalah jumlah dokter, jumlah bidan, persentase RT atas akses terhadap sanitasi layak, cakupan daerah UCI, penolong persalinan, dan IPM. Harapannya dengan diketahuinya indikator terburuk tersebut dapat menjadi perhatian lebih bagi pemerintah Provinsi Jawa Timur dan dapat menambah jumlah tenaga kesehatan di kelompok ini, dapat memberikan fasilitas sanitasi yang lebih layak, dapat melakukan imunisasi dasar lengkap untuk balita di kelompok ini dengan lebih gencar, dapat melakukan edukasi tentang pentingnya proses persalinan yang ditolong oleh tenaga medis, dan dapat melakukan pelatihan bagi masyarakat yang berada di kelompok ini agar IPM kelompok ini bisa meningkat.

Kemudian variabel yang menjadi perhatian untuk ditingkatkan pada kelompok kesehatan masyarakat yang sedang adalah jumlah perawat, karena jumlah perawat dalam kelompok ini adalah yang terendah. Harapannya dengan ditingkatkannya jumlah perawat di kelompok ini mampu meningkatkan kualitas pelayanan kesehatan masyarakat sehingga masyarakat mendapat pelayanan kesehatan yang maksimal.

Sedangkan variabel yang menjadi perhatian khusus untuk kelompok kabupaten/kota dengan kualitas kesehatan yang baik adalah perlu lebih diperhatikan lagi terkait akses terhadap air minum yang layak karena kelompok ini menghasilkan persentase terendah untuk variabel persentase RT akses terhadap air minum layak. Harapannya dengan diketahuinya hal tersebut, Dinas Kesehatan Provinsi Jawa Timur mampu menyediakan lebih banyak akses terhadap air minum layak untuk kelompok kabupaten/kota ini, sehingga kualitas kesehatan masyarakat akan meningkat.

V. KESIMPULAN

Kesimpulan yang dapat ditarik dari penelitian ini adalah bahwa terjadi disparitas kondisi kesehatan masyarakat antar kabupaten/kota di Jawa Timur sehingga diperlukan

pengelompokan kualitas kesehatan masyarakat yang menghasilkan bahwa kabupaten/kota di Jawa Timur terbagi kedalam 3 kelompok kesehatan masyarakat yang dikategorikan dengan kualitas kesehatan masyarakat yang buruk, sedang, dan baik. Adapun variabel yang perlu ditingkatkan pada kelompok kesehatan masyarakat dengan kualitas buruk adalah jumlah dokter, jumlah bidan, persentase RT atas akses terhadap sanitasi layak, cakupan daerah UCI, penolong persalinan, dan IPM. Kemudian untuk kelompok kesehatan masyarakat yang berkualitas sedang perlu ditingkatkan persebaran jumlah perawat, serta kelompok kesehatan masyarakat dengan kualitas baik perlu ditingkatkan akses terhadap air minum layak.

DAFTAR PUSTAKA

- [1] Dinas Kesehatan, *Profil Kesehatan Provinsi Jawa Timur Tahun 2019*, 1st ed. Surabaya: Dinas Kesehatan Provinsi Jawa Timur, 2019.
- [2] Kementerian Kesehatan RI, *Indeks Pembangunan Kesehatan Masyarakat 2018*, 1st ed. Jakarta: Badan Penelitian dan Pengembangan Kesehatan, 2019.
- [3] Z. He, X. Xu, and S. Deng, "A cluster ensemble method for clustering categorical data," *Inf. Fusion*, vol. 6, no. 2, pp. 143–151, 2005.
- [4] I. A. Sari and D. R. S. Saputro, "Algoritme Quick ROBust Clustering using linKs (QROCK) untuk Clustering Data Kategorik," in *PRISMA, Prosiding Seminar Nasional Matematika*, 2021, vol. 4, pp. 640–644.
- [5] A. Maully, "Faktor-Faktor yang Memengaruhi Indeks Kesehatan Kabupaten dan Kota di Provinsi Jawa Timur," Institut Teknologi Sepuluh Nopember, Surabaya, 2014.
- [6] N. Khotimah, "Perubahan sosial masyarakat tebuireng setelah adanya wisata religi makam Gus Dur," *Paradigma*, vol. 5, no. 3, 2017.
- [7] L. I. Ichsan, "Pengelompokan Kabupaten/Kota di Jawa Timur Berdasarkan Pembangunan Kualitas Sumber Daya Manusia dan Pembangunan Ekonomi dengan Metode Ensemble ROCK," Institut Teknologi Sepuluh Nopember, Surabaya, 2018.
- [8] R. E. Walpole, *Pengantar Metode Statistika, Diterjemahkan: Bambang Sumantri*, 3rd ed. Jakarta: PT. Gramedia Pustaka Utama, 1955.
- [9] R. A. Johnson and D. W. Wicherin, *Applied Multivariate Statistical Analysis*, 6th ed. New Jersey: Person Prentice Hall, 2007.
- [10] M. Dutta, A. K. Mahanta, and A. K. Pujari, "QROCK: A quick version of the ROCK algorithm for clustering of categorical data," *Pattern Recognit. Lett.*, vol. 26, no. 15, pp. 2364–2373, 2005.
- [11] J. Hair, *Multivariate Data Analysis*, 7th ed. New York: Pearson Education, Inc, 2010.
- [12] M. J. Bunkers, J. R. Miller, and A. T. DeGaetano, "Definition of climate regions in the northern plains using an objective cluster modification technique," *J. Clim.*, vol. 9, no. 1, pp. 130–146, 1996.
- [13] F. M. D., *Multivariate Statistical Methods*, 3rd ed. United States of America: McGraw-Hill, Inc, 1990.
- [14] W. Daniel, *Statistika Non Parametrik, Alih Bahasa: Alex Tri Kuncoro*, 1st ed. Jakarta: PT. Gramedia, 1989.
- [15] T. S. Yanti, "Perluasan uji kruskal wallis untuk data multivariat," *Statistika*, vol. 10, no. 1, 2010.
- [16] F. He, "Nonparametric MANOVA Approaches for Non-Normal Multivariate Outcomes," University of Pittsburgh, 2013.
- [17] Dinas Kesehatan Provinsi Jawa Timur, *Rencana Strategis Dinas Kesehatan Provinsi Jawa Timur 2019 - 2024*, 1st ed. Surabaya: Dinas Kesehatan Provinsi Jawa Timur, 2019.