

ANALISIS CLUSTER TERHADAP TINDAKAN KEKERASAN DI INDONESIA DENGAN ALGORITMA CLUSTER HIRARKI MENGGUNAKAN METODE WARD

Muhammad Taufiqurrahman Alghani¹, Marhaeni^{2*}

Program Studi Teknik Informatika, Fakultas Sains dan Teknologi Informasi
Institut Sains dan Teknologi Nasional
email: gadsitaufig@gmail.com, marhaeni@istn.ac.id

ABSTRAK

Penelitian ini dilatarbelakangi oleh viralnya kasus-kasus terjadinya tindakan kekerasan oleh para artis atau public figure pada tahun 2022. Ramainya tindakan kekerasan ini menunjukkan bahwa masyarakat Indonesia sangat menentang adanya tindakan kekerasan sehingga dapat dirumuskan dalam penelitian ini bagaimana tingkatan tindak kekerasan yang terjadi diseluruh provinsi di Indonesia. Tujuan penelitian ini adalah melakukan proses pengelompokan provinsi berdasarkan kategori lingkungan terjadinya tindak kekerasan. Penelitian ini akan melakukan analisa terhadap 34 provinsi yang ada di Indonesia dengan menggunakan metode cluster hirarki. Hasil yang didapatkan dari penelitian ini adalah dapat mengetahui tingkat keamanan dan kondisi keamanan dan ketertiban keseluruhan provinsi di Indonesia. Kata kunci: Tindakan kekerasan, clustering, metode hirarki, Indonesia.

ABSTRACT

This research was motivated by the virality of cases of acts of violence by artists or public figures in 2022. The prevalence of these acts of violence shows that the Indonesian people are very opposed to acts of violence so that this research can formulate how the level of acts of violence occurred in all provinces in Indonesia. The purpose of this study was to carry out the process of grouping provinces based on the environmental categories where acts of violence occurred. This study will analyze 34 provinces in Indonesia using the hierarchical cluster method. The results obtained from this study are able to determine the level of security and the condition of security and order in all provinces in Indonesia.

Kata kunci: Acts of violence, clustering, hierarchical method, Indonesia

I. PENDAHULUAN

Pada tahun 2022, begitu ramainya terjadi kasus-kasus tindakan kekerasan rumah tangga yang sudah dilakukan oleh para artis atau public figure. Berdasarkan berita dan informasi yang sudah tersebar masyarakat merasakan kegeraman terhadap pelaku tindakan kekerasan tersebut. Akan tetapi tentu saja tindakan kekerasan bukan hanya bisa dilakukan oleh para artis atau *public figure* saja, bahkan masyarakat biasa bisa melakukan tindakan kekerasan apabila sedang tidak berkepal dingin.

Proses pengolahan data menjadi satu hal yang sangat penting di era persaingan bisnis yang mengharuskan mendapatkan informasi secara cepat. Informasi terkait data yang berguna dalam proses bisnis, hingga penentuan strategi

kedepannya, karena itulah penggunaan penambangan data atau disebut dengan data mining sangat penting bagi kelangsungan bisnis yang berjalan. Pesatnya penggunaan internet membuat data yang dihasilkan dalam jumlah data yang sangat besar, hal ini memberi manfaat dan bisa dipakai dalam membuat keputusan. Selain itu juga pemrosesan dengan menggunakan data tersebut. Penambangan data sebagai proses pengumpulan dan pengolahan data yang dilakukan memakai perangkat lunak dengan bantuan perhitungan statistika.

II. TINJAUAN PUSTAKA

Dalam Kamus Besar Bahasa Indonesia (KBBI) kekerasan adalah perbuatan seseorang atau kelompok orang yang menyebabkan cedera

atau matinya orang lain atau menyebabkan kerusakan fisik atau barang orang lain. Sedangkan menurut World Health Organization (WHO) kekerasan adalah penggunaan seluruh kekuatan fisik demi mendapatkan kekuasaan yang biasanya disertai dengan ancaman, sehingga mengakibatkan kerugian bagi pihak lain, seperti luka memar, kematian, kerugian secara psikologis, dan lain sebagainya.

Clustering adalah proses membuat pengelompokan sehingga semua anggota dari setiap partisi mempunyai persamaan berdasarkan matriks tertentu. Analisis cluster atau analisis kelompok merupakan teknik analisa data yang bertujuan untuk mengelompokkan individu atau objek ke dalam beberapa kelompok yang memiliki sifat berbeda antar kelompok, sehingga individu atau objek yang terletak di dalam satu kelompok akan mempunyai sifat relatif homogen (Talukua .M.W et al., 2017).

Data mengenai ukuran kesamaan tersebut dapat dianalisis dengan analisis cluster sehingga dapat ditentukan siapa yang masuk cluster mana (Gudono, 2011).

2.1. Merumuskan Masalah

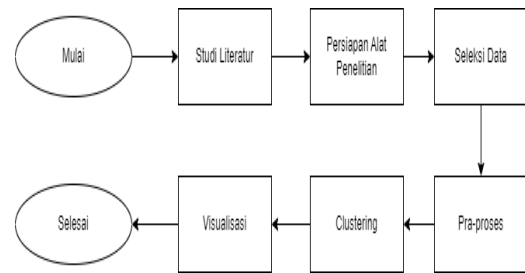
Hal yang paling penting di dalam masalah analisis cluster adalah pemilihan variabel-variabel yang akan dipergunakan untuk pengclusteran (pembentukan cluster). Memasukkan satu atau dua variabel yang tidak relevan dengan masalah pengclusteran sehingga akan menyebabkan penyimpangan hasil pengclusteran yang kemungkinan besar sangat bermanfaat (Supranto, 2004).

2.2. Metode Ward

Metode varians bertujuan untuk memperoleh cluster yang memiliki varians internal cluster yang sekecil mungkin. Metode varians yang umum dipakai adalah metode Ward dimana rata-rata untuk setiap cluster dihitung. Lalu, dihitung jarak Euclidean antara setiap objek dan nilai rata-rata itu, lalu jarak itu dihitung semua. Pada setiap tahap, dua cluster yang memiliki kenaikan 'sum of squares dalam cluster' yang terkecil digabungkan (Simamora, 2005.)

III. METODE PENELITIAN

Tahapan penelitian dibentuk dengan tujuan penelitian berjalan secara baik dan terstruktur. Penelitian ini memiliki beberapa tahapan yang akan dilakukan, dan tahapan tahapan tersebut dapat dicermati pada kerangka penelitian yang ditunjukkan pada Gambar 3.1.



Gambar 3.1 Tahapan Penelitian

3.1. Persiapan Alat Penelitian

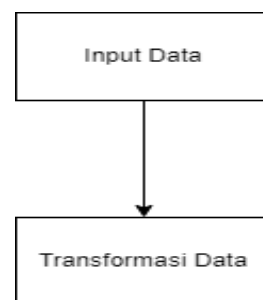
Pada tahap persiapan alat penelitian dilakukan pemasangan beberapa perangkat lunak yang akan digunakan untuk melakukan penelitian. Beberapa perangkat lunak yang akan dipasang adalah seperti bahasa pemrograman Python, aplikasi interpreter Jupyter Notebook, dan library yang digunakan seperti Pandas, Numpy, Matplotlib, Seaborn, dan juga sklearn atau scikit learn. Persiapan persiapan alat ini merupakan langkah awal untuk melakukan pengolahan data dan proses analisis clustering.

3.2. Seleksi Data

Pada bagian ini peneliti akan melakukan seleksi data yang nantinya akan digunakan untuk penelitian. Tahapan ini dibagi menjadi beberapa tahapan lain, yaitu memilih data yang akan digunakan, menghitung rasio kasus yang terjadi, menghitung rasio jumlah korban keseluruhan, serta menyimpan data yang sudah didapatkan.

3.3. Pra Proses

Pada tahapan pra-proses ini peneliti akan mulai menggunakan aplikasi jupyter notebook, gambar 3.2 adalah tahapan pra-proses pada penelitian kali ini



Gambar 3.2 Tahapan Pra-Proses

Tahapan ini dimulai dengan input data, pada bagian ini peneliti akan memasukkan data yang sebelumnya sudah disimpan kedalam program dan akan melanjutkan dengan transformasi data terhadap data yang sudah di input tersebut.

3.3.1. Input Data

Tahapan input data akan dikerjakan dengan menggunakan library pandas. Pandas akan membaca file csv dan menyimpan data tersebut kedalam variabel dataset yang akan digunakan. Data tersebut kemudian akan diseleksi dengan memilih tingkat korban dan crime clock dari masing-masing kategori tempat kejadian tindak kekerasan dan dibentuk menjadi sebuah array dan disimpan ke dalam variabel X.

3.3.2. Transformasi Data

Data tingkat korban dan *crime clock* belum memiliki kesamaan nilai, sehingga hal tersebut dapat menyebabkan hasil cluster yang akan dilakukan menjadi berantakan. Untuk mencegah hal tersebut maka data yang akan digunakan harus dilakukan proses normalisasi. Pada penelitian kali ini peneliti akan menggunakan perhitungan *autoscale* atau *standard scaller*. *Autoscale* akan melakukan penyamaan data dengan cara mengurangi nilai yang akan dinormalisasi dengan rata-rata dari keseluruhan data yang ada. Kemudian dibagi dengan standar deviasi keseluruhan data yang akan diukur kemudian diakarkan. Untuk rumus standar deviasi bisa dilihat seperti yang dijelaskan di bawah ini.

$$\text{Standar Deviasi} = \sqrt{\frac{\sum (x_i - \bar{x})^2}{N}}$$

x_i = nilai dari populasi yang dihitung

\bar{x} = Rata-rata nilai populasi

N = jumlah populasi yang dihitung

3.4. Clustering

Tahapan berikutnya adalah melakukan clustering, clustering akan dilakukan dengan menggunakan metode hirarki cluster, dalam melakukan hal ini pengclusteran akan dilakukan dengan menggunakan fungsi atau rumus yang sudah disediakan oleh library scipy, yaitu fungsi hirarki clustering. Berdasarkan teori-teori yang sudah dijabarkan pada bab sebelumnya, setiap objek akan dianggap sebagai satu cluster sehingga menjadi 34 cluster karena pada penelitian kali ini menggunakan data 34 provinsi yang akan dibuat cluster nantinya. Setelah itu langkah selanjutnya adalah mencari jarak terkecil pada setiap cluster dan menghubungkannya sehingga menjadi cluster baru. Sehingga jumlah cluster yang akan

terbentuk nantinya adalah N-1 yaitu 33 cluster yang terbentuk. Setelah itu lakukan perhitungan pada masing-masing cluster sehingga berubah hanya menjadi 1 cluster. Masing-masing dari cluster tersebut akan dihubungkan dengan dendogram, dibawah ini adalah rumus perhitungan dalam mencari jarak pada metode ward.

$$d(u, v) = \frac{|v| + |s|}{T} d(v, s)^2 + \frac{|v| + |t|}{T} d(v, t)^2 - \frac{|v|}{T} d(s, t)^2$$

$d(u, v)$: Jarak Antar dua cluster

$|v|$: Jumlah Objek dalam cluster v

$|s|$: Jumlah Objek dalam cluster s

$|t|$: Jumlah Objek dalam cluster t

T : Jumlah Objek Keseluruhan

$d(v, s)$: Jarak antar cluster v dengan cluster s

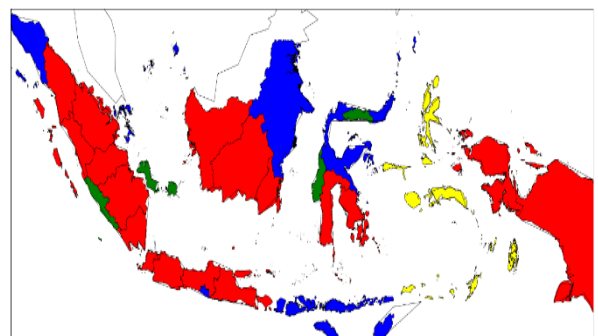
$d(u, v)$: Jarak antar cluster v dengan cluster t

$d(s, t)$: Jarak antar cluster s dengan t

3.5. Visualisasi

Tahap ini akan menampilkan hasil clustering dengan menampilkan grafik peta. Grafik peta inilah yang akan memberikan informasi provinsi apa saja yang ada pada suatu cluster tersebut.

Grafik peta ditampilkan dengan menggunakan library basemap dengan membaca file yang berekstensi *shapefile*. Grafik kemudian dibuat menggunakan library matplotlib. Matplotlib akan melakukan penggambaran berdasarkan koordinat yang disediakan oleh shapefile. Gambar 3.3 adalah contoh grafik peta yang dibuat berdasarkan tindak kekerasan dalam rumah tangga.



Gambar 3.3 Grafik Peta

IV. HASIL PENELITIAN DAN PEMBAHASAN

Berdasarkan hasil penelitian yang sudah didapat dari keseluruhan analisa diatas, sudah didapatkan hasil mengenai tingkat keamanan serta kondisi keamanan dan ketertiban dalam aspek tindakan kekerasan. Pada bagian ini akan menampilkan hasil analisa clustering tersebut secara keseluruhan pada masing masing lingkungan tempat kejadian. Tabel 4.1 menyajikan hasil clustering keseluruhan.

Tabel 4.1 Hasil Clustering Keseluruhan

Lingkungan Tempat Kejadian	Cluster	Jumlah Provinsi	Rata-rata	
			Crime Rate	Crime Clock
Rumah Tangga	1	18	6.707	21.851
	2	8	16.83	16.3487
	3	4	8.945	71.392
	4	4	17.665	47.622
Tempat Kerja	1	26	0.165	1325.706
	2	3	0.636	498.406
	3	4	0.107	5475
	4	1	1.42	1251.42
Sekolah	1	4	0.347	1918.855
	2	16	0.909	371.747
	3	3	2.223	327.703
	4	11	0.37	272.166
Tempat Umum	1	19	1.295	121.618
	2	7	1.364	465.388
	3	7	2.898	138.482
	4	1	5.47	96.26
LPK	1	7	0.054	2815.714
	2	22	0.003	106.181
	3	4	0.025	8760
	4	1	1.54	796.359
Lainnya	1	5	1.812	294.14
	2	3	13.043	56.463
	3	20	2.663	56.221
	4	6	7.229	65.011

Berdasarkan tabel 4.1 diatas dapat diketahui tingkat keamanan serta tingkat keamanan dan ketertiban pada aspek tindak kekerasan dari setiap cluster dengan melihat nilai tingkatan korban tindak kekerasan dan selang waktu kejadian tindak kekerasan. Tabel 4.2 berisi informasi yang bisa didapatkan setelah mengetahui nilai-nilai tersebut.

Tabel 4.2 Tingkat & Kondisi Keamanan masing – masing Cluster

Lingkungan Tempat Kejadian	Cluster	Tingkat Keamanan	Kondisi Keamanan dan Ketertiban
Rumah Tangga	1	T	R
	2	R	R
	3	T	T
	4	R	T
Tempat Kerja	1		
	2		
	3		
	4	T	S
Sekolah	1	S	R
	2	T	T
	3	R	S
	4	T	T
Tempat Umum	1	S	S
	2	R	S
	3	T	S
	4	T	R
LPK	1	T	T
	2	T	R
	3	R	R
	4	T	T
Lainnya	1	T	R
	2	T	T
	3	R	S
	4	T	T

Pada Tabel 4.2 menampilkan tingkatan hasil pada tingkat keamanan serta kondisi keamanan dan ketertiban pada aspek tindak kekerasan pada masing masing tempat kejadian dengan kategori sebagai berikut :

T/Tinggi = Dinilai berdasarkan tingkat korban yang lebih rendah dari cluster lainnya serta selang waktu kejadian yang lebih lama dari cluster lainnya.

S/Sedang = Dinilai berdasarkan tingkat korban yang tidak tinggi namun tidak rendah juga serta selang waktu kejadian yang tidak cepat dan tidak lama dari cluster lainnya.

R/Rendah/Kurang Baik = Dinilai berdasarkan tingkat korban yang lebih tinggi dari cluster lainnya serta selang waktu kejadian yang lebih cepat dari cluster lainnya.

Dengan diketahuinya hasil dari penelitian clustering dan wilayah grafik yang sudah dijabarkan, didapatkan cluster yang paling aman dan paling tidak aman atau paling rendah keamanannya. Tabel 4.3 akan menampilkan cluster apa saja yang termasuk sebagai cluster paling aman di setiap kategori tempat kejadian tindak kekerasan.

Tabel 4.3 Cluster Paling Aman

Tempat Kejadian Tindak Kekerasan	Cluster	Provinsi
Rumah Tangga	3	'Bengkulu', 'Bangka Belitung', 'Sulawesi Barat', 'Gorontalo'
Tempat Kerja	3	'Bengkulu', 'Nusa Tenggara Timur', 'Sulawesi Barat', 'Sulawesi Tenggara'
Sekolah	1	'Bangka Belitung', 'Sulawesi Barat', 'Gorontalo', 'Papua'
Tempat Umum	2	'Bengkulu', 'Bali', 'Kalimantan Utara', 'Sulawesi Barat', 'Gorontalo',

		'Papua Barat', 'Papua'
LPK	3	'Sumatera Barat', 'Kepulauan Riau', 'Bengkulu', 'Sumatera Selatan'
Lainnya	1	'Bengkulu', 'Bangka Belitung', 'Kalimantan Tengah', 'Sulawesi Barat', 'Papua'

Pada tabel 4.3 sudah ditampilkan cluster yang paling aman berdasarkan tempat kejadian tindak kekerasan. Maka selanjutnya akan ditampilkan cluster mana saja yang dicap sebagai cluster paling tidak aman / kurang baik. Tabel 4.4 akan menampilkan informasi tersebut.

Tabel 4.4 Cluster Paling Tidak Aman / Kurang Baik

Tempat Kejadian Tindak Kekerasan	Cluster	Provinsi
Rumah Tangga	2	'Aceh', 'Kepulauan Riau', 'Yogyakarta', 'Nusa Tenggara Barat', 'Nusa Tenggara Timur', 'Kalimantan Timur', 'Sulawesi Tengah', 'Sulawesi Utara'
Tempat Kerja	2	'Riau', 'Kalimantan Tengah', 'Kalimantan Selatan'
Sekolah	3	'Kepulauan Riau', 'Kalimantan Utara', 'Sulawesi Utara'
Tempat Umum	4	'Maluku'
LPK	4	'Kalimantan Utara'

Lainnya	3	'Aceh', 'Sumatera Utara', 'Sumatera Barat', 'Riau', 'Jambi', 'Sumatera Selatan', 'Lampung', 'Banten', 'Jawa Barat', 'Jakarta Raya', 'Jawa Tengah', 'Jawa Timur', 'Bali', 'Nusa Tenggara Barat', 'Nusa Tenggara Timur', 'Kalimantan Barat', 'Kalimantan Selatan', 'Sulawesi Selatan', 'Sulawesi Tenggara', 'Maluku'
---------	---	--

V. KESIMPULAN DAN SARAN

5.1. Kesimpulan

penelitian ini berhasil menampilkan hasil analisa pada setiap cluster. Penelitian ini berhasil menyajikan cluster mana saja yang paling aman serta cluster mana saja yang paling tidak aman di setiap kategori lingkungan terjadinya tindak kejahatan. Provinsi paling aman dipilih berdasarkan provinsi yang memiliki tingkat keamanan serta kondisi keamanan dan ketertiban dalam aspek tindakan kekerasan yang paling tinggi. Sedangkan provinsi paling tidak aman dipilih berdasarkan provinsi yang memiliki tingkat keamanan serta kondisi keamanan dan ketertiban dalam aspek tindakan kekerasan yang paling rendah diantara cluster yang lain pada kategori tersebut. Semisal nya pada kategori lingkungan terjadinya tindak kejahatan di lingkungan rumah tangga memiliki cluster dengan daerah paling aman adalah cluster 3 yang berisi provinsi Bengkulu, Bangka Belitung, Sulawesi Barat, dan Gorontalo. Dan cluster paling tidak aman adalah cluster 2 yang berisi provinsi Aceh, Kepulauan Riau, Yogyakarta, Nusa Tenggara Barat, Nusa Tenggara Timur, Kalimantan Timur, Sulawesi Tengah, dan Sulawesi Utara.

5.2. Saran

Adapun penelitian kali ini masih bisa dikembangkan lagi, dikarenakan data yang digunakan kali ini adalah data tindakan kekerasan yang terjadi di Indonesia pada tahun 2022. Diharapkan pada penelitian selanjutnya menggunakan data yang paling baru seperti 2023, 2024, dan tahun-tahun berikutnya. Selanjutnya metode yang digunakan pada penelitian ini adalah metode ward, diharapkan pada penelitian selanjutnya dapat menggunakan metode-metode yang lain seperti metode *single linkage*, *Complete linkage*, *Average linkage*, dan metode *centroid*.

DAFTAR PUSTAKA

Talakua .M.W, Leleury .Z.A, & Talluta .A.W. (2017). Analisis Cluster Dengan Menggunakan Metode K-Means Untuk Pengelompokan Kabupaten/Kota Di Provinsi Maluku Berdasarkan Indikator Indeks Pembangunan Manusia Tahun 2014 (Vol. 11).

Windasari Renta. (2020). analisis cluster hirarki metode average linkage jumlah kriminalitas.

Syahrudin Akbar N, & Kurniawan Tedi. (2018). Jurnal Dasar Pemrograman Python STMIK.

Anonim. Introduction

<https://docs.scipy.org/doc/scipy/referenc/utorial/general.html>. 10 januari 2023