

## Penerapan Metode OPTICS dan ST-DBSCAN untuk Klasterisasi Data Kesehatan

Siti Hariati Hastuti<sup>1</sup>, Ayu Septiani<sup>1,\*</sup>, Hendrayani<sup>1</sup>, Wiwit Pura Nurmayanti<sup>2</sup>

<sup>1</sup> Program Studi Statistika, Universitas Hamzanwadi, Indonesia

<sup>2</sup> Program Studi Statistika, Universitas Mulawarman, Indonesia

\* Correspondence: ayuseptiani@hamzanwadi.ac.id.

**Copyright:** © 2024 by the authors

Received: 22 April 2024 | Revised: 14 Mei 2024 | Accepted: 5 Mei 2024 | Published: 20 Juni 2024

### Abstrak

*Clustering analysis* merupakan metode statistika yang digunakan untuk melakukan proses pengelompokan satu set objek data. Dalam praktiknya, data yang terkumpul seringkali mengandung derau dan *outlier*, khususnya pada data kesehatan. Hal ini dapat memberikan pengaruh yang tidak proporsional terhadap hasil kluster yang terbentuk. Oleh karena itu, penelitian ini bertujuan untuk mengelompokkan kecamatan di Kabupaten Lombok Timur berdasarkan distribusi tenaga kesehatan menggunakan metode OPTICS dan ST-DBSCAN, kemudian membandingkan keefektifan kedua metode tersebut dalam menangani derau dan *outlier*. Penelitian ini menggunakan analisis kuantitatif dengan 6 variabel yang mewakili data tenaga kesehatan di 21 kecamatan di Lombok Timur. Teknik analisis data yang digunakan meliputi analisis deskriptif dilanjutkan analisis cluster metode OPTICS dan ST-DBSCAN. Hasil pengelompokan metode OPTICS menunjukkan terdapat 5 kluster yang terbentuk yakni kluster-1 memiliki dua anggota kecamatan, kluster-2 memiliki tiga anggota kecamatan, kluster-3 memiliki dua anggota kecamatan, kluster-4 memiliki tiga anggota kecamatan, dan kluster-5 memiliki tujuh anggota kecamatan. Hasil pengelompokan menggunakan ST-DBSCAN menunjukkan hanya terdapat 2 kluster yang terbentuk yakni kluster-1 dengan enam anggota kecamatan dan kluster-2 memiliki empat anggota kecamatan. Berdasarkan hasil validasi internal, metode OPTICS memiliki nilai *silhouette* lebih tinggi dibandingkan metode ST-DBSCAN, hal ini menunjukkan metode yang lebih baik dalam mengelompokkan tenaga kesehatan di Lombok Timur ialah metode OPTICS.

**Kata kunci:** *clustering*; optics; st-dbscan; tenaga kesehatan

### Abstract

*Clustering analysis* is a statistical method used to group a set of data objects. In practical applications, the collected data often contains noise and outliers, particularly in health data, which can significantly influence the resulting clusters. Therefore, this study aims to cluster sub-districts in East Lombok Regency based on the distribution of health workers using the OPTICS and ST-DBSCAN methods, and subsequently compare the effectiveness of the two methods in handling noise and outliers. This research employs quantitative analysis with 6 variables representing data on health workers in 21 sub-districts in East Lombok. The data analysis techniques include descriptive analysis followed by cluster analysis using the OPTICS and ST-DBSCAN methods. The results from the OPTICS clustering method reveal the formation of 5 clusters: cluster-1 comprises two sub-district members, cluster-2 has three sub-district members, cluster-3 consists of two sub-district members, cluster-4 includes three sub-district members, and cluster-5 encompasses seven sub-district members. Conversely, the results from the ST-DBSCAN grouping indicate the formation of only 2 clusters: cluster-1 with six sub-district members and cluster-2 with four sub-district members. Based on internal validation results, the OPTICS method exhibits a higher silhouette value compared to the ST-DBSCAN method, indicating that the OPTICS method is superior for clustering health workers in East Lombok.



**Keywords:** *clustering; health workers; optics; st-dbscan*

## PENDAHULUAN

Perkembangan *Big Data* saat ini terus mengalami kemajuan yang sangat pesat. Penyebab utama dari perkembangan ini adalah meningkatnya kebutuhan manusia akan efisiensi dan kemudahan dalam mengakses data. *Big Data* menawarkan potensi besar untuk berbagai bidang, termasuk sektor kesehatan, di mana analisis data yang akurat dan tepat waktu sangat penting. Salah satu metode yang memudahkan dalam mengakses dan menganalisis data adalah *clustering*. *Clustering* merupakan metode statistika yang bertujuan untuk mengelompokkan satu set objek data sehingga objek dalam satu kelompok memiliki kesamaan yang tinggi tetapi sangat berbeda dengan objek di kelompok lain (Elmayati, 2017).

Masalah kekurangan tenaga kesehatan di berbagai daerah, termasuk di Kabupaten Lombok Timur, menjadi perhatian penting karena berdampak pada kualitas layanan kesehatan (Fahzirah, 2021). Pada awal pandemi COVID-19, jumlah pasien yang sakit lebih banyak dibandingkan dengan jumlah tenaga kesehatan yang tersedia, sehingga masyarakat mengalami kesulitan dalam mendapatkan layanan kesehatan yang memadai. Kekurangan tenaga kesehatan ini menunjukkan adanya kesenjangan antara kebutuhan layanan kesehatan dan ketersediaan sumber daya manusia. Berdasarkan data BPS Provinsi NTB tahun 2022, Kabupaten Lombok Timur memiliki jumlah penduduk terbanyak di provinsi tersebut, yaitu sebanyak 1.325.240 jiwa. Tingginya jumlah penduduk ini menuntut peningkatan kualitas layanan kesehatan dengan jumlah tenaga kesehatan yang memadai. Dalam hal ini, analisis *clustering* dapat digunakan untuk mengidentifikasi daerah-daerah yang paling membutuhkan penambahan tenaga kesehatan dan membantu dalam perencanaan distribusi sumber daya yang lebih efektif.

Beberapa metode *clustering* yang biasa digunakan seperti partisi, hirarki, kisi, dan berbasis model memiliki masalah utama, yakni ketidakmampuan mereka dalam menangani derau dan outlier dalam data (Mustikasari & Salman, 2023). Adanya *outlier* dapat menghasilkan hasil cluster yang tidak sesuai dengan gambaran sebenarnya (Silvi, 2018). Oleh karena itu, diperlukan metode *clustering* yang mampu menangani derau dan *outlier* dengan efektif khususnya pada data kesehatan.

Algoritma berbasis densitas, seperti *Spatio-Temporal Density-Based Clustering Algorithm* (ST-DBSCAN) dan *Ordering Points to Identify the Clustering Structure* (OPTICS), merupakan solusi yang efektif untuk masalah ini (Sholikhah, 2022). Algoritma ST-DBSCAN merupakan pengembangan dari algoritma DBSCAN yang mampu menemukan cluster berdasarkan nilai-nilai spasial dan temporal objek (Nisrina *et al.*, 2022); Ansari *et al.*, 2020). Algoritma ini memiliki kemampuan untuk menemukan cluster dalam data spasial yang besar serta tidak memerlukan asumsi dalam pengerjaannya (Gaonkar & Sawant, 2013; Yu *et al.*, 2022). Di sisi lain, OPTICS melakukan klusterisasi dengan pembelajaran berbasis kepadatan yang tinggi berdasarkan jarak jangkauan dan dapat mendeteksi cluster dengan kepadatan beragam, cocok untuk data campuran serta data berdimensi tinggi (Abid *et al.*, 2017; Anshori, 2022).

Beberapa penelitian sebelumnya telah menunjukkan keunggulan masing-masing metode *clustering*. Ahmed *et al.*, (2020) menemukan bahwa OPTICS lebih efektif dibandingkan K-means dalam konteks tertentu karena kemampuannya untuk menangani data dengan kepadatan yang bervariasi dan menemukan cluster dengan bentuk yang arbitrer. OPTICS tidak memerlukan penentuan jumlah cluster di awal, yang membuatnya lebih fleksibel dalam aplikasi praktis (Ankerst *et al.*, 1999). Sementara itu, Mustikasari & Salman (2023) menunjukkan bahwa DBSCAN lebih efisien dalam implementasinya menggunakan *software R*, karena DBSCAN mampu mendeteksi *outlier* dengan baik dan efektif dalam mengelompokkan data berdimensi tinggi (Ester *et al.*, 1996). Namun, DBSCAN sangat bergantung pada parameter *epsilon* dan *minPts* yang harus ditentukan secara tepat untuk

mendapatkan hasil yang optimal (Schubert *et al.*, 2017). Eriansya & Syafrullah (2018) menemukan bahwa K-Means menghasilkan nilai *Silhouette Coefficient* yang lebih tinggi dibandingkan ST-DBSCAN dalam studi kasus tertentu. Namun, K-Means memiliki keterbatasan dalam menangani data dengan bentuk cluster yang tidak bulat dan sangat sensitif terhadap *outlier* (MacQueen, 1967). Meskipun banyak penelitian telah membahas keunggulan masing-masing metode clustering, penelitian yang secara khusus membandingkan efektivitas ST-DBSCAN dan OPTICS dalam konteks data kesehatan di Kabupaten Lombok Timur masih terbatas, terutama dalam hal analisis distribusi tenaga kesehatan di tingkat kecamatan dan penanganan data yang mengandung derau dan *outlier*. Penelitian ini mengisi kekosongan tersebut dengan memberikan analisis komprehensif mengenai performa kedua metode dalam konteks yang spesifik ini.

Penelitian ini penting dilakukan karena dapat memberikan solusi konkret untuk mengatasi kekurangan tenaga kesehatan di Kabupaten Lombok Timur melalui analisis *clustering* yang akurat. Oleh karena itu, tujuan dari penelitian ini adalah untuk melakukan pengelompokan kecamatan di Kabupaten Lombok Timur berdasarkan jumlah dan distribusi tenaga kesehatan menggunakan metode OPTICS dan ST-DBSCAN, serta membandingkan keefektifan kedua metode tersebut dalam menangani data yang mengandung derau dan *outlier*. Penelitian ini juga bertujuan untuk mengidentifikasi kecamatan-kecamatan yang memiliki jumlah tenaga kesehatan yang belum maksimal dan yang sudah memadai, serta memberikan rekomendasi kebijakan kepada pihak terkait untuk penambahan tenaga kesehatan di kecamatan yang membutuhkan berdasarkan hasil *clustering*. Melalui penelitian ini, diharapkan dapat diperoleh pemahaman yang lebih baik mengenai distribusi tenaga kesehatan di Kabupaten Lombok Timur dan memberikan kontribusi dalam perencanaan kebijakan kesehatan yang lebih tepat sasaran.

## METODE

Penelitian ini menggunakan pendekatan kuantitatif berdasarkan pengumpulan data informasi yang berupa simbol atau angka. Sumber data penelitian diperoleh dari *website* resmi Badan Pusat Statistik Kabupaten Lombok Timur tahun 2023 mengenai data tenaga kesehatan. Populasi dalam penelitian ini adalah seluruh tenaga kesehatan di Kabupaten Lombok Timur. Sampel yang digunakan mencakup data dari 21 kecamatan di Kabupaten Lombok Timur. Aspek atau variabel yang diteliti mencakup jumlah penduduk, jumlah dokter, jumlah bidan, jumlah tenaga kesehatan masyarakat, jumlah tenaga farmasi, dan jumlah tenaga gizi. Variabel-variabel ini dipilih karena masing-masing berperan penting dalam menentukan distribusi tenaga kesehatan yang ideal di setiap kecamatan.

Teknik analisis yang digunakan dalam penelitian ini adalah analisis deskriptif dan analisis kluster menggunakan algoritma OPTICS dan ST-DBSCAN. Tahapan analisis yang dilakukan meliputi beberapa langkah penting. Pertama, dilakukan input data, diikuti dengan eksplorasi data. Selanjutnya, data distandarisasi menggunakan *Z-score* untuk memastikan bahwa semua variabel berada pada skala yang sama (Han *et al.*, 2006). Setelah itu, jarak *Euclidean* dihitung untuk mengukur kedekatan antara objek-objek dalam data tersebut

Analisis *clustering* menggunakan algoritma OPTICS mencakup beberapa langkah. Pada tahap awal, distribusi masing-masing variabel dianalisis untuk mengidentifikasi potensi *outlier*. Selanjutnya, model *clustering* dibangun menggunakan algoritma OPTICS untuk mengidentifikasi struktur kluster dalam data. OPTICS memesan titik data berdasarkan kerapatan lokal mereka untuk menemukan kluster dengan berbagai bentuk dan ukuran tanpa menentukan jumlah kluster sebelumnya. Setelah model *clustering* terbentuk, kerangka kerja untuk visualisasi hasil *clustering* dibuat, seperti plot dendrogram atau plot kerapatan, untuk membantu memahami struktur kluster yang terbentuk (Ester *et al.*, 1996).

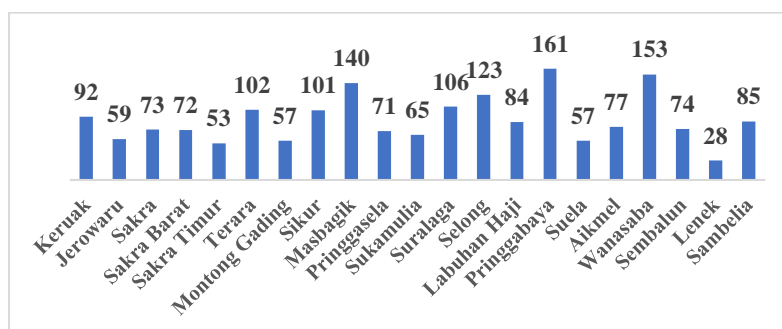
Analisis *clustering* dengan algoritma ST-DBSCAN juga melibatkan beberapa langkah penting. Inisialisasi parameter dilakukan dengan menentukan nilai awal untuk parameter *epsilon* (radius lingkungan) dan *minimum points* (jumlah minimum titik yang diperlukan untuk membentuk kluster). Setelah itu, nilai *epsilon* dan *minimum points* yang terbaik dipilih berdasarkan analisis jumlah *noise* yang paling sedikit dan nilai indeks *silhouette* tertinggi. Model *clustering* kemudian dibangun menggunakan algoritma ST-DBSCAN untuk membentuk kluster dengan mempertimbangkan dimensi spasial dan temporal. Algoritma ini efektif dalam mengidentifikasi kluster di data yang memiliki aspek spasial dan temporal, serta dapat menangani *noise* dengan lebih baik dibandingkan metode *clustering* tradisional (Ester *et al.*, 1996).

Terakhir, evaluasi dilakukan dengan menghitung nilai indeks *silhouette* untuk menilai kualitas kluster yang terbentuk. Indeks *silhouette* mengukur seberapa mirip suatu objek dengan kluster yang sama dibandingkan dengan kluster lain, yang membantu dalam menilai validitas kluster (Rousseeuw, 1987). Indeks ini dihitung untuk menentukan tingkat kepercayaan dalam proses *clustering* pada suatu pengamatan. Indeks *silhouette* juga digunakan untuk membuktikan ketepatan jumlah kelompok yang terbentuk dari suatu algoritma *clustering* (Hastuti *et al.*, 2023).

## HASIL DAN PEMBAHASAN

### Hasil

Tahap awal analisis adalah melakukan analisis deskriptif terhadap data yang digunakan. Analisis ini mencakup pemahaman mengenai sebaran data untuk mendapatkan gambaran awal tentang karakteristik data tersebut. Diagram batang digunakan untuk menggambarkan frekuensi distribusi tenaga kesehatan. Selain itu, boxplot digunakan untuk mengidentifikasi ada tidaknya *outlier* dalam data. Analisis ini membantu dalam menentukan langkah-langkah lanjutan yang diperlukan dalam proses *clustering*.

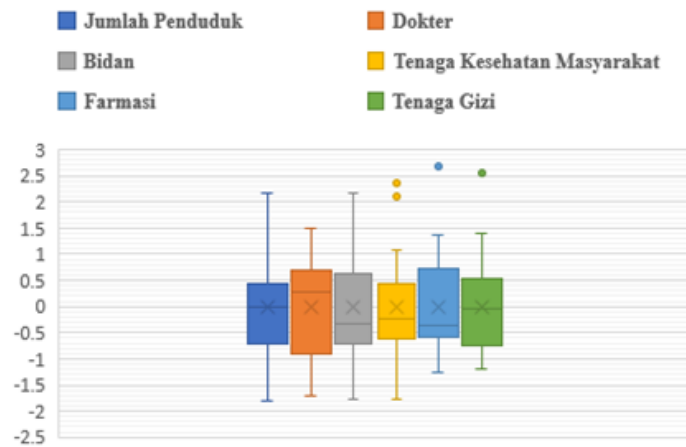


**Gambar 1.** Sebaran tenaga kesehatan setiap kecamatan di kabupaten lombok timur 2023

Berdasarkan Gambar 1, hasil analisis sebaran tenaga kesehatan di Kabupaten Lombok Timur pada tahun 2023 menunjukkan bahwa Kecamatan Pringgabaya memiliki jumlah tenaga kesehatan tertinggi, yaitu sebanyak 161 tenaga kesehatan. Tingginya jumlah tenaga kesehatan di Kecamatan Pringgabaya dapat dijelaskan oleh jumlah penduduk yang juga tertinggi di Kabupaten Lombok Timur, mencapai 110.813 jiwa berdasarkan data dari Badan Pusat Statistik (BPS) Lombok Timur. Sebaliknya, Kecamatan Lenek memiliki jumlah tenaga kesehatan terendah, yaitu sebanyak 28 tenaga kesehatan. Jumlah tenaga kesehatan yang rendah di Kecamatan Lenek menunjukkan bahwa kecamatan ini mungkin memiliki kebutuhan layanan kesehatan yang lebih rendah dibandingkan kecamatan lainnya, atau disebabkan oleh distribusi sumber daya kesehatan yang kurang merata.

Gambar 2 menunjukkan menunjukkan distribusi tenaga kesehatan di Kabupaten Lombok Timur melalui representasi *boxplot*. Pada gambar ini, terlihat adanya *outlier* pada variabel

tenaga farmasi, tenaga kesehatan masyarakat, dan tenaga gizi, yang mengindikasikan variasi yang signifikan dalam distribusi tenaga kesehatan di kecamatan tersebut. *Outlier* ini menunjukkan adanya kecamatan dengan jumlah tenaga kesehatan yang jauh lebih tinggi dibandingkan dengan kecamatan lainnya. Sebelum analisis kluster dilakukan, data asli distandarisasi, sebagaimana ditunjukkan pada Tabel 1. Standarisasi ini bertujuan untuk menyamakan skala variabel, memudahkan perbandingan, dan memastikan akurasi dalam analisis kluster. Nilai-nilai dari hasil standarisasi ini merefleksikan skala standar dengan rata-rata nol dan simpangan baku satu.



Gambar 2. Boxplot tenaga kesehatan

Tabel 1. Hasil standarisasi data

No	Kecamatan	Jumlah penduduk k	Dokter	Bidan	Kesehatan Masyarakat	Farmasi	Tenaga Gizi
1	Keruak	-0,245	-0,114	0,117	1,075	-0,374	-0,328
2	Jerowaru	-0,077	-1,719	-0,702	-0,482	-0,374	-1,189
21	Sambelia	-1,1706	-0,286	-0,037	-0,370	-0,374	-0,041

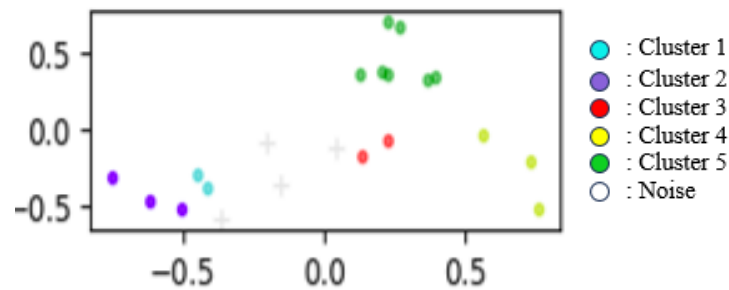
Klasterisasi menggunakan algoritma OPTICS dengan *minimum points* = 2 dan *alpha* = 0,5 menghasilkan lima *cluster* dan satu *noise*, seperti yang ditunjukkan pada Tabel 2. *Cluster-5* memiliki anggota terbanyak dengan tujuh kecamatan, sementara *Cluster-1* dan *Cluster-3* memiliki anggota terendah dengan masing-masing hanya memiliki dua kecamatan. Analisis dari Tabel 3 mengungkapkan bahwa *Cluster-4* memiliki jumlah tenaga kesehatan tertinggi dengan tiga kecamatan yang memiliki nilai rata-rata semua variabel lebih tinggi dibandingkan kluster lainnya. Ini menunjukkan bahwa kecamatan dalam *Cluster-4* memiliki jumlah tenaga kesehatan yang paling belum maksimal dan membutuhkan penambahan tenaga kesehatan.

Tabel 2. Label *cluster* algoritma optics

Label <i>cluster</i>	Jumlah Anggota	Anggota <i>cluster</i>
<i>Noise</i>	4	Keruak, Montong Gading, Sikur, Sakra
<i>Cluster-1</i>	2	Selong, Masbagik
<i>Cluster-2</i>	3	Sakra Barat, Suralaga, Terara
<i>Cluster-3</i>	2	Pringgasela, Sukamulia
<i>Cluster-4</i>	3	Jerowaru, Suela, Sakra Timur
<i>Cluster-5</i>	7	Pringgabaya, Sembalun, Aikmel, Wanasaba, Lenek, Sambelia, Labuhan Haji

**Tabel 3.** Rata-rata cluster optics

	Banyak penduduk	Dokter	Bidan	Kesehatan Masyarakat	Tenaga Farmasi	Tenaga Gizi
Cluster-1	-0,016	-0,916	-0,538	-0,374	-0,612	-0,758
Cluster-2	-0,049	-0,515	-0,292	-0,083	-0,309	-0,519
Cluster-3	-0,608	0,085	-0,661	-0,374	-0,352	-0,328
Cluster-4	0,490	0,687	0,540	0,645	-0,222	0,437
Cluster-5	-0,277	0,057	0,070	0,062	0,148	0,246



**Gambar 3.** Visualisasi kluster algoritma optics

Hasil kluster pada gambar 3 menunjukkan bahwa metode OPTICS menghasilkan lima *cluster* dengan beberapa *outlier*. Metode OPTICS menerapkan konsep *reachability distance* dan tidak mempertimbangkan nilai *epsilon* sehingga metode OPTICS mampu menjangkau objek-objek yang berpotensi untuk dikategorikan sebagai cluster. Pada metode OPTICS, rata-rata nilai cluster yang tinggi mengindikasikan kelompok pengamatan yang memiliki jumlah penduduk yang tinggi, namun memiliki jumlah tenaga kesehatan yang relatif rendah. Ini menunjukkan adanya kesenjangan antara kebutuhan layanan kesehatan yang tinggi dengan ketersediaan tenaga kesehatan yang kurang memadai dalam kelompok tersebut. Dengan kata lain, meskipun terdapat banyak penduduk yang memerlukan layanan kesehatan, namun jumlah tenaga kesehatan yang tersedia belum cukup untuk memenuhi kebutuhan tersebut.

**Tabel 4.** Label cluster algoritma st-dbscan

Label cluster	Jumlah Anggota	Anggota cluster
Noise	11	Keruak, Terara, Masbagik, Sukamulia, Selong, Pringgabaya, Aikmel, Wanasaba, Sembalun, Lenek, Sambelia
Cluster-1	6	Jerowaru, Sakra, Sakra Barat, Sakra Timur, Montong Gading, Suela
Cluster-2	4	Sikur, Pringgasela, Suralaga, Labuhan Haji.

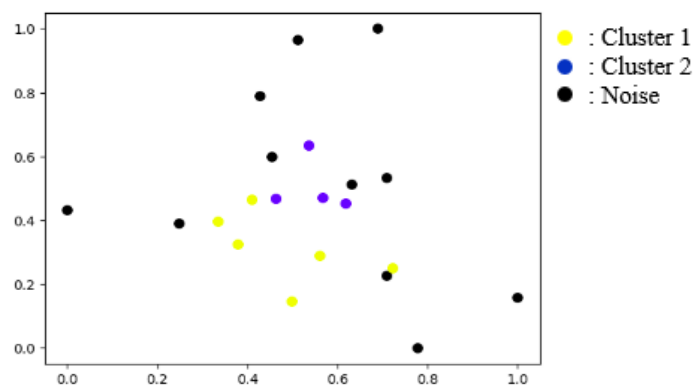
Adapun pengelompokan dengan algoritma ST-DBSCAN sangat tergantung pada peentuan parameter. Parameter ini ialah *epsilon* 1 (jarak atribut spasial), *epsilon* 2 (jarak atribut temporal), *minimum points* (jumlah minimum objek yang berada dalam jarak *epsilon* 1 dan *epsilon* 2) dan  $\Delta\epsilon$  adalah batas selisih nilai non-spasial. Pada penelitian ini menggunakan parameter *epsilon* 1 = 0.40 , *epsilon* 2 = 8, *minimum points* = 2 dan  $\Delta\epsilon$  = 2 yang ditentukan oleh peneliti untuk mendapatkan hasil *cluster* yang optimal. Hasil klusterisasi algoritma ST-DBSCAN yang dapat dilihat pada tabel 4. Berdasarkan Tabel 4, dapat dilihat bahwa terbentuk dua kluster/kelompok serta satu *noise* pada tenaga kesehatan di Kabupaten Lombok Timur

menggunakan algoritma ST-DBSCAN. Anggota dengan cluster terbanyak terdapat pada cluster-1 dengan total enam kecamatan sedangkan cluster-2 memiliki empat anggota kecamatan. Analisis identifikasi dari cluster ST-DBSCAN dapat dilihat pada tabel 5.

**Tabel 5.** Rata-rata cluster st-dbscan

	Banyak Penduduk	Dokter	Bidan	Kesehatan Masyarakat	Farmasi	Tenaga Gizi
<i>Cluster-1</i>	-0,36	-0,98	-0,70	-0,73	-0,39	-0,75
<i>Cluster-2</i>	0,21	0,58	0,11	-0,06	-0,61	0,31

Berdasarkan tabel 5, jumlah tenaga kesehatan tertinggi berada di *cluster 2* dengan empat kecamatan dan memiliki nilai rata-rata semua variabel lebih banyak dibandingkan *cluster 1*. Hal ini dapat dikatakan bahwa pada kecamatan-kecamatan tersebut belum maksimal dalam jumlah tenaga kesehatan. Adapun cluster dengan nilai rata-rata rendah terdapat pada cluster 1 yang terdiri dari enam kecamatan. Hal ini menunjukkan bahwa kecamatan pada cluster tersebut memiliki tenaga kesehatan yang memadai.



**Gambar 4.** Visualisasi kluster algoritma st-dbscan

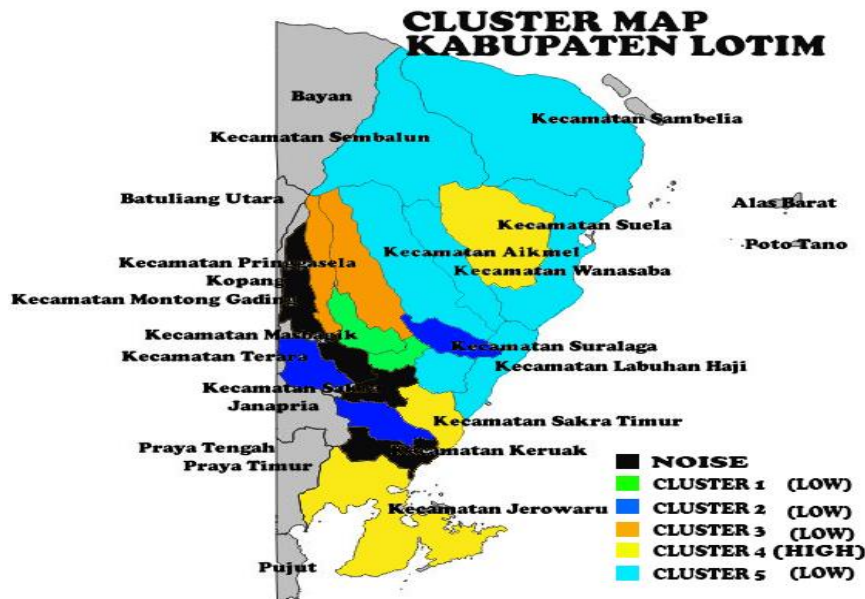
Selanjutnya, hasil pada gambar 4 terdapat dua cluster yang dihasilkan oleh algoritma ST-DBSCAN. *Cluster* pertama adalah titik yang berwarna kuning dengan enam anggota kecamatan dan cluster kedua adalah titik yang berwarna biru dengan empat kecamatan. Adapun titik yang berwarna hitam merupakan *noise* yang berjumlah sebelas kecamatan. Sementara itu, untuk melihat jumlah cluster optimal dapat dilakukan dengan menguji jumlah cluster pada masing-masing algoritma. Perbandingan metode cluster terbaik dapat dilihat dari validasi internal pada tabel 6. Berdasarkan Tabel 6, didapatkan hasil nilai *silhouette* tertinggi terdapat pada metode OPTICS sebesar 0.607 dibandingkan dengan metode DBSCAN. Sehingga dapat disimpulkan bahwa metode yang optimal untuk mengelompokkan tenaga kesehatan di Kabupaten Lombok Timur yaitu metode OPTICS. Hal ini dikarenakan metode OPTICS memiliki performa dalam mendeteksi *outlier* dan memiliki nilai *silhouette* paling tinggi.

**Tabel 6.** Perbandingan algoritma *cluster*

Metode	Jumlah Cluster	<i>Silhouette</i>
OPTICS	5	0,607
ST-DBSCAN	2	0,329

Selanjutnya hasil yang disajikan pada gambar 5 menunjukkan bahwa lima *cluster* terbagi menjadi dua kategori yaitu rendah dan tinggi. *Cluster* pertama, kedua, ketiga, dan kelima menunjukkan cluster dengan kategori rendah. Sementara cluster keempat merupakan cluster

dengan kategori tinggi atau yang memiliki data tenaga kesehatan yang masih kurang. Sementara *noise* merupakan cluster yang tidak masuk kedalam *cluster* manapun.



**Gambar 5.** Peta pengelompokan tenaga kesehatan dengan algoritma optics

## Pembahasan

Hasil penelitian menunjukkan bahwa penggunaan algoritma OPTICS dan ST-DBSCAN dalam mengelompokkan jumlah tenaga kesehatan di Kabupaten Lombok Timur menghasilkan temuan yang signifikan. Berdasarkan hasil klasterisasi menggunakan OPTICS, terbentuk lima kluster/kelompok dan satu *noise*, sementara penggunaan ST-DBSCAN menghasilkan dua kluster/kelompok dan satu *noise*. Analisis lebih lanjut terhadap kluster yang dihasilkan menunjukkan perbedaan yang signifikan antara keduanya.

Pertama, metode OPTICS mampu mengidentifikasi lima kluster dengan lebih baik daripada ST-DBSCAN, yang hanya mampu menghasilkan dua kluster. Hal ini disebabkan oleh konsep reachability distance yang diterapkan oleh OPTICS, yang memungkinkan metode ini untuk menjangkau objek-objek yang berpotensi untuk dikategorikan sebagai kluster tanpa mempertimbangkan nilai *epsilon*. Sebaliknya, ST-DBSCAN sangat bergantung pada penentuan parameter seperti *epsilon* 1, *epsilon* 2, *minimum points*, dan  $\Delta\epsilon$ , yang dapat memengaruhi hasil klasterisasi secara signifikan.

Kedua, hasil klasterisasi OPTICS menunjukkan adanya kesenjangan antara kebutuhan layanan kesehatan yang tinggi dengan ketersediaan tenaga kesehatan yang kurang memadai dalam beberapa kelompok pengamatan. Cluster dengan nilai rata-rata yang tinggi mengindikasikan kelompok pengamatan yang memiliki jumlah penduduk yang tinggi, namun memiliki jumlah tenaga kesehatan yang relatif rendah. Sebaliknya, ST-DBSCAN tidak menunjukkan kesenjangan yang sama secara jelas.

Terakhir, nilai Indeks *silhouette* tertinggi terdapat pada metode OPTICS, menunjukkan bahwa metode ini lebih optimal dalam mengelompokkan data tenaga kesehatan di Kabupaten Lombok Timur dibandingkan dengan ST-DBSCAN. Nilai *Silhouette* yang tinggi mengindikasikan bahwa kluster yang dihasilkan oleh OPTICS lebih padat dan terpisah secara jelas dibandingkan dengan ST-DBSCAN.

Hasil penelitian juga menunjukkan bahwa metode klasterisasi seperti OPTICS dapat menjadi alat yang efektif dalam mengidentifikasi pola distribusi tenaga kesehatan di tingkat kecamatan. Berdasarkan nilai *Indeks Silhouette* tertinggi yang diperoleh, OPTICS mampu menghasilkan kluster yang lebih optimal dan informatif dibandingkan dengan ST-DBSCAN.



Temuan ini sejalan dengan penelitian sebelumnya yang menunjukkan keunggulan metode berbasis kepadatan dalam menangani data dengan *outlier* dan *noise*, namun penelitian ini menambahkan kontribusi baru dengan membandingkan langsung performa OPTICS dan ST-DBSCAN dalam konteks distribusi tenaga kesehatan.

Penelitian sebelumnya yang menggunakan ST-DBSCAN menunjukkan bahwa algoritma ini efektif dalam menangani *outlier*, tetapi kurang optimal dalam interpretasi kluster yang terbentuk (Eriansya & Syafrullah, 2018; Mustikasari & Salman, 2023; Schubert et al., 2017). Dalam penelitian ini, OPTICS tidak hanya menangani *outlier* dengan baik tetapi juga memberikan kluster yang lebih jelas dan terdefinisi, yang ditunjukkan oleh nilai *Indeks Silhouette* yang lebih tinggi. Dengan demikian, penelitian ini menegaskan bahwa OPTICS lebih unggul dalam menganalisis distribusi tenaga kesehatan di tingkat kecamatan, memberikan dasar kuat untuk pengembangan kebijakan distribusi tenaga kesehatan yang lebih efektif dan membuka jalan bagi penelitian lebih lanjut dalam penggunaan algoritma klusterisasi untuk data kesehatan yang kompleks.

## SIMPULAN

Hasil penelitian menunjukkan bahwa metode OPTICS dan ST-DBSCAN efektif dalam mengidentifikasi kecamatan dengan jumlah tenaga kesehatan yang belum maksimal dan yang sudah memadai. Dalam perbandingan keefektifan, OPTICS menunjukkan hasil yang lebih konsisten dalam mengatasi derau dan *outlier* dengan Indeks *silhouette* yang lebih tinggi dibandingkan ST-DBSCAN. Analisis ini mengidentifikasi beberapa kecamatan yang sangat membutuhkan penambahan tenaga kesehatan. Berdasarkan hasil *clustering* ini, rekomendasi kebijakan dapat diberikan kepada pihak terkait untuk meningkatkan distribusi tenaga kesehatan di kecamatan-kecamatan yang membutuhkan, sehingga dapat memperbaiki kualitas layanan kesehatan di Kabupaten Lombok Timur.

## REFERENSI

- Abid, A., Masmoudi, A., Kachouri, A., & Mahfoudhi, A. (2017). Outlier Detection in Wireless Sensor Networks Based on OPTICS Method for Events and Errors Identification. *Wireless Personal Communications*, 97(1), 1503–1515. <https://doi.org/10.1007/s11277-017-4583-7>
- Ahmed, M. A., Baharin, H., & Nohuddin, P. N. E. (2020). Analysis of K-means, DBSCAN and OPTICS Cluster Algorithms on Al-Quran Verses. *International Journal of Advanced Computer Science and Applications*, 11(8), 248–254. <https://doi.org/10.14569/IJACSA.2020.0110832>
- Ankerst, M., Breunig, M. M., Kriegel, H.-P., & Sander, J. (1999). OPTICS: Ordering points to identify the clustering structure. *ACM Sigmod Record*, 28(2), 49–60. <https://doi.org/10.1145/304181.304187>
- Ansari, M. Y., Ahmad, A., Khan, S. S., Bhushan, G., & Mainuddin. (2020). Spatiotemporal clustering: A review. *Artificial Intelligence Review*, 53, 2381–2423. <https://doi.org/10.1007/s10462-019-09736-1>
- Anshori, M. L. (2022). *Implementasi Metode Ordering Points To Identify The Clustering Structure (Optics) Dengan Algoritma Db-Scan Dalam Pengelompokan Pemesanan Indihome (Studi Kasus: Kluster Data Pemesanan Pemasangan Indihome Berdasarkan Jenis Fiber 3p Dan Fiber 2p Di Provinsi Lampung Pada Tahun 2020)*. [Unpublished Undergraduate Thesis] Universitas Islam Indonesia.
- Elmayati, E. (2017). Data Mining dengan Metode Clustering untuk Pengolahan Informasi Persediaan Obat pada Klinik Srikandi Medika Berbasis Web. *Jurnal Pelita Informatika*, 6(2), 159–164.

- Eriansya, M. I. P., & Syafrullah, M. (2018). Implementasi Algoritma ST-DBSCAN dan K-MEANS Untuk Pengelompokan Indeks Pembangunan Manusia Kabupaten/Kota Pulau Jawa Tahun 2014-2016 Berbasis Web Di Badan Pusat Statistik. *SKANIKA: Sistem Komputer Dan Teknik Informatika*, 1(3), 1026–1032.
- Ester, M., Kriegel, H.-P., Sander, J., & Xu, X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. *Kdd*, 96(34), 226–231.
- Fahzirah, R. (2021). Analisis Ketersediaan Tenaga Kesehatan Berdasarkan Standar Ketenagaan Minimal PERMENKES Nomor 43 Tahun 2019 di Puskesmas Tanjung Rejo Kecamatan Percut Sei Tuan Tahun 2021. *Journal Transformation of Mandalika*, 2(3), 106–111.
- Gaonkar, M. N., & Sawant, K. (2013). AutoEpsDBSCAN: DBSCAN with Eps automatic for large dataset. *International Journal on Advanced Computer Theory and Engineering*, 2(2), 11–16.
- Han, J., Kamber, M., & Mining, D. (2006). Concepts and techniques. *Morgan Kaufmann*, 340, 94104–3205.
- Hastuti, S. H., Nurmawanti, W. P., & Saputri, A. A. (2023). Penerapan Metode Clustering Self Organizing Maps (SOM) dan K-Affinity Propagation (K-AP) dalam Mengelompokkan Nilai Tukar Petani di Indonesia 2022. *VARIANCE: Journal of Statistics and Its Applications*, 5(1), 79–88. <https://doi.org/10.30598/variancevol5iss1page79-88>
- MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, 1(14), 281–297.
- Mustikasari, M., & Salman, N. (2023). Density-Based Clustering Analysis with DBSCAN and OPTICS. *Jurnal INSYPRO (Information System and Processing)*, 8(1), 1–8. <https://doi.org/10.24252/insypro.v8i1.36347>
- Nisrina, S., Nurmawanti, W. P., & Gazali, M. (2022). Penerapan Metode Clustering SOM dan DBSCAN dalam Mengelompokkan Unmet Need Keluarga Berencana di Nusa Tenggara Barat. *J Statistika: Jurnal Ilmiah Teori dan Aplikasi Statistika*, 15(2), 237–244. <https://doi.org/10.36456/jstat.vol15.no2.a5549>
- Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 20, 53–65. [https://doi.org/10.1016/0377-0427\(87\)90125-7](https://doi.org/10.1016/0377-0427(87)90125-7)
- Schubert, E., Sander, J., Ester, M., Kriegel, H. P., & Xu, X. (2017). DBSCAN revisited, revisited: Why and how you should (still) use DBSCAN. *ACM Transactions on Database Systems (TODS)*, 42(3), 1–21. <https://doi.org/10.1145/3068335>
- Sholikhah, N. A. (2022). Studi Perbandingan Clustering Kecamatan di Kabupaten Bojonegoro Berdasarkan Keaktifan Penduduk Dalam Kepemilikan Dokumen Kependudukan. *Jurnal Statistika Dan Komputasi*, 1(1), 42–53. <https://doi.org/10.32665/statkom.v1i1.443>
- Silvi, R. (2018). Analisis Cluster dengan Data Outlier Menggunakan Centroid Linkage dan K-Means Clustering untuk Pengelompokan Indikator HIV/AIDS di Indonesia. *Jurnal Matematika "MANTIK"*, 4(1), 22–31. <https://doi.org/10.15642/mantik.2018.4.1.22-31>
- Yu, H., Lu, N., Fu, B., Zhang, L., Wang, M., & Tian, H. (2022). Hotspots, co-occurrence, and shifts of compound and cascading extreme climate events in Eurasian drylands. *Environment International*, 169, 107509. <https://doi.org/10.1016/j.envint.2022.107509>