

ARTIKEL
PENERAPAN METODE DBSCAN
UNTUK JUMLAH SARANA KEBERSIHAN
BERDASARKAN BANYAKNYA TIMBUNAN SAMPAH
DI KABUPATEN LOMBOK TIMUR TAHUN
2023



AHMAD ABDA'I TO'MIN FAHRUDDIN
NPM. 180401001

PROGRAM STUDI STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS HAMZANWADI
2023

HALAMAN PERSETUJUAN

**PENERAPAN METODE DBSCAN UNTUK JUMLAH SARANA
KEBERSIHAN BERDASARKAN BANYAKNYA TIMBUNAN SAMPAH DI
KABUPATEN LOMBOK TIMUR TAHUN 2023**

ARTIKEL

Yang Diajukan Oleh :

AHMAD ABDA'I TO'MIN FAHRUDDIN

180401001

Telah Disetujui Oleh :

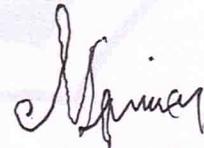
Pembimbing I



Kertanah, M.Si

NIDN. 0831129010

Pembimbing II



Ayu Septiani, M.Si

NIDN. 0812099203

**PENERAPAN METODE DBSCAN UNTUK JUMLAH SARANA
KEBERSIHAN BERDASARKAN BANYAKNYA TIMBUNAN SAMPAH DI
KABUPATEN LOMBOK TIMUR TAHUN 2023**

Kertanah¹, Ayu Septiani², Ahmad Abda'i To'min Fahrudin³

^{1,2,3} Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas
Hamzanwadi

Abstract

The development of Bigdata at this time continues to progress very rapidly. One easy way to access data and information can be using the clustering method. Clustering was a statistical method with the purpose of grouping a set of data objects so that objects in a group have high similarity (or have a relationship) but are very different from objects in other groups. The clustering method was applied to all fields, including data on the number of cleaning suggestions and the amount of garbage. The clustering method used in this research was DBSCAN. The purpose of this study is first to determine the data description of the number of cleaning facilities in East Lombok Regency; and the second is to determine the clustering results of the DBSCAN method for data on the number of cleaning facilities in East Lombok Regency. The results of the DBSCAN method with epsilon 1.0 and Minimum Points 2 obtained 2 clusters, namely cluster 1 with 11 sub-districts where the average cause of waste accumulation is low and cluster 2 with 3 sub-districts where the average cause of waste accumulation is high.

Keywords: Clustering, DBSCAN, Hygiene Facilities

1. Pendahuluan

Sarana kebersihan mempunyai peranan penting untuk menurunkan jumlah timbunan sampah dengan tujuan meningkatkan kesadaran masyarakat yang rendah dalam menjaga kebersihan lingkungan, oleh karena itu perlunya sarana Kebersihan yang memadai pada setiap wilayah demi terwujudnya lingkungan yang sehat dan bersih. Semakin lengkap sarana kebersihan maka kebersihan lingkungan akan lebih baik (Mokodompis *et al.*, 2019:4). Berdasarkan data dari

BPS Kabupaten Lombok Timur tahun 2023 (<https://lomboktimurkab.bps.go.id/>). Penumpukan sampah di Kabupaten Lombok timur masih sangat tinggi pada setiap kecamatan, Salah satu penyebabnya yaitu berdasarkan peningkatan jumlah penduduk dan sarana kebersihan yang kurang memadai. Data tentang timbunan sampah perlu dianalisis karena sampah merupakan sisa kegiatan manusia yang tidak diinginkan lagi, baik sifatnya organik maupun non organik yang bersifat dapat terurai dan tidak dapat terurai yang menimbulkan pencemaran lingkungan (Kai *et al.*, 2018:1)

Perkembangan *bigdata* dari tahun ke tahun terus mengalami peningkatan secara pesat. Penyebab dari perkembangan *bigdata* adalah kebutuhan manusia akan efisiensi dan kemudahan dalam mengakses data. Salah satu cara dari mengakses data dan informasi adalah menggunakan metode *cluster*. Menurut Maukar *et al.*, (2022:143). *Clustering* adalah pengelompokan sejumlah data atau objek ke dalam *cluster* (group) agar setiap *cluster* tersebut berisi data yang semirip mungkin dan berbeda dengan data atau objek dalam *cluster* lainnya. Analisis *cluster* dapat diklasifikasikan kedalam lima jenis utama yaitu metode partisi, hirarki, kisi dan berbasis model (Mustikasari & Salman, 2023:2). Metode partisi, hirarki, kisi dan berbasis model sangat terpengaruh oleh adanya drau dan *outlier* dalam data (Mustikasari & Salman, 2023:2). Oleh karena itu di perlukan metode cluster yang mampu menangani drau dan *outlier* yaitu menggunakan algoritma berbasis kepadatan atau desitas yang biasa kita kenal dengan *Density-Based Spatial clustering Algorithm with Noise* (DBSCAN).

Density-Based Spatial clustering Algorithm with Noise (DBSCAN) merupakan salah satu metode yang membentuk *cluster* berdasarkan kepadatan yang cukup tinggi ke dalam suatu *cluster*, sementara untuk objek yang tidak masuk ke dalam suatu kelompok manapun dianggap sebagai data penciran atau *noise* (Nisrina *et al.*, 2022:238). Metode *clustering* dapat diterapkan hampir pada semua bidang, salah satunya data jumlah timbunan sampah yang ada di setiap kecamatan di Kabupaten Lombok Timur..

Semakin banyak penduduk maka peningkatan jumlah timbunan sampah semakin banyak oleh sebab itu sarana kebersihan sangat dibutuhkan. Berdasarkan hal tersebut, maka penelitian ini bertujuan untuk melakukan pengelompokkan terhadap Kecamatan yang memiliki sarana kebersihan yang masih kurang sehingga data tersebut dapat digunakan untuk peningkatan kualitas kebersihan. Adapun metode yang digunakan dalam pengelompokkan kecamatan yang memiliki jumlah sarana kebersihan di Kabupaten Lombok Timur adalah metode DBSCAN.

Beberapa penelitian sebelumnya terkait dengan studi kasus yang sama diantaranya penelitian pertama yang dilakukan oleh (Magriaty, 2022) dengan judul” Analisis *K-Means Cluster* untuk Identifikasi Kawasan Pengelolaan Sampah di Kabupaten Tapin Provinsi Kalimantan Selatan” didapatkan hasil Pengembangan zona pengelolaan persampahan di Kabupaten Tapinakan dilakukan pada kawasan prioritas yaitu Zona Tipe 1 dan Tipe 2 dengan cakupan 73.33% daris eluruh Desa/Kelurahan yang ada di Kabupaten Tapin untuk perluasan cakupan pelayanan. Pengelolaan sampah secara mandiri oleh

masyarakat dilakukan pada Zona Tipe 3 yang memiliki cakupan 26.67% dari wilayah di Kabupaten Tapin. Penelitian kedua yang dilakukan oleh (Alia, 2023) dengan judul “Klasifikasi Jenis Sampah Menggunakan *Image Classification Convolutional Neural Network*” diperoleh hasil penerapan model *image classification convolutional neural network* terhadap klasifikasi jenis sampah adalah baik dengan nilai akurasi penerapan model berdasarkan data prediktif mencapai 89.44% hingga 92.74% dan tingkat keberhasilan model dalam mengklasifikasi secara benar adalah 93.33% dari 15 data citra baru. Penelitian ketiga yang dilakukan oleh (Nurrohman *et al.*, 2023) dengan judul “Sistem *Klasterisasi* Volume Sampah Organik di Kota Magelang menggunakan *K-Means*” diperoleh hasil jumlah *cluster* yang optimal adalah 2 *cluster*, *cluster* pertama merupakan kategori penghasil volume sampah rendah, sedangkan *cluster* kedua merupakan kategori penghasil volume sampah tinggi. Setelah diperoleh hasil *clustering* dengan menggunakan algoritma *K-Means*, maka dilakukan evaluasi terhadap hasil tersebut dengan menggunakan metode skor *silhouette* yang menghasilkan nilai skor sebesar 0,66.

Berdasarkan hal tersebut peneliti mencoba melakukan penelitian lanjutan guna melengkapi penelitian sebelumnya dengan judul “Penerapan Metode DBSCAN untuk Jumlah Sarana Kebersihan Berdasarkan Banyaknya Timbunan Sampah di Kabupaten Lombok Timur Tahun 2023”.

2. Metodologi

2.1 *Clustering*

Clustering merupakan suatu proses pembagian data ke dalam kelas atau disebut *cluster* yang berdasarkan tingkat kemiripan, dalam *cluster*, data-data yang mempunyai kemiripan akan digabungkan ke dalam *cluster* yang sama, begitu juga sebaliknya data yang tidak mempunyai kemiripan akan dimasukkan ke dalam *cluster* yang berbeda (Irawan, 2019:14). Analisis *cluster* mengklasifikasi objek sehingga setiap objek yang paling dekat kesamaannya dengan objek lain berada dalam *cluster* yang sama. *Cluster* yang baik adalah *cluster* yang terbentuk memiliki kriteria sebagai berikut (Mhlwatika & Aderibigbe, 2018:25)

1. Homogenitas (kesamaan) yang tinggi antara anggota dalam satu *cluster* (*within cluster*)
2. Heterogenitas (perbedaan) yang tinggi antara *cluster* yang satu dengan *cluster* lain (*between-cluster*).

Terdapat beberapa kebutuhan *clustering* dalam *data mining* seperti untuk menangani tipe atribut yang berbeda, skalabilitas, menangani dimensionalitas yang tinggi, mampu menangani data yang mempunyai *noise* serta dapat dijelaskan dan diinterpretasikan dengan mudah (Defiyanti *et al.*, 2017:257)

2.2 Euclidean Distance

Euclidean distance atau jarak euclid adalah tipe pengukuran jarak yang digunakan dalam analisis cluster yang paling umum sebagai pengukur jarak dari obyek data ke pusat *cluster*. Semakin dekat jarak maka semakin dekat atau mirip suatu obyek tersebut (Dani *et al.*, 2021). *euclidean space* diperkenalkan oleh seorang matematikawan dari Yunani sekitar tahun 300,

untuk mempelajari hubungan antara sudut dan jarak *euclidean* ini biasanya diterapkan pada 2 dimensi, kemudian juga bisa sederhana jika diterapkan pada dimensi lain yang lebih tinggi (Suparmi & Soeheri, 2020:107).

Untuk mengukur tingkat kemiripan data dengan *euclidean distance* digunakan rumus berikut:

$$d(x, y) = |x - y| = \sqrt{\sum_{i=1}^n (x_{ik} - y_{ik})^2} \quad (1)$$

yang dimana:

d = jarak antara x dan y

x = data pusat *cluster*

y = data pada atribut

i = setiap data

n = jumlah data,

x_i = data pada pusat *cluster* ke- k

2.3 Density-Based spatial Clustering of Application with Noise (DBSCAN)

Density-Based Spatial Clustering of Applications with Noise (DBSCAN) adalah algoritma dasar untuk *clustering* berbasis kepadatan. Konsep yang dimaksud dalam DBSCAN adalah banyaknya data (*Minimum Points*) yang berada dalam radius *Epsilon* (ϵ) dan menghasilkan tiga macam status dari setiap data, yaitu inti (*core*), batas (*border*), dan (*noise*) (Nisrina *et al.*, 2022:238). DBSCAN adalah algoritma dasar untuk *clustering* berbasis kepadatan. DBSCAN dapat menemukan *cluster* dengan berbagai bentuk dan

ukuran dari sejumlah besar data, yang mengandung *noise* dan *outlier* (Cinderatama *et al.*, 2022:174). Algoritma DBSCAN menggunakan dua parameter:

- a. *Minimum Points*: Jumlah minimum titik (ambang batas) yang dikelompokkan bersama untuk suatu wilayah yang dianggap padat.
- b. *epsilon* (ϵ): Ukuran jarak yang akan digunakan untuk menemukan titik-titik ke tetangga terdekatnya.

Berikut algoritma DBSCAN menurut (Nisrina *et al.*, 2022:239) :

- 1) Inisialisasi parameter *Minimum Points* dan *Epsilon*.
- 2) Tentukan titik awal atau p secara acak.
- 3) Hitung *Epsilon* atau semua jarak titik atau kepadatan yang dapat dijangkau terhadap p .
- 4) Jika titik yang memenuhi *Epsilon* lebih dari *Minimum Points* maka titik p adalah *core point* dan *cluster* terbentuk. Jika iya maka *cluster* terbentuk, jika tidak maka dianggap sebagai *noise*.

Beberapa istilah dalam metode DBSCAN (Harjanto & Sumarni, 2021:34) :

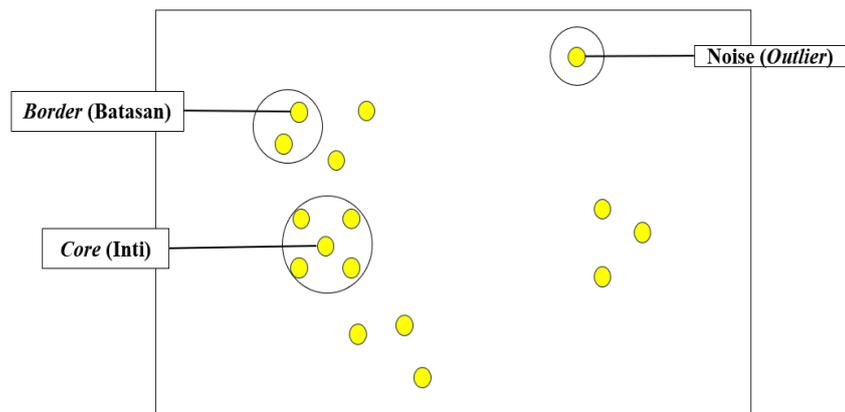
- 1) *Core*

Core merupakan titik pusat dalam *cluster* yang didasarkan pada densitas di mana terdapat sejumlah titik yang harus berada dalam *Epsilon* (radius atau nilai ambang batas) dan *Minimum Point* (minimal titik dalam *Cluster*) yang ditentukan pengguna.

- 2) *Border*

Border merupakan titik yang menjadi batasan dalam kawasan titik pusat (*core*).

Berikut adalah gambar yang menunjukkan mana yang merupakan *border point* dan mana yang merupakan *core point* dengan contoh menggunakan *Minimum Points* = 5 dan *Epsilon* = 1.



Gambar 1. *Border Point*, *Core Point* dan *Noise*

3) *Noise*

Noise merupakan titik yang tidak dapat dijangkau oleh *core* dan bukan merupakan *border*.

$$Noise = \{x \in X \mid \forall i : x \notin C_i\} \quad (2)$$

Dimana:

X : Gugus data

C_i : merupakan *cluster* ke- i

4) Densitas

- 1) Densitas terjangkau merupakan sebuah titik yang dimana dikatakan titik terjangkau apabila titik tersebut secara langsung berdekatan dengan titik pusat (*core*)
- 2) *Densitas* terjangkau langsung merupakan sebuah titik terjangkau langsung apabila titik tersebut terhubung secara langsung dengan titik pusat (*core*)

$$x \in N_{Eps}(y) \mid |N_{Eps}(y)| \geq MinPts. \quad (3)$$

Dimana :

$N_{Eps}(y)$: Titik sekitar dari y
dalam radius Eps

$MinPts$: Minimal titik dalam
cluster

- 3) Densitas terhubung merupakan sebuah titik yang saling terhubung satu sama lain. DBSCAN memerlukan dua parameter *input* yaitu (*Epsilon*) dan titik minimum (*Minimum Point*). *Epsilon* titik sekitar didefinisikan sebagai:

$$N_{eps}(x) = \{y \in D \mid dist(x,y) \leq Eps\} \quad (4)$$

Dimana :

$N_{epsilon}(x)$: Titik sekitar dari x dalam radius Eps

D : Gugus data

Distance : Jarak *euclidean* dari objek x dan y

(x,y)

Eps : Radius atau ambang batas

2.4 Validasi Cluster

Validasi *cluster* merupakan tahapan evaluasi hasil analisis *cluster* yang memiliki seperangkat ukuran karakteristik sebagai acuan dalam menentukan jumlah *cluster* optimal (Brock *et al.*, 2008:87). Pada penelitian ini menggunakan validasi internal sebagai evaluasi pemilihan *cluster* optimal. Validasi *internal* menggunakan informasi internal pada data untuk menilai kualitas *clustering*. Validasi *internal* menggambarkan hubungan yang merujuk pada penempatan beberapa data pengamatan dalam sebuah *cluster*. Berikutnya, kepadatan yang berhubungan dengan mengevaluasi *homogenitas* (kedekatan) *cluster*, ditinjau dari varian *intra-cluster*. Kemudian, pemisahan partisi *cluster* yang menunjukkan tingkat pengukuran jarak antar dua *cluster* (Murjani, 2022:29).

Terdapat beberapa metode atau cara pada validasi *internal*, diantaranya nilai indeks *silhouette*, indeks *dunn* dan indeks *connectivity*.

1) Indeks *Dunn*

Indeks *Dunn* adalah rasio jarak terkecil antara *observasi* pada *cluster* yang berbeda dengan jarak terbesar pada masing-masing *cluster* data. Semakin tinggi nilai indeks *dunn*, semakin baik *cluster* yang

dihasilkan. Indeks *dunn* diperoleh dari hasil pembagian antara d_{\min} dengan d_{\max} (Irwansyah & Faisal, 2015:16)

$$Dunn = \frac{d_{\min}}{d_{\max}} \quad (5)$$

Dimana :

d_{\min} : jarak terkecil antara obsevasi pada *cluster* yang berbeda

d_{\max} : jarak terbesar pada masing-masing *cluster* data

2) Indeks *Silhouette*

Indeks *Silhouette* dihitung sebagai drajad kepercayaan dalam proses *clustering* pada suatu pengamatan dengan *cluster* yang dikatakan terbentuk baik bila nilai indeks mendekati 1 dan kondisi sebaliknya jika nilai indeks mendekati angka -1 (Irwansyah & Faisal, 2015:29).

$$S_{(i)} = \frac{b_{(i)} - a_{(i)}}{\max(a_{(i)}, b_{(i)})} \quad (6)$$

Dimana :

$a_{(i)}$: jarak rata-rata i dan seluruh pengamatan lainnya.

$b_{(i)}$:jarak rata-rata i dengn pengamatan pada *cluster* terdekat.

3) Indeks *Connectivity*

Connectivity menunjukkan tingkat hubungan *cluster*, ditentukan dengan jumlah tetangga terdekat (Murjani, 2022:29). Indeks *connectivity* memiliki kriteria nilai diantara 0 sampai tak hingga. Semakin kecil indeks *connectivity cluster* yang terbentuk semakin baik (Afira & Wijayanto, 2021:104)

$$Conn(c) = \sum_{j=1}^N \sum_{i=1}^N X_{i,nni(i)} \quad (7)$$

Dengan :

$nn_{i(j)}$: pengamatan tetangga terdekat dengan (*nearest neighbor*)
i ke *j*

$nn_{i(j)}$: sebagai parameter yang menentukan jumlah tetangga yang berkontribusi pada pengukuran *connectivity*

3. Hasil dan Pembahasan

3.1 Hasil penelitian

a. Analisis deskriptif



Gambar 2. Jumlah timbunan Sampah berdasarkan Kecamatan di Kabupaten Lombok Timur

Kecamatan yang ada di Kabupaten Lombok Timur sangatlah tinggi, berdasarkan Gambar di atas juga kita bisa mengetahui jumlah tumpukan sampah terbanyak berada di Kecamatan Masbagek dengan jumlah 15.014 (Ton). Kemudian di susul oleh Kecamatan Pringgabaya dan Kecamatan Selong yang dimana kedua kecamatan tersebut menduduki urutan kedua dan ketiga dengan timbunan sampah terbanyak di Kabupaten Lombok Timur dengan jumlah

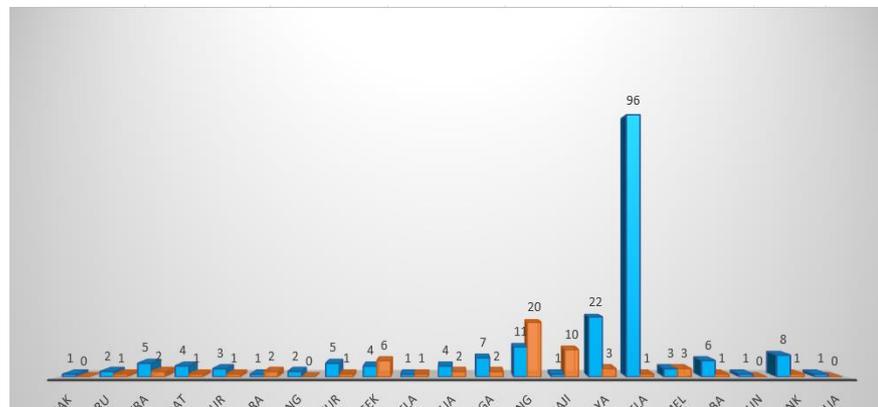
14.078 dan 13.774 (Ton). Sedangkan jumlah timbunan paling rendah berada di Kecamatan Sembalun dengan jumlah penumpukan seberat



Gambar 3. Jumlah Sarana TPS3R dan TPA Kecamatan di Kabupaten Lombok Timur

2.988 (Ton).

Berdasarkan Gambar 3, dapat kita lihat keadaan sarana pembuangan sampah (TPA DAN TPS3R) sangatlah minim adanya yang dimana TPA di Kabupaten Lombok Timur hanya ada 1 yakni di Kecamatan Labuhan Haji. Sedangkan TPS3R hanya ada 7



Gambar 4. Jumlah Sarana Bank Sampah dan TPS Desa Kecamatan di Kabupaten Lombok Timur

kecamatan yang memiliki, sisanya kebanyakan kecamatan tidak memiliki sarana TPS3R.

Berdasarkan Gambar 4, sarana kebersihan khususnya Bank Sampah, Kecamatan Suela menjadi penyumbang terbanyak Bank Sampah yakni berkisar di angka 96, selanjutnya penyumbang terbanyak kedua ditempati oleh Kecamatan Pringgabaya dengan jumlah Bank Sampah sebanyak 22 sarana. Sedangkan Kecamatan Sambalia, Sembalun, Labuhan Haji, Terara dan Pringgasele memiliki sarana bank sampah paling sedikit yakni 1 sarana.

b. Analisis *Cluster*

Proses *Clustering* dilakukan dengan menggunakan satu metode yakni dengan menggunakan metode DBSCAN. Adapun proses *clustering* pada penelitian ini adalah sebagai berikut:

1) Standarisasi Data

Standarisasi dilakukan jika data yang digunakan dalam sebuah penelitian mempunyai satuan yang bervariasi atau berbeda. Adapun data yang digunakan dalam penelitian ini yaitu data dengan satuan yang bervariasi atau berbeda, sehingga proses pengelompokan tidak langsung dilakukan. Untuk itu, selanjutnya dilakukan proses standarisasi dengan melakukan standarisasi

pada data sebelum dilakukan analisis lebih lanjut. Adapun hasil standarisasi tersebut dapat dilihat pada Tabel di bawah ini.

Tabel 1. Hasil Standarisasi Data

No	Kecamatan	Timbunan Sampah	Bank Sampah	TPS Desa	TPS3R	TPA
1	Keruak	-0.347	-0.397	-0.619	-0.649	-0.223
2	Jerowaru	-0.208	-0.347	-0.395	-0.649	-0.223
.
.
.
21	Sambelia	-1.152	-0.397	-0.619	-0.649	-0.223

2) Membentuk Matriks Jarak

Dalam membentuk ukuran jarak yang biasa digunakan dalam analisis *cluster* adalah jarak *Euclid*. Jika X_1 dan X_2 merupakan 2 observasi yang berada di jarak R^n , maka jarak *Euclidean* disajikan dalam persama (1). Perhitungan pada persamaan (1) untuk menentukan jarak atau kemiripan antara obyek satu dengan obyek lainnya, atau kecamatan satu dengan kecamatan yang lainnya. Hasil perhitungan untuk jarak kemiripan dari Kecamatan satu dengan kecamatan lainnya dapat dilihat pada matriks berikut:

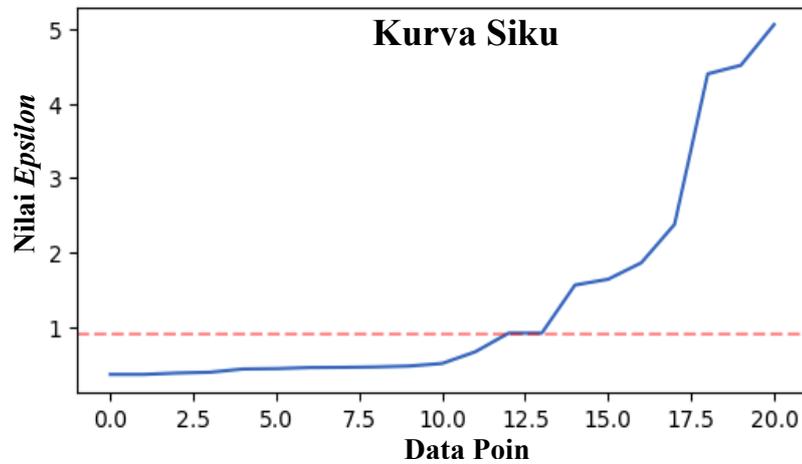
$$D = \begin{bmatrix} 0.0000000 & 0.4281178 & 3.002204 & 1.5059928 & \dots & 0.9959136 \\ 0.4281178 & 0.0000000 & 2.975238 & 1.5552216 & \dots & 1.3789998 \\ 3.0022039 & 2.9752382 & 0.0000000 & 1.5237192 & \dots & 3.2203903 \\ 1.5059928 & 1.5552216 & 1.523719 & 0.0000000 & \dots & 1.7459947 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0.9959136 & 1.3789998 & 3.220390 & 1.7459947 & \dots & 0.0000000 \end{bmatrix}$$

Dari matriks tersebut dapat kita ambil contoh hasil dari jarak *Euclidean* antara data 1 Kecamatan Keruak dan data 2 Kecamatan Jerowaru adalah 0.4281178, sedangkan jarak *Euclidean* antara data 1 Kecamatan Keruak dengan data 2 Kecamatan Sakra adalah 3.002204. Demikian pula untuk penafsiran obyek yang lainnya, semakin kecil jarak kedua obyek maka akan semakin mirip karakteristik dari kedua obyek tersebut.

c. Pengelompokan DBSCAN

1) Menentukan *Epsilon*

Penentuan awal nilai dari parameter *epsilon* tersebut ditunjukkan agar menghasilkan *cluster* optimal. Penentuan nilai *epsilon* dan *minimum points* sangat mempengaruhi jumlah *cluster* yang akan dihasilkan dan ditentukan dari nilai *epsilon* dan *minimum points* pada titik dalam sebuah penggerombolan. Titik yang dimaksud yaitu *k*-tetangga terdekatnya memiliki jarak yang sama, titik *noise* memiliki jarak terjauh dari *k*-tetangga terdekat sehingga dilakukan plot jarak secara terurut pada setiap titik pada *k*-tetangga terdekat (Kim *et al.*, 2020). Nilai *epsilon* dan *minimum points* yang optimal dipilih berdasarkan garis yang mendekati garis menaik kemudian dipotong secara vertikal pada plot *k*-tetangga terdekat.



Gambar 5. Kurva siku penentuan *epsilon*

Gambar 5 menunjukkan *plot knn* yang dimana nilai *epsilon* yang optimal berada pada rentang titik 1.0 melalui pengamatan *knee of a curve* atau kurva siku adalah titik dimana kurva terlihat tikungan, khususnya dari lereng tinggi ke lereng rendah atau ke arah lain. Digunakan dalam optimasi, misalnya ketika ada peningkatan dari *vertikal* ke *horizontal*.

2) Evaluasi Jumlah *Cluster optimal*

Pemilihan parameter *epsilon* dan *minimum points* yang optimal dilihat dari nilai *Silhouette* yang tertinggi dengan memakai beberapa percobaan yang diimplementasikan pada *software python jupyter notebook*. Eksperimen nilai *epsilon* dan *minimum points* pada data sarana kebersihan di Kabupaten Lombok Timur

Tabel 2. Evaluasi Jumlah *Cluster* Optimal

<i>Epsilon</i>	<u><i>MinPts</i></u>	<i>Cluster</i>	<i>Noise</i>	SI
1.15	3	2	10	0.076
1.10	2	2	10	0.076
1.20	4	2	11	0.085
1.5	4	2	11	0.075
1.0	2	2	7	0.197

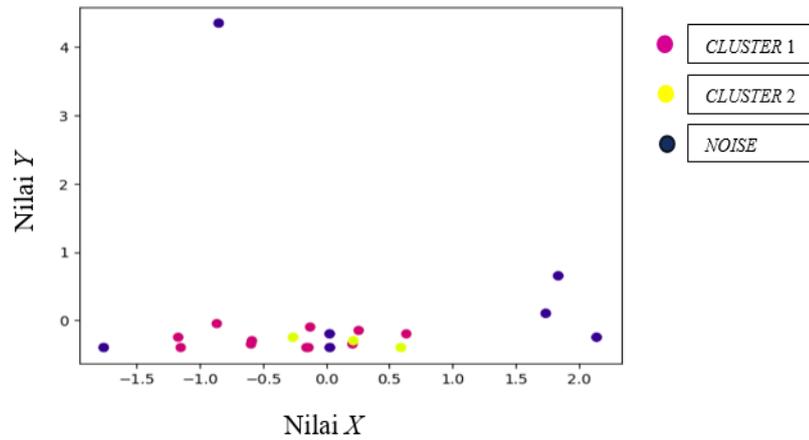
Berdasarkan hasil eksperimen pada Tabel 5, dipilih nilai *epsilon* 1.0 dan *minimum points* 2 karena memiliki nilai *silhouette* tertinggi yaitu sebesar 0.197. Hasil perhitungan menggunakan *Sillhouette Index* (SI) berkisar antara nilai -1 hingga 1, dengan -1 merupakan nilai terburuk dan nilai yang mendekati 1 merupakan nilai *cluster* terbaik. Untuk *Sillhouette Index* (SI) di dapatkan nilai 0.197 dapat dilihat dari Tabel di bawah ini.

Tabel 3. Nilai *Silhouette*, *Dunn* dan *Connectivity*

<u>Jumlah cluster</u>	<i>Silhouette</i>	<i>Dunn</i>	<i>Connectivity</i>
2	0.197	0.819	21.55

3) Pengelompokan dengan metode DBSCAN

Setelah melakukan analisis penentuan *epsilon*, *mini point* dan penentuan jumlah *cluster* optimal dengan melihat nilai *indeks silhouette*, kemudian dilakukan pengolahan data dalam plot visualisasi sebagai berikut.



Gambar 6. Visualisasi *cluster*

Pada Gambar 6 di atas terdapat 2 *cluster* yang dihasilkan, *cluster* pertama adalah titik berwarna pink dengan 11 Kecamatan dan *cluster* kedua pada pada titik yang berwarna kuning dengan 3 Kecamatan. Sedangkan titik biru merupakan *noise* yang berjumlah 7 kecamatan, Adapun pembagian *cluster* untuk Kecamatan di Kabupaten Lombok Timur berdasarkan Gambar di atas adalah sebagai berikut:

Tabel 4. Anggota *Cluster* DBSCAN

Label	Jumlah cluster	Anggota cluster
<i>Cluster 1</i>	11	<u>Keruak</u> , <u>Jerowaru</u> , <u>Sakra Timur</u> , <u>Montong</u> , <u>Gading</u> , <u>Sikur</u> , <u>Pringgasela</u> , <u>Sukamulia</u> , <u>Suralaga</u> , <u>Wanasaba</u> , <u>Lenek</u> , <u>Sambalia</u>
<i>Cluster 2</i>	3	<u>Sakra Barat</u> , <u>Terara</u> , <u>Aikmel</u>
<i>Noise</i>	7	<u>Sakra</u> , <u>Masbagik</u> , <u>Selong</u> , <u>Labuhan Haji</u> , <u>Suela</u> , <u>Pringgabaya</u> , <u>Sembalun</u>

Hasil *cluster* pada penelitian ini menunjukkan adanya karakteristik yang berbeda antar *cluster*, analisis identifikasinya ditampilkan pada Tabel 5.

Tabel 5. Rata-Rata *Cluster*

<i>CLUSTER</i>	TIMBUNAN SAMPAH (TON)	BANK SAMPAH	TPS DESA	TPS3R	TPA
0	-0,336474959	-0,26553	-0,41539	-0,64952	-0,22361
1	0,180967025	-0,31397	-0,17081	0,866025	-0,22361

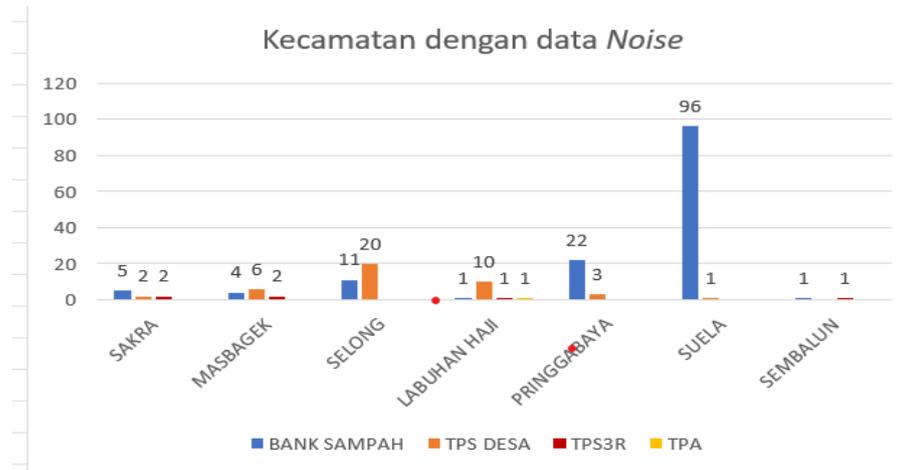
Tabel 5 menggambarkan karakteristik dari masing-masing variabel. Berikut adalah interpretasi karakteristik dari masing-masing variabel

- a) Penyebab penumpukan sampah tertinggi berada di *cluster* 2 dengan 3 kecamatan memiliki nilai rata-rata semua variabel lebih banyak di bandingkan *cluster* 1. Sehingga dapat dikatakan bahwa pada kecamatan tersebut belum maksimal dikarenakan sarana kebersihan masih kurang. Seperti salah satunya Kecamatan Sakra Barat hanya memiliki 4 Bank Sampah, 1 TPS Desa, 1 TPS3R dan tidak memiliki Tempat Pembuangan Akhir (TPA).
- b) *Cluster* terendah yaitu *cluster* 1 dengan 11 Kecamatan dengan rata-rata penyebab penumpukan sampah rendah jika dibandingkan dengan *cluster* 2, sehingga dapat dikatakan bahwa pada Kecamatan tersebut cukup maksimal dikarenakan jumlah tumpukan sampah dan sarana kebersihan cukup mengimbangi

3.2 Pembahasan

Berdasarkan hasil analisis data perhitungan yang telah dilakukan dengan menggunakan metode DBSCAN menghasilkan dua *cluster* yang optimal menggunakan validasi internal (*silhouette*). Pada metode DBSCAN, pemilihan nilai *epsilon* dan *minimum point* terbaik dengan menggunakan *knee of curva* ditetapkan nilai *epsilon* 1.00. lalu dilakukan eskperimen terhadap nilai *epsilon*, *minimum point* dan dievaluasi dengan melihat nilai *indeks silhouette*, sehingga jumlah *cluster* yang terbentuk yaitu sebanyak dua *cluster* dengan nilai *epsilon* = 1.00, *minimum points* = 2 dan 7 *noise*.

Dalam jumlah *cluster* yang terbentuk dengan metode peneilitian ini sebanyak dua *cluster* dengan jumlah *noise* 7. Penyebab penumpukan sampah tertinggi berada pada *cluster* 2 dengan tiga kecamatan memiliki nilai rata-rata semua variabel lebih banyak di bandingkan *cluster* 1. Sehingga dapat dikatakan bahwa pada Kecamatan tersebut belum maksimal dikarenakan sarana kebersihan masih kurang. *Cluster* terendah yaitu *cluster* 1 dengan 11 Kecamatan dimana rata-rata penyebab penumpukan sampah rendah jika dibandingkan dengan *cluster* 2, sehingga dapat dikatakana bahwa pada Kecamatan tersebut cukup maksimal. Sedangkan untuk data *noise* bukan sebuah *cluster*, melainkan tidak masuk ke dalam *cluster* manapun. Hal ini dikarenakan pada *noise*, memiliki nilai yang jauh dari rata-rata.



Gambar 7. Kecamatan dengan data *noise*

Berdasarkan Gambar 7, dapat dikatakan bahwa Kecamatan yang termasuk *noise* memiliki nilai yang ekstrim, seperti pada Kecamatan Suela untuk variabel Bank Sampah memiliki angka sampai 96, hal ini dapat diartikan bahwa sebanyak 96 tempat Bank Sampah berada pada Kecamatan Suela. Hal ini disebabkan oleh beberapa factor, seperti Pendidikan juga budaya serta kesadaran kaum muda untuk menunjang kebersihan lingkungan. Berbeda dengan Kecamatan Sembalun, untuk variabel Bank Sampah dan TPS3R memiliki rata-rata yang sama. Dapat diartikan bahwa Kecamatan Sembalun terdapat disalah satu tempat yang sadar akan kebersihan lingkungan dengan mampu membuat sarana kebersihan.

4. Simpulan Dan Saran

4.1 Simpulan

Berdasarkan hasil analisis *clustering* dengan metode DBSCAN untuk pengelompokkan jumlah sarana kebersihan berdasarkan jumlah penumpukan sampah di Kabupaten Lombok Timur didapatkan kesimpulan sebagai berikut:

- a. *Cluster* terendah yaitu *cluster* 1 dengan 11 Kecamatan dimana rata-rata penyebab penumpukan sampah rendah jika di bandingkan dengan *cluster* 2, sehingga dapat di katakana bahwa pada Kecamatan tersebut cukup maksimal dikarenakan jumlah tumpukan sampah dan sarana kebersihan cukup mengimbangi.

Cluster 1 terdiri dari Kecamatan Keruak, Jerowaru, Sakra Timur, Montong Gading, Sikur, Pringgasela, Sukamulia, Suralaga, Wanasaba, Lenek dan Sambalia.

- b. Penyebab penumpukan sampah tertinggi berada di *cluster* 2 dengan 3 kecamatan memiliki nilai rata-rata semua variabel lebih banyak dibandingkan *cluster* 1. Sehingga dapat dikatakan bahwa pada Kecamatan tersebut belum maksimal dikarenakan sarana kebersihan masih kurang. Seperti salah satunya Kecamatan Sakra Barat hanya memiliki 4 Bank Sampah, 1 TPS Desa, 1 TPS3R dan tidak memiliki Tempat Pembuangan Akhir (TPA).

Cluster 2 terdiri dari Kecamatan Sakra Barat, Terara dan Aikmel.

- c. Kecamatan yang termasuk *noise* memiliki nilai yang ekstrem. Seperti pada Kecamatan Suela untuk variabel Bank Sampah memiliki angka sampai 96, hal ini dapat diartikan bahwa sebanyak

96 tempat Bank Sampah berada pada Kecamatan Suela. Hal ini disebabkan oleh beberapa factor, seperti Pendidikan juga budaya serta kesadaran kaum muda untuk menunjang kebersihan lingkungan. Berbeda dengan Kecamatan Sembalun, untuk variabel Bank Sampah dan TPS3R memiliki rata-rata yang sama. Dapat diartikan bahwa Kecamatan Sembalun terdapat di salah satu tempat yang sadar akan kebersihan lingkungan dengan mampu membuat sarana kebersihan.

Noise terdiri dari Kecamatan Sakra, Masbagek, Selong, Labuhan Haji, Pringgabaya, Suela dan Sembalun.

4.2 Saran

Penelitian ini bertujuan untuk mengetahui kelompo-kelompok penyebab penumpukan sampah berdasarkan sarana yang tersedia, penelitian ini menggunakan metode DBSCAN *clustering* dengan menggunakan validasi Intenal untuk menentukan jumlah *cluster* yang optimal. Dapat menjadi gambaran baik pembaca dan bagi masyarakat agar dapat mengetahui pentingnya memiliki sarana kebersihan disetiap Kecamatan agar dapat mengurangi penumpukan sampah. Diharapkan juga penelitian ini dapat dikembangkan dengan lebih baik dari variabel hingga metode yang digunakan, serta menjadi informasi agar kebijakan tepat pada sasaran.

Daftar Pustaka

- Afira, N., & Wijayanto, A. W. (2021). Optimization of Waste Transportation Routes using Multi-objective Non-dominated Sorting Genetic Algorithm II (MNSGA-II) in the Eastern and Southern Regions of Bandung City, Indonesia. *Proceedings of The International Conference on Data Science and Official Statistics, 2021*(1), 20–30.
- ALIA, L. S. (2023). *KLASIFIKASI JENIS SAMPAH MENGGUNAKAN IMAGE CLASSIFICATION CONVOLUTIONAL NEURAL NETWORK.*

UNIVERSITAS ISLAM NEGERI SULTAN SYARIF KASIM RIAU.

- Brock, G., Pihur, V., Datta, S., & Datta, S. (2008). clValid: An R package for cluster validation. *Journal of Statistical Software*, 25, 1–22.
- Cinderatama, T. A., Alhamri, R. Z., & Yunhasnawa, Y. (2022). Implementasi Metode K-Means, DbSCAN, dan Meanshift Untuk Analisis Jenis Ancaman Jaringan Pada Intrusion Detection System. *INOVTEK Polbeng-Seri Informatika*, 7(1), 169–184.
- Dani, A. T. R., Wahyuningsih, S., & Rizki, N. A. (2021). Pengelompokan Data Runtun Waktu menggunakan Analisis Cluster. *EKSPONENSIAL*, 11(1), 29–38.
- Defiyanti, S., Jajuli, M., & Rohmawati, N. (2017). Optimalisasi K-Medoid Dalam Pengklasteran Mahasiswa Pelamar Beasiswa Dengan Cubic Clustering Criterion. *Jurnal Nasional Teknologi Dan Sistem Informasi*, 3(1), 211–218.
- Harjanto, A. S., & Sumarni, S. (2021). TEACHERS' EXPERIENCES ON THE USE OF GOOGLE CLASSROOM. *English Language and Literature International Conference (ELLiC) Proceedings*, 3, 172–178.
- Irawan, Y. (2019). Penerapan Data Mining Untuk Evaluasi Data Penjualan Menggunakan Metode Clustering Dan Algoritma Hirarki Divisive Di Perusahaan Media World Pekanbaru. *Jurnal Teknologi Informasi Universitas Lambung Mangkurat (JTIULM)*, 4(1), 13–20.
- Irwansyah, E., & Faisal, M. (2015). *Advanced Clustering: Teori dan Aplikasi*. Deepublish.
- Kai, H. N., Sompie, S. R. U. A., & Sambul, A. M. (2018). Aplikasi Layanan Pengangkutan Sampah Berbasis Android. *Jurnal Teknik Informatika*, 13(4).
- Kim, L., Whitaker, M., O'Halloran, A., Kambhampati, A., Chai, S. J., Reingold, A., Armistead, I., Kawasaki, B., Meek, J., & Yousey-Hindes, K. (2020).

- Hospitalization rates and characteristics of children aged < 18 years hospitalized with laboratory-confirmed COVID-19—COVID-NET, 14 states, March 1–July 25, 2020. *Morbidity and Mortality Weekly Report*, 69(32), 1081.
- Magriaty, R., Murtilaksono, K., & Anwar, S. (2023). Analisis K-Means Cluster untuk Identifikasi Kawasan Pengelolaan Sampah di Kabupaten Tapin Provinsi Kalimantan Selatan. *Journal of Regional and Rural Development Planning (Jurnal Perencanaan Pembangunan Wilayah Dan Perdesaan)*, 7(1), 79–90.
- Maukar, A. L., Marisa, F., & Widodo, A. A. (2022). Analisis Data Penerimaan Mahasiswa Baru Berbasis K-Means. *JIKO (Jurnal Informatika Dan Komputer)*, 6(2), 142–147.
- Mhlwatika, Z., & Aderibigbe, B. A. (2018). Application of dendrimers for the treatment of infectious diseases. *Molecules*, 23(9), 2205.
- Mokodompis, Y., Kaunang, M., & Kasenda, V. (2019). Implementasi Kebijakan Pengelolaan Sampah Di Kota Manado. *Jurnal Eksekutif*, 3(3).
- Murjani, M. (2022). Pendidikan masa depan indonesia (analisis swot, som, inter dan multidisipliner). *Educational Journal: General and Specific Research*, 2(1), 19–33.
- Mustikasari, M., & Salman, N. (2023). ANALISIS KLASSTER BERBASIS KEPADATAN DENGAN DBSCAN DAN OPTICS. *Jurnal INSYPRO (Information System and Processing)*, 8(1), 1.
- Nisrina, S., Nurmayanti, 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.
- Nurrohman, M., Maimunah, M., & Sukmasetya, P. (2023). Sistem Klasterisasi Volume Sampah Organik di Kota Magelang menggunakan K-Means.

TEMATIK, 10(1), 146–153.

Suparmi, S., & Soeheri, S. (2020). Sistem Informasi Geografis Pemetaan Tempat Kost Berbasis Web Menggunakan Metode Euclidean Distance. *INFOSYS (INFORMATION SYSTEM) JOURNAL*, 5(1), 105–113.