
Application of the DBSCAN Algorithm for Clustering Patient Satisfaction Data at the Dental and Oral Hospital of Mulawarman University

Nurussyifaa¹⁾, Ita Arfyanti²⁾, dan Hanifah Ekawati³⁾

^{1,2,3}Teknik Informatika, STMIK Widya Cipta Dharma

^{1,2,3}Jalan M. Yamin No.25, Samarinda, 75123

E-mail: 2443905@wicida.ac.id¹⁾, ita@wicida.ac.id²⁾, hanifah@wicida.ac.id³⁾

ABSTRACT

Patient satisfaction is an important indicator for evaluating the quality of hospital services, yet aggregate averages or percentages may not reveal groups of patients with different assessment patterns. This study aims to apply the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm to cluster patient satisfaction data at the Dental and Oral Hospital of Mulawarman University. The study used secondary questionnaire data collected from July 2025 to April 2026. From 690 initial records, 75 records using an earlier five-indicator instrument were excluded, three duplicate records were removed, and 75 incomplete records were excluded, resulting in 612 respondents with complete answers to nine service indicators. The categorical responses were converted into numerical Likert scores, followed by the construction of two clustering variables: the mode of each respondent's answers and the average response score. Both variables were standardized before DBSCAN was applied. Parameter testing produced a final configuration of epsilon 0.25 and MinPts 4, with a Silhouette Score of 0.894071. The clustering formed two main clusters and four noise observations. Cluster 0 contained 320 respondents (52.29%) with an average response score of 3.93, while Cluster 1 contained 288 respondents (47.06%) with an average of 3.06. The results indicate generally positive assessments, while waiting time was the lowest-rated indicator in both main clusters. The identified noise observations also showed notably low ratings for cost, service, and waiting time, providing additional information for service evaluation.

Keywords: DBSCAN, clustering, patient satisfaction, hospital services, Silhouette Score

Penerapan Algoritma DBSCAN untuk Pengelompokan Data Kepuasan Pasien pada Rumah Sakit Gigi dan Mulut Universitas Mulawarman

ABSTRAK

Kepuasan pasien merupakan salah satu indikator penting dalam mengevaluasi kualitas pelayanan rumah sakit, tetapi penyajian hasil dalam bentuk rata-rata atau persentase belum sepenuhnya menunjukkan adanya kelompok pasien dengan pola penilaian yang berbeda. Penelitian ini bertujuan menerapkan algoritma Density-Based Spatial Clustering of Applications with Noise (DBSCAN) untuk mengelompokkan data kepuasan pasien di Rumah Sakit Gigi dan Mulut Universitas Mulawarman. Data yang digunakan berupa data sekunder kuesioner kepuasan pasien periode Juli 2025 sampai April 2026. Dari 690 data awal, sebanyak 75 data dengan instrumen lima indikator dieliminasi, tiga data duplikat dihapus, dan 75 data yang tidak lengkap dikeluarkan sehingga diperoleh 612 responden dengan jawaban lengkap pada sembilan indikator pelayanan. Jawaban kategorikal dikonversi menjadi skor numerik skala Likert, kemudian dibentuk dua variabel Clustering, yaitu modus jawaban setiap responden dan rata-rata jawaban. Kedua variabel distandardisasi sebelum penerapan DBSCAN. Hasil pengujian parameter menetapkan epsilon sebesar 0,25 dan MinPts sebesar 4 dengan nilai Silhouette Score 0,894071. Clustering menghasilkan dua cluster utama dan empat data noise. Cluster 0 terdiri atas 320 responden (52,29%) dengan rata-rata jawaban 3,93, sedangkan Cluster 1 terdiri atas 288 responden (47,06%) dengan rata-rata 3,06. Hasil menunjukkan bahwa penilaian responden secara umum positif, sedangkan indikator waktu tunggu menjadi indikator dengan nilai relatif terendah pada kedua cluster utama. Data noise juga menunjukkan penilaian rendah terutama pada biaya, pelayanan, dan waktu tunggu sehingga dapat menjadi informasi tambahan dalam evaluasi pelayanan.

Kata Kunci: DBSCAN, Clustering, kepuasan pasien, pelayanan rumah sakit, Silhouette Score

1. PENDAHULUAN

Salah satu aspek penting dalam pelayanan kesehatan adalah kualitas pelayanan yang diterima pasien. Kepuasan pasien terbentuk dari pengalaman selama menerima

pelayanan, baik dari aspek administratif, teknis, interpersonal, maupun fasilitas pendukung. Oleh karena itu, kepuasan pasien dapat digunakan sebagai indikator untuk menilai apakah pelayanan telah sesuai dengan harapan

pasien atau masih memerlukan perbaikan (Zulvia et al., 2022; Harini et al., 2023).

Pada Rumah Sakit Gigi dan Mulut (RSGM) Universitas Mulawarman, evaluasi kepuasan pasien menjadi relevan karena rumah sakit ini merupakan institusi pelayanan yang relatif baru dan mulai beroperasi pada November 2024. Penilaian kepuasan pada penelitian ini mengacu pada sembilan unsur Survei Kepuasan Masyarakat (SKM), yaitu persyaratan, sistem/mekanisme/prosedur, waktu penyelesaian, biaya atau tarif, jenis pelayanan, kompetensi pelaksana, perilaku pelaksana, penanganan pengaduan, serta sarana dan prasarana (Kementerian PANRB, 2017).

Pengukuran kepuasan yang hanya disajikan melalui nilai rata-rata atau persentase dapat memberikan gambaran umum, tetapi belum tentu memperlihatkan kelompok pasien yang mempunyai pola penilaian berbeda. Penelitian Frangky et al. (2025) menunjukkan bahwa persepsi pasien terhadap pelayanan dapat membentuk segmentasi yang berbeda. Sementara itu, Susilo (2024) menunjukkan bahwa pemilihan metode Clustering perlu disesuaikan dengan karakteristik data, terutama ketika data memiliki pola yang heterogen.

DBSCAN dipilih karena membentuk cluster berdasarkan kepadatan data, tidak mensyaratkan jumlah cluster ditentukan sejak awal, dan dapat mengenali data yang tidak memenuhi kepadatan tertentu sebagai noise. Karakteristik tersebut sesuai dengan tujuan penelitian untuk menemukan kelompok pasien berdasarkan pola penilaian sekaligus mengenali responden dengan pola yang tidak umum (Simbolon & Friskila, 2024; Badawi et al., 2025).

Kebaruan penelitian ini terletak pada pengintegrasian sembilan indikator SKM dengan Clustering berbasis kepadatan untuk mengelompokkan data kepuasan pasien RSGM Universitas Mulawarman. Data indikator diolah menjadi dua variabel yang merepresentasikan jawaban dominan dan kecenderungan rata-rata penilaian setiap responden, kemudian hasil cluster diinterpretasikan berdasarkan indikator pelayanan dan karakteristik responden.

Tujuan penelitian adalah menerapkan DBSCAN pada data kepuasan pasien, menganalisis hasil pengelompokan, dan mengevaluasi hasil tersebut sebagai informasi pendukung dalam upaya peningkatan kualitas pelayanan di RSGM Universitas Mulawarman.

2. RUANG LINGKUP

Penelitian menggunakan data sekunder kuesioner kepuasan pasien RSGM Universitas Mulawarman pada periode Juli 2025 sampai April 2026.

Analisis dibatasi pada sembilan indikator pelayanan yang tersedia dalam instrumen kuesioner dan menggunakan data responden yang mempunyai jawaban lengkap pada seluruh indikator tersebut.

Variabel pembentuk cluster adalah modus jawaban dan rata-rata jawaban setiap responden. Usia, jenis kelamin, dan pekerjaan digunakan sebagai informasi pendukung untuk interpretasi, bukan sebagai fitur utama DBSCAN.

3. BAHAN DAN METODE

3.1 Data Penelitian

Data penelitian berupa data sekunder kuesioner kepuasan pasien. Dataset awal berjumlah 690 responden yang terdiri atas 75 responden dengan instrumen lama berisi lima pertanyaan dan 615 responden dengan instrumen baru berisi sembilan pertanyaan. Karena penelitian menetapkan sembilan indikator sebagai dasar analisis, 75 data dengan instrumen lama dieliminasi. Selanjutnya ditemukan tiga data duplikat yang dihapus dan 75 data dengan nilai kosong pada indikator kepuasan, sehingga diperoleh 612 responden untuk proses transformasi dan Clustering.

Tabel 1. Ringkasan Seleksi dan Pembersihan Data

Table 1. Summary of Data Selection and Cleaning

Tahap	Jumlah data	Keterangan
Data awal	690	75 data instrumen lama + 615 data instrumen baru
Seleksi instrumen	615	75 data lama dieliminasi; tersisa data dengan 9 indikator
Pemeriksaan duplikat	687	Dari 690, 3 duplikat dihapus
Pemeriksaan missing value	612	75 data tidak lengkap dihapus dari 687
Data akhir	612	Lengkap pada 9 indikator dan digunakan untuk DBSCAN

3.2 Pra-pemrosesan dan Transformasi

Pra-pemrosesan meliputi seleksi data, penyederhanaan nama fitur, pembersihan kesalahan penulisan, pemeriksaan duplikat, pemeriksaan missing value, serta anonimisasi dan pengkodean karakteristik responden. Kesalahan penulisan pada jawaban kepuasan diseragamkan tanpa mengubah makna tingkat penilaian. Identitas responden tidak digunakan secara langsung dan diganti dengan kode responden.

Sembilan indikator kepuasan terdiri atas Persyaratan, Prosedur, Waktu Tunggu, Biaya, Fasilitas, Pelayanan, Keterampilan Petugas, Sikap Petugas, dan Kritik/Saran. Jawaban dikonversi ke skala Likert: Sangat Tidak Puas = 1, Tidak Puas = 2, Puas = 3, dan Sangat Puas = 4.

Tabel 2. Skala Likert Penilaian Kepuasan

Table 2. Likert Scale for Satisfaction Assessment

Skor	Kategori
1	Sangat Tidak Puas
2	Tidak Puas
3	Puas
4	Sangat Puas

Setelah konversi numerik, dibentuk dua fitur Clustering. Variabel X adalah modus sembilan jawaban setiap responden, sedangkan variabel Y adalah rata-rata jawaban sembilan indikator. Kedua fitur kemudian distandardisasi menggunakan StandardScaler agar kontribusi masing-masing fitur seimbang dalam perhitungan jarak.

3.3 Penerapan DBSCAN dan Evaluasi

MinPts awal ditentukan menggunakan aturan praktis $\text{MinPts} = 2 \times d$. Karena digunakan dua fitur, nilai awal MinPts adalah 4. Kandidat epsilon ditentukan melalui k-

distance graph dengan jarak terhadap tetangga terdekat keempat. Titik elbow pada grafik berada sekitar 0,25 sehingga digunakan sebagai nilai acuan. Selanjutnya dilakukan pengujian beberapa kombinasi epsilon dan MinPts untuk membandingkan jumlah cluster, noise, dan kualitas pengelompokan.

Tabel 3. Hasil Pengujian Parameter DBSCAN

Table 3. DBSCAN Parameter Testing Results

Skenario	Epsilon	MinPts	Cluster	Nois	Silhouette
0				e	
1	0,20	4	11	16	1,000000
2	0,25	4	2	4	0,894071
3	0,30	4	2	4	0,894071
4	0,25	3	2	4	0,894071
5	0,25	5	2	4	0,894071

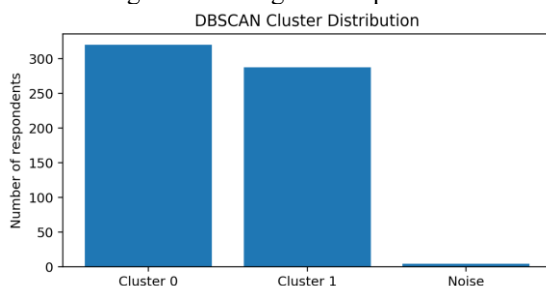
Skenario 2 dipilih karena menghasilkan dua cluster utama, hanya empat data noise, dan nilai Silhouette Score yang tinggi. Skenario 1 menghasilkan 11 cluster meskipun memiliki Silhouette Score 1,000000, sehingga struktur pengelompokannya dianggap terlalu terfragmentasi untuk tujuan interpretasi penelitian. Model final menggunakan epsilon 0,25 dan MinPts 4 dengan jarak Euclidean.

4. PEMBAHASAN

4.1 Hasil Clustering

Penerapan DBSCAN pada 612 responden menghasilkan dua cluster utama dan empat data noise. Cluster 0 terdiri atas 320 responden atau 52,29%, sedangkan Cluster 1 terdiri atas 288 responden atau 47,06%. Empat responden atau 0,65% diberi label -1 sebagai noise. Hasil ini menunjukkan bahwa sebagian besar data dapat dikelompokkan ke dalam dua pola kepuasan, sementara hanya sebagian kecil responden memiliki pola penilaian yang berbeda dari kepadatan utama.

*Nilai modus pada kelompok noise merupakan ringkasan dari hasil data yang teridentifikasi sebagai noise dan tidak digunakan sebagai dasar penamaan cluster.

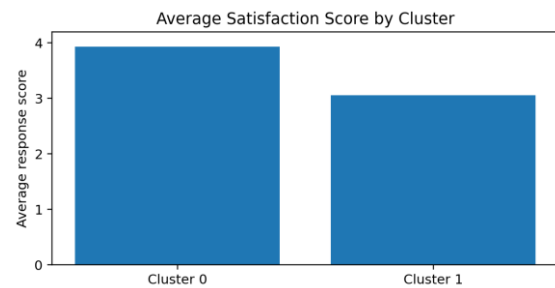


Gambar 1. Distribusi Hasil Cluster DBSCAN

Figure 1. DBSCAN Cluster Distribution

4.2 Kualitas Cluster

Evaluasi internal menghasilkan Silhouette Score sebesar 0,894071. Nilai tersebut menunjukkan struktur cluster yang kuat. Rata-rata Silhouette Coefficient Cluster 0 adalah 0,9064 dan Cluster 1 adalah 0,8804. Tidak ditemukan nilai Silhouette negatif pada kedua cluster, sehingga tidak terdapat anggota yang secara internal lebih dekat ke cluster lain dibandingkan cluster tempatnya berada.



Gambar 2. Rata-rata Skor Kepuasan Setiap Cluster

Figure 2. Average Satisfaction Score of Each Cluster

4.3 Profil Kepuasan dan Karakteristik Responden

Cluster 0 merupakan kelompok dengan kepuasan relatif lebih tinggi. Rata-rata jawaban sebesar 3,93 dan modus 4 menunjukkan kecenderungan penilaian sangat puas. Indikator dengan nilai tertinggi adalah Sikap Petugas sebesar 3,98, sedangkan nilai terendah adalah Waktu Tunggu sebesar 3,81. Cluster ini didominasi responden berusia 18–25 tahun, perempuan, dan bekerja sebagai pelajar/mahasiswa.

Cluster 1 memiliki kepuasan relatif lebih rendah dibandingkan Cluster 0, dengan rata-rata jawaban 3,06 dan modus 3. Nilai tertinggi terdapat pada Persyaratan sebesar 3,19, sedangkan Waktu Tunggu menjadi nilai terendah sebesar 2,99. Karakteristik dominan Cluster 1 juga adalah responden berusia 18–25 tahun, perempuan, dan pelajar/mahasiswa.

Kelompok noise berjumlah empat responden atau 0,65% dengan rata-rata jawaban 2,69. Kelompok ini didominasi responden berusia 26–35 tahun, laki-laki, dan pegawai/karyawan. Nilai terendah terdapat pada indikator Biaya sebesar 1,50. Perbandingan noise dan non-noise juga menunjukkan perbedaan terbesar pada Biaya, Pelayanan, dan Waktu Tunggu. Noise tidak menunjukkan kesalahan data, tetapi pola jawaban yang tidak umum dibandingkan kepadatan mayoritas responden.

Tabel 5. Profil Indikator pada Setiap Kelompok

Table 5. Indicator Profile of Each Group

Kelompok	Indikator tertinggi	Nilai	Indikator terendah	Nilai
Cluster 0	Sikap Petugas	3,98	Waktu Tunggu	3,81
Cluster 1	Persyaratan	3,19	Waktu Tunggu	2,99
Noise	Persyaratan	3,50	Biaya	1,50

4.4 Implikasi Hasil

Hasil Clustering memperlihatkan bahwa mayoritas responden memberikan penilaian positif, tetapi terdapat perbedaan tingkat kepuasan yang cukup jelas antara dua cluster utama. Temuan yang konsisten pada kedua cluster adalah Waktu Tunggu sebagai indikator dengan nilai relatif terendah. Selain itu, kelompok noise menunjukkan penilaian sangat rendah pada Biaya, Pelayanan, dan Waktu Tunggu. Dengan demikian, hasil analisis dapat digunakan sebagai informasi pendukung untuk mempertahankan aspek pelayanan yang sudah dinilai baik sekaligus

memprioritaskan evaluasi waktu tunggu dan aspek yang menonjol pada kelompok dengan penilaian berbeda.

5. KESIMPULAN

Algoritma DBSCAN berhasil diterapkan pada 612 data kepuasan pasien menggunakan dua fitur, yaitu modus jawaban dan rata-rata jawaban yang telah distandardisasi. Parameter final epsilon 0,25 dan MinPts 4 menghasilkan dua cluster utama, empat data noise, serta Silhouette Score 0,894071.

Cluster 0 berjumlah 320 responden (52,29%) dengan rata-rata jawaban 3,93 dan diinterpretasikan sebagai kelompok dengan kepuasan relatif lebih tinggi. Cluster 1 berjumlah 288 responden (47,06%) dengan rata-rata 3,06 dan diinterpretasikan sebagai kelompok dengan kepuasan relatif lebih rendah. Waktu Tunggu menjadi indikator dengan nilai relatif terendah pada kedua cluster utama.

Empat data noise (0,65%) menunjukkan pola penilaian yang berbeda dan memiliki penilaian sangat rendah terutama pada Biaya, Pelayanan, dan Waktu Tunggu. Temuan tersebut dapat menjadi informasi tambahan dalam evaluasi kualitas pelayanan RSGM Universitas Mulawarman.

6. SARAN

Rumah Sakit Gigi dan Mulut Universitas Mulawarman dapat mempertahankan aspek pelayanan yang memperoleh penilaian positif dan melakukan evaluasi berkelanjutan terhadap waktu tunggu.

Survei kepuasan dapat dilakukan secara berkala agar perubahan pola penilaian pasien dapat dipantau dari waktu ke waktu.

Penelitian selanjutnya dapat menggunakan periode data yang lebih panjang dan jumlah responden yang lebih besar, membandingkan DBSCAN dengan metode Clustering lain, serta menggunakan Grid Search untuk memperoleh parameter yang lebih optimal.

Pengembangan berikutnya dapat diarahkan pada sistem pemantauan kepuasan secara berkala atau real-time, penerapan Explainable Artificial Intelligence, penggabungan Clustering dengan prediksi, dan analisis sentimen terhadap komentar, kritik, serta saran pasien.

7. REFERENSI

- Badawi, M. S., Suria, F., Elfiani, E., Tsunami, W. D., & Saputra, R. A. (2025). Analisis klaster pasien Rumah Sakit Jiwa Kota Kendari tahun 2022–2023: Identifikasi pola dan karakteristik menggunakan metode DBSCAN. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 9(2), 3072–3079.
- Frangky, F., Sinaga, R., & Raihansyah, M. (2025). Analisis segmentasi pasien berdasarkan persepsi kualitas pelayanan dengan algoritma clustering. *Explorer*, 5(1), 52–58. <https://doi.org/10.47065/explorer.v5i1.1818>
- Harini, T. L., Mantjoro, E. M., & Napoleon Tatura, S. N. (2023). Analisis mutu pelayanan kesehatan dengan kepuasan pasien rawat jalan di Rumah Sakit Umum Pusat Ratatotok Buyat. *Jurnal Sains dan Teknologi*, 4(3), 189–198. <https://doi.org/10.55338/saintek.v4i3.1338>
- Istiqomah, H., Nisa, K., & Sandi A., A. S. (2025). Penerapan algoritma DBSCAN dalam mengidentifikasi risiko stroke.

METHOMIKA Jurnal Manajemen Informatika dan Komputerisasi Akuntansi, 9(2), 340–349. <https://doi.org/10.46880/jmika.Vol9No2.pp340-349>

- Kementerian Pendayagunaan Aparatur Negara dan Reformasi Birokrasi Republik Indonesia. (2017). Peraturan Menteri PANRB Nomor 14 Tahun 2017 tentang Pedoman Penyusunan Survei Kepuasan Masyarakat Unit Penyelenggara Pelayanan Publik.
- Rizana, N., Wahyuni, L., Nadiya, S., Zahara, R., & Elfira, C. (2024). Hubungan kualitas pelayanan home care dengan kepuasan pasien. *Jurnal Assyifa' Ilmu Kesehatan*, 9(1).
- Simbolon, I. N., & Friskila, P. D. (2024). Analisis dan evaluasi algoritma DBSCAN (Density-Based Spatial Clustering of Applications with Noise) pada tuberkulosis. *Jurnal Informatika dan Teknik Elektro Terapan*, 12(3S1). <https://doi.org/10.23960/jitet.v12i3s1.5206>
- Susilo, A. (2024). Evaluasi efektifitas algoritma K-Means dan DBSCAN dalam clustering data rumah sakit untuk optimalisasi layanan kesehatan. *Jurnal Komputer dan Informatika*, 19, 43–51.
- Zulvia, I., Hidayatulloh, A., & Rahmawati, D. E. (2022). Analisis kepuasan pasien terhadap pelayanan kesehatan di Klinik Alkindi Herbal menggunakan metode K-Means clustering. *Jurnal Informatika dan Komputer*, 6(2), 261–272.