Journal of Computing and Informatics Research
ISSN 2808-375X (Media Online)
Vol 5, No 3, July 2026
Hal 473−481
DOI 10.47065/comforch.v5i3.3072

Analisis Kondisi Sosial Ekonomi Santri Menggunakan Metode K-Means *Clustering*

Putri Nuraini Qolbiati*, Irfan Pratama

Fakultas Teknologi Informasi, Program Studi Sistem Informasi, Universitas Mercu Buana Yogyakarta, Yogyakarta, Indonesia
Email: [1,*]221220076@student.mercubuana-yogya.ac.id, [2]Irfanp@mercubuana-yogya.ac.id
Email Correspondencial: 221220076@student.mercubuana-yogya.ac.id

Abstrak−Analisis kondisi sosial ekonomi santri di MI Al Huda Kota Malang dilakukan berdasarkan indikator tingkat penghasilan, jenis pekerjaan, latar belakang pendidikan, dan kepemilikan tempat tinggal. Belum adanya sistem klasifikasi yang terorganisir menyebabkan penentuan tingkat kesejahteraan belum dapat dilakukan secara objektif. Penelitian ini menerapkan pendekatan data mining menggunakan algoritma K-Means Clustering untuk mengelompokkan data sosial ekonomi. Tahap awal meliputi preprocessing data berupa pemilihan variabel, penanganan data, dan normalisasi menggunakan StandardScaler. Selanjutnya, hasil clustering dievaluasi menggunakan Elbow Method, Silhouette Score, dan Davies-Bouldin Index untuk menentukan jumlah cluster yang optimal. Hasil evaluasi menunjukkan bahwa jumlah cluster terbaik adalah empat (k = 4) dengan nilai Silhouette Score sebesar 0,3341 dan Davies-Bouldin Index sebesar 1,1277. Sebagai pembanding, algoritma DBSCAN juga diterapkan, namun menghasilkan kualitas clustering yang lebih rendah dibandingkan K-Means. Kontribusi penelitian ini adalah membandingkan performa K-Means dan DBSCAN pada data sosial ekonomi santri menggunakan tiga metode evaluasi sehingga memberikan dasar yang lebih objektif dalam mendukung penentuan prioritas bantuan pendidikan. Secara keseluruhan, metode K-Means mampu menghasilkan pengelompokan yang lebih konsisten dan representatif sebagai dasar pengambilan keputusan terkait kondisi sosial ekonomi santri.

Kata kunci: Klasterisasi K-Means; Klasifikasi Sosial Ekonomi; *Data mining*; *Silhouette Score*; *Davies-Bouldin Index*; DBSCAN

Abstract−Analysis of the social conditions of students at MI Al Huda Malang City was conducted based on several important indicators, including income level, type of occupation, educational background, and home ownership. The absence of an organized classification system has caused the process of determining the level of welfare to be less objective. To overcome this problem, a data mining approach was applied through the K-Means Clustering algorithm as a data grouping method. The initial stage consisted of data preprocessing, including relevant variable selection, data handling, and normalization using StandardScaler to make the data distribution more uniform. Furthermore, the clustering results were evaluated using the Elbow Method, Silhouette Score, and Davies-Bouldin Index to determine the optimal number of clusters. Based on the evaluation results, the optimal number of clusters was four (k = 4), with a Silhouette Score of 0.3341 and a Davies-Bouldin Index of 1.1277. For comparison, the DBSCAN algorithm was also applied, but it showed lower clustering performance than K-Means. The contribution of this study is the comparison of the performance of K-Means and DBSCAN on students' socio-economic data using three evaluation methods, thereby providing a more objective basis for determining the priority of educational assistance. Overall, the K-Means method was proven to produce more consistent and representative clustering results, making it suitable as a basis for decision-making related to the socio-economic conditions of students.

Keywords: K-Means *Clustering*; Socio-Economic Classification; *Data mining*; *Silhouette Score*; *Davies-Bouldin Index*; DBSCAN

1. PENDAHULUAN

Kemajuan di bidang teknologi informasi turut memperluas penggunaan data dalam berbagai sektor, salah satunya pendidikan. Data yang dihasilkan tidak hanya mencakup aspek akademik, tetapi juga aspek sosial ekonomi siswa yang memiliki pengaruh terhadap proses pembelajaran. Kondisi sosial ekonomi dapat memengaruhi akses terhadap fasilitas pendidikan, motivasi belajar, serta tingkat keberhasilan akademik siswa. Namun, pemanfaatan data sosial ekonomi dalam lingkungan pendidikan masih belum optimal, sehingga menyulitkan dalam mengidentifikasi kondisi siswa secara objektif dan sistematis [1], [2].

Kondisi ini terjadi di MI AL Huda, dimana identifikasi status sosial ekonomi murid masih dilakukan secara konvensional. Oleh karena itu, pihak sekolah mengalami kesulitan dalam memetakan kondisi ekonomi siswa secara, yang berimbas pada kurang tepatnya pengambilan keputusan administratif, seperti penerimaan bantuan untuk siswa kurang mampu dan bantuan pendidikan lainnya. Untuk mengatasi permasalahan tersebut, diperlukan pendekatan berbasis data mining yang mampu mengekstraksi pola, hubungan, dan pengetahuan dari kumpulan data sehingga menghasilkan informasi yang bermanfaat sebagai dasar pengambilan keputusan [1], [3]. Dalam bidang pendidikan, data mining telah dimanfaatkan untuk menganalisis karakteristik siswa, perilaku penggunaan media sosial, serta mengelompokkan siswa berdasarkan karakteristik tertentu guna mendukung pengambilan keputusan di bidang pendidikan [2], [4].

Clustering merupakan salah satu teknik data mining yang digunakan untuk mengelompokkan objek berdasarkan tingkat kemiripan karakteristik sehingga objek dalam satu kelompok memiliki tingkat kesamaan yang lebih tinggi dibandingkan kelompok lainnya [3]. Algoritma K-means merupakan salah satu metode clustering yang banyak digunakan karena memiliki proses yang sederhana, efisien, dan mampu menghasilkan pengelompokan yang baik pada berbagai jenis data [3], [5]. Selain K-means, algortima DBSCAN digunakan sebagai metode clustering berbasis kepadatan yang mamu mengidentifikasi noise serta membentuk cluster tanpa menentukan jumlah cluster di awal [6], [7]. Berbagai penelitian telah membandingkan algortima K-Means dan DBSCAN pada berbagai jenis dataset untuk memperoleh algoritma dengan kualitas clustering terbaik [8], [9].

Berbagai penelitian telah menerapkan metode clustering dalam analisis sosial ekonomi dan pendidikan guna mendukung pengambilan keputusan. Penelitian oleh Putri Dkk. [10] mengindikasikan bahwa algoritma K-Means mampu melakukan pengelompokan siswa sesuai kondisi sosial ekonomi secara efektif. Temuan ini sejalan dengan penelitian oleh Sari dan Yasin menunjukkan bahwa metode clustering mampu mengelompokkan kondisi sosial ekonomi masyarakat secara efektif berdasarkan karakteristik yang dimiliki [11]. Selain itu, Khudin Dkk. [5] menunjukkan bahwa K-Means memiliki performa yang stabil pada berbagai dataset sosial ekonomi. Selain pada bidang sosial ekonomi, metode K-Means juga diterapkan dalam proses penentuan penerima Bantuan Langsung Tunai (BLT) dan pengelompokan prestasi belajar siswa sebagai pendukung pengambilan keputusan [4],

Journal of Computing and Informatics Research
ISSN 2808-375X (Media Online)
Vol 5, No 3, July 2026
Hal 473−481
DOI 10.47065/comforch.v5i3.3072

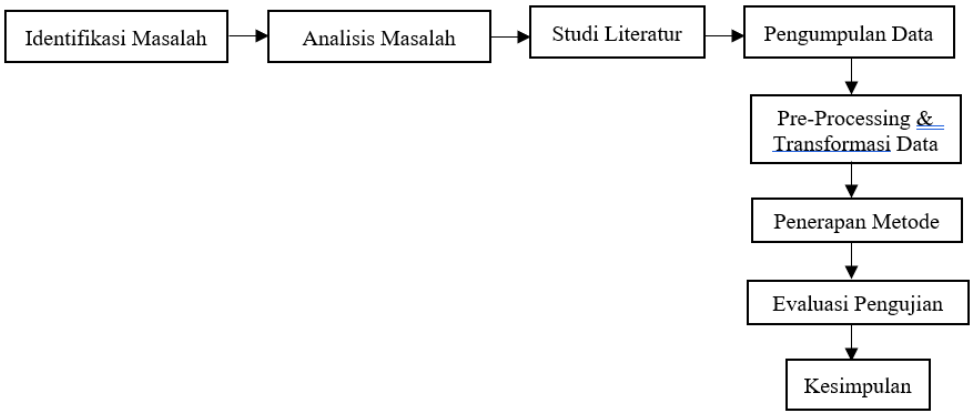
[12]. Akan tetapi, K-Means memiliki kelemahan ketika dihadapkan pada data yang mengandung outlier atau sebaran data yang tidak merata. Padahal, kondisi seperti ini umum ditemukan pada data sosial ekonomi. Sebagai solusinya, algoritma berbasis density-based mulai banyak digunakan. Penelitian lain membandingkan algoritma K-Means dan DBSCAN pada berbagai bidang, seperti pengelompokan data makanan [6], tren belanja [8], segmentasi UMKM pariwisata [13], pengelompokan siswa terbaik [14], segmentasi wilayah berdasarkan indikator mortalitas dan fertilitas [15], analisis profil siswa [16], serta data travel review [9].

Berdasarkan penelitian sebelumnya, sebagian besar studi hanya berfokus pada penerapan metode clustering tanpa melakukan analisis mendalam terhadap kondisi sosial ekonomi dalam konteks pendidikan tertentu. Selain itu, penelitian yang secara langsung membandingkan performa algoritma K-Means dan DBSCAN dalam mengelompokkan kondisi sosial ekonomi santri masih relatif terbatas. Oleh karena itu, penelitian ini dilakukan untuk memberikan gambaran mengenai karakteristik sosial ekonomi santri melalui proses clustering sekaligus mengevaluasi performa kedua algoritma berdasarkan kualitas cluster yang dihasilkan. Kontribusi utama penelitian ini adalah menyajikan analisis komparatif antara algoritma K-Means dan DBSCAN pada data sosial ekonomi santri yang dikombinasikan dengan evaluasi menggunakan *Elbow Method*, *Silhouette Score*, dan *Davies-Bouldin Index*. Selain menghasilkan pengelompokan kondisi sosial ekonomi yang representative, penelitian ini juga memberikan rekomendasi pemanfaatan hasil clustering sebagai dasar pengambilan keputusan dalam penentuan prioritas bantuan pendidikan di sekolah.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Metodologi penelitian ini menjelaskan tahapan yang dilakukan dalam proses analisis data untuk mengelompokkan kondisi sosial ekonomi santri melalui penerapan K-Means *Clustering* dengan pembanding metode DBSCAN. Pemilihan K-Means sebagai algoritma utama didasarkan pada efisiensi komputasinya terhadap data tabular berdimensi rendah dan kemampuannya menghasilkan batas klaster yang diskrit (hard clustering). Karakteristik ini sangat krusial untuk memberikan kepastian pelabelan profil sosial ekonomi siswa secara aplikatif. Sebagai pembanding spasial dan mekanisme penanganan data ekstrem, metode DBSCAN diimplementasikan untuk mengevaluasi kepadatan data dan keberadaan outliers secara inheren [1], [3]. Diagram pada Gambar 1 menggambarkan tahapan yang disusun mulai dari identifikasi masalah hingga evaluasi hasil. Langkah awal penelitian dilakukan melalui pengumpulan data dilanjutkan dengan preprocessing, selanjutnya dilakukan proses pengelompokan data menggunakan metode K-Means, serta pada tahap akhir dilakukan evaluasi hasil menggunakan metode *Silhouette Score* dan *Davies-Bouldin Index*.



Gambar 1. Tahapan Penelitian

a. Identifikasi Masalah
Proses ini bertujuan untuk mengidentifikasi masalah utama dalam penelitian, yaitu bagaimana mengelompokkan data sosial ekonomi santri berdasarkan karakteristik tertentu.

b. Analisis Masalah
Tahap ini dilakukan untuk memahami lebih dalam permasalahan yang telah diidentifikasi, termasuk penentuan variabel yang digunakan dan metode yang sesuai.

c. Studi Literatur
Studi literatur dilakukan dengan mengkaji berbagai referensi yang berkaitan dengan konsep data mining, algoritma K-Means, algoritma DBSCAN, serta metode evaluasi clustering menggunakan Silhouette Score dan Davies-Bouldin Index sebagai landasan dalam pelaksanaan penelitian [17]–[23].

d. Pengumpulan Data
Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari MI AL Huda Kota Malang. Dataset tersebut berisi informasi terkait kondisi sosial ekonomi santri yang meliputi penghasilan orang tua, pekerjaan orang tua, tingkat pendidikan orang tua, dan status kepemilikan rumah. Jumlah data yang berhasil dikumpulkan sebanyak 363 data santri. Sebelum digunakan pada tahap *clustering*, dilakukan pemeriksaan terhadap kualitas data untuk memastikan data layak dianalisis. Berdasarkan hasil pemeriksaan, seluruh data telah terisi dengan lengkap sehingga tidak ditemukan nilai yang hilang (*missing value*s). Selain itu, terdapat beberapa data yang memiliki kombinasi nilai atribut yang sama. Kondisi tersebut tidak dianggap sebagai kesalahan data karena masih memungkinkan beberapa santri memiliki karakteristik sosial ekonomi yang serupa. Oleh karena itu, seluruh data tetap digunakan dalam proses penelitian sehingga jumlah data yang dianalisis sebanyak 363 data.

e. Pre-Processing & Transformasi Data
Langkah prapemrosesan (preprocessing) bertujuan untuk mempersiapkan dataset sebelum memasuki fase pemodelan guna meningkatkan kualitas data sehingga hasil pengelompokan menjadi lebih [1][3]. Proses ini diawali dengan seleksi variabel

ISSN 2808-375X (Media Online)
Vol 5, No 3, July 2026
Hal 473−481
DOI 10.47065/comforch.v5i3.3072

penelitian yang memfokuskan analisis pada atribut penghasilan, pekerjaan, pendidikan, dan status kepemilikan rumah. Selanjutnya, hasil pemeriksaan kelengkapan data memastikan bahwa seluruh entri telah terisi penuh sehingga tidak memerlukan penanganan nilai kosong (missing values). Untuk mengidentifikasi pola persebaran dan keberadaan nilai ekstrem, deteksi outliers dievaluasi secara visual menggunakan boxplot. Mengingat keempat variabel tersebut memiliki rentang skala numerik yang bervariasi, tahap akhir menerapkan metode StandardScaler untuk menyeragamkan skala distribusi data. Meskipun metode normalisasi ini secara teoritis sensitif terhadap distorsi akibat outliers, kelemahan tersebut secara sengaja dikompensasi melalui pengujian komparatif dengan algoritma DBSCAN yang secara inheren mampu memisahkan outliers sebagai noise.

f. Penerapan Metode
Setelah melalui tahap pengolahan, selanjutnya dikelompokkan melalui penerapan K-Means. Proses ini melibatkan penentuan jumlah *cluster* (K), pemilihan centroid awal, penghitungan jarak, dan iterasi hingga *cluster* stabil.

g. Evaluasi Pengujian
Hasil pengelompokan dievaluasi untuk menentukan kualitasnya. Evaluasi dilakukan menggunakan metode seperti:
1) Elbow Method (bertujuan untuk menentukan jumlah cluster optimal).
2) Perbandingan hasil pengelompokan.
3) Evaluasi kualitas cluster yang terbentuk dilakukan menggunakan dua metrik, yaitu Silhouette Coefficient dan Davies-Bouldin Index (DBI). Secara khusus, Silhouette Coefficient berfungsi untuk mengukur derajat kemiripan data di dalam satu cluster yang sama jika dibandingkan dengan data pada cluster lainnya, di mana nilai yang lebih tinggi mengindikasikan pemisahan antar cluster yang lebih baik [24].

h. Kesimpulan/Hasil
Hasil pengelompokan menggunakan metode K-Means menunjukkan bahwa data sosial ekonomi siswa dapat dipetakan ke dalam tiga cluster yang memiliki karakteristik berbeda. Perbedaan tersebut mencerminkan adanya keragaman kondisi sosial ekonomi dilingkungan sekolah sehingga setiap kelompok memerlukan pendekatan yang sesuai dengan karakteristiknya. Temuan ini memberikan informasi yang dapat mendukung sekolah dalam Menyusun kebijakan, menetapkan prioritas penerima bantuan, serta merancang program pembinaan secara lebih objektif. Dengan demikian, penerapan metode K-Means berhasil mencapai tujuan penelitian, yaitu menghasilkan pengelompokan siswa berdasarkan tingkat kemiripan karakteristik sosial ekonomi.

2.2 K-Means Clustering

K-Means termasuk dalam salah satu metode *clustering* yang digunakan untuk membagi data ke dalam beberapa kelompok (*cluster*) berdasarkan tingkat kesamaan karakteristik. Dalam penerapannya, metode ini bertujuan untuk mengurangi jarak antara setiap data dengan pusat *cluster* (centroid), sehingga data yang berada pada satu kelompok memiliki tingkat kemiripan yang lebih tinggi dibandingkan dengan kelompok lainnya. Pada tahap awal, jumlah *cluster* (k) ditentukan sebagai parameter utama. Setelah itu, dilakukan penentuan centroid awal sebagai titik pusat dari masing-masing *cluster*. Selanjutnya, setiap data dihitung jaraknya terhadap centroid menggunakan jarak Euclidean, kemudian data ditempatkan pada *cluster* yang memiliki jarak paling dekat. Proses ini berlangsung secara berulang (iteratif) hingga posisi centroid tidak lagi mengalami perubahan yang signifikan. Tahapan dalam penerapan K-Means adalah sebagai berikut :

a. Penentuan Jumlah *Cluster*
Jumlah *cluster* (k) ditetapkan sebagai parameter awal dalam proses *clustering*. Penentuan nilai k dapat dilakukan dengan memanfaatkan beberapa metode evaluasi, seperti Elbow Method, *Silhouette Score*, dan *Davies-Bouldin Index*, sehingga diperoleh hasil pengelompokan yang optimal.

b. Normalisasi Data
Normalisasi dilakukan guna menyetarakan skala antar parameter guna meminimalkan pengaruh dominan oleh parameter yang memiliki nilai lebih besar. Proses normalisasi menggunakan metode standardisasi sebagai berikut :

$$z = \frac{x - \mu}{\sigma} \quad (1)$$

Pada Persamaan (1), *x* menyatakan nilai data yang akan dinormalisasi, μ merupakan nilai rata-rata dari setiap variabel, sedangkan σ merupakan nilai standar deviasi yang digunakan untuk menyesuaikan skala data sehingga setiap variabel memiliki distribusi yang seragam sebelum dilakukan proses clustering.

c. Perhitungan Jarak (Euclidean Distance)
Perhitungan jarak digunakan untuk menentukan kedekatan antara data dengan centroid dalam proses pengelompokan. Jarak perhitungan jarak dilakukan menggunakan rumus Euclidean Distance sebagai berikut :

$$D = \sqrt{\sum_{i=1}^{n} (x_i - c_i)^2} \quad (2)$$

Pada Persamaan (2), $\boldsymbol{D}$ menunjukkan jarak antara suatu data terhadap centroid, x_i merupakan nilai data ke-i, sedangkan c_i merupakan nilai centroid ke-i yang digunakan sebagai pusat cluster dalam proses perhitungan jarak *Euclidean*.

d. Fungsi Objektif K-Means
Tujuan dari algoritma K-Means adalah untuk mengurangi total jarak antara dataset dan centroid dalam setiap *cluster*.

$$J = \sum_{i=1}^{k} \sum_{x \in C_i} ||x - c_i||^2 \quad (3)$$

Pada Persamaan (3), *J* menyatakan total fungsi objektif yang diminimalkan dalam algoritma K-Means, c_i merupakan cluster ke-i, sedangkan c_i merupakan centroid dari cluster tersebut. Fungsi objektif digunakan untuk memperoleh pembagian cluster dengan total jarak terkecil antara data dan centroid.

Journal of Computing and Informatics Research
ISSN 2808-375X (Media Online)
Vol 5, No 3, July 2026
Hal 473−481
DOI 10.47065/comforch.v5i3.3072

2.3 DBSCAN

DBSCAN (Density-Based Spatial *Clustering* of Applications with *Noise*) merupakan metode c*lustering* yang membentuk kelompok data berdasarkan tingkat kepadatan data pada suatu wilayah. Metode ini mengelompokkan data yang memiliki kedekatan jarak dan kepadatan yang serupa sehingga mampu membentuk *cluster* secara alami. Berbeda dengan *K-Means* yang memerlukan jumlah *cluster* pada tahap awal, DBSCAN membentuk *cluster* berdasarkan distribusi data yang ada. Dalam Proses pengelompokan, DBSCAN menggunakan dua parameter utama, yaitu *epsilon (ε)* dan *min_samples*. Parameter *epsilon (ε)* digunakan untuk menentukan batas jarak maksimum antar data yang masih dianggap berada dalam satu lingkaran (*neighborhood)*, sedangkan *min_samples* digunakan untuk menentukan jumlah minimum data yang harus berada dalam radius *epsilon* agar dapat membentuk suatu *cluster*. Pemilihan kedua parameter tersebut berpengaruh terhadap hasil pengelompokan karena menentukan tingkat kepadatan minimum yang digunakan dalam pembentukan c*luster*. Selain mampu membentuk *cluster* berdasarkan kepadatan data, DBSCAN juga dapat mengidentifikasi data yang tidak termasuk ke dalam kelompok tertentu *(noise* atau *outlier).* Oleh karena itu, pada penelitian ini DBSCAN digunakan sebagai metode pembanding untuk mengevaluasi kualitas hasil pengelompokan yang diperoleh dari algoritma *K-Means*.

3. HASIL DAN PEMBAHASAN

Bagian ini memaparkan hasil pengolahan data menggunakan metode K-Means beserta penilaian terhadap hasil *clustering*.

3.1 Data Sampel Penelitian

Sumber data dalam penelitian ini berupa data sosial ekonomi seperti yang ditunjukan Tabel 1 santri MI Al Huda Kota Malang yang terdiri dari variabel penghasilan, pekerjaan, pendidikan, dan status kepemilikan rumah. Data ini digunakan sebagai dasar dalam proses pengelompokan menggunakan metode *clustering*.

Tabel 1. Atribut yang Digunakan dalam Penelitian

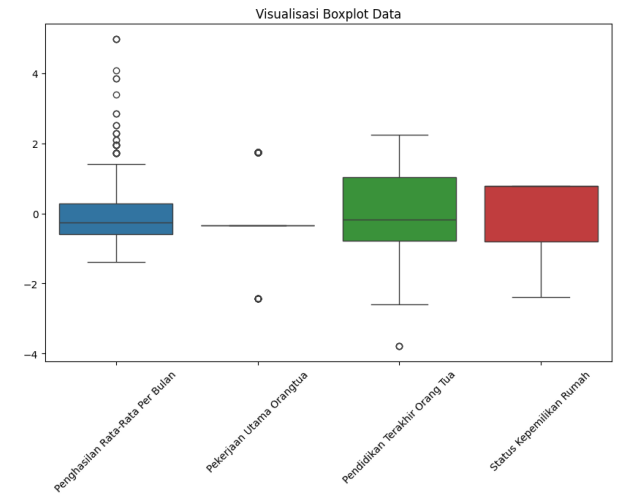
No.	Atribut	Deskripsi
1.	Penghasilan	Pendapatan Per Bulan
2.	Pekerjaan	Jenis Pekerjaan
3.	Pendidikan	Tingkat Pendidikan
4.	Kepemilikan Rumah	Kepemilikan Tempat Tinggal

3.2 Hasil Preprocessing

Data telah melalui tahap preprocessing yang meliputi seleksi variabel, penanganan *missing value*, deteksi *outlier*, serta normalisasi menggunakan *StandardScaler*. Berdasarkan hasil pemeriksaan data pada Tabel 2, jumlah data yang digunakan dalam penelitian ini sebanyak 363 data santri. Seluruh atribut yang digunakan telah terisi dengan lengkap sehingga tidak ditemukan *missing value*. Selain itu, terdapat 107 data yang memiliki kombinasi atribut yang sama. Data tersebut tetap dipertahankan karena tidak menunjukkan adanya kesalahan pencatatan dan masih merepresentasikan karakteristik sosial ekonomi yang sesuai dengan objek penelitian. Dengan demikian, seluruh data digunakan pada tahap *clustering*.

Tabel 2. Hasil Pemeriksaan Data

Tahapan Pemeriksaan Data	Hasil
Jumlah Data Awal	363
Missing value	0
Data dengan kombinasi atribut yang sama	107
Data yang dihapus	0
Jumlah Data yang digunakan	363



Gambar 2. Visualisasi Boxplot Data

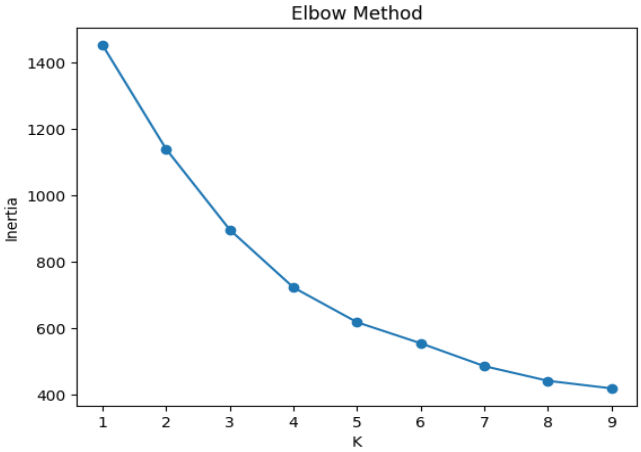
Mengacu Gambar 2, dapat dilihat bahwa sebagian besar data berada dalam rentang distribusi yang wajar. Namun, terdapat

Journal of Computing and Informatics Research
ISSN 2808-375X (Media Online)
Vol 5, No 3, July 2026
Hal 473−481
DOI 10.47065/comforch.v5i3.3072

beberapa nilai yang berada di luar batas (*outlier*) pada beberapa variabel. Keberadaan *outlier* tersebut masih dalam batas yang dapat diterima dan tidak terlalu ekstrem sehingga tidak mengganggu proses *clustering*. Selain itu, hasil normalisasi menunjukkan bahwa data telah memiliki skala yang seragam antar variabel, sehingga siap digunakan dalam proses pengelompokan.

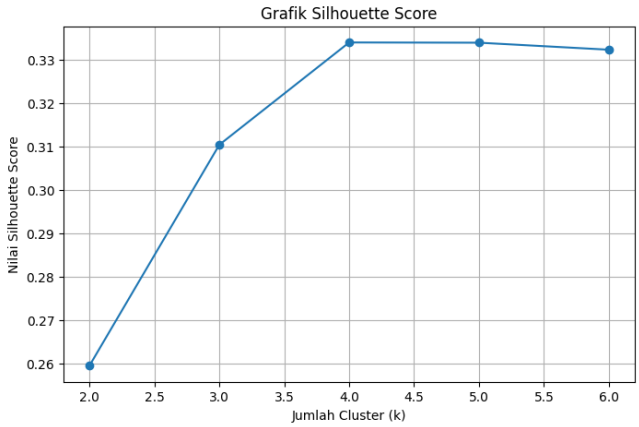
3.3 Penentuan Jumlah *Cluster*

Jumlah kelompok ditentukan melalui penerapan metode Elbow dan *Silhouette Score* [9]. Guna menentukan banyaknya kelompok data optimal, digunakan grafik Elbow Method seperti pada Gambar 3.



Gambar 3. Grafik Elbow Method

Mengacu pada Gambar 3, terlihat bahwa nilai inertia mengalami penurunan yang cukup besar pada k = 1 hingga k = 4. Selanjutnya k = 4, nilai yang mengalami penurunan cukup besar. Kondisi tersebut mengindikasikan bahwa titik siku (elbow) terletak pada k = 4, sehingga jumlah *cluster* yang paling sesuai untuk digunakan dalam penelitian ini adalah 4. Selain itu, digunakan *Silhouette Score* untuk mengevaluasi kualitas *cluster* seperti pada Gambar 4.



Gambar 4. Grafik Silhouette Score

Berdasarkan Gambar 4, nilai *Silhouette Score* tertinggi ditemukan pada k = 4. Meskipun nilai pada k = 5 memiliki nilai yang mendekati, namun k = 4 dipilih karena memberikan nilai tertinggi. Hal ini memperkuat hasil metode Elbow yang mengindikasikan bahwa jumlah *cluster* yang optimal adalah 4.

3.4 Hasil Pengelompokan K-Means

Setelah dilakukan proses pengelompokan menggunakan algoritma K-Means, diperoleh hasil pengelompokan data dibagi ke dalam 4 kelompok berdasarkan tingkat kesamaan karakteristik [1].

3.4.1 Distribusi Cluster

Distribusi *cluster* pada Tabel 3 menunjukkan jumlah data yang masuk ke dalam masing-masing kelompok hasil pengelompokan data dilakukan menggunakan algoritma K-Means. Hasil pengelompokan menunjukkan bahwa *cluster* 0 menampung data paling banyak, yaitu 154 data, diikuti oleh *cluster* 3 sebanyak 104 data, *cluster* 1 sebanyak 77 data, sedangkan *cluster* 2 memiliki jumlah paling sedikit, yaitu 28 data. Perbedaan jumlah anggota pada setiap cluster menunjukkan bahwa kondisi sosial ekonomi santri tidak tersebar secara merata. Dominasi jumlah data pada Cluster 0 mengindikasikan bahwa sebagian besar santri berada pada kelompok ekonomi menengah bawah. Informasi ini dapat menjadi dasar bagi pihak sekolah untuk memperkirakan proporsi kebutuhan bantuan maupun penyusunan program yang lebih sesuai dengan kondisi mayoritas santri.

Tabel 3. Distribusi *Cluster*

Journal of Computing and Informatics Research
ISSN 2808-375X (Media Online)
Vol 5, No 3, July 2026
Hal 473−481
DOI 10.47065/comforch.v5i3.3072

Cluster	Jumlah Data
0	154
1	77
2	28
3	104

Distribusi *cluster* menunjukkan bahwa *cluster* 0 memiliki total data terbanyak, sedangkan *cluster* 2 merupakan *cluster* dengan jumlah data paling sedikit. Hal ini menunjukkan adanya variasi jumlah anggota pada setiap *cluster*.

3.4.2 Nilai Centroid

Nilai centroid seperti yang ditunjukan pada Tabel 4 merupakan nilai rata-rata setiap atribut pada masing-masing kelompok yang dihasilkan dari pengelompokan data dilakukan dengan algoritma K-Means. Centroid digunakan untuk merepresentasikan karakteristik dari setiap *cluster* berdasarkan variabel yang digunakan dalam penelitian. Nilai centroid dihitung menggunakan Persamaan (4) untuk menentukan rata-rata setiap variabel dalam *cluster*.

$$C = \frac{1}{n}\sum_{i=1}^{n} x_i \quad (4)$$

Tabel 4. Nilai Centroid

Cluster	Penghasilan	Pekerjaan	Pendidikan	Status Rumah
0	6.558.117	1.97	4.24	4.95
1	6.825.974	3.00	4.09	3.88
2	19.217.857	1.89	2.27	4.61
3	5.568.750	1.91	4.00	2.57

Berdasarkan nilai centroid pada Tabel 4, setiap cluster menunjukkan karakteristik sosial ekonomi yang berbeda sehingga dapat digunakan untuk menggambarkan kondisi ekonomi santri secara lebih spesifik. Cluster 3 didominasi oleh santri dengan tingkat penghasilan keluarga yang paling rendah disertai kondisi kepemilikan rumah yang relatif kurang baik. Karakteristik tersebut menunjukkan bahwa kelompok ini dapat dikategorikan sebagai kelompok ekonomi rentan sehingga berpotensi menjadi prioritas dalam pemberian bantuan pendidikan, keringanan biaya sekolah, maupun program kesejahteraan yang disediakan oleh MI Al Huda Kota Malang. Cluster 0 menggambarkan kelompok santri dengan kondisi sosial ekonomi menengah bawah. Meskipun memiliki pendapatan yang lebih baik dibandingkan Cluster 3, karakteristik ekonomi kelompok ini masih memerlukan perhatian karena sebagian keluarga berada pada kondisi ekonomi yang belum sepenuhnya stabil. Cluster 1 menunjukkan kondisi sosial ekonomi yang relatif lebih stabil. Kelompok ini memiliki karakteristik pekerjaan orang tua yang lebih baik dibandingkan cluster lainnya sehingga dapat dikategorikan sebagai kelompok ekonomi menengah yang umumnya telah mampu memenuhi kebutuhan pendidikan secara mandiri. Sementara itu, Cluster 2 memiliki rata-rata pendapatan keluarga paling tinggi dibandingkan cluster lainnya. Kondisi tersebut menunjukkan bahwa kelompok ini memiliki kemampuan ekonomi yang lebih baik sehingga kebutuhan pendidikan cenderung dapat dipenuhi tanpa memerlukan prioritas bantuan finansial. Perbedaan karakteristik antarcluster menunjukkan bahwa metode K-Means mampu mengidentifikasi variasi kondisi sosial ekonomi santri yang dapat dimanfaatkan sebagai dasar dalam penyusunan kebijakan sekolah secara lebih objektif.

3.4.3 Iterasi dan Konvergensi

Proses *clustering* dilakukan secara iteratif hingga tidak terdapat perubahan yang signifikan pada centroid. Hal ini mengindikasikan bahwa proses telah mencapai kondisi konvergen dan hasil *clustering* sudah stabil untuk digunakan dalam analisis.

3.5 Evaluasi Cluster

Evaluasi dilakukan melalui penggunaan *Silhouette Score* dan *Davies-Bouldin Index* [9], [19].
Hasil evaluasi menunjukkan bahwa metode K-Means menghasilkan nilai *Silhouette Score* sebesar 0,3341 dan *Davies-Bouldin Index* sebesar 1,1277. Nilai tersebut menunjukkan bahwa hasil clustering telah mampu membentuk kelompok data dengan tingkat pemisahan yang cukup baik meskipun masih terdapat sedikit kemiripan antarcluster yang merupakan karakteristik umum pada data sosial ekonomi. Nilai *Silhouette Score* sebesar 0,3341 menunjukkan bahwa masih terdapat kedekatan karakteristik antar beberapa cluster sehingga batas pemisahan antarkelompok belum sepenuhnya tegas. Kondisi tersebut merupakan hal yang wajar pada data sosial ekonomi karena karakteristik antar keluarga, seperti tingkat penghasilan, pekerjaan, maupun pendidikan orang tua, sering kali memiliki perbedaan yang tidak terlalu jauh sehingga menyebabkan adanya irisan (overlap) antar kelompok. Sementara itu, nilai Davies-Bouldin Index sebesar 1,1277 menunjukkan bahwa masing-masing cluster masih memiliki tingkat kemiripan tertentu dengan cluster lainnya. Meskipun demikian, hasil evaluasi ini tetap menunjukkan bahwa proses clustering berhasil membentuk kelompok-kelompok yang memiliki karakteristik sosial ekonomi yang berbeda dan masih dapat digunakan sebagai dasar dalam proses pengelompokan santri sesuai tujuan penelitian.

3.6 Perbandingan dengan DBSCAN

Sebagai metode pembanding, digunakan algoritma DBSCAN. Pada penelitian ini digunakan parameter *epsilon (ε)* sebesar 1,5 dan *min_samples* sebesar 5. Nilai *epsilon* digunakan sebagai batas jarak maksimum antar data yang masih dianggap berada dalam satu lingkungan, sedangkan *min_samples* digunakan untuk menentukan jumlah minimum data yang diperlukan dalam pembentukan *cluster*. Pemilihan parameter tersebut dilakukan untuk menyesuaikan karakteristik data yang telah melalui proses normalisasi menggunakan *StandardScaler*. Selanjutnya, hasil pengelompokan dievaluasi menggunakan *Silhouette Score* dan *Davies-Bouldin Index* untuk dibandingkan dengan hasil yang diperoleh dari metode *K-Means*. Berdasarkan hasil seperti yang ditunjukan pada Tabel 5, metode K-Means menghasilkan nilai Silhouette yang lebih tinggi dan nilai DBI yang lebih rendah dibandingkan DBSCAN. Hal ini menunjukkan

Journal of Computing and Informatics Research
ISSN 2808-375X (Media Online)
Vol 5, No 3, July 2026
Hal 473−481
DOI 10.47065/comforch.v5i3.3072

bahwa K-Means memiliki kualitas *clustering* yang lebih baik pada dataset penelitian ini.

Tabel 5. Perbandingan

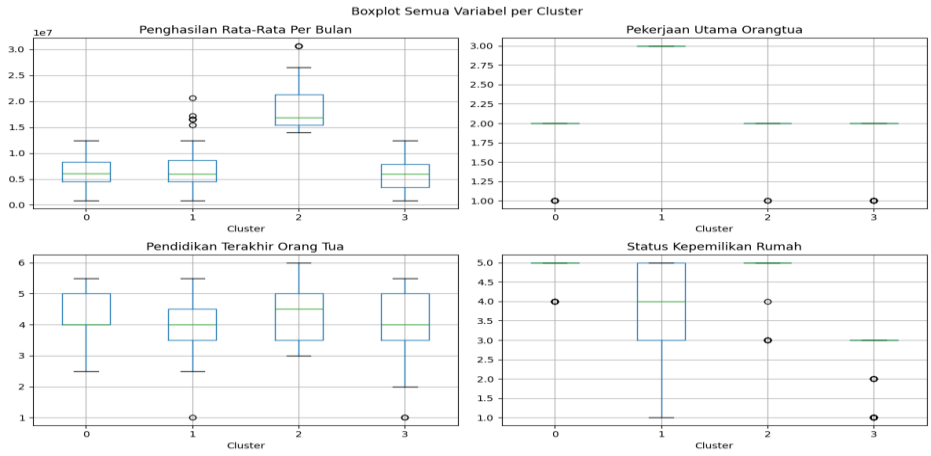
Metode	Silhouette	DBI
K-Means	0.3341	1.1277
DBSCAN	0.2809	1.3382

3.7 Interpretasi Cluster

Berdasarkan nilai centroid, dilakukan interpretasi terhadap masing-masing *cluster*. Hasil interpretasi pada Tabel 6 menunjukkan bahwa setiap cluster merepresentasikan kelompok santri dengan karakteristik sosial ekonomi yang berbeda. Cluster 3 dikategorikan sebagai kelompok ekonomi rendah karena memiliki rata-rata pendapatan keluarga paling rendah serta kondisi kepemilikan rumah yang relatif kurang baik. Kelompok ini dapat menjadi prioritas utama dalam pemberian bantuan pendidikan maupun program keringanan biaya sekolah. Cluster 0 termasuk kategori menengah bawah. Kelompok ini memiliki kondisi ekonomi yang lebih baik dibandingkan Cluster 3, namun masih berpotensi mengalami keterbatasan dalam memenuhi kebutuhan pendidikan sehingga tetap memerlukan perhatian dari pihak sekolah. Cluster 1 merepresentasikan kelompok ekonomi menengah dengan kondisi pekerjaan dan pendapatan keluarga yang relatif stabil. Sementara itu, Cluster 2 menunjukkan kelompok dengan kondisi ekonomi tinggi yang memiliki kemampuan finansial lebih baik dibandingkan cluster lainnya. Pengelompokan ini memberikan informasi yang dapat dimanfaatkan oleh pihak MI Al Huda Kota Malang dalam menyusun kebijakan yang lebih tepat sasaran, seperti penentuan prioritas penerima bantuan pendidikan, pemberian subsidi biaya sekolah, maupun penyusunan program pembinaan sesuai kondisi sosial ekonomi masing-masing kelompok santri. Untuk melihat perbedaan distribusi data pada setiap *cluster*, digunakan visualisasi boxplot seperti pada Gambar 5.

Tabel 6. Interpretasi Hasil *Clustering*

Cluster	Karakteristik	Implikasi bagi Sekolah
0	Ekonomi menengah bawah dengan pendapatan relative stabil namun masih terbatas	Perlu pemantauan dan dapat dipertimbangkan sebagai penerima bantuan apabila memenuhi kriteria lain
1	Ekonomi menengah dengan kondisi pekerjaan orang tua relative stabil	Tidak menjadi prioritas utama penerima bantuan
2	Ekonomi tinggi dengan kemampuan finansial lebih baik	Bukan prioritas bantuan pendidikan
3	Ekonomi rentan dengan pendapatan rendah dan kondisi tempat tinggal relative kurang baik	Menjadi prioritas utama untuk program bantuan atau keringanan biaya pendidikan



Gambar 5. Boxplot Data Berdasarkan *Cluster*

Berdasarkan Gambar 5, terlihat bahwa setiap *cluster* memiliki distribusi nilai yang berbeda pada masing-masing variabel. Perbedaan ini mengindikasikan bahwa hasil *clustering* mampu membagi data ke dalam kelompok dengan karakteristik yang berbeda, oleh karena itu memperkuat hasil interpretasi *cluster* yang telah dilakukan.

Secara keseluruhan, penerapan metode K-Means berhasil mengelompokkan data sosial ekonomi santri ke dalam empat kelompok yang memiliki karakteristik berbeda. Hasil pengelompokan menunjukkan bahwa kondisi sosial ekonomi santri bersifat heterogen sehingga memerlukan pendekatan yang disesuaikan dengan karakteristik masing-masing kelompok. Informasi yang dihasilkan dari proses clustering ini dapat dimanfaatkan oleh MI Al Huda Kota Malang sebagai dasar dalam menentukan prioritas penerima bantuan pendidikan, penyusunan kebijakan pembiayaan sekolah, maupun program pembinaan yang lebih tepat sasaran. Dengan demikian, tujuan penelitian untuk memetakan kondisi sosial ekonomi santri berdasarkan tingkat kemiripan karakteristik telah tercapai.

3.8 Pembahasan

Hasil penelitian menunjukkan bahwa algoritma K-Means mampu menghasilkan kualitas clustering yang lebih baik dibandingkan DBSCAN berdasarkan nilai *Silhouette Score* sebesar 0,3341 dan *Davies-Bouldin Index* sebesar 1,1277. Temuan ini menunjukkan

bahwa K-Means lebih sesuai digunakan pada data sosial ekonomi santri yang memiliki karakteristik relative homogen setelah melalui proses normalisasi menggunakan *StandarScaler*.

Hasil penelitian ini sejalan dengan penelitian Putri dan Mandala [10] yang menunjukkan bahwa algoritma K-Means mampu mengelompokkan siswa berdasarkan kondisi sosial ekonomi secara efektif sehingga dapat dimanfaatkan sebagai dasar pengambilan keputusan dalam bidang pendidikan. Temuan ini juga mendukung penelitian Khudin dkk. [5] yang menyatakan bahwa K-Means memiliki performa yang stabil pada berbagai dataset sosial ekonomi.

Selain itu, hasil penelitian ini konsisten dengan penelitian Insany dkk. [6] dan Nafandra dkk. [15] yang membandingkan algoritma K-Means dan DBSCAN. Kedua penelitian tersebut menunjukkan bahwa performa masing-masing algoritma sangat dipengaruhi oleh karakteristik dataset yang digunakan. Pada penelitian ini, K-Means memberikan hasil evaluasi yang lebih baik dibandingkan DBSCAN karena data yang telah dinormalisasi memiliki pola penyebaran yang lebih sesuai dengan pendekatan berbasis centroid dibandingkan pendekatan berbasis kepadatan.

Dengan demikian, penelitian ini memperkuat temuan penelitian sebelumnya bahwa algoritma K-Means merupakan metode yang efektif untuk analisis data sosial ekonomi pada lingkungan pendidikan, sekaligus memberikan kontribusi berupa evaluasi komparatif terhadap DBSCAN sehingga dapat menjadi referensi dalam pemilihan algoritma clustering pada penelitian sejenis.

4. KESIMPULAN

Penelitian ini telah berhasil mengelompokkan data kondisi sosial ekonomi terhadap 363 santri di MI AL Huda Kota Malang menggunakan metode K-Means Clustering menjadi empat klaster terbaik (k=4). K-Means terbukti lebih stabil dibandingkan algoritma DBSCAN. Secara praktis, hasil pengelompokan ini dapat dimanfaatkan langsung oleh pihak sekolah sebagai dasar pembuatan kebijakan administratif. Pihak pengelola sekolah dapat menjadikan klaster santri dengan tingkat ekonomi terendah sebagai prioritas utama penerima beasiswa atau keringanan biaya pendidikan. Sebaliknya, klaster santri dari keluarga menengah ke atas dapat diarahkan untuk program subsidi silang fasilitas sekolah, sehingga keputusan manajerial lebih tepat sasaran.Namun, penelitian ini juga memiliki keterbatasan pada hasil evaluasinya. Pengujian menghasilkan nilai Silhouette Score sebesar 0,3341 dan Davies-Bouldin Index 1,1277. Angka Silhouette tersebut menunjukkan adanya sebaran data yang saling tumpang tindih (overlap) antar kelompok. Kondisi tumpang tindih ini sangat wajar terjadi karena batas perbedaan penghasilan masyarakat di lapangan selisih hanya sedikit. Metode K-Means yang membagi data memiliki kesulitan memisahkan data sosial yang samar. Oleh karena itu, penelitian selanjutnya disarankan untuk menggunakan algoritma seperti Fuzzy C-Means yang mampu menangani data secara fleksibel, sehingga hasil pengelompokan bisa lebih akurat. Kontribusi penelitian ini tidak hanya menunjukkan bahwa algoritma K-Means memberikan performa yang lebih baik dibandingkan DBSCAN pada dataset sosial ekonomi santri, tetapi juga menghasilkan model pengelompokan yang dapat digunakan sebagai dasar pengambilan keputusan dalam penyaluran bantuan pendidikan secara lebih objektif dan tepat sasaran.

REFERENCES

[1] A. A. Baskara., N. Hidayati., and D. Kartini., "Framework Data Mining : Sebuah Survei," *J. Media Inform. Budidarma*, vol. 9, no. 2, pp. 899–911, 2025, doi: 10.30865/mib.v9i2.8879.

[2] Utari. and M. Iqbal, "Implementasi Data Mining dalam Pengelompokan Perilaku Penggunaan Media Sosial Siswa SMA Menerapkan Algoritma K-Means," *JIMAT*, vol. 6, no. 1, pp. 44–51, 2026, doi: 10.47065/jimat.v6i1.959.

[3] I. Pii., N. Suarna., and N. Rahaningsih., "Penerapan Data Mining pada Penjualan Produk Pakaian Dameyra Fashion Menggunakan Metode K-Means Clustering," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 7, no. 1, pp. 423–430, 2023, doi: 10.36040/jati.v7i1.6336.

[4] R. Dila, S. Defit, and S. Arlis, "Analisis Algoritma K-Means Clustering dalam Pengelompokan Prestasi Belajar Siswa Menengah Atas (SMA)," *Bull. Comput. Sci. Res.*, vol. 5, no. 5, pp. 1113–1119, 2025, doi: 10.47065/bulletincsr.v5i5.751.

[5] F. Khudin, I. D. Saputra, M. Ilham, M. S. H. Fadilah, and D. S. Rahmadan, "Analisis Komparatif Penerapan K-Means Clustering pada Lima Dataset Nyata untuk Evaluasi Sosial Ekonomi dan Finansial," *J. Artif. Intell. Digit. Bus.*, vol. 4, no. 2, pp. 5081–5091, 2025, doi: 10.31004/riggs.v4i2.1361.

[6] G. P. Insany., J. L. Buliali., and A. Affandi., "Comparison of K-Means and DBSCAN Web-Based Food Clustering Based on Nutritional Content," *J. Nas. Tek. Elektro dan Teknol. Inf.*, vol. 14, no. 2, pp. 194–202, 2025, doi: 10.22146/jnteti.v14i2.11275.

[7] M. S. Abidin, Y. Kustiyahningsih, E. Rahmanita, B. D. Satoto, and M. I. Firmansyah, "Application of the DBSCAN Algorithm in MSME Clustering using the Silhouette Coefficient Method," *J. Sist. Cerdas*, vol. 7, no. 3, pp. 260–270, 2024, doi: 10.37396/jsc.v7i3.472.

[8] V. Wulandari., Y. Syarif., Z. Alfian., M. A. Althof., and M. Mufidah., "Comparison of Density-Based Spatial Clustering of Applications with Noise (DBSCAN), K-Means and X-Means Algorithms on Shopping Trends Data," *Int. J. Adv. Technol. Inf. Syst.*, vol. 1, no. 1, pp. 11–18, 2024, doi: 10.57152/ijatis.v1i1.1135.

[9] N. A. Yolandari, L. E. Butarbutar, G. C. H. Rajagukguk, M. F. Zulfi, and F. Ramadhani, "Analisis Perbandingan K-Means dan DBSCAN dalam Pengelompokan Data Travel Review Ratings Menggunakan Evaluasi Silhouette Index dan Davies-Bouldin Index," *JITET (Jurnal Inform. dan Tek. Elektro Ter.*, vol. 13, no. 3, pp. 470–481, 2023, doi: 10.23960/jitet.v13i3.6884.

[10] D. E. Putri, E. Praja, and W. Mandala, "Penerapan K-Means Clustering dalam Segmentasi Siswa Berdasarkan Status Sosial Ekonomi," *Progresif J. Ilm. Komput.*, vol. 21, no. 2, pp. 483–494, 2025, doi: 10.35889/progresif.v21i2.2809.

[11] R. Sari and M. Yasin, "Penerapan Data Mining untuk Clustering Kondisi Sosial Ekonomi Berdasarkan Kepemilikan Jaminan Kesehatan," *J. Tek. Inform. dan Teknol. Inf.*, vol. 5, no. 3, pp. 44–55, 2025, doi: 10.55606/jutiti.v5i3.6082.

[12] R. L. Budiarti., L. Puad., and Rismawati., "Penerapan Metode K-Means Clustering Untuk Proses Penentuan Bantuan Langsung Tunai (BLT) Pada Desa Mekar Sari Kecamatan Kumpeh Kabupaten Muaro Jambi," *J. Akad.*, vol. 16, no. 2, pp. 57–63, 2024, doi: 10.53564/akademika.v16i2.1190.

[13] R. Ardiansyah., A. Fandi., Z. Ibnu., and A. Masayoshi., "Evaluation of K-Means, DBSCAN, and Hierarchical Clustering for Strategic Segmentation of Tourism SMEs in Rembang, Indonesia," *J. Tek. Inform.*, vol. 6, no. 3, pp. 1605–1630, 2025, doi: 10.52436/1.jutif.2025.6.3.4602.

ISSN 2808-375X (Media Online)
Vol 5, No 3, July 2026
Hal 473−481
DOI 10.47065/comforch.v5i3.3072

[14] M. S. Hasibuan, A. H. Lubis, and M. N. Sari, "Perbandingan Algoritma Clustering DBSCAN dan K-Means dalam Pengelompokan Siswa Terbaik," *J. Inform. dan Teknol.*, vol. 5, no. 2, pp. 301–309, 2024, doi: 10.37373/infotech.v5i2.1457.
[15] B. Nafandra., D. Sulistiowati., and N. Amalita., "Comparison of K-Means and DBSCAN Algorithms in Regional Segmentation in Indonesia Based on Mortality and Fertility Indicators," *J. MSA (Matematika dan Stat. serta Apl.*, vol. 14, no. 1, pp. 24–35, 2026, doi: 10.24252/msa.v14i1.61961.
[16] A. Fauziah., "Perbandingan Metode K-Means dan DBSCAN pada Analaisis Klaster Multivariat Profil Siswa," *J. Ilm. Multidisiplin Ilmu*, vol. 3, no. 1, pp. 48–56, 2026, doi: 10.69714/1kays705.
[17] A. Ardhi Baskara, N. Maharani Piranti, and M. Fahrury Romdendine, "Framework Data Mining: Sebuah Survei," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 9, no. 3, pp. 4886–4895, 2025, doi: 10.36040/jati.v9i3.13803.
[18] I. Pii, N. Suarna, and N. Rahaningsih, "Penerapan Data Mining Pada Penjualan Produk Pakaian Dameyra Fashion Menggunakan Metode K-Means Clustering," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 7, no. 1, pp. 423–430, 2023, doi: 10.36040/jati.v7i1.6336.
[19] G. P. Insany, A. Fergina, and M. I. Juardi, "Comparison of K-means and DBSCAN Web- Based Food in Clustering Based on Nutritional Content," *bit-Tech*, vol. 8, no. 2, pp. 1245–1255, 2025, doi: 10.32877/bt.v8i2.2813.
[20] M. S. Abidin, Y. Kustiyahningsih, E. Rahmanita, B. D. Satoto, and M. I. Firmansyah, "Application of the DBSCAN Algorithm in MSME Clustering using the Silhouette Coefficient Method," *J. Sist. Cerdas*, vol. 7, no. 3, pp. 260–270, 2024, doi: 10.37396/jsc.v7i3.472.
[21] S. Paembonan. and H. Abduh., "Penerapan Metode Silhouette Coefficient Untuk Evaluasi Clustering Obat," *PENA Tek. J. Ilm. Ilmu-Ilmu Tek.*, vol. 6, no. 2, pp. 48–54, 2021, doi: 10.51557/pt_jiit.v6i2.690.
[22] Y. Hasan, "Pengukuran Silhouette Score dan Davies-Bouldin Index pada Hasil Cluster K-Means dan DBSCAN," *KAKIFIKOM (Kumpulan Artik. Karya Ilm. Fak. Ilmu Komputer)*, vol. 6, no. 1, pp. 60–74, 2024, [Online]. Available: https://ejournal.ust.ac.id/index.php/KAKIFIKOM/article/view/3938
[23] I. F. Ashari., E. D. Nugroho., D. D. Andrianto., and M. A. N. M. Y. R. Liwardana., "Analysis of Elbow, Silhouette, Davies-Bouldin, Calinski-Harabasz, and Rand-Index Evaluation on K-Means Algorithm for Classifying Flood-Affected Areas in Jakarta," *J. Appl. Informatics Comput.*, vol. 7, no. 1, pp. 95–103, 2023, doi: 10.30871/jaic.v7i1.4947.
[24] Muslim, N. Mayasari, and W. Fitriani, "Clustering K-Means and DBSCAN for Network Traffic Anomaly Detection in Digital Islamic Community Systems," in *2nd International Conference on Islamic Community Studies (ICICS)*, 2025.