

## PENERAPAN ALGORITMA *K-MEANS CLUSTERING* UNTUK SEGMENTASI PENGGUNA DISCORD BERDASARKAN POLA PENGGUNAAN DAN TINGKAT KEPUASAN

Dea Alya<sup>1</sup>, Tiara Bela Harahap<sup>2</sup>, Salsabila Mahfuza<sup>3</sup>, Naina Nazwa Hasibuan<sup>4</sup>, Mhd. Furqan<sup>5</sup>

<sup>1,2,3,4,5</sup> Ilmu Komputer, Fakultas Sains dan Teknologi, Universitas Islam Negeri Sumatera Utara  
<sup>1</sup>alya0701232102@uinsu.ac.id, <sup>2</sup>tiarabelahrp@gmail.com, <sup>3</sup>salsabilamahfuza@gmail.com,  
<sup>4</sup>nainanazwa212@gmail.com, <sup>5</sup>mfurqan@uinsu.ac.id

### ABSTRAK

Discord merupakan platform komunikasi digital yang digunakan untuk berbagai kebutuhan seperti komunitas, hiburan, pembelajaran, dan komunikasi daring. Perbedaan pola penggunaan Discord menyebabkan munculnya karakteristik pengguna dan tingkat kepuasan yang berbeda sehingga diperlukan proses segmentasi pengguna. Penelitian ini bertujuan untuk melakukan segmentasi pengguna Discord menggunakan metode *K-Means Clustering* berdasarkan pola penggunaan dan tingkat kepuasan pengguna. *Dataset* penelitian diperoleh melalui penyebaran kuesioner daring kepada 200 responden. Proses penelitian meliputi *preprocessing* data, pengujian reliabilitas menggunakan Cronbach Alpha, transformasi data, normalisasi menggunakan StandardScaler, penentuan jumlah *cluster* menggunakan *Elbow Method*, serta evaluasi model menggunakan *Silhouette Score*. Seluruh proses pengolahan data dilakukan menggunakan Google Colab berbasis Python. Hasil pengujian reliabilitas memperoleh nilai *Cronbach Alpha* sebesar 0,861 yang menunjukkan bahwa data penelitian memiliki tingkat konsistensi yang baik. Hasil penelitian menunjukkan bahwa jumlah *cluster* optimal diperoleh pada  $K=2$  dengan nilai *Silhouette Score* sebesar 0,33. Hasil clustering berhasil membagi pengguna Discord ke dalam dua kelompok utama, yaitu kelompok pengguna aktif dengan frekuensi penggunaan, interaksi sosial, dan tingkat kepuasan yang tinggi serta kelompok pengguna moderat dengan frekuensi penggunaan dan tingkat kepuasan yang relatif lebih rendah. Visualisasi menggunakan *Principal Component Analysis (PCA)* menunjukkan persebaran *cluster* yang cukup baik.

**Kata Kunci**— Discord, *K-Means Clustering*, Segmentasi Pengguna, *Silhouette Score*

### ABSTRACT

Discord is a digital communication platform used for various purposes such as community activities, entertainment, learning, and online communication. Differences in Discord usage patterns lead to varying user characteristics and satisfaction levels, making user segmentation necessary. This study aims to segment Discord users using the *K-Means Clustering* method based on usage patterns and user satisfaction levels. The research dataset was obtained through an online questionnaire distributed to 200 Discord users. The research process included data preprocessing, reliability testing using Cronbach Alpha, data transformation, normalization using StandardScaler, determining the optimal number of clusters using the *Elbow Method*, and model evaluation using the *Silhouette Score*. All data processing was conducted using Python-based Google Colab. The reliability test obtained a Cronbach Alpha value of 0.861, indicating that the research data had good consistency. The results showed that the optimal number of clusters was obtained at  $K=2$  with a *Silhouette Score* of 0.33. The clustering process successfully divided Discord users into two main groups, namely active users with high usage frequency, social interaction, and satisfaction levels, and moderate users with relatively lower usage frequency and satisfaction levels. Visualization using *Principal Component Analysis (PCA)* showed a fairly good distribution of the clusters.

**Keywords**— Discord, *K-Means Clustering*, User Segmentation, *Silhouette Score*

## I. PENDAHULUAN

Perkembangan teknologi komunikasi digital telah mendorong meningkatnya penggunaan platform komunikasi berbasis komunitas secara daring. Salah satu platform yang mengalami pertumbuhan pengguna cukup pesat adalah Discord [1]. Discord awalnya dikenal sebagai platform komunikasi untuk komunitas gim (*gaming*), namun saat ini penggunaannya telah berkembang ke berbagai bidang seperti pendidikan, hiburan, diskusi komunitas, hingga komunikasi sosial sehari-hari. Fitur seperti *voice channel*, *text chat*, serta pembentukan komunitas (*server*) menjadikan Discord sebagai media komunikasi yang fleksibel dan interaktif bagi penggunanya [2].

Meningkatnya jumlah pengguna Discord menyebabkan munculnya berbagai pola penggunaan yang berbeda pada setiap pengguna. Sebagian pengguna memanfaatkan Discord secara aktif untuk berkomunikasi melalui *voice channel* dan chat, sementara sebagian lainnya hanya menggunakan Discord untuk mengikuti komunitas tertentu atau sebagai media hiburan [3]. Perbedaan pola penggunaan tersebut juga dapat memengaruhi tingkat kepuasan pengguna terhadap platform Discord. Oleh karena itu, diperlukan suatu metode yang mampu mengelompokkan pengguna berdasarkan karakteristik penggunaan agar dapat diketahui pola perilaku pengguna Discord secara lebih jelas [4].

Salah satu metode yang dapat digunakan untuk melakukan segmentasi pengguna adalah *K-Means Clustering*. Metode *K-Means Clustering* merupakan salah satu algoritma unsupervised learning yang digunakan untuk mengelompokkan data berdasarkan tingkat kemiripan karakteristik antar data dengan membagi data ke dalam beberapa kelompok (*cluster*) dan didapatkan data dalam satu *cluster* memiliki karakteristik yang lebih mirip dibandingkan dengan data pada *cluster* lainnya [5].

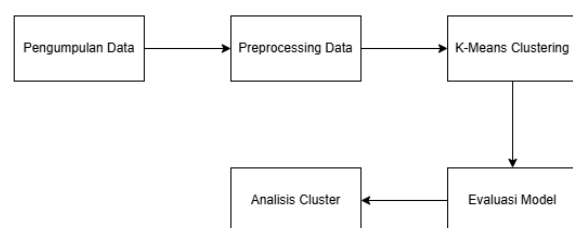
Penelitian terdahulu menunjukkan bahwa metode *K-Means Clustering* efektif digunakan dalam menganalisis tingkat kepuasan pengguna pada berbagai platform digital. Penelitian yang dilakukan oleh Syahrul Anwar menggunakan metode *K-Means Clustering* untuk menganalisis kepuasan pengguna aplikasi NyariGawe berdasarkan pendekatan *Technology Acceptance Model* (TAM) dan berhasil mengelompokkan pengguna ke dalam beberapa segmen kepuasan berdasarkan persepsi kemudahan dan kegunaan aplikasi [6]. Selain itu, penelitian oleh Dicha Mutia Dhani dkk. menerapkan metode *K-Means Clustering* untuk menentukan tingkat kepuasan mahasiswa terhadap fasilitas sarana dan prasarana kampus di STMIK Kaputama Binjai dan menghasilkan

pengelompokan mahasiswa berdasarkan tingkat kepuasan terhadap fasilitas kampus [7]. Penelitian lain yang dilakukan oleh Kayyisu Willyani juga menggunakan metode *Clustering* untuk menganalisis ulasan pengguna aplikasi Gojek berdasarkan versi aplikasi dan tingkat kepuasan pengguna, sehingga menghasilkan segmentasi pengguna berdasarkan pola kepuasan yang berbeda pada setiap versi aplikasi [8].

Penelitian terdahulu yang menggunakan metode *K-Means Clustering* umumnya berfokus pada analisis kepuasan pengguna aplikasi tertentu seperti aplikasi pencarian kerja, fasilitas kampus, maupun layanan transportasi digital [6][7][8]. Namun, penelitian tersebut belum secara khusus mengkaji segmentasi pengguna berdasarkan kombinasi pola penggunaan platform dan tingkat kepuasan pengguna. Selain itu, sebagian besar penelitian terdahulu hanya berfokus pada aspek kepuasan tanpa mempertimbangkan karakteristik aktivitas pengguna dalam memanfaatkan fitur-fitur komunikasi digital. Oleh karena itu, penelitian ini dilakukan dengan menerapkan algoritma *K-Means Clustering* yang dievaluasi menggunakan *Silhouette Score* dan divisualisasikan menggunakan PCA pada pengguna Discord berdasarkan pola penggunaan dan tingkat kepuasan pengguna dengan demikian dapat menghasilkan segmentasi yang lebih representatif terhadap perilaku pengguna platform Discord.

## II. METODE PENELITIAN

Penelitian ini menerapkan pendekatan data mining untuk melakukan segmentasi pengguna Discord berdasarkan pola penggunaan dan tingkat kepuasan pengguna. Seluruh proses penelitian dilakukan menggunakan *tools* Google Colab berbasis Python, mulai dari pengumpulan data, *preprocessing* data, *ordinal encoding*, uji reliabilitas, normalisasi data, hingga proses *Clustering* menggunakan algoritma *K-Means Clustering*. Evaluasi model dilakukan menggunakan *Silhouette Score* untuk mengetahui kualitas *cluster* yang dihasilkan. Tahapan penelitian yang dilakukan ditunjukkan pada Gambar 1.



Gbr 1. Tahapan Alur Penelitian

### A. Pengumpulan Data

Tahap pengumpulan data dilakukan melalui penyebaran kuesioner secara daring menggunakan Google Form kepada pengguna Discord. Kuesioner disebarakan kepada responden yang aktif menggunakan platform Discord untuk berbagai kebutuhan.

Teknik pengambilan sampel yang digunakan pada penelitian ini adalah purposive sampling, yaitu responden dipilih berdasarkan kriteria tertentu, yaitu pengguna yang pernah atau sedang aktif menggunakan platform Discord. Teknik ini dipilih karena penelitian berfokus pada pengguna yang memiliki pengalaman dalam menggunakan Discord

Data yang diperoleh pada penelitian ini berjumlah 200 responden. Data hasil kuesioner kemudian disimpan dalam format CSV dan diolah menggunakan Google Colab berbasis Python. Variabel yang digunakan pada penelitian berkaitan dengan pola penggunaan Discord dan tingkat kepuasan pengguna terhadap platform Discord.

### B. Preprocessing Data

Tahap *preprocessing data* bertujuan untuk meningkatkan kualitas data agar hasil *cluster* yang dihasilkan lebih optimal [9]. Pada penelitian ini, *preprocessing data* dilakukan melalui beberapa tahapan, yaitu *Data Selection*, *Data Cleaning*, *data transformation* dan normalisasi data.

#### 1. Data Selection

Tahap *Data Selection* dilakukan untuk memilih atribut yang relevan dengan tujuan penelitian. *Data Selection* bertujuan untuk mengurangi atribut yang tidak memiliki keterkaitan terhadap proses *Clustering* agar data yang digunakan dapat lebih fokus dalam menggambarkan pola penggunaan dan tingkat kepuasan pengguna Discord [10].

Pada tahap ini, atribut seperti nama pengguna dihapus karena tidak memiliki pengaruh terhadap proses segmentasi pengguna. Selain itu, penelitian ini lebih berfokus pada perilaku penggunaan pengguna Discord dibandingkan karakteristik identitas atau demografis pengguna. Oleh karena itu, atribut utama yang digunakan dalam proses *Clustering* meliputi durasi penggunaan Discord, frekuensi penggunaan Discord, frekuensi penggunaan fitur chat, frekuensi penggunaan voice channel, tingkat interaksi sosial, dan tingkat kepuasan pengguna.

#### 2. Data Cleaning

Tahap berikutnya adalah *Data Cleaning* yang dilakukan untuk membersihkan data dari kesalahan penulisan (*typo*), perbedaan huruf kapital, spasi berlebih, serta menangani *missing value* pada data kuesioner. Proses ini dilakukan agar data memiliki

format yang konsisten sehingga dapat diproses pada tahap selanjutnya .

#### 3. Data Transformation

Pada tahapan ini dilakukan transformasi data untuk mempersiapkan data agar lebih optimal dan siap digunakan pada proses pemodelan menggunakan algoritma *K-Means Clustering* [11]. Pada tahap *Clustering*. Pada penelitian ini, proses *data transformation* dilakukan dengan *ordinal encoding*.

Tahap *ordinal encoding* dilakukan untuk mengubah data kategorikal bertingkat menjadi data numerik [12]. Pada penelitian ini, *ordinal encoding* diterapkan pada atribut yang memiliki tingkatan nilai seperti frekuensi penggunaan Discord, frekuensi penggunaan fitur chat, frekuensi penggunaan voice channel, tingkat interaksi sosial, dan tingkat kepuasan pengguna.

#### 4. Pengujian reliabilitas

Pengujian reliabilitas data dilakukan menggunakan metode *Cronbach Alpha* untuk mengetahui tingkat konsistensi data kuesioner yang digunakan pada penelitian. Metode *Cronbach Alpha* digunakan karena mampu mengukur tingkat konsistensi antar atribut pada data kuesioner. Instrumen penelitian dinyatakan reliabel apabila nilai *Cronbach Alpha* lebih besar dari 0,70 [11].

#### 5. Normalisasi Data

Tahap selanjutnya adalah data dinormalisasi menggunakan *StandardScaler* dengan pendekatan standardisasi data. Proses standardisasi dilakukan menggunakan rumus berikut [12]:

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

Dimana:

- $z$  = nilai hasil normalisasi
- $x$  = nilai data asli
- $\mu$  = rata-rata data
- $\sigma$  = standar deviasi

Setiap atribut diubah ke dalam skala standar berdasarkan nilai rata-rata dan standar deviasi data. Tujuan normalisasi adalah untuk menghindari dominasi atribut tertentu yang memiliki rentang nilai lebih besar [12].

### C. Penerapan K-Means Clustering

Pada tahap ini, algoritma *K-Means Clustering* diterapkan pada data yang telah melalui proses *preprocessing* dan normalisasi. Sebelum proses *Clustering* dilakukan, penentuan jumlah *cluster* terlebih dahulu dilakukan menggunakan metode *Elbow* untuk memperoleh jumlah *cluster* yang optimal. Metode *Elbow* digunakan dengan melihat nilai inerti pada beberapa jumlah *cluster* yang memperoleh titik penurunan paling optimal sebagai jumlah *cluster* terbaik yang digunakan pada proses *Clustering* [13].

Setelah jumlah *cluster* ditentukan, proses *Clustering* dilakukan dengan menghitung jarak antara data dan centroid menggunakan *Euclidean Distance*. Data kemudian dikelompokkan ke dalam *cluster* dengan jarak centroid terdekat [14].

Rumus Euclidean Distance yang digunakan adalah sebagai berikut [15]:

$$d(x, y) = \sqrt{\sum_{j=1}^n (x_j - y_j)^2}$$

Dimana:

- $d$  = jarak antar data
- $x$  = pusat *cluster* (centroid)
- $y$  = data
- $x_j$  = nilai data
- $y_j$  = nilai centroid
- $n$  = jumlah atribut

Setelah proses perhitungan jarak dilakukan, setiap data akan dikelompokkan berdasarkan centroid terdekat. Selanjutnya dilakukan pembaruan nilai centroid menggunakan rata-rata data pada masing-masing *cluster* hingga proses *Clustering* mencapai kondisi stabil.

Rumus pembaruan centroid adalah sebagai berikut [16]:

$$c_k = \left(\frac{1}{n_k}\right) * \sum d_1$$

Dimana:

- $ck$  = centroid *cluster* ke- $k$
- $nk$  = jumlah data pada *cluster* ke- $k$
- $di$  = data pada *cluster* ke- $k$

#### D. Evaluasi Model

Tahap selanjutnya adalah evaluasi model *Clustering* yang dilakukan untuk mengetahui kualitas *cluster* yang dihasilkan oleh algoritma *K-Means Clustering*. Pada penelitian ini, evaluasi model dilakukan menggunakan metode *Silhouette Score*. Metode ini digunakan untuk mengukur tingkat kemiripan data dalam satu *cluster* dibandingkan dengan *cluster* lainnya. Nilai *Silhouette Score* berada pada rentang -1 hingga 1, di mana semakin tinggi nilai yang diperoleh maka kualitas *cluster* yang dihasilkan semakin baik [17].

Perhitungan *Silhouette Score* dilakukan berdasarkan rata-rata jarak antar data dalam *cluster* yang sama dan jarak data terhadap *cluster* terdekat lainnya. Rumus *Silhouette Score* yang digunakan adalah sebagai berikut [18]:

$$s_i = \frac{(b_i - a_i)}{\max(a_i, b_i)}$$

Dimana:

- $s_i$  = nilai silhouette data ke- $i$
- $a_i$  = rata-rata jarak data ke- $i$  dengan data lain dalam *cluster* yang sama

- $b_i$  = rata-rata jarak data ke- $i$  dengan *cluster* terdekat lainnya

Hasil evaluasi menggunakan *Silhouette Score* kemudian digunakan untuk membantu proses analisis *cluster*. Analisis dilakukan secara deskriptif berdasarkan nilai rata-rata setiap atribut pada masing-masing *cluster* serta didukung melalui visualisasi *cluster* menggunakan PCA dan grafik perbandingan *cluster*.

### III. HASIL DAN PEMBAHASAN

#### A. Dataset

*Dataset* yang digunakan pada penelitian ini terdiri dari 200 data responden pengguna Discord yang diperoleh melalui kuesioner. *Dataset* memiliki beberapa atribut yang berkaitan dengan pola penggunaan dan tingkat kepuasan pengguna Discord. Adapun atribut yang digunakan pada *dataset* penelitian dapat dilihat pada Tabel 1.

Tabel 1. Atribut pada *dataset*

no	atribut
1	usia
2	perangkat_penggunaan
3	durasi_penggunaan
4	frekuensi_penggunaan
5	frekuensi_chat
6	frekuensi_voice
7	tujuan_penggunaan
8	interaksi_teman
9	Kepuasan_penggunaan

Pada penelitian ini, tidak seluruh atribut digunakan pada proses *Clustering*. Penelitian berfokus pada atribut yang paling merepresentasikan pola penggunaan dan tingkat kepuasan pengguna Discord. Oleh karena itu, atribut utama yang digunakan pada proses *Clustering* meliputi durasi penggunaan Discord, frekuensi penggunaan Discord, frekuensi penggunaan fitur chat, frekuensi penggunaan voice channel, tingkat interaksi sosial, dan tingkat kepuasan pengguna.

#### B. Preprocessing Data

Tahap *preprocessing data* dilakukan proses pembersihan dan transformasi data untuk mempersiapkan *dataset* sebelum dilakukan proses *Clustering* menggunakan algoritma *K-Means Clustering*. Tahap *preprocessing* diawali dengan proses pembersihan data yang meliputi perbaikan kesalahan penulisan (*typo*), penyamaan penggunaan huruf kapital, penghapusan spasi berlebih, serta pengecekan *missing value* pada *dataset* hasil kuesioner.

durasi_penggunaan	0
frekuensi_penggunaan	0
frekuensi_chat	0
frekuensi_voice	0
tujuan_penggunaan	0
interaksi_teman	0
kepuasan	0

Gbr 2. Hasil pengecekan missing value

Pada gambar 2, merupakan hasil pengecekan *missing value* yang menunjukkan bahwa setelah dilakukan pembersihan data, *dataset* tidak memiliki data kosong.

Selanjutnya dilakukan proses *ordinal encoding* untuk mengubah data kategorikal bertingkat menjadi data numerik yang diterapkan pada atribut yang memiliki tingkatan nilai seperti frekuensi penggunaan Discord, frekuensi penggunaan fitur chat, frekuensi penggunaan voice channel, tingkat interaksi sosial, dan tingkat kepuasan pengguna.

Hasil transformasi menunjukkan bahwa seluruh atribut berhasil diubah ke dalam bentuk numerik berdasarkan urutan tingkatannya sehingga data dapat digunakan pada proses *Clustering* menggunakan algoritma *K-Means Clustering*. Contoh mapping nilai *ordinal encoding* dapat dilihat pada Tabel 2

Tabel II. Contoh mapping nilai *ordinal encoding*

kategori	nilai
tidak pernah	0
sangat jarang	1
jarang	2
sering	3
sangat sering	4

durasi_penggunaan	frekuensi_penggunaan	frekuensi_chat	frekuensi_voice	interaksi_teman	kepuasan
1	2	3	3	3	1
2	3	3	3	2	1
3	3	2	3	2	1
1	1	1	2	1	1
2	3	3	3	2	2

Gbr 3. Hasil *ordinal encoding* dataset

Pada gambar 3 hasil *ordinal encoding*, seluruh atribut kategorikal bertingkat berhasil diubah menjadi data numerik sesuai urutan tingkatannya. Proses transformasi ini menghasilkan *dataset* yang telah siap digunakan pada tahap selanjutnya.

Tahap selanjutnya, penelitian ini terlebih dahulu melakukan pengujian reliabilitas data menggunakan metode *Cronbach Alpha* untuk mengetahui tingkat konsistensi data kuesioner yang digunakan pada penelitian.

Tabel III. Hasil Pengujian Reliabilitas Menggunakan *Cronbach Alpha*

pengujian	nilai
Cronbach Alpha	0.861

Berdasarkan hasil pengujian reliabilitas menggunakan *Cronbach Alpha* diperoleh nilai sebesar 0,861. Nilai tersebut menunjukkan bahwa instrumen kuesioner pada penelitian ini memiliki tingkat reliabilitas yang baik karena melebihi batas minimum reliabilitas sebesar 0,70. Dengan demikian, data penelitian dinilai konsisten dan layak digunakan pada proses *Clustering*.

Selanjutnya dilakukan proses normalisasi data menggunakan metode *StandardScaler* untuk menyamakan skala antar atribut sebelum proses *Clustering* dilakukan.

Tabel IV. Contoh normalisasi data

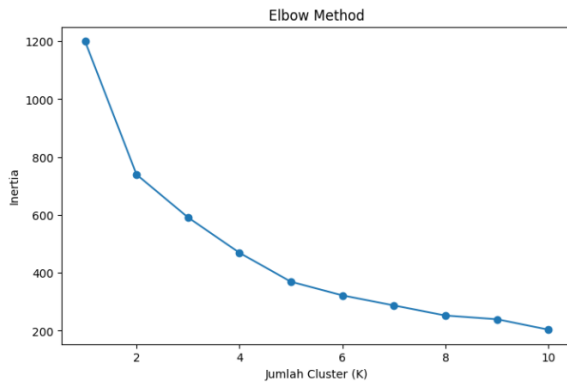
Durasi	Frek peng-unaan	chat	voice	Interaksi	Ke-puasan
-1.41	-0.23	0.21	0.14	1.10	-1.19
-0.10	0.72	0.21	0.14	0.05	-1.19
1.20	0.72	-0.90	0.14	0.05	-1.19
-1.41	-1.20	-2.02	-0.80	-0.99	-1.19
-0.10	0.72	0.21	0.14	0.05	-0.062

Hasil normalisasi menunjukkan bahwa seluruh atribut berhasil diubah ke dalam skala standar sehingga tidak terdapat atribut yang mendominasi proses *Clustering*. Data hasil normalisasi memiliki rentang nilai yang lebih seimbang dan siap digunakan pada proses *Clustering* menggunakan algoritma *K-Means Clustering*. Contoh hasil normalisasi data dapat dilihat pada Tabel 4.

### C. Evaluasi Model *K-Means Clustering*

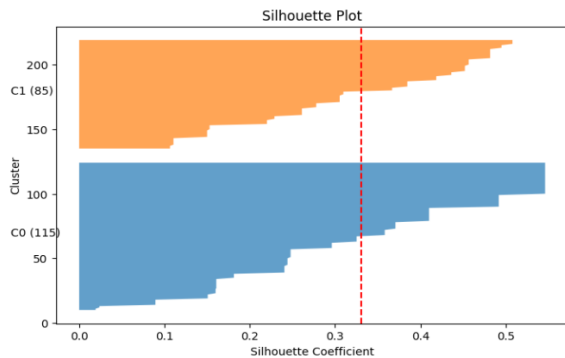
Sebelum proses *Clustering* dilakukan, terlebih dahulu dilakukan penentuan jumlah *cluster* optimal menggunakan metode *Elbow*. Metode ini digunakan untuk melihat penurunan nilai inerti pada beberapa jumlah *cluster* dengan demikian dapat diketahui jumlah *cluster* yang paling optimal untuk digunakan pada proses *Clustering*.

Hasil pengujian menggunakan *Elbow Method* menunjukkan bahwa titik penurunan inerti paling optimal berada pada nilai  $K = 2$ . Berdasarkan hasil tersebut, penelitian ini menggunakan dua *cluster* pada proses *Clustering* pengguna Discord. Hasil visualisasi *Elbow Method* dapat dilihat pada Gambar 4.



Gbr 4. Hasil elbow method

Untuk mengevaluasi kualitas *cluster* yang dihasilkan, penelitian ini menggunakan metode *Silhouette Score*. Metode ini digunakan untuk mengukur tingkat kemiripan data dalam satu *cluster* dibandingkan dengan *cluster* lainnya.



Gbr 5. Hasil silhouette plot

Pada gambar 5. Menunjukkan hasil evaluasi menggunakan *Silhouette Score* pada penelitian ini memperoleh nilai sebesar 0,33.

Nilai *Silhouette Score* sebesar 0,33 menunjukkan bahwa *cluster* yang terbentuk memiliki tingkat pemisahan yang cukup, meskipun masih terdapat beberapa data yang memiliki kedekatan antar *cluster*. Hal ini dapat disebabkan oleh karakteristik perilaku pengguna Discord yang cenderung heterogen dan memiliki pola penggunaan yang saling beririsan antar pengguna.

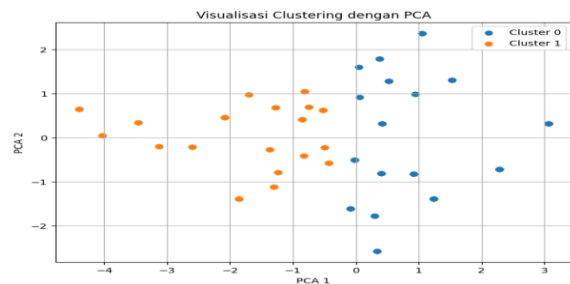
Hasil clustering menghasilkan dua cluster, yaitu Cluster 0 (Pengguna Aktif dan Puas) sebanyak 115 responden dan Cluster 1 (Pengguna Moderat) sebanyak 85 responden. Berdasarkan karakteristik yang terbentuk, Cluster 0 diberi nama "Pengguna Aktif dan Puas" karena memiliki nilai rata-rata yang tinggi pada frekuensi penggunaan Discord, penggunaan fitur chat dan voice channel, tingkat interaksi sosial, serta tingkat kepuasan pengguna. Sementara itu, Cluster 1 diberi nama "Pengguna Moderat" karena memiliki tingkat

aktivitas penggunaan dan kepuasan yang relatif lebih rendah dibandingkan Cluster 0.

Distribusi tersebut menunjukkan bahwa pengguna Discord tidak terpusat pada satu kelompok tertentu, melainkan tersebar ke dalam beberapa kelompok dengan pola penggunaan dan tingkat kepuasan yang berbeda. Selain itu, sebagian besar data pada masing-masing *cluster* memiliki nilai silhouette positif sehingga menunjukkan bahwa algoritma *K-Means Clustering* mampu melakukan pemisahan *cluster* dengan cukup baik pada penelitian ini.

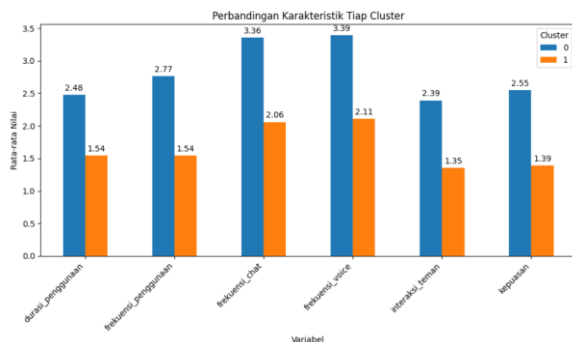
#### D. Analisis Hasil Clustering

Untuk melihat pola persebaran pengguna Discord pada masing-masing *cluster*, penelitian ini menggunakan *Principal Component Analysis (PCA)* sebagai alat bantu visualisasi *cluster*. Metode PCA digunakan untuk membantu menunjukkan apakah hasil pembentukan *cluster* menggunakan algoritma *K-Means Clustering* telah terpisah dengan cukup baik atau masih mengalami tumpang tindih antar *cluster*.



Gbr 6. Hasil PCA

Pada Gambar 6 tersebut terlihat bahwa persebaran data pada masing-masing *cluster* menunjukkan adanya perbedaan karakteristik antar kelompok Pengguna. Aktif dan Puas menunjukkan kelompok pengguna dengan pola aktivitas penggunaan Discord yang berbeda dibandingkan *Pengguna Moderat*. Perbedaan tersebut terlihat dari posisi persebaran data yang membentuk area *cluster* masing-masing. Meskipun demikian, masih terdapat beberapa titik data yang saling berdekatan antar *cluster* yang mengindikasikan adanya kemiripan karakteristik pada sebagian pengguna.



Gbr 7. Perbandingan rata-rata tiap cluster

Berdasarkan hasil perbandingan karakteristik pada masing-masing *cluster* pada gambar 7, terlihat bahwa setiap *cluster* memiliki karakteristik aktivitas Discord dan tingkat kepuasan yang berbeda. Pengguna Aktif dan Puas memiliki nilai rata-rata yang lebih tinggi pada hampir seluruh atribut dibandingkan *Pengguna Moderat*. Pada atribut frekuensi penggunaan Discord, Pengguna Aktif dan Puas memperoleh nilai rata-rata sebesar 2,77 sedangkan Pengguna Moderat sebesar 1,54. Selain itu, pada atribut frekuensi penggunaan fitur chat dan voice channel, Pengguna Aktif dan Puas masing-masing memiliki nilai rata-rata sebesar 3,36 dan 3,39, lebih tinggi dibandingkan Pengguna Moderat yang memperoleh nilai rata-rata sebesar 2,06 dan 2,11.

Tingginya aktivitas penggunaan pada Pengguna Aktif dan Puas menunjukkan bahwa pengguna yang aktif memanfaatkan fitur komunikasi seperti voice channel dan chat cenderung memiliki tingkat kepuasan yang lebih tinggi terhadap platform Discord. Hasil ini sejalan dengan penelitian Des Christian dan Sakaria [4] yang menyatakan bahwa intensitas interaksi dan kualitas layanan komunikasi digital memengaruhi kepuasan pengguna platform Discord.

Sementara itu, Pengguna Moderat menunjukkan nilai rata-rata yang lebih rendah pada sebagian besar atribut. Pengguna pada *cluster* ini cenderung memiliki frekuensi penggunaan Discord yang lebih rendah serta tingkat aktivitas komunikasi dan interaksi sosial yang tidak terlalu tinggi. Selain itu, tingkat kepuasan pengguna pada Pengguna Moderat juga berada di bawah Pengguna Aktif dan Puas. Hal tersebut menunjukkan bahwa pengguna pada Pengguna Moderat merupakan kelompok pengguna dengan tingkat aktivitas penggunaan Discord yang relatif lebih rendah dibandingkan pengguna pada Pengguna Aktif dan Puas.

Perbedaan karakteristik antar *cluster* menunjukkan bahwa atribut frekuensi penggunaan Discord, frekuensi penggunaan fitur chat, frekuensi penggunaan voice channel, dan tingkat kepuasan pengguna menjadi indikator utama dalam proses segmentasi pengguna Discord. Hasil tersebut menunjukkan bahwa algoritma

*K-Means Clustering* mampu mengelompokkan pengguna berdasarkan karakteristik aktivitas dan tingkat kepuasan pengguna dengan cukup baik.

Hasil segmentasi yang diperoleh menunjukkan bahwa pengguna Discord memiliki karakteristik penggunaan yang berbeda berdasarkan tingkat aktivitas komunikasi dan kepuasan pengguna. Informasi hasil segmentasi tersebut dapat dimanfaatkan sebagai dasar dalam memahami perilaku pengguna platform komunikasi digital serta mendukung pengembangan strategi peningkatan layanan dan pengalaman pengguna pada platform Discord.

## KESIMPULAN

Metode *K-Means Clustering* berhasil diterapkan untuk melakukan segmentasi pengguna Discord berdasarkan pola penggunaan dan tingkat kepuasan pengguna. Hasil penelitian menunjukkan bahwa jumlah *cluster* optimal diperoleh pada  $K=2$  dengan nilai *Silhouette Score* sebesar 0,33 yang menunjukkan kualitas pemisahan *cluster* yang cukup baik.

Hasil segmentasi menghasilkan dua kelompok pengguna utama, yaitu kelompok Pengguna Aktif dan Puas (*Cluster 0*) yang memiliki frekuensi penggunaan, interaksi sosial, serta tingkat kepuasan yang tinggi, dan kelompok Pengguna Moderat yang memiliki tingkat aktivitas dan kepuasan yang relatif lebih rendah. Perbedaan karakteristik tersebut menunjukkan bahwa intensitas penggunaan fitur komunikasi seperti chat dan voice channel memiliki hubungan dengan tingkat kepuasan pengguna terhadap platform Discord.

Hasil penelitian ini dapat dimanfaatkan sebagai dasar dalam memahami perilaku pengguna Discord serta mendukung pengembangan strategi peningkatan layanan dan pengalaman pengguna pada platform komunikasi digital.

## REFERENSI

- [1] M. S. Rizal and N. M. Aesthetika, "Efektivitas Penggunaan Aplikasi Discord Dalam Meningkatkan Komunikasi Interpersonal di Kalangan Pecinta Film," *Medium*, vol. 10, no. 1, pp. 19–27, 2022, [Online]. Available: [https://doi.org/10.25299/medium.2022.vol10\(1\).8882](https://doi.org/10.25299/medium.2022.vol10(1).8882)
- [2] A. K. Nayla, E. W. B. Prasetyo, and B. H. Novaludin, "Pola Komunikasi Masyarakat dalam Sosial Media Discord," *Semin. Nas. Univ. Negeri Surabaya*, pp. 1085–098, 2024, [Online]. Available: <https://proceeding.unesa.ac.id/index.php/sniis/article/view/3852>
- [3] M. Z. Al Hilal, D. Sari, S. D. K. A. Harmadi, S. R. Utami, and R. Novianti, "Phenomenological Study Of Interpersonal Communication Among Online Game Players In The Indonesian Pubg Mobile Discord Community," vol. 4, no. 4, pp. 57–61, 2025, [Online]. Available: <https://doi.org/10.56127/jukim.v4i04.2168>
- [4] T. Des Christian and S. Sakaria, "Analisis Pengaruh e-Service Quality, Sense of Belonging dan User Satisfaction terhadap Intention to Use Aplikasi Discord," *TeknoIS J. Ilm.*

- Tekno. Inf. dan Sains*, vol. 15, no. 1, pp. 51–61, 2025, [Online]. Available: <https://doi.org/10.36350/jbs.v15i1.287>
- [5] L. R. Iyohu, Ismail Djakaria, and La Ode Nashar, “Perbandingan Metode K-Means Clustering dengan Self-Organizing Maps (SOM) untuk Pengelompokan Provinsi di Indonesia Berdasarkan Data Potensi Desa,” *J. Stat. dan Apl.*, vol. 7, no. 2, pp. 195–206, 2023, doi: 10.21009/jsa.07208.
- [6] S. Anwar, “Analisis Kepuasan Pengguna Aplikasi Nyari Gawe Menggunakan Model Technology Acceptance Model (TAM) dan K-Means Clustering,” *J. Sos. dan Tekno.*, vol. 6, no. 4, pp. 1514–1526, 2026.
- [7] Dicha Mutia Dhani, Relita Buatun, and I Gusti Prahmana, “Penerapan Metode K – Means Clustering untuk Menentukan Kepuasan Mahasiswa terhadap Fasilitas Sarana dan Prasarana Kampus di STMIK Kaputama Binjai,” *Bridg. J. Publ. Sist. Inf. dan Telekomun.*, vol. 2, no. 3, pp. 229–243, 2024, doi: 10.62951/bridge.v2i3.170.
- [8] K. Willyani, “Analisis Ulasan Pengguna Aplikasi Gojek Berdasarkan Versi Aplikasi Menggunakan Metode Clustering untuk Evaluasi Kepuasan,” vol. 7, no. 1, pp. 1–6, 2025.
- [9] Baromim Triwijaya, Setyoningsih Wibowo, and Nur Latifah Dwi Mutiara Sari, “Performance Comparison of K-Means Algorithm and BIRCH Algorithm in Clustering Earthquake Data in Indonesia with Web-Based Map Visualization,” *J. Tekno. Dan Open Source*, vol. 8, no. 1, pp. 278–287, 2025, doi: 10.36378/jtos.v8i1.4400.
- [10] I. P. Muhammad Fikri Fadhillah, Aldo Lovely Arief Suyoso, “Customer Segmentation with Clustering Algorithm Based on Recency, Frequency, and Monetary (RFM) Attributes,” *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 5, no. 1, pp. 48–56, 2025.
- [11] E. W. Muhammad Harlanu, Agus Suryanto, Syahrul Ramadhan, “Construct validity of the instrument of digital skill literacy Muhammad,” *Cakrawala Pendidik. J. Ilm. Pendidik.*, vol. 42, no. 3, pp. 781–790, 2023.
- [12] I. U. Muhammad Hidayatullah, Afif Budi Andy B, “Analisis Komponen Utama dan Klasterisasi K-Means pada Data Pengangguran Terbuka Kabupaten / Kota di Provinsi Sulawesi Selatan,” *Techno Com*, vol. 25, no. 2, pp. 355–366, 2026.
- [13] N. A. Maori and E. Evanita, “Metode Elbow dalam Optimasi Jumlah Cluster pada K-Means Clustering,” *Simetris J. Tek. Mesin, Elektro dan Ilmu Komput.*, vol. 14, no. 2, pp. 277–288, 2023, doi: 10.24176/simet.v14i2.9630.
- [14] Matius Ivan Bimasena, I Gede Aris Gunadi, and I Made Agus Wirawan, “Studi Perbandingan Algoritma Euclidean, Manhattan Dan Chebysev Distance Untuk Optimasi Metode K-Means Clustering Pada Pengelompokan Data Penerimaan Mahasiswa Baru,” *J. Pendidik. Tekno. dan Kejuru.*, vol. 22, no. 2, pp. 116–127, 2025.
- [15] W. Wirdawati, S. Yulihartati, and A. Ramadhanu, “Implementasi Euclidean Distance dan Segmentasi K-Means Clustering Pada Identifikasi Citra Jeruk Nipis,” *Indones. J. Comput. Sci.*, vol. 3, no. 2, pp. 72–79, 2024, doi: 10.31294/ijcs.v3i2.5600.
- [16] A. Primandana, S. Adinugroho, and C. Dewi, “Optimasi Penentuan Centroid pada Algoritma K-Means Menggunakan Algoritma Pillar (Studi Kasus: Penyandang Masalah Kesejahteraan Sosial di Provinsi ...),” *Tekno. Inf. dan Ilmu ...*, vol. 3, no. 11, pp. 10678–10683, 2020, [Online]. Available: <http://j-ptiik.ub.ac.id/index.php/j-ptiik/article/download/6748/3264>
- [17] H. Mulyani, R. A. Setiawan, and H. Fathi, “Optimization of K Value in Clustering Using Silhouette Score (Case Study: Mall Customers Data),” *J. Inf. Technol. Its Util.*, vol. 6, no. 2, pp. 45–50, 2023, doi: 10.56873/jitu.6.2.5243.
- [18] M. D. Ananda, K. N. Malik, A. F. N. Masruriyah, and M. Mardiah, “Studi Komparatif Algoritma K-Means dan K-Medoids untuk Segmentasi Informasi Kesehatan,” *Comput. Sci.*, vol. 5, no. 2, pp. 103–112, 2025, doi: 10.31294/coscience.v5i2.9207.