
METODE SINGLE LINKAGE PADA AGGLOMERATIVE HIERARCHICAL CLUSTERING DALAM PENENTUAN TINGKAT KEPUASAN MAHASISWA TERHADAP LAYANAN AKADEMIK

Nadia Annisa Maori

Fakultas Sains dan Teknologi, Program Studi Teknik Informatika
Universitas Islam Nahdlatul Ulama Jepara
Email: nadia@unisnu.ac.id

Harminto Mulyo

Fakultas Sains dan Teknologi, Program Studi Teknik Informatika
Universitas Islam Nahdlatul Ulama Jepara
Email: minto@unisnu.ac.id

ABSTRAK

Dalam era pendidikan modern, penting bagi institusi akademik untuk memahami dan meningkatkan kepuasan mahasiswa terhadap layanan akademik yang diberikan. Tingkat kepuasan ini berperan penting dalam menilai kualitas pendidikan, pengalaman belajar, serta reputasi dan daya saing institusi. Dengan kemajuan teknologi dan peningkatan pengumpulan data, analisis data menjadi krusial untuk memahami persepsi dan kebutuhan mahasiswa. Penelitian ini mengadopsi teknik *Agglomerative Hierarchical Clustering* (AHC) dengan metode *Single Linkage* untuk mengelompokkan data survei kepuasan mahasiswa terhadap layanan akademik di Program Studi Teknik Informatika UNISNU Jepara. Metode ini dipilih karena tidak memerlukan penentuan jumlah klaster sebelumnya dan cocok untuk data dengan struktur yang tidak teratur.

Hasil penelitian menunjukkan bahwa AHC dengan *Single Linkage* efektif dalam mengidentifikasi dua kelompok mahasiswa berdasarkan tingkat kepuasan mereka, yaitu **puas** dan **tidak puas**. Evaluasi menggunakan *Silhouette Coefficient* menunjukkan nilai tertinggi 0.80 untuk dua klaster mengindikasikan bahwa pengelompokan ini cukup baik. Visualisasi dendrogram memberikan wawasan tambahan tentang struktur klaster dan hubungan antar data. Penelitian ini memberikan kontribusi penting untuk pemahaman tentang kepuasan mahasiswa dan dasar untuk pengembangan strategi peningkatan kualitas layanan akademik di masa depan. Metode AHC dengan pendekatan *Single Linkage* terbukti efisien dalam mengelompokkan data berdasarkan jarak terdekat antar objek dalam klaster, meskipun sensitif terhadap *outlier* dan efek *chaining*.

Kata kunci: *single linkage, agglomerative hierarchical clustering, distance, silhouette coefficient*

ABSTRACT

In the modern educational era, it is important for academic institutions to understand and increase student satisfaction with the academic services provided. This level of satisfaction plays an important role in assessing the quality of education, learning experience, as well as the reputation and competitiveness of the institution. With advances in technology and increased data collection, data analysis has become crucial to understanding student perceptions and needs. This research uses the Agglomerative Hierarchical Clustering (AHC) technique with the Single Linkage method to group survey data on student satisfaction with academic services in the UNISNU Jepara Informatics Engineering Study Program. This method was chosen because it does not require prior determination of the number of clusters and is suitable for data with an irregular structure.

The research results show that AHC with Single Linkage is effective in identifying two groups of students based on their level of satisfaction, namely satisfied and dissatisfied. Evaluation using the Silhouette Coefficient shows the highest value of 0.80 for the two clusters, indicating that this grouping is quite good. Dendrogram visualization provides additional insight into cluster structure and relationships between data. This research provides an important contribution to the

understanding of student satisfaction and the basis for developing strategies for improving the quality of academic services in the future. The AHC method with the Single Linkage approach has proven to be efficient in grouping data based on the closest distance between objects in the cluster, although it is sensitive to outliers and chaining effects.

Keywords: *single linkage, agglomerative hierarchical clustering, distance, silhouette coefficient*

1. PENDAHULUAN

Dalam era pendidikan yang semakin berkembang, penting bagi institusi akademik untuk memahami dan meningkatkan tingkat kepuasan mahasiswa terhadap layanan akademik yang ditawarkan. Penilaian mahasiswa memainkan peran kunci dalam menilai kualitas pendidikan [1] dan pengalaman belajar mereka, serta dapat mempengaruhi reputasi dan daya saing institusi pendidikan. Seiring kemajuan teknologi dan pengumpulan data yang semakin besar, pendekatan analisis data menjadi lebih penting dalam memberikan wawasan yang mendalam tentang persepsi dan kebutuhan mahasiswa terhadap layanan akademik. Teknik *clustering* adalah sebuah pendekatan dalam analisis data yang digunakan untuk mengatasi masalah pengelompokan data. Salah satu pendekatan analisis data yang digunakan untuk mengelompokkan data dan mengidentifikasi pola adalah *Agglomerative Hierarchical Clustering* (AHC) [2]. Algoritma ini merupakan salah satu teknik *clustering* yang cukup populer dikalangan praktisi analisis data. Tersedia metode yang dapat diterapkan pada AHC, diantaranya, Metode *Single Linkage*. Metode *Single Linkage* merupakan salah satu pendekatan dalam penggabungan kluster yang digunakan dalam AHC, di mana jarak antara dua kluster diukur sebagai jarak terdekat antara titik-titik di kedua kluster tersebut [3].

Agglomerative Hierarchical Clustering (AHC) dilengkapi dengan beberapa kelebihan yang membuatnya menjadi salah satu metode *clustering* yang populer dan sering digunakan dalam pengolahan data. Pertama, AHC tidak memerlukan jumlah kluster yang telah ditentukan sebelumnya. Ini memungkinkan algoritma untuk mengelompokkan data tanpa asumsi awal tentang jumlah kluster yang optimal, sehingga cocok untuk situasi di mana kita tidak memiliki informasi sebelumnya tentang jumlah kluster. Berbeda dengan Algoritms *K-Means*, dimana penentuan jumlah kluster k harus ditentukan sebelumnya, sehingga membutuhkan metode optimalisasi dalam penentuan nilai k [4]. AHC menggunakan pendekatan hierarki, dimana setiap data dimulai sebagai kluster tunggal dan kemudian secara berangsur-angsur digabungkan bersama berdasarkan kedekatan mereka, hingga semua data menjadi satu kluster atau sampai batas tertentu. Sementara, *K-means* adalah pendekatan non-hierarkis yang memisahkan data ke dalam k kluster yang telah ditentukan sebelumnya. Pusat kluster diinisialisasi secara acak, dan setiap data ditempatkan ke dalam kluster terdekat, kemudian pusat kluster diperbarui secara iteratif hingga konvergen.

Penelitian terdahulu yang dilakukan oleh Komang Ariasi dalam penelitiannya tentang pengelompokan kinerja akademik mahasiswa menggunakan Algoritma *K-Means* menyatakan bahwa hasil akhir dari *K-Means Clustering* bisa sangat bergantung pada inisialisasi awal centroid. Inisialisasi yang berbeda bisa menghasilkan kluster yang sangat berbeda [5]. Ketergantungan *K-Means clustering* pada inisialisasi awal centroid merupakan salah satu kelemahan signifikan dari algoritma ini. Karena *K-Means* menggabungkan data ke dalam kluster berdasarkan kedekatan dengan centroid, pemilihan titik awal centroid dapat sangat mempengaruhi jalannya proses *clustering*. Jika centroid awal dipilih secara tidak representatif, hasil akhir mungkin tidak optimal dan bahkan bisa menyesatkan.

Penelitian selanjutnya yang dilakukan oleh Nandya Arifa Wulandari, dkk. yang bertemakan komparasi antara algoritma *K-Means* dan *Agglomerative Nesting* dalam mengelompokkan data digital marketing pada platform Twitter menyatakan bahwa hasil uji coba antara *K-Means* dan *Agglomerative Nesting* membuktikan bahwa *silhouette_score* tertinggi dicapai oleh metode *Agglomerative Nesting* dengan *complete linkage*, dengan skor 0.754428965 dan jumlah kluster $K=2$ [6]. Dalam hal ini *K-means* menunjukkan bahwa dalam skenario ini, metode hierarkis lebih cocok untuk pengelompokan data yang dianalisis. *K-means* mungkin tidak berhasil menemukan struktur

alami dalam data dengan baik, kemungkinan karena ketergantungan pada inisialisasi awal centroid dan asumsi bentuk klaster yang bulat.

Selain itu, AHC menghasilkan hierarki klaster yang informatif melalui visualisasi dendrogram. Dendrogram memberikan representasi visual yang intuitif tentang struktur pengelompokan data, yang memungkinkan untuk memahami hubungan antar klaster dan tingkat kemiripan di antara mereka. Ini memungkinkan pengguna untuk memahami struktur data secara holistik dan memperoleh wawasan yang mendalam tentang bagaimana data dikelompokkan. Kelebihan ini menjadikan AHC sebagai pilihan yang kuat untuk analisis eksploratif data dan pengambilan keputusan yang didasarkan pada pemahaman yang komprehensif tentang struktur data.

Dalam penelitian ini, peneliti akan menerapkan metode *Single Linkage* pada *Agglomerative Hierarchical Clustering* menggunakan data survei kepuasan mahasiswa untuk mengidentifikasi kelompok-kelompok mahasiswa dengan tingkat kepuasan yang serupa terhadap layanan akademik. Visualisasi pengelompokan klaster akan ditampilkan melalui dendrogram. Melalui penelitian ini, diharapkan dapat memberikan wawasan penting mengenai kepuasan mahasiswa dan menjadi dasar untuk pengembangan strategi yang lebih efektif dalam meningkatkan kualitas layanan akademik di institusi pendidikan.

2. METODOLOGI PENELITIAN

2.1 Pengumpulan Data

Penelitian ini berfokus pada data hasil survei kepuasan mahasiswa terhadap layanan akademik di Program Studi Teknik Informatika UNISNU Jepara pada tahun akademik pada TA. 2022/2023. Terdapat 323 data yang terkumpul dan 5 aspek layanan akademik yang menjadi indikator penilaian, yaitu Aspek *Reliability*, Aspek *Responsiveness*, Aspek *Emphaty*, Aspek *Assurance*, dan Aspek *Tangible*. Setiap aspek memberikan 3 pernyataan penilaian yang berbeda, sehingga pengelompokan akan ditentukan berdasarkan 5 aspek tersebut. Setiap pernyataan pada setiap aspek memberikan tingkat kepuasan mahasiswa, dimana nilai 1 menyatakan sangat tidak puas, nilai 2 menyatakan tidak puas, nilai 3 menyatakan puas dan nilai 4 menyatakan sangat puas. Tabel 1 memuat data yang digunakan dalam penelitian ini.

Tabel 1. Data Survei Kepuasan Mahasiswa terhadap Layanan Akademik

<i>NIM</i>	<i>Nama</i>	<i>Reliability</i>			<i>..</i>	<i>Tangible</i>		
191240000868	Amril Hafid	3	4	3	..	2	3	3
191240000871	Muhammad	4	3	3	..	3	4	4
	Julio Wildan							
191240000872	Noor	3	3	3	..	3	3	3
	Septiyanto							
..	..	..	..	..	..	..	..	..
191240000882	Santika	3	3	3	..	3	3	3
	Agustianti							

2.2 Pre-Processing

Tahapan *preprocessing* (pra-pemrosesan) adalah serangkaian langkah yang dilakukan pada data sebelum masuk ke dalam proses analisis atau pemodelan. Tujuan dari *preprocessing* adalah untuk membersihkan, mempersiapkan, dan mengubah data mentah menjadi bentuk yang lebih sesuai untuk analisis atau pemodelan selanjutnya [7]. Terdapat beberapa tahapan *preprocessing*, yaitu proses *pembersihan data*, *pengurangan data*, *transformasi data*, dan *integrasi data* [8].

Dataset memiliki 3 *variable* yang berbeda pada setiap aspek, sehingga akan dilakukan *preprocessing* dengan penggabungan terhadap *variable* yang serupa dengan menggunakan nilai rata-rata. Hasil pengolahan awal *dataset* tersaji pada Tabel 2.

Tabel 2. Hasil Pengolahan Awal

<i>Index</i>	<i>Reliability</i>	<i>..</i>	<i>Tangible</i>
1	3.33	..	2.67
2	3.33	..	3.67
3	3	..	3
..	...	..	...
323	3	..	3

2.3 Metode Single Linkage

Metode *Single Linkage* merupakan salah satu metode penggabungan kluster dalam algoritma *Agglomerative Hierarchical Clustering* (AHC). Dalam metode ini, jarak antara dua kluster diukur sebagai jarak terdekat antara titik-titik di kedua kluster tersebut [9]. Metode ini juga dikenal dengan nama "*nearest neighbor*" karena menggabungkan dua kluster yang memiliki titik-titik terdekat satu sama lain. Pertama-tama, tentukan jarak terpendek $D = \{d_{ik}\}$ dan mengumpulkan objek yang relevan. Langkah ketiga pada Algoritma AHC perhitungan jarak antara kluster UV dan W dapat dilakukan dengan menghitung

$$d_{(uv)w} = \min \{d_{uw}, d_{vw}\} \quad (1)[10]$$

Pada persamaan tersebut, d_{uw} menyatakan jarak terpendek antara U dan W , sedangkan d_{vw} menyatakan jarak terpendek antara V dan W .

Metode *Single Linkage* memiliki beberapa kelebihan dan kelemahan, diantaranya Kelebihan [11]:

1. Mudah diimplementasikan dan dipahami.
2. Cocok untuk data dengan struktur berbentuk rantai atau *cluster* dengan bentuk yang tidak teratur.
3. Biasanya menghasilkan kluster dengan batasan yang jelas.

Kelemahan :

1. Rentan terhadap efek "*chaining*", di mana kluster yang panjang dapat terbentuk jika terdapat titik-titik yang berdekatan satu sama lain.
2. Rentan terhadap "efek cincin" (*ring effect*), di mana kluster memiliki bentuk seperti cincin yang tidak alami.
3. Sensitif terhadap pencilan (*outliers*), yang dapat mempengaruhi jarak terdekat antara kluster.

Pemilihan metode *Single Linkage*, seringkali tergantung pada sifat data dan tujuan analisis. Metode ini merupakan salah satu opsi yang bisa dipilih dari berbagai pilihan yang ada dalam AHC untuk mengkluster data dan membentuk struktur kluster hierarkis.

2.4 Agglomerative Hierarchical Clustering (AHC)

Agglomerative Hierarchical Clustering (AHC) adalah metode *clustering* hierarkis yang dimulai dari setiap sampel sebagai kluster tunggal dan secara bertahap menggabungkan kluster yang serupa satu sama lain hingga semua sampel menjadi satu kluster besar. Ini adalah proses *bottom-up* di mana penggabungan berlangsung dari kluster individu ke kluster yang lebih besar [12].

Berikut langkah-langkah untuk mengimplementasikan *Agglomerative Hierarchical Clustering*:

1. Inisialisasi Kluster
Setiap titik data diasumsikan sebagai sebuah kelompok tunggal pada awalnya. Misalkan, kita memiliki n sebagai titik data. Pada tahap inisialisasi, setiap titik data diinterpretasikan sebagai 1 kluster. Apabila kita memiliki n titik data, maka kita pasti memiliki n kluster pada mulanya.

2. Perhitungan Matriks Jarak

Hitung matriks jarak antara semua pasangan klaster. Jarak antara dua klaster dapat diukur menggunakan berbagai metrik jarak, seperti jarak *Euclidean* atau jarak *Manhattan*. Pada penelitian ini menggunakan metrik jarak dengan *Euclidean Distance*, dengan rumus berikut

$$D = \sqrt{\sum_{i=1}^n (b_i - a_i)^2} \quad (2) \quad [13]$$

3. Penggabungan Klaster Terdekat

Gabungkan dua klaster yang memiliki jarak terdekat satu sama lain, sehingga membentuk satu klaster baru. Pemilihan pasangan klaster yang akan digabungkan dapat dilakukan menggunakan berbagai metode *linkage*, seperti *Single Linkage*, *Complete Linkage*, atau *Average Linkage*. Penelitian ini menggunakan metode *Single Linkage* dan perhitungan jarak 2 klaster menggunakan Persamaan (1).

4. Update Matriks Jarak

Perbarui matriks jarak untuk mencerminkan penggabungan klaster baru. Jarak antara klaster yang baru terbentuk dengan klaster lain dihitung berdasarkan metode *single linkage*.

5. Ulangi Langkah 3 dan 4

Ulangi proses penggabungan klaster terdekat dan pembaruan matriks jarak sampai semua titik data menjadi satu klaster besar.

6. Dendrogram

Dendrogram dapat digunakan untuk merepresentasikan proses penggabungan secara visual, di mana sumbu x menunjukkan klaster dan sumbu y menunjukkan jarak antara klaster. Dendrogram ini dapat membantu dalam menentukan jumlah klaster yang optimal dan memahami struktur hierarkis dari klaster-klaster yang terbentuk.

2.5 Silhouette Coefficient

Silhouette Coefficient adalah metrik evaluasi yang umum diimplementasikan untuk mengukur efektivitas klaster telah dibentuk dalam proses *clustering*. Nilai *silhouette coefficient* rentang antara -1 hingga 1, di mana nilai optimal mencerminkan klaster yang lebih baik terpisah dan lebih kompak [14]. Dalam penelitian ini, *silhouette coefficient* digunakan untuk mengevaluasi kualitas klaster yang dihasilkan oleh algoritma AHC dengan metode *single linkage*. *Silhouette coefficient* dihitung untuk setiap titik data dan kemudian nilai rata-rata dari semua titik data digunakan sebagai metrik evaluasi.

Hasil *silhouette coefficient* yang mendekati 1 menunjukkan bahwa setiap titik data berada di dalam klaster yang sesuai, sedangkan nilai yang mendekati -1 menunjukkan bahwa titik data mungkin ditempatkan di klaster yang salah. Selain itu, nilai sekitar 0 menunjukkan adanya tumpang tindih antar klaster. Untuk setiap k cluster, *silhouette value* ($sil(k)$) dihitung sebagai rata-rata *silhouette value* dari semua klaster, sesuai Persamaan (3).

$$sil(c) = sil(k) \frac{1}{|k|} \sum_{i=1}^k sil(c_i) \quad (3) \quad [15]$$

Dimana, $sil(k)$ menyatakan nilai siluet semua klaster, $|k|$ banyaknya cluster k , dan $sil(c_i)$ merupakan rata-rata nilai *silhouette*.

3. HASIL DAN PEMBAHASAN

Dataset yang telah diolah melalui pemrosesan data awal selanjutnya akan dilakukan pengelompokkan menggunakan *Agglomerative Hierarchical Clustering* untuk mendapatkan visualisasi dendrogram yang dapat membantu menentukan jumlah klaster optimal berdasarkan seberapa dekat atau jauhnya klaster satu sama lain. Tingkat kedekatan antar klaster dapat dilihat berdasarkan panjang garis vertikal yang menghubungkan dua klaster.

3.1 Euclidean Distance

Pengukuran jarak antara titik data menggunakan *Euclidean Distance*. Perhitungan metrik jarak menggunakan Persamaan (2), Gambar 1 menggambarkan hasil perhitungan tersebut.

```
# Tampilkan matriks jarak
print("Matriks Jarak Euclidean:")
print(distance_matrix)

Matriks Jarak Euclidean:
[[0.          1.59861051 1.29099445 ... 1.29099445 1.82574186 1.29099445]
 [1.59861051 0.          0.81649658 ... 0.81649658 1.73205081 0.81649658]
 [1.29099445 0.81649658 0.          ... 0.          2.23606798 0.          ]
 ...
 [1.29099445 0.81649658 0.          ... 0.          2.23606798 0.          ]
 [1.82574186 1.73205081 2.23606798 ... 2.23606798 0.          2.23606798]
 [1.29099445 0.81649658 0.          ... 0.          2.23606798 0.          ]]]
```

Gambar 1. Hasil Perhitungan Jarak Antar Data Menggunakan Euclidean Distance

3.2 Metode Single Linkage

Setelah mendapatkan jarak antar data menggunakan jarak *Euclidean*, selanjutnya mengukur jarak antar kluster menggunakan *single linkage*. Dalam *Single Linkage* dalam *Agglomerative Hierarchical Clustering*, jarak antara dua kluster diukur sebagai jarak terdekat antara satu titik dalam kluster pertama dengan satu titik dalam kluster kedua. Hasil perhitungan jarak antar kluster dapat terlihat pada Gambar 2.

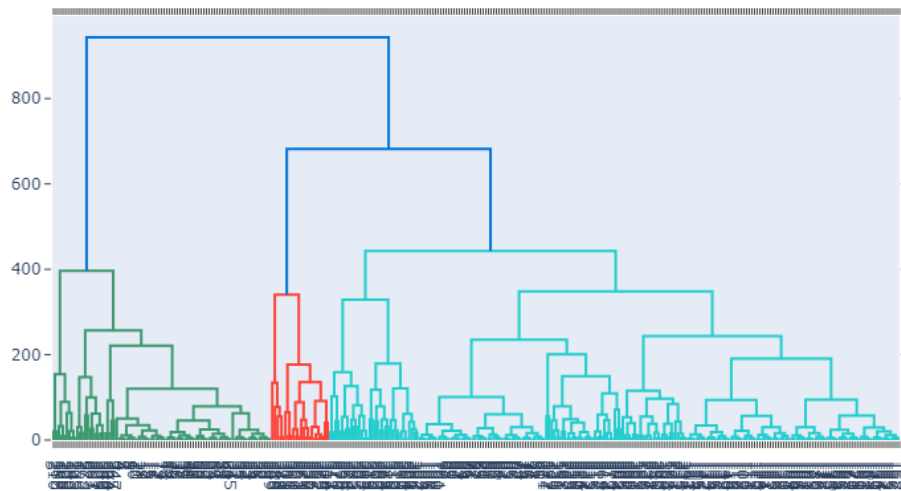
```
# Tampilkan hasil linkage
print("\nHasil Linkage (Single Linkage):")
print(linkage_matrix)

Hasil Linkage (Single Linkage):
[[198.      216.         0.         2.         ]
 [217.      323.         0.         3.         ]
 [234.      324.         0.         4.         ]
 ...
 [313.      641.         1.29099445 320.         ]
 [ 24.      642.         1.37436854 321.         ]
 [565.      643.         2.23606798 323.         ]]
```

Gambar 2. Hasil Perhitungan Jarak Antar Cluster Menggunakan Single Linkage

3.3 Visualisasi Dendrogram

Dendrogram yang dihasilkan dari AHC memberikan visualisasi bagaimana setiap titik data digabungkan menjadi kluster. Sumbu horizontal (x) pada dendrogram mewakili data, sedangkan sumbu vertikal (y) mewakili jarak atau kemiripan antara kluster. Proses penggabungan dimulai dari bawah, di mana setiap titik data individu adalah kluster tersendiri, dan berlanjut ke atas, di mana kluster digabungkan berdasarkan jarak kemiripan terdekat hingga seluruh data menjadi satu kluster besar di puncak dendrogram. Hasil dendrogram tersaji pada Gambar 3.



Gambar 3. Visualisasi Dendrogram

Dendrogram tidak secara langsung menunjukkan jumlah kluster yang terbentuk. Namun, dalam dendrogram memberikan wawasan tentang bagaimana objek dikelompokkan berdasarkan kesamaan karakteristik. Setiap cabang pada dendrogram mewakili penggabungan dua kluster. Tinggi cabang menunjukkan seberapa berbeda dua kluster yang digabungkan. Semakin tinggi, semakin berbeda. Jumlah kluster dapat ditentukan dengan memotong dendrogram pada tingkat tertentu. Pemotongan ini dapat dilakukan dengan mengamati panjang cabang atau menggunakan evaluasi *Silhouette coefficient*.

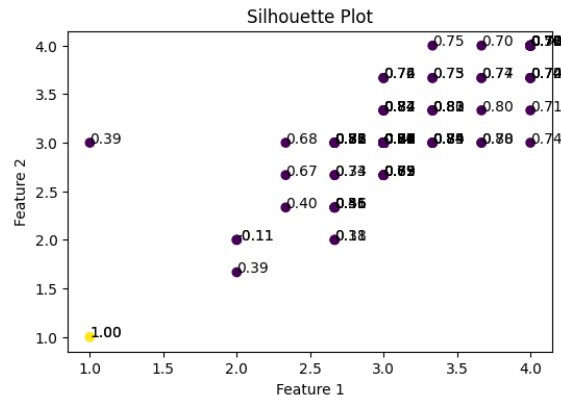
3.4 Hasil Evaluasi *Silhouette Coefficient*

Silhouette Coefficient akan mengukur seberapa baik setiap objek data dikelompokkan dalam kluster tertentu. Nilai berkisar antara -1 hingga 1. Nilai yang lebih tinggi menunjukkan pengelompokan yang lebih baik. Hasil evaluasi penentuan kluster menggunakan koefisien siluet dapat terlihat pada Tabel 3.

Tabel 3. Hasil Evaluasi *Silhouette Coefficient*

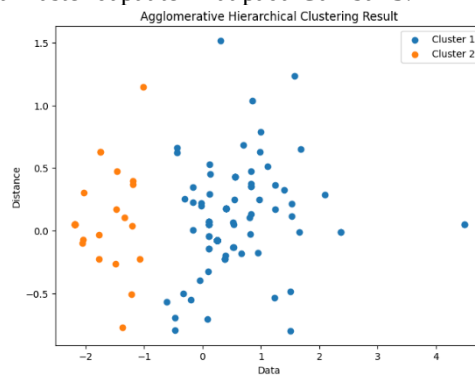
<i>Jumlah Cluster</i>	<i>Nilai Silhouette Coefficient</i>
2	0.80
3	0.63
4	0.57
5	0.54
6	0.52

Dari hasil evaluasi tersebut, nilai *silhouette Coefficient* yang tertinggi mendekati nilai 1 adalah 0.80 dengan jumlah kluster optimal sejumlah 2 kluster.



Gambar 4. Hasil Plot Silhouette Coefficient

Pada Gambar 4. menampilkan *silhouette coefficient* untuk setiap data berdasarkan kluster yang dikelompokkan. Plot ini membantu secara visual mengevaluasi kepadatan dan pemisahan antara kluster. Persebaran anggota kluster dapat terlihat pada Gambar 5.



Gambar 5. Persebaran Anggota Kluster

Pada Gambar 5, terdapat beberapa anggota kluster terlihat sebagai *outlier*. Dalam AHC, metode pengukuran jarak yang digunakan, seperti *Euclidean distance* dapat mempengaruhi hasil klustering. *Euclidean distance* mengukur jarak lurus antara dua titik dalam ruang multidimensi. Penggunaan jarak kuadrat, jarak *euclidean* sangat sensitif terhadap perbedaan yang besar dalam satu atau beberapa fitur. Jika terdapat titik data yang sangat jauh dari titik-titik lainnya (*outlier*), jaraknya menjadi sangat besar, sehingga sangat mempengaruhi struktur kluster yang terbentuk. *Outlier* sering kali memiliki nilai yang ekstrem pada satu atau lebih dimensi, sehingga ketika dihitung dengan *Euclidean distance*, *outlier* tampak sangat jauh dari kelompok utama data, membuatnya terisolasi atau terlihat sebagai kluster tersendiri.

Pembentukan anggota tiap *cluster* adalah *cluster* 0 sebanyak 279 anggota dan *cluster* 1 sebanyak 44 anggota. *Cluster* 0 menyatakan mahasiswa **Puas** dan *cluster* 1 menyatakan mahasiswa **Tidak Puas** dengan layanan akademik di Program Studi Teknik informatika UNISNU Jepara. Detail anggota tiap *cluster* ditunjukkan pada Tabel dibawah ini

Tabel 4. Anggota Cluster

Cluster	Jumlah Anggota	Anggota Cluster
0	279	Amril Hafid, Muhammad Julio Wildan, Noor Septiyanto Santika Agustianti
1	44	Muhammad Rizal Syarif, Zulian Syahreza Latie, Ahmad Aminuddin Yahya Raras Kumala Ragillina

4. KESIMPULAN

Analisis yang telah dilakukan mengarah pada kesimpulan bahwa algoritma *agglomerative hierarchical clustering* merupakan algoritma berbasis hirarki, dimana setiap data dimulai sebagai kluster tunggal dan kemudian secara iteratif digabungkan berdasarkan kedekatan jarak yang sama. Algoritma ini dapat diterapkan tanpa perlu menentukan jumlah kluster terlebih dahulu, namun dengan bantuan dendrogram dapat membantu memberikan secara visual bagaimana hubungan antar data berdasarkan jarak atau kesamaan antara kluster yang digabungkan setiap tahap, walaupun tidak secara langsung dapat menentukan jumlah kluster optimal.

Penentuan kluster optimal dapat melakukan pemotongan dendrogram berdasarkan *threshold* setiap jumlah *cluster* yang ditentukan. Pengujian kluster optimal dapat menggunakan pengujian *silhouette coefficient*, dimana penelitian ini menghasilkan nilai *silhouette* tertinggi di 0.80 dengan 2 jumlah kluster optimal. Kluster pertama memiliki sebanyak 279 anggota menyatakan **Puas** dan kluster kedua sebanyak 44 menyatakan **Tidak Puas** terhadap layanan akademik. Sehingga hasil pengelompokkan ini dapat membantu dalam meningkatkan layanan akademik di masa yang akan datang.

Penerapan metode *linkage* pada algoritma hirarki klustering seperti *Agglomerative Hierarchical Clustering* juga dapat mempengaruhi dalam menentukan kriteria untuk mengukur jarak antar kluster dan bagaimana kluster digabungkan dalam setiap iterasi sehingga berpengaruh pada bentuk dan karakteristik dari dendrogram yang dihasilkan dalam proses clustering. Nilai *Silhouette Coefficient* yang tinggi mengindikasikan bahwa metode *Agglomerative Hierarchical Clustering* dengan pendekatan *Single Linkage* berhasil mengelompokkan data dengan baik dan efisien dengan membentuk kluster yang berbentuk rantai panjang berdasarkan jarak terdekat antar obyek dari kluster tersebut. Pada penelitian selanjutnya dapat mengevaluasi hasil klusterisasi dengan menggunakan berbagai metode *linkage*.

DAFTAR PUSTAKA

- [1] D. Nyoman and B. Kusyana. 2020. "Peran Kualitas Layanan Dalam Menciptakan Loyalitas Mahasiswa Putu Atim Purwaningrat (2) Milla Permata Sunny (3)(1)(2)(3) Fakultas Ekonomi Bisnis dan Pariwisata Universitas Hindu Indonesia," vol. 2, no. 1, pp. 10–27. doi: <https://dx.doi.org/10.32795/widyamanajemen.v2i1>
- [2] K. Atma Wijaya and D. Swanjaya, "Integrasi Metode Agglomerative Hierarchical Clustering dan Backpropagation Pada Model Peramalan Penjualan."
- [3] Z. Arifin, D. S. Santosa, and M. A. Soeleman, "KLAUSTERISASI GENRE CERPEN KOMPAS MENGGUNAKAN AGGLOMERATIVE HIERARCHICAL CLUSTERING-SINGLE LINKAGE," 2017. [Online]. Available: <http://research>.
- [4] N. A. Maori. 2023. "METODE ELBOW DALAM OPTIMASI JUMLAH CLUSTER PADA K-MEANS CLUSTERING," *Jurnal SIMETRIS*, vol. 14.
- [5] K. Ariasa, I. Gede, A. Gunadi, and I. Made Candiasa, "OPTIMASI ALGORITMA KLAUSTER DINAMIS PADA K-MEANS DALAM PENGELOMPOKAN KINERJA AKADEMIK MAHASISWA (STUDI KASUS: UNIVERSITAS PENDIDIKAN GANESHA)."
- [6] N. Arifa Wulandari, H. Pratiwi, S. Sulistijowati Handayani, and F. Matematika dan Ilmu Pengetahuan Alam, "Program Studi Teknik Informatika," 2023. [Online]. Available: <https://netlytic.org/>.

- [7] F. Alghifari and D. Juardi, "Fauzan Alghifari Penerapan Data Mining Pada Penerapan Data Mining Pada Penjualan Makanan Dan Minuman Menggunakan Metode Algoritma Naïve Bayes."
- [8] H. Junaedi *et al.*, "Prosiding Konferensi Nasional 'Inovasi dalam Desain dan Teknologi'-IDEaTech 2011 DATA TRANSFORMATION PADA DATA MINING".
- [9] N. Afira and A. W. Wijayanto. 2021. "Analisis Cluster dengan Metode Partitioning dan Hierarki pada Data Informasi Kemiskinan Provinsi di Indonesia Tahun 2019," *Komputika: Jurnal Sistem Komputer*, vol. 10, no. 2, pp. 101–109. doi: <https://dx.doi.org/10.34010/komputika.v10i2.4317>
- [10] A. F. Dewi and K. Ahadiyah. 2022. "Agglomerative Hierarchy Clustering Pada Penentuan Kelompok Kabupaten/Kota di Jawa Timur Berdasarkan Indikator Pendidikan," *Zeta - Math Journal*, vol. 7, no. 2, pp. 57–63. doi: <https://dx.doi.org/10.31102/zeta.2022.7.2.57-63>
- [11] Y. Reinaldi, N. Ulinnuha, and Moh. Hafiyusholeh. 2021. "Comparison of Single Linkage, Complete Linkage, and Average Linkage Methods on Community Welfare Analysis in Cities and Regencies in East Java," *Jurnal Matematika, Statistika dan Komputasi*, vol. 18, no. 1, pp. 130–140. doi: <https://dx.doi.org/10.20956/j.v18i1.14228>
- [12] R. Jannah Alfirdausy *et al.*, "SISTEMASI: Jurnal Sistem Informasi Analisis Data Indeks Pembangunan Manusia Kabupaten/ Kota di Jawa Timur melalui Pengelompokan menggunakan Metode Clustering Aglomeratif Hirarki Analysis of Regency/City Human Development Index Data in East Java Through Grouping Using Hierarchical Agglomerative Clustering Method." [Online]. Available: <http://sistemasi.ftik.unisi.ac.id>
- [13] R. Kusumastuti, E. Bayunanda, A. Muhammad Rifa, M. Ryandy Ghonim Asgar, and F. Inti Ilmawati, "Clustering Titik Panas Menggunakan Algoritma Agglomerative Hierarchical Clustering (AHC) Hot Spot Clustering Using Agglomerative Hierarchical Clustering (AHC) Algorithm," *Cogito Smart Journal* |, vol. 8, no. 2.
- [14] S. Mutdilah, I. Cholissodin, and F. A. Bachtiar, "Pengelompokan Daerah Berpotensi Transmigrasi Menggunakan Metode Improved k-Means (Studi Kasus: Kabupaten Malang)," 2020. [Online]. Available: <http://j-ptiik.ub.ac.id>
- [15] S. Paembonan, H. Abduh, and K. Kunci, "Penerapan Metode Silhouette Coefficient Untuk Evaluasi Clustering Obat Clustering; K-means; Silhouette coefficient," 2021. [Online]. Available: <https://ojs.unanda.ac.id/index.php/jiit/index>