

IMPLEMENTASI ALGORITMA *CLUSTERING* UNTUK PENGELOMPOKAN MAHASISWA BERDASARKAN RESPON TERHADAP METODE PEMBELAJARAN BAHASA INGGRIS KOMPUTER

Justin Eduardo Simarmata^{1*)}, Iis Aprianti²⁾, Debora Chrisinta³⁾, Miko Purnomo⁴⁾
Universitas Timor¹²³⁴⁾

justinesimarmata@unimor.ac.id¹⁾, iisaprianti@unimor.ac.id²⁾, deborachrisinta@unimor.ac.id³⁾,
mikopurnomo@unimor.ac.id⁴⁾

Abstrak

Permasalahan yang dihadapi dalam proses pembelajaran Bahasa Inggris adalah rendahnya minat dan motivasi belajar mahasiswa yang dipengaruhi oleh persepsi mahasiswa terhadap metode pembelajaran yang digunakan, khususnya media video. Perbedaan persepsi ini dapat menyebabkan ketidakefektifan dalam penyampaian materi serta pencapaian hasil belajar yang tidak merata. Penelitian ini bertujuan untuk mengelompokkan mahasiswa berdasarkan respon mahasiswa terhadap media video dalam pembelajaran Bahasa Inggris, sehingga dapat diketahui karakteristik masing-masing kelompok dan strategi pembelajaran yang sesuai. Metode yang digunakan adalah analisis kluster dengan pendekatan Hierarchical Agglomerative Clustering menggunakan metode Ward. Data diperoleh dari kuesioner yang mencakup enam variabel, kemudian dilakukan pra-pemrosesan meliputi pembersihan data dan transformasi data kategorik menjadi numerik. Hasil analisis menunjukkan terbentuknya tiga kluster mahasiswa dengan distribusi yang relatif merata, di mana setiap kluster memiliki karakteristik persepsi dan motivasi belajar yang berbeda. Kluster pertama didominasi oleh sikap netral, kluster kedua menunjukkan respon positif terhadap video pembelajaran, dan kluster ketiga memperlihatkan preferensi terhadap video dengan konten percakapan. Hasil penelitian menunjukkan bahwa pengajar dapat menyesuaikan strategi dan konten pembelajaran dengan karakteristik masing-masing kelompok untuk meningkatkan efektivitas dan keterlibatan mahasiswa.

Kata kunci: Clustering, Metode Ward, Pembelajaran Bahasa Inggris, Persepsi Mahasiswa.

1. Pendahuluan

Inovasi dalam teknologi informasi mendorong pembaruan dalam dunia pendidikan, termasuk dalam mengajarkan Bahasa Inggris. Pembelajaran bahasa Inggris menggunakan komputer adalah salah satu inovasi yang kini banyak digunakan. (Haydar et al., 2022) mengemukakan bahwa dengan komputer, interaksi dan partisipasi siswa dalam belajar dapat lebih baik. Namun, efektivitas metode ini sangat bervariasi di antara mahasiswa. Sementara sebagian mahasiswa mengalami peningkatan yang sangat baik, ada sebagian yang mengalami kesulitan mengenal materi yang disajikan. Hal itu disebabkan oleh berbagai hal diantaranya adalah gaya belajar, motivasi, dan literasi teknologi. Penelitian (Contrino et al., 2024) menggaris bawahi bahwa semua usaha dalam pembelajaran komputer harus disertai dengan perubahan untuk

karakteristik khas masing-masing mahasiswa agar mengoptimalkan proses pembelajaran.

Oleh karena itu untuk mengatasi tantangan diatas, pendekatan yang menggambarkan siswa berdasarkan reaksi mahasiswa terhadap metode pembelajaran perlu diterapkan. Salah satu pendekatan yang digunakan adalah dengan yaitu teknik penambangan data (*data mining*), khususnya analisis kluster. Dengan mengelompokkan mahasiswa berdasarkan pola respons mahasiswa, mahasiswa dapat merancang strategi pengajaran yang lebih responsif terhadap kebutuhan setiap kelompok. (Niu, 2024) dalam studinya menggunakan algoritma K-Means untuk mengevaluasi kualitas pengajaran bahasa Inggris dan mampu mengelompokkan mahasiswa berdasarkan karakteristik tertentu untuk perbaikan metode pengajaran selanjutnya. Selain itu, (Jia, 2024) menunjukkan bahwa kombinasi

metode AHP dan K-Means dalam menilai pembelajaran hibrida meningkatkan akurasi dan efisiensi dalam mengevaluasi proses pengajaran dan pembelajaran mahasiswa.

Namun dalam studi ini, algoritma Ward akan digunakan sebagai pendekatan untuk analisis kluster. Berbeda dengan K-Means yang lebih populer di kalangan praktisi penelitian pendidikan, metode Ward bersifat hierarkis dan berkonsentrasi pada minimisasi varians dalam setiap kluster. Metode ini diharapkan dapat menghasilkan kelompok yang lebih homogen dan menangkap gambaran komposit karakteristik mahasiswa dengan lebih lengkap. Selain itu, dalam penelitian ini data yang digunakan tergolong dalam ukuran data yang sedikit sehingga lebih cocok diterapkan pada algoritma hirarki seperti metode Ward.

Pendekatan pembelajaran bahasa Inggris dengan analisis kluster telah diadopsi dalam beberapa penelitian dan menunjukkan hasil yang efektif. (Guo et al., 2024) menerapkan teknik *clustering* pada pengajaran Bahasa Inggris, sedangkan (Niu, 2024) (2022) dan (Jia, 2024) berhasil mengevaluasi pembelajaran serta merancang metodenya secara lebih efisien. (Zhang, 2025) menganalisis skor Bahasa Inggris menggunakan algoritma K-Means dan berhasil mengidentifikasi pola belajar mahasiswa, sementara (Yu, 2025) menggunakan fuzzy K-Means untuk mengukur kompetensi pengajaran. Berdasarkan penelitian terdahulu yang mengkaji algoritma *clustering* pada pembelajaran Bahasa Inggris, belum ditemukan penerapan algoritma Ward sebagai metode utama pengelompokan. Oleh karena itu, pada penelitian ini berfokus pada penerapan algoritma Ward untuk mengetahui pola respon mahasiswa dalam menerima pembelajaran Bahasa Inggris menggunakan *Multimedia-Based Learning* (video). Diharapkan penelitian ini menjadi kontribusi dalam pengembangan strategi pembelajaran yang lebih profesional di tingkat pendidikan tinggi, terutama yang berkaitan dengan penggunaan teknologi.

2. Metode Penelitian

Studi ini menggunakan pendekatan kuantitatif dengan metode penelitian eksploratif deskriptif (Simarmata & Chrisinta, 2024). Tujuan utama dari studi ini adalah untuk mengklasifikasikan mahasiswa berdasarkan respons kelas mahasiswa

terhadap teknik pembelajaran bahasa Inggris yang dibantu komputer menggunakan algoritma kluster Ward. Penelitian ini dilakukan dalam langkah-langkah sistematis yang melibatkan pengumpulan data, pengolahan data, implementasi algoritma, dan interpretasi hasil kluster dengan tujuan untuk menemukan karakteristik mahasiswa dalam belajar bahasa Inggris dengan teknologi.

2.1 Data dan Teknik Pengumpulan Data

Dataset dalam studi ini terdiri dari data primer yang dikumpulkan melalui pemberian kuesioner kepada mahasiswa yang telah mengambil mata kuliah Bahasa Inggris Komputer 1 Program Studi Teknologi Informasi Universitas Timor. Kuesioner disusun menggunakan skala Likert 5 poin untuk menilai respons mahasiswa terhadap pemahaman topik yang diajarkan.

2.2 Definisi Operasional Variabel

Variabel yang digunakan dalam penelitian ini terdiri dari Pemahaman, Ketertarikan, Kejelasan Video, *Subtitle*, Motivasi Belajar dan Preferensi Video. Penjelasan definisi tiap variabel diberikan pada Tabel 1 berikut:

Tabel 1. Definisi Operasional Variabel

Variabel	Definisi
Pemahaman	Seberapa paham mahasiswa terhadap materi dalam video.
Ketertarikan	Apakah mahasiswa menikmati metode pembelajaran dengan video.
Kejelasan Video	Apakah penyampaian dalam video jelas.
<i>Subtitle</i>	Seberapa besar subtitle membantu pemahaman.
Motivasi Belajar	Apakah video membuat mahasiswa lebih semangat belajar.
Preferensi Video	Tipe video yang paling disukai, bisa dikodekan untuk klustering juga.

2.3 Subjek Penelitian dan Teknik Pengambilan Sampel

Subjek penelitian adalah mahasiswa yang mengikuti mata kuliah Bahasa Inggris Komputer 1 Program Studi Teknologi Informasi Universitas

Timor. Teknik pengambilan sampel menggunakan *purposive sampling* (Bait Bifel et al., 2025; Simarmata, 2020) dengan kriteria mahasiswa yang mengikuti mata kuliah Bahasa Inggris Komputer 1 Program Studi Teknologi Informasi Universitas Timor. Jumlah sampel minimal ditetapkan sebanyak 150 mahasiswa untuk mendapatkan hasil analisis kluster yang representatif.

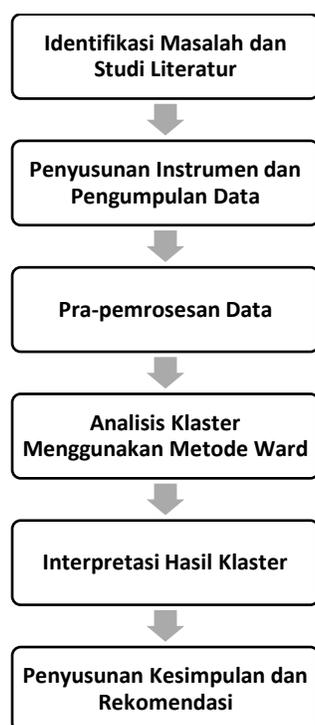
2.4 Perangkat Lunak dan Platform Penelitian

Dalam proses analisis data, penelitian ini menggunakan perangkat lunak Google Form untuk pengumpulan data menggunakan kuisoner, serta Google Colab yang menyediakan Notebook bahasa pemrograman Python (scikit-learn, scipy, matplotlib) untuk proses analisis kluster menggunakan algoritma Ward dan visualisasi dendrogram.

2.5 Tahapan Penelitian

Sub ini dilanjutkan sampai dengan keseluruhan tahapan dalam Penelitian dilaksanakan. Apabila Penelitian memiliki 5 tahapan, maka sub metode harus berjumlah 5.

Adapun alur tahapan penelitian dapat dilihat pada Gambar 1 berikut:



Gambar 1. Tahapan Alur Penelitian

1. Identifikasi Masalah dan Studi Literatur

Pada tahap awal, peneliti mengidentifikasi permasalahan dalam pembelajaran Bahasa Inggris berbasis komputer, yaitu adanya perbedaan respons mahasiswa terhadap metode tersebut. Untuk memahami lebih dalam, dilakukan studi literatur dari jurnal-jurnal terkini (2020–2025) yang berkaitan dengan evaluasi pembelajaran, data mining, dan teknik analisis kluster. Tahap ini bertujuan membangun dasar teori dan menyusun arah metodologi penelitian secara jelas.

2. Penyusunan Instrumen dan Pengumpulan Data

Peneliti menyusun instrumen berupa kuisoner dengan skala Likert untuk mengukur respons mahasiswa terhadap berbagai aspek pembelajaran Bahasa Inggris berbasis komputer, seperti kenyamanan penggunaan teknologi, keterlibatan, pemahaman materi, dan motivasi belajar. Kuisoner didistribusikan secara daring menggunakan Google Form. Data tambahan seperti nilai akademik mahasiswa juga dikumpulkan untuk melengkapi analisis.

3. Pra-pemrosesan Data

Data yang terkumpul selanjutnya melalui proses data cleaning untuk menghilangkan duplikasi, mengisi atau membuang data kosong, dan mengecek konsistensi jawaban (Chrisinta & Simarmata, 2023; Simarmata et al., 2024). Setelah itu dilakukan transformasi data agar setiap variabel memiliki skala yang sebanding. Tahapan ini sangat penting agar hasil analisis kluster tidak bias karena skala data yang berbeda.

4. Analisis Kluster Menggunakan Metode Ward

Pada tahap inti, analisis kluster dilakukan menggunakan algoritma hierarkis metode Ward dengan pengukuran jarak Euclidean. Berikut persamaan yang digunakan:

$$d(A, B) = \sqrt{(x_1 - y_1)^2 + \dots + (x_n - y_n)^2}$$

dimana objek $A = (x_1, x_2, \dots, x_n)$ dan $B = (y_1, y_2, \dots, y_n)$. Proses perhitungan dilakukan dengan bantuan bahasa pemrograman Python. Dendrogram akan divisualisasikan untuk menentukan jumlah kluster optimal. Metode Ward dipilih karena kemampuannya meminimalkan variansi intra-kluster, sehingga

menghasilkan kelompok mahasiswa yang lebih homogen. Adapun langkah umum dalam metode Ward adalah sebagai berikut:

- Asumsikan bahwa setiap objek pengamatan adalah satu kluster.
- Menghitung jarak semua pasang kluster.
- Gabungan dua kluster dengan peningkatan variasi terkecil dengan menggunakan persamaan berikut:

$$SSE = \sum_{j=1}^p \left(\sum_{i=1}^n x_{ij}^2 - \frac{1}{2} \left(\sum_{i=1}^n x_{ij} \right)^2 \right)$$

dimana x_{ij} adalah nilai objek ke- i pada kluster ke- j , p adalah banyaknya variabel yang diukur, dan n adalah banyaknya objek dalam kluster yang terbentuk.

- Ulangi proses sampai semua objek tergabung dalam satu kluster besar atau samapai jumlah kluster yang diinginkan tercapai.
- Visualisasikan hasil dalam bentuk dendrogram.

5. Interpretasi Hasil Kluster

Setelah terbentuk kluster, masing-masing kluster dianalisis berdasarkan nilai persentase

kemunculan kategori. Tahap ini menjadi dasar dalam menyusun strategi pembelajaran yang lebih adaptif untuk setiap kelompok.

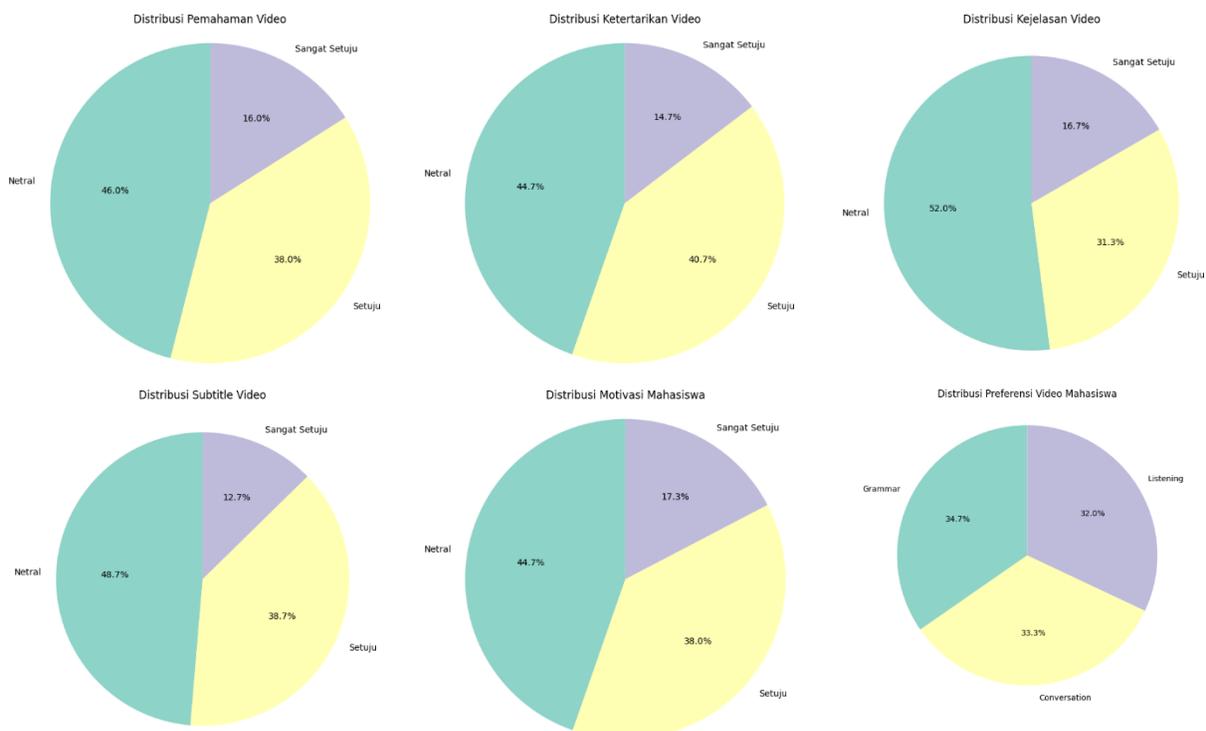
6. Penyusunan Kesimpulan dan Rekomendasi

Tahap akhir adalah menyimpulkan hasil penelitian dan memberikan rekomendasi konkret kepada pengajar atau lembaga pendidikan, berdasarkan pola yang ditemukan dari hasil kluster. Peneliti juga menguraikan kontribusi penelitian dan saran untuk pengembangan studi lanjutan. Dokumentasi dan pelaporan dilakukan secara sistematis dalam bentuk skripsi atau artikel ilmiah.

3. Hasil dan Pembahasan

3.1 Deskripsi Data Responden

Penelitian ini melibatkan 150 mahasiswa yang telah mengikuti pembelajaran Mata Kuliah Bahasa Inggris Komputer. Karakteristik responden berdasarkan distribusi jawaban kuisioner diberikan pada Gambar 2 berikut:



Gambar 2. Distribusi Jawaban Responden Pada Variabel Penelitian

Distribusi jawaban responden terhadap variabel-variabel dalam penelitian ini menunjukkan kecenderungan mayoritas berada pada kategori netral dan setuju. Pada variabel Pemahaman Video, sebanyak 46% responden memberikan jawaban netral, diikuti oleh 38% yang menyatakan setuju, dan hanya 16% yang sangat setuju. Pola serupa terlihat pada variabel Ketertarikan terhadap Video, di mana 44,7% menyatakan netral dan 40,7% setuju, sementara 14,7% sangat setuju. Hal ini menunjukkan bahwa sebagian besar mahasiswa memiliki ketertarikan sedang hingga tinggi terhadap video pembelajaran, meskipun belum menunjukkan antusiasme yang sangat kuat.

Pada variabel Kejelasan Video, kecenderungan netral lebih dominan dengan 52%, disusul 31,3% setuju, dan 16,7% sangat setuju. Ini mengindikasikan bahwa kejelasan penyampaian materi video perlu ditingkatkan agar pemahaman mahasiswa lebih optimal. Pada aspek Subtitle Video, sebanyak 48,7% responden memberikan tanggapan netral, 38,7% setuju, dan hanya 12,7% sangat setuju. Ini menandakan bahwa keberadaan subtitle belum sepenuhnya dimanfaatkan atau dirasakan efektif oleh mahasiswa.

Variabel Motivasi Mahasiswa menunjukkan bahwa 44,7% berada dalam posisi netral, 38% setuju, dan 17,3% sangat setuju. Artinya, motivasi mahasiswa dalam pembelajaran bahasa Inggris berbasis video masih tergolong sedang dan perlu didorong lebih lanjut. Terakhir, pada variabel Preferensi Jenis Video, terdapat distribusi yang cukup seimbang antara tiga pilihan utama: Grammar disukai oleh 34,7% responden, Conversation oleh 33,3%, dan Listening oleh 32%. Hal ini mencerminkan bahwa minat mahasiswa cukup beragam dan tersebar merata pada ketiga jenis materi tersebut.

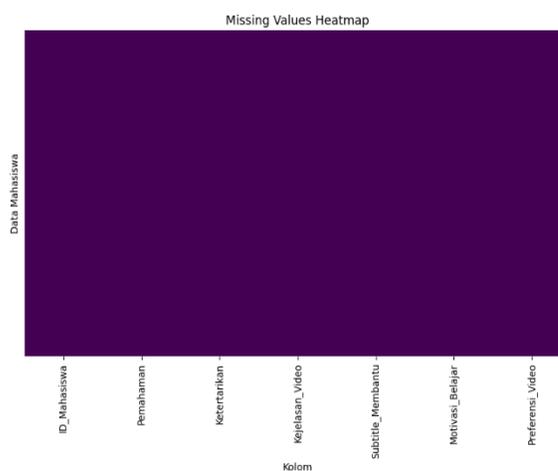
Secara keseluruhan, hasil ini menunjukkan bahwa meskipun respon mahasiswa terhadap video pembelajaran cenderung positif, masih terdapat ruang untuk peningkatan kualitas konten, kejelasan, serta strategi penyajian yang lebih menarik dan memotivasi.

3.2 Hasil Uji Validitas dan Reliabilitas Instrumen

Seperti halnya tabel, pastikan setiap gambar mempunyai nomor urut dan judul. Buatlah gambar yang Anda gunakan nampak seperti buatan profesional dan tidak perlu diberi bingkai. Pastikan gunakan gambar hitam-putih. Peletakan gambar berada di tengah halaman, dengan judul

3.3 Pra-pemrosesan Data

Sebelum dilakukan analisis lebih lanjut, data yang telah dikumpulkan melalui kuesioner online mengalami tahap pra-pemrosesan untuk memastikan kualitas dan kelengkapan data. Berdasarkan hasil visualisasi dalam bentuk heatmap pada Gambar 3, seluruh kolom tidak mengandung missing values atau data kosong, yang ditunjukkan dengan warna ungu pekat yang seragam pada seluruh bagian grafik. Ini menandakan bahwa setiap responden mengisi seluruh bagian kuesioner dengan lengkap, sehingga data dapat langsung digunakan tanpa perlu melakukan imputasi atau penghapusan baris.



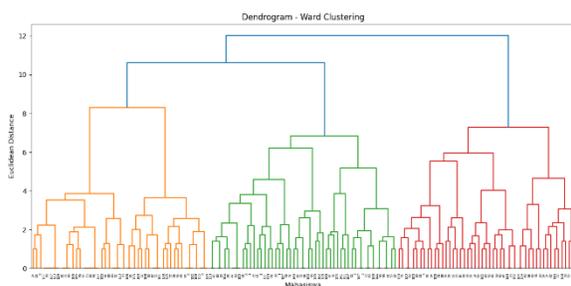
Gambar 3. Pemeriksaan Data Hilang

Selain itu, dilakukan juga transformasi data kategorik menjadi numerik untuk keperluan analisis kluster. Misalnya, pada variabel skala Likert seperti “Sangat Setuju”, “Setuju”, dan “Netral” dikonversi ke dalam bentuk angka, masing-masing menjadi 3, 2, dan 1. Demikian pula, pada variabel preferensi video yang bersifat

kategorik (Grammar, Conversation, dan Listening) juga dikodekan dalam bentuk numerik agar dapat diolah secara matematis. Tahapan ini penting dilakukan agar algoritma kluster seperti Ward's Method dapat bekerja secara optimal terhadap data numerik dan menghasilkan pengelompokan yang akurat.

3.4 Hasil Analisis Kluster dengan Metode Ward

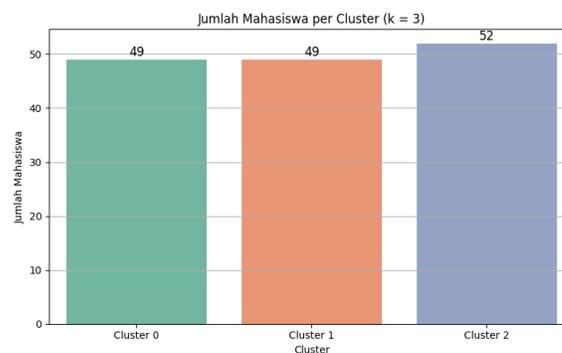
Hasil analisis kluster menggunakan metode Ward divisualisasikan melalui dendrogram pada Gambar 1. Dendrogram ini menunjukkan struktur hierarki pengelompokan mahasiswa berdasarkan kesamaan pola jawaban terhadap variabel-variabel dalam penelitian. Pada dendrogram tersebut, garis vertikal menunjukkan jarak antar kluster (*Euclidean Distance*), sedangkan percabangan horizontal menggambarkan proses penggabungan antar data atau kluster. Titik potong pada level sekitar 12 digunakan untuk menentukan jumlah kluster optimal, yang menghasilkan 3 kluster utama. Pemilihan titik ini berdasarkan pengamatan terhadap jarak penggabungan yang signifikan, di mana semakin besar jarak antar penggabungan, semakin berbeda pula karakteristik antar kluster.



Gambar 4. Dendrogram Kluster Optimal

Selanjutnya, pada Gambar 2 ditampilkan visualisasi distribusi jumlah mahasiswa pada tiap kluster. Terlihat bahwa jumlah mahasiswa dalam tiap kluster relatif seimbang, yaitu 49 mahasiswa pada Kluster 0, 49 mahasiswa pada Kluster 1, dan 52 mahasiswa pada Kluster 2. Hal ini menunjukkan bahwa metode Ward berhasil membentuk kluster dengan distribusi yang proporsional, tanpa dominasi jumlah responden yang ekstrem di salah satu kelompok. Pembagian ini memberikan dasar yang kuat untuk melakukan

interpretasi lebih lanjut terhadap karakteristik masing-masing kluster, guna mengetahui preferensi dan respons mahasiswa yang berbeda terhadap metode pembelajaran Bahasa Inggris berbasis komputer.



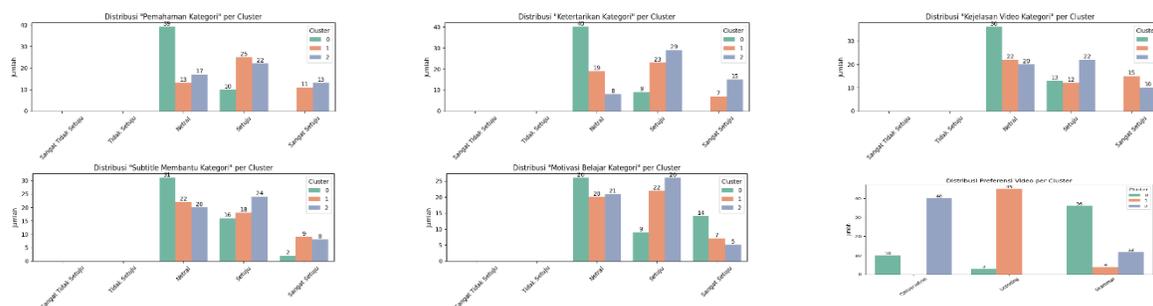
Gambar 5. Distribusi Frekuensi Jumlah Mahasiswa dalam Kluster Optimal

Distribusi kategori pada variabel penelitian berdasarkan kluster optimal diberikan pada Gambar 6. Seperti halnya tabel, pastikan setiap gambar mempunyai nomor urut dan judul. Buatlah gambar yang Anda gunakan nampak seperti buatan profesional dan tidak perlu diberi bingkai. Setiap grafik batang menunjukkan jumlah mahasiswa dalam tiap kluster (*Cluster 0*, *Cluster 1*, dan *Cluster 2*) berdasarkan kategori jawaban mahasiswa terhadap masing-masing variabel. Terlihat bahwa *Cluster 0* didominasi oleh responden yang cenderung netral terhadap sebagian besar variabel seperti pemahaman, ketertarikan, kejelasan video, dan bantuan subtitle. Sebaliknya, *Cluster 1* menunjukkan dominasi respon “setuju” dan “sangat setuju”, khususnya pada variabel motivasi belajar dan subtitle membantu pemahaman. Sementara itu, *Cluster 2* memiliki kecenderungan yang cukup tinggi pada kategori “setuju” dan menunjukkan distribusi yang relatif merata di berbagai kategori untuk beberapa variabel seperti ketertarikan dan preferensi video.

Pada distribusi preferensi video, masing-masing kluster memiliki kecenderungan yang berbeda. *Cluster 0* lebih banyak memilih grammar, *Cluster 1* didominasi oleh preferensi listening, sedangkan *Cluster 2* menunjukkan kecenderungan terhadap conversation. Hal ini menunjukkan bahwa tiap kluster memiliki karakteristik pembelajaran yang berbeda, baik

dari segi persepsi terhadap media pembelajaran video maupun jenis materi yang lebih diminati. Informasi ini sangat berguna dalam menyusun

strategi pembelajaran Bahasa Inggris berbasis komputer yang lebih personal dan tepat sasaran bagi masing-masing kelompok mahasiswa.



Gambar 6. Distribusi Kategori pada Variabel Penelitian Berdasarkan Kluster Optimal

3.5 Pembahasan

Berdasarkan hasil penelitian, ditemukan bahwa persepsi mahasiswa terhadap penggunaan media video sebagai alat bantu pembelajaran Bahasa Inggris menunjukkan variasi yang cukup menarik. Secara umum, mahasiswa cenderung memberikan respon pada kategori netral hingga setuju, menunjukkan penerimaan yang cukup baik terhadap media pembelajaran berbasis video. Hasil pra-pemrosesan data menunjukkan bahwa data yang digunakan dalam penelitian ini tidak mengandung nilai kosong (missing values), sebagaimana ditunjukkan oleh heatmap yang bersih dari area terang. Ini menandakan kualitas data yang baik untuk analisis lebih lanjut. Selain itu, data kategorik yang ada telah berhasil ditransformasi menjadi numerik agar sesuai dengan kebutuhan metode klusterisasi.

Melalui metode Hierarchical *Clustering* dengan pendekatan Ward, diperoleh struktur dendrogram yang memperlihatkan tiga kelompok (kluster) utama mahasiswa berdasarkan kemiripan persepsi mahasiswa terhadap beberapa aspek video pembelajaran. Distribusi mahasiswa dalam tiap kluster juga cukup merata, yaitu 49 mahasiswa pada *Cluster 0* dan *Cluster 1*, serta 52 mahasiswa pada *Cluster 2*. Keseimbangan ini menunjukkan bahwa pengelompokan cukup stabil dan mampu merepresentasikan heterogenitas persepsi secara adil. Hasil ini sejalan dengan studi oleh (Wang et al., 2024) yang mengungkap bahwa pendekatan klusterisasi dalam analisis persepsi dapat mengungkapkan

preferensi pembelajaran yang tersembunyi dan tidak terdeteksi melalui analisis deskriptif biasa.

Distribusi kategori variabel penelitian berdasarkan kluster optimal menunjukkan karakteristik berbeda-beda. *Cluster 0* didominasi oleh jawaban netral, yang mencerminkan sikap ragu-ragu atau ketidakpastian mahasiswa terhadap efektivitas video pembelajaran. Ini bisa jadi disebabkan oleh kurangnya paparan terhadap media video atau preferensi terhadap metode pembelajaran konvensional. *Cluster 1* menunjukkan mayoritas respon setuju dan sangat setuju, mencerminkan sikap positif dan pengalaman belajar yang lebih baik melalui video. Hal ini sesuai dengan temuan (Sitinjak, 2022) yang menyatakan bahwa video pembelajaran yang dikemas secara menarik mampu meningkatkan minat dan pemahaman mahasiswa dalam pembelajaran daring.

Cluster 2 memperlihatkan keragaman persepsi yang unik, dengan jumlah tertinggi dalam preferensi video kategori conversation, menunjukkan bahwa mahasiswa dalam kluster ini lebih tertarik pada video yang bersifat komunikatif dan aplikatif. Ini diperkuat oleh studi (Wijayanti et al., 2024) yang menemukan bahwa mahasiswa cenderung merasa lebih tertarik dan termotivasi belajar ketika disajikan materi dalam bentuk video yang menampilkan percakapan sehari-hari, karena materi tersebut dianggap relevan dan kontekstual dalam kehidupan nyata.

Lebih lanjut, analisis terhadap variabel motivasi belajar menunjukkan bahwa mahasiswa dalam *Cluster 1* memiliki tingkat motivasi yang lebih tinggi dibanding kluster lainnya. Hal ini

mengindikasikan bahwa persepsi positif terhadap media video berbanding lurus dengan peningkatan motivasi belajar. Penelitian oleh Sari & Nugroho (2020) mendukung hal ini dengan menyatakan bahwa pembelajaran berbasis video dapat menciptakan pengalaman belajar yang lebih personal dan engaging, sehingga mampu meningkatkan motivasi intrinsik peserta didik.

Dengan demikian, hasil penelitian ini menegaskan bahwa media video memiliki potensi besar dalam mendukung pembelajaran Bahasa Inggris, terutama jika dirancang dengan mempertimbangkan preferensi mahasiswa. Oleh karena itu, disarankan agar pengembang media pembelajaran memperhatikan aspek visual, kejelasan narasi, keberadaan subtitle, serta konten yang kontekstual seperti percakapan. Strategi ini selaras dengan pendekatan student-centered learning, di mana materi pembelajaran disesuaikan dengan kebutuhan dan karakteristik peserta didik secara lebih personal.

4. Kesimpulan dan Saran

Penelitian ini berhasil mengelompokkan mahasiswa ke dalam tiga klaster berdasarkan persepsi mahasiswa terhadap pembelajaran menggunakan media video. Proses pra-pemrosesan data menunjukkan bahwa tidak terdapat data yang hilang, sehingga seluruh data dapat digunakan dalam analisis. Hasil klusterisasi menggunakan metode Ward menghasilkan dendrogram yang menggambarkan hubungan hierarkis antar mahasiswa dan menunjukkan struktur klaster yang optimal. Distribusi mahasiswa yang hampir merata dalam ketiga klaster—*Cluster 0* (49 mahasiswa), *Cluster 1* (49 mahasiswa), dan *Cluster 2* (52 mahasiswa)—menandakan adanya perbedaan preferensi dan persepsi yang cukup signifikan di antara kelompok-kelompok tersebut. Analisis lebih lanjut terhadap karakteristik masing-masing klaster menunjukkan perbedaan dominan dalam tingkat pemahaman, ketertarikan, kejelasan video, kebermanfaatan subtitle, serta motivasi belajar. Hal ini mengindikasikan bahwa setiap klaster memiliki kebutuhan dan pendekatan pembelajaran yang berbeda. penelitian selanjutnya, disarankan untuk memperluas cakupan data dan mempertimbangkan faktor tambahan seperti latar belakang pendidikan, gaya belajar, atau tingkat penguasaan teknologi guna

mendapatkan segmentasi yang lebih komprehensif. Selain itu, dapat pula melakukan studi komparasi algoritma data mining untuk memperoleh hasil yang lebih representatif dengan mempertimbangkan kualitas klaster yang diperoleh.

DAFTAR PUSTAKA

- Bait Bifel, M. G., Timor, U., & Simarmata, J. E. (2025). Identifikasi Pengaruh Fasilitas Belajar Terhadap Hasil Belajar Matematika Peserta Didik Menggunakan Pendekatan Structural Equation Modeling. *Jurnal Jendela Matematika*, 3.
- Chrisinta, D., & Simarmata, J. E. (2023). Analisis Sentimen Penilaian Masyarakat Terhadap Pejabat Publik Menggunakan Algoritma Naïve Bayes Classifier. *Komputika: Jurnal Sistem Komputer*, 12(1), 93–101. <https://doi.org/10.34010/KOMPUTIKA.V12I1.9638>.
- Contrino, M. F., Reyes-Millán, M., Vázquez-Villegas, P., & Membrillo-Hernández, J. (2024). Using an adaptive learning tool to improve student performance and satisfaction in online and face-to-face education for a more personalized approach. *Smart Learning Environments*, 11(1), 6. <https://doi.org/0.1186/s40561-024-00292-y>.
- Guo, S., Zheng, Y., & Zhai, X. (2024). Artificial intelligence in education research during 2013–2023: A review based on bibliometric analysis. *Education and Information Technologies*, 29(13), 16387–16409. <https://doi.org/10.1007/s10639-024-12491-8>.
- Haydar, A., Priharsari, D., & Wicaksono, S. A. (2022). Analisis Learning Management System terhadap Pengalaman Mahasiswa pada Pembelajaran Jarak Jauh (Studi Kasus: FILKOM UB). *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 6(10), 4638–4645.
- Jia, H. (2024). A study on evaluation of english hybrid teaching courses based on AHP and

- K-means. *PeerJ Computer Science*, 10(e2074). <https://doi.org/10.7717/peerj-cs.2074>.
- Niu, M. (2024). Design and Application of the DPC-K-Means Clustering Algorithm for Evaluation of English Teaching Proficiency. *International Journal of Advanced Computer Science & Applications*, 15(8). <https://doi.org/10.14569/ijacsa.2024.0150851>.
- Simarmata, J. E. (2020). Pemanfaatan Aplikasi Geogebra Dalam Pembelajaran Kalkulus I Pada Mahasiswa Program Studi Pendidikan Matematika Universitas Timor. *MES: Journal of Mathematics Education and Science*, 6(1), 40–47. <https://doi.org/10.30743/mes.v6i1.2624>.
- Simarmata, J. E., & Chrisinta, D. (2024). Implementation of Completely Randomized Design (CRD) using R Software to Evaluate Linear Algebra Learning Systems Pasca the Covid-19 Pandemic. *SITEKIN: Jurnal Sains, Teknologi Dan Industri*, 22(1), 66–74. <https://doi.org/10.24014/sitekin.v22i1.22226>.
- Simarmata, J. E., Chrisinta, D., & Purnomo, M. (2024). Implementation of K-Means Clustering to Human Development Indicators in East Nusa Tenggara. *Journal of Research in Mathematics Trends and Technology*, 6(2), 46–56. <https://doi.org/10.32734/jormtt.v6i2.17066>
- Sitinjak, E. K. (2022). Penggunaan Video Pembelajaran untuk Meningkatkan Kemampuan Pemecahan Masalah Fisika. *Jurnal Penelitian Dan Pengembangan Pendidikan*, 6(1), 19–25. <https://doi.org/10.23887/jppp.v6i1.45006>
- Wang, Z., Sun, Y., Cao, Y., Yang, J., Shi, W., Zhang, A., & Yang, X. (2024). Adaptive Hierarchical Clustering Based Student Group Exercise Recommendation via Multi-objective Evolutionary Method. *International Conference on Neural Computing for Advanced Applications*, 186–200. https://doi.org/10.1007/978-981-97-7001-4_14.
- Wijayanti, R. A., Rahwanda, D., Efendi, Y., & Syaputri, W. (2024). Challenging The Students by Using Video Movie to Improve Students' Vocabulary on Independent Curriculum. *ELT-Lectura*, 11(1), 13–24. <https://doi.org/10.31849/elt-lectura.v11i1.16396>.
- Yu, X. (2025). A method for evaluating the learning effectiveness of MOOC English online education based on fuzzy clustering decision tree. *International Journal of Continuing Engineering Education and Life Long Learning*, 35(1–2), 46–61. <https://doi.org/10.1504/IJCEELL.2025.143798>.
- Zhang, D. (2025). Quality Evaluation of College English Classroom Teaching based on K-Means Clustering Algorithm. *International Conference on Intelligent Systems and Computational Networks (ICISCN)*, 1(1), 1–6. <https://doi.org/10.1109/ICISCN64258.2025.10934355>.