



Clustering Analysis Using the K-Means Method to Identify Alumni Satisfaction Pattern

Analisis Clustering Menggunakan Metode K-Means untuk Mengidentifikasi Pola Kepuasan Alumni

**Muhammad Ilham Ramadhan¹, Alwis Nazir^{2*}, Muhammad Irsyad³,
Suwanto Sanjaya⁴, Fadhilah Syafria⁵**

^{1,2,3,4,5}Program Studi Teknik Informatika, Fakultas Sains dan Teknologi,
Universitas Islam Negeri Sultan Syarif Kasim Riau, Indonesia

E-Mail: ¹12150114705@students.uin-suska.ac.id, ²alwis.nazir@uin-suska.ac.id,
³irsyadtech@uin-suska.ac.id, ⁴suwantosanjaya@uin-suska.ac.id, ⁵fadhilah.syafria@uin-suska.ac.id

Received Nov 19th 2025; Revised Dec 28th 2025; Accepted Jan 04th 2025; Available Online Jan 05th 2026

Corresponding Author: Alwis Nazir

Copyright ©2026 by Authors, Published by Institut Riset dan Publikasi Indonesia (IRPI)

Abstract

Tracer studies play an important role in evaluating the quality of educational services based on alumni experiences. Analyses of alumni satisfaction with learning facilities are generally still limited to descriptive statistics, and therefore have not been able to reveal segmented satisfaction patterns in large-scale data. This study aims to identify the segmentation patterns of alumni satisfaction with learning facilities at Sultan Syarif Kasim State Islamic University Riau (UIN Suska Riau) as an important indicator in evaluating the quality of educational services. The method used is K-Means Clustering, implemented through the stages of Knowledge Discovery in Database (KDD) on 6,508 tracer study data of undergraduate (S1) alumni who graduated between 2010 and 2023. The preprocessing stage includes numerical data normalization using Min-Max Scaling to equalize the scale of six satisfaction indicators (Library, Information Technology, Learning Modules, Classroom Facilities, Laboratories, and Variety of Courses), thereby minimizing bias in the calculation of Euclidean distance. Based on the Elbow Method, the optimal number of clusters obtained is K=3, and the quality of clustering is validated using a Davies-Bouldin Index (DBI) value of 0.874, confirming the stability of the formed clusters. The analysis yields three distinct clusters: Cluster 0 (High Satisfaction Level), Cluster 1 (Low Satisfaction Level), and Cluster 2 (Very High Satisfaction Level). These results provide explicit satisfaction segmentation as a basis for universities to formulate targeted and sustainable facility improvement strategies.

Keywords: Alumni Satisfaction, Data Mining, K-Means Algorithm, Learning Facilities, Tracer Study

Abstrak

Tracer study berperan penting dalam mengevaluasi kualitas layanan pendidikan berdasarkan pengalaman alumni. Analisis kepuasan alumni terhadap fasilitas pembelajaran umumnya masih terbatas pada statistik deskriptif, sehingga belum mampu mengungkap pola kepuasan secara tersegmentasi pada data berskala besar. Penelitian ini bertujuan untuk mengidentifikasi pola segmentasi kepuasan alumni terhadap fasilitas pembelajaran di Universitas Islam Negeri Sultan Syarif Kasim Riau (UIN Suska Riau) sebagai indikator penting dalam evaluasi kualitas layanan pendidikan. Metode yang digunakan adalah K-Means Clustering, diimplementasikan melalui tahapan Knowledge Discovery in Database (KDD) pada 6.508 data tracer study alumni S1 lulusan 2010–2023. Proses preprocessing mencakup normalisasi data numerik menggunakan Min-Max Scaling untuk menyamakan skala enam indikator kepuasan (Perpustakaan, Teknologi Informasi, Modul Belajar, Ruang Belajar, Laboratorium, dan Variasi Mata Kuliah), sehingga meminimalkan bias dalam perhitungan jarak Euclidean. Berdasarkan Elbow Method, diperoleh jumlah kluster optimal adalah K=3, dan kualitas pengelompokan divalidasi dengan nilai Davies-Bouldin Index (DBI) sebesar 0,874, mengonfirmasi stabilitas kluster yang terbentuk. Analisis menghasilkan tiga kluster berbeda: Kluster 0 (Tingkat Kepuasan Tinggi) yang dominan, Kluster 1 (Tingkat Kepuasan Rendah), dan Kluster 2 (Tingkat Kepuasan Sangat Tinggi). Hasil ini memberikan segmentasi kepuasan yang eksplisit sebagai dasar bagi universitas untuk merumuskan strategi peningkatan fasilitas secara terarah dan berkelanjutan.

Kata Kunci: Algoritma K-Means, Data Mining, Fasilitas Pembelajaran, Kepuasan Alumni, Tracer Study

1. PENDAHULUAN

Tracer study merupakan sebuah survei penelitian yang diselenggarakan institusi pendidikan tinggi untuk mendapatkan data tentang para lulusan alumni, mulai dari saat selesai menempuh pendidikan hingga saat mereka mengisi survei tersebut. Informasi yang diperoleh dari *tracer study* dimanfaatkan guna mengevaluasi kualitas pelayanan pendidikan dan hubungannya dengan dunia kerja. *Tracer study* menjadi satu di antara tugas yang harus dilaksanakan oleh perguruan tinggi untuk mengetahui bermacam aspek alumni butuhkan, mulai dari informasi pribadi saat masih menjadi mahasiswa, kualitas pelayanan selama masa studi, lama waktu mencari pekerjaan, kesinkronan antara program studi dengan pekerjaan yang hendak diambil, serta informasi mengenai status pekerjaan yang diperlukan [1].

Perguruan tinggi berperan krusial dalam mempersiapkan mahasiswanya menjadi tenaga kerja yang ahli serta dapat memberikan sumbangsih untuk kemajuan perekonomian dan masyarakat. Namun, satu di antara kendala yang lembaga pendidikan hadapi ialah kesulitan dalam memahami tingkat kepuasan alumni setelah mereka lulus. Kepuasan alumni merupakan indikator penting dalam mengevaluasi kualitas pendidikan yang diberikan, relevansi kurikulum yang diterapkan, dan kesiapan lulusan untuk memasuki dunia kerja. Dengan demikian, perguruan tinggi wajib melakukan evaluasi terhadap fasilitas pembelajaran yang ada seperti perpustakaan, ruang belajar, teknologi informasi, dan laboratorium untuk meningkatkan kualitas layanan yang disediakan [2].

Pengukuran kepuasan alumni terhadap fasilitas pembelajaran masih dianalisis menggunakan pendekatan statistik deskriptif dan penyajian grafik, yang cenderung hanya menghasilkan gambaran rata-rata dan belum mampu mengungkap heterogenitas kepuasan alumni secara mendalam. Keterbatasan ini berpotensi menyebabkan pengelola kampus memperoleh informasi yang bersifat umum dan kurang tepat sasaran dalam merumuskan strategi peningkatan fasilitas. Meskipun data kepuasan alumni yang digunakan tidak bersifat realtime, analisis dan histori tetap relevan karena bertujuan mengidentifikasi pola dan struktur kepuasan alumni dalam jangka panjang. Pendekatan data mining menjadi peran penting karena mampu mengekstraksi pola dan membentuk segmentasi kepuasan alumni yang tidak dapat diperoleh secara optimal melalui metode statistik konvensional. Tujuan dari studi ini yakni guna mengidentifikasi pola segmentasi kepuasan alumni terhadap fasilitas pembelajaran di Universitas Islam Negeri Sultan Syarif Kasim Riau (UIN Suska Riau) sebagai indikator penting dalam evaluasi kualitas layanan pendidikan.

Pemilahan informasi yang berharga dari himpunan data berukuran besar dilakukan melalui serangkaian tahapan sistematis, yang dikenal sebagai data mining [3]. Data mining adalah satu di antara fase dalam proses dari *Knowledge Discovery in Database* (KDD). Pendekatan ini memfasilitasi proses penyaringan data secara menyeluruh untuk menghasilkan wawasan yang bermakna dan mudah diaplikasikan [4]. Dalam mekanisme KDD, sejumlah konsep dan teknik diaplikasikan secara bertahap guna memperoleh informasi yang relevan, meliputi pembersihan data, penggabungan data, pemilihan data, transformasi data, proses penambangan data, evaluasi pola, hingga penyajian pengetahuan [5]. Data mining juga berperan krusial dalam *tracer study*, khususnya untuk menggali pola-pola tersembunyi dalam pengalaman alumni. Data yang dianalisis pada studi ini bersumber dari hasil survei *tracer study* atau kuesioner yang dihimpun dari alumni. Adapun satu di antara teknik penambangan data yang dimanfaatkan dalam penelitian ini adalah algoritma *K-Means*, yang digunakan untuk mengelompokkan data berdasarkan karakteristik kesamaan tertentu.

Algoritma *K-Means* menjadi satu di antara algoritma *clustering* yang dimanfaatkan dalam pengelompokkan data ke dalam sejumlah kelompok homogen yang dilakukan melalui pengukuran kedekatan terhadap titik pusat tertentu [6]. Pendekatan ini bekerja dengan cara mengoptimalkan posisi pusat kluster (*centroid*) melalui proses berulang, sehingga jarak antara setiap data dan pusat kelompoknya menjadi seminimal mungkin hingga struktur kluster yang stabil terbentuk [7]. Adapun satu di antara algoritma yang digolongkan pada metode *unsupervised learning* atau pendekatan pembelajaran tanpa pengawas, ialah *K-Means clustering*. *K-Means clustering* tidak membutuhkan informasi kategori awal atau label yang telah ditentukan sebelumnya. Proses pengelompokan sepenuhnya didasarkan pada kesamaan karakteristik antar data. Dalam praktiknya, metode ini termasuk ke dalam teknik *clustering* non-hirarki, di mana sekumpulan data dikelompokkan ke dalam satu atau lebih kluster yang saling terpisah. Keserupaan atribut menjadi dasar penempatan data dalam kelompok yang sama, sementara perbedaan karakteristik mendorong pemisahan ke dalam kluster lain. Dengan demikian, tingkat variasi pada setiap kluster cenderung rendah, sedangkan perbedaan antar kluster menjadi lebih jelas dan terdefinisi [8].

Algoritma *K-Means* membagi data ke dalam sejumlah himpunan yang relatif homogen dilakukan dengan menempatkan objek yang memiliki kemiripan atribut ke dalam satu kelompok yang sama, sementara objek dengan ciri yang tidak sepadan dialihkan ke kelompok lain [9]. Pemilihan metode ini didasarkan pada ketepatan klasifikasi yang tinggi, kemudahan penerapan, serta efisiensinya dalam mengolah himpunan data berukuran besar. Keunggulan lainnya adalah fleksibilitasnya, yang memungkinkan pengguna untuk menentukan jumlah kluster sesuai dengan kebutuhan [10]. Beragam persoalan telah ditangani menggunakan *K-Means*, salah satunya ditunjukkan dalam penelitian Joko Sutrisno, Arief Wibowo, dan Bayu Satria Pratama tahun 2023 yang mengkaji klusterisasi data *tracer study* lulusan perguruan tinggi terkait aspek karier serta

pekerjaannya. Hasil analisis tersebut memperlihatkan terbentuknya beberapa kluster lulusan dengan kualitas pemisahan yang baik, ditunjukkan oleh angka 0,287 selaku nilai *Davies-Bouldin Index* (DBI) pada percobaan pertama serta angka 0,291 pada percobaan keduanya [11]. Algoritma *K-Means* dimanfaatkan guna mengetahui tingkat kepuasan mahasiswa terhadap layanan akademik di STMIK Pringsewu berdasarkan lima kriteria dari model *ServQual*. Berdasarkan hasil penelitian, nilai uji sensitivitas metode algoritma *K-Means* adalah 77,00% untuk *variable Tangiable*, 78,83% untuk *Reliability*, 86,49% untuk *Responsiveness*, 70,07% untuk *Assurance* dan 79,19% untuk *Empathy* [12]. Dalam penelitian ini, perhitungan K-Means dilakukan menggunakan perangkat lunak *GoogleColab*, dengan DBI menjadi parameter guna mengevaluasi kuliatas cluster yang terbentuk.

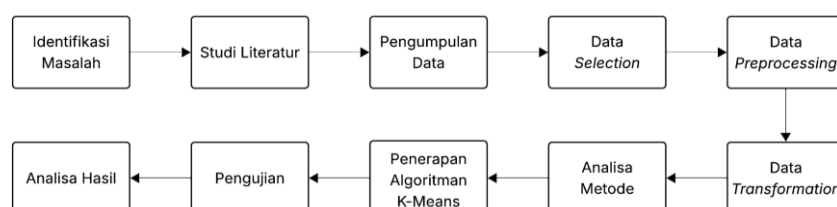
Melihat sejauh mana kepuasan alumni terhadap suatu *tracer study* dapat dilakukan dengan menerapkan pendekatan analisis *clustering*, seperti studi yang mengkaji analisis *clustering* terhadap kelulusan Mahasiswa di Unisba dengan memanfaatkan metode algoritma *K-Means*. Dalam studi ini mendapatkan hasil pengelompokkan yang membentuk 3 kluster dengan karakteristik yang berbeda. Kluster 1 terdiri dari 8 mahasiswa dengan menyelesaikan studi lebih cepat dengan IPK yang tinggi. Kluster 2 berjumlah 37 mahasiswa yang menyelesaikan studi dalam waktu normal dengan IPK yang sedang dan diprediksi lulus tepat waktu. Sementara itu, Kluster 3 berisi 13 mahasiswa dengan menyelesaikan studi lebih dari 4 tahun dengan IPK yang lebih rendah dan diprediksi lulus dengan hasil yang sangat baik. Hasil tingkat akurasi 96,55% dan tingkat error 3,45%. Hasil tersebut memperlihatkan bahwasanya guna memprediksi kelulusan mahasiswa berdasarkan lamanya masa studi dan nilai IPK, maka penggunaan algoritma K-Means dinilai cukup efektif, selain itu, algoritma ini membantu pihak kampus dalam merencanakan bentuk dukungan yang lebih tepat untuk mahasiswa, sehingga dapat menyelesaikan studi tepat waktu [13].

Analisis *clustering* dapat digunakan sebagai pendekatan untuk mengukur tingkat kepuasan alumni dalam mengisi *tracer study*, seperti pada penilaian yang menilai kepuasan siswa di SMA Swasta Bani Adam AS. Studi tersebut mendapatkan 3 kluster penilaian alumni dimana kluster 1 menjadi tingkat kepuasan yang terbanyak, di mana metode *clustering* dengan memanfaatkan algoritma *K-Means* yang diterapkan pada data kepuasan mahasiswa menjadi metode yang diterapkan pada studi ini. Hasil dari penelitian ini memperlihatkan bahwasanya terdapat sejumlah 86 siswa yang kurang puas, 69 siswa yang cukup puas, serta 45 siswa yang puas dari total siswa sebanyak 200 [14]. Penelitian yang dilakukan oleh Ayunia Khaerunnisa pada Tahun 2022 dan penelitian yang dilakukan oleh Juliansyah Putra Tanjung, Bayu Angga Wijaya, dan Muhammad Ridho pada tahun 2023 dengan menggunakan metode *clustering* untuk mengelompokkan data kelulusan mahasiswa dan kepuasan mahasiswa berdasarkan berbagai kriteria. Namun penelitian yang menggunakan metode *clustering* untuk menganalisis tingkat kepuasan alumni terhadap fasilitas pembelajaran ini masih terbatas.

Fokus dari studi ini ialah kepada penerapan *K-Means clustering* dalam kerangka KDD guna mengidentifikasi pola segmentasi kepuasan alumni UIN Suska Riau terhadap fasilitas pembelajaran berdasarkan data *tracer study*, sebagai upaya mengisi keterbatasan penelitian *tracer study* dan kepuasan alumni yang umumnya masih bersifat deskriptif serta menyediakan segmentasi kepuasan yang terukur dan tervalidasi sebagai dasar perumusan kebijakan peningkatan fasilitas pembelajaran, dengan memperhatikan bahwa alumni yang mengisi *tracer study* ini adalah mereka yang telah mengalami dan memanfaatkan fasilitas pembelajaran tersebut selama masa studi sebagai mahasiswa di UIN Suska Riau.

2. METODE PENELITIAN

Bagian metodologi dalam sebuah penelitian memiliki peran penting untuk menjelaskan secara rinci terkait prosedur yang dapat diterapkan guna menyelesaikan persoalan yang diangkat dalam penelitian tersebut. Penjabaran ini mencakup proses-proses yang sistematis dan terstruktur, mulai dari proses mengumpulkan data, menganalisis data, sampai menyajikan hasil. Urutan langkah-langkah tersebut dapat divisualisasikan dengan jelas untuk memudahkan pemahaman, sebagaimana ditampilkan pada Gambar 1.



Gambar 1. Metodologi Penelitian

2.1. Identifikasi Masalah

Tracer study di UIN Suska Riau telah dilaksanakan dengan tujuan mengumpulkan data mengenai pengalaman serta tingkat kepuasan alumni. Namun, hasil yang diperoleh belum dimanfaatkan secara maksimal untuk mengidentifikasi pola kepuasan terhadap fasilitas pembelajaran. Data yang ada masih dalam

bentuk mentah dan belum dianalisis menggunakan metode yang mampu mengungkapkan informasi secara komprehensif. Oleh karena itu, implementasi metode *K-Means* diperlukan untuk mengelompokkan alumni berdasarkan tingkat kepuasan selama menjadi mahasiswa, sehingga universitas dapat memperoleh pola kepuasan secara lebih terstruktur dan menjadikannya sebagai dasar dalam upaya peningkatan kualitas fasilitas pembelajaran.

Atribut yang digunakan dalam penelitian ini merupakan indikator kepuasan alumni terhadap fasilitas pembelajaran yang diperoleh dari *tracer study* alumni UIN Suska Riau. Data tersebut dikumpulkan secara resmi, adapun atribut yang dianalisis meliputi kepuasan terhadap perpustakaan, teknologi informasi dan komunikasi, modul belajar, ruang belajar, laboratorium, dan variasi mata kuliah yang ditawarkan. Dengan menggunakan atribut tersebut, penelitian ini mampu menangkap dimensi kepuasan alumni secara komprehensif dan relevan untuk proses pengelompokkan menggunakan metode *K-Means*.

2.2. Studi Literatur

Studi literatur menjadi tahap guna mengumpulkan sumber-sumber teori yang dapat mendukung studi ini. Studi literatur pada studi ini difokuskan pada teori *data mining*, algoritma *K-Means*, dan *tracer study*. Ketiga konsep tersebut digunakan sebagai dasar dalam menganalisis dan mengelompokkan tingkat kepuasan alumni UIN Suska Riau terhadap fasilitas pembelajaran.

2.3. Pengumpulan Data

Landasan utama dalam kajian yang berfokus pada analisis data dibangun melalui ketersediaan informasi yang memadai, sehingga tahap awal penelitian diarahkan pada proses pengumpulan data yang sistematis [15]. Adapun sumber informasi pada studi ini diperoleh dari hasil *tracer study* sejumlah 7.727 data alumni lulusan tahun 2010 hingga 2023 dan setelah melalui proses pembersihan data tersisa 6.508 data yang layak olah. Data tersebut mencakup berbagai aspek penilaian alumni terhadap fasilitas pembelajaran, meliputi perpustakaan, teknologi informasi dan komunikasi, modul belajar, ruang belajar, laboratorium, serta variasi mata kuliah yang ditawarkan. Data yang dikumpulkan bersifat kuantitatif dengan skala penilaian Likert 1–5, di mana skor 1 mencerminkan tingkat ketidakpuasan paling rendah dan skor 5 mencerminkan tingkat kepuasan paling tinggi. Kumpulan data ini kemudian dimanfaatkan sebagai pijakan analitis dalam penerapan algoritma *K-Means clustering* guna mengidentifikasi pola kepuasan alumni.

2.4. Data Selection

Penyaringan informasi yang relevan dilakukan sebagai tahap awal dalam kerangka KDD dengan tujuan memastikan bahwa hanya data operasional yang memiliki keterkaitan langsung dengan tujuan analisis yang dipertahankan. Informasi terpilih pada fase ini berfungsi sebagai pondasi bagi proses data mining, sehingga elemen yang tidak berkontribusi terhadap analisis dikeluarkan dari pengolahan [16]. Dalam penelitian ini, proses data *selection* dilakukan dengan menyeleksi variabel yang berkaitan langsung dengan kepuasan alumni terhadap fasilitas pembelajaran. *Dataset* awal *tracer study* memuat atribut seperti nama, program studi, fakultas, dan tahun lulus. Atribut tersebut tidak digunakan karena bersifat deskriptif dan tidak berpengaruh terhadap penghitungan jarak pada algoritma *K-Means*. Hanya menggunakan enam indikator penilaian fasilitas pembelajaran seperti perpustakaan, teknologi informasi dan komunikasi, modul belajar, ruang belajar, laboratorium, serta variasi mata kuliah yang seluruhnya menggunakan skala Likert penilaian lima tingkat. Hasil seleksi tersebut melahirkan subset data yang siap digunakan untuk tahapan selanjutnya.

Enam indikator fasilitas pembelajaran yang digunakan dipilih karena merepresentasikan aspek utama yang membentuk pengalaman belajar. Seluruh indikator diukur secara numerik menggunakan skala Likert yang seragam, sehingga sesuai untuk analisis berbasis jarak pada algoritma *K-Means clustering*. Dengan menggunakan indikator yang relevan, hasil pengelompokan diharapkan benar-benar mencerminkan perbedaan tingkat kepuasan alumni terhadap fasilitas pembelajaran.

2.5. Data Preprocessing

Setelah pengumpulan *dataset* dan data *selection*, langkah selanjutnya adalah tahap *preprocessing*. Dalam proses ini, data dibersihkan dari data duplikat, nilai kosong, atau ketidakkonsistenan [17]. *Preprocessing* bertujuan untuk mempersiapkan data supaya dapat dipakai dalam proses pengelompokan yang memanfaatkan algoritma *K-Means*. Dalam penelitian ini, *preprocessing* secara khusus meliputi penghapusan data duplikat atau berulang, penghapusan kolom-kolom tidak relevan untuk analisis *clustering* berbasis numerik, yaitu nama, angkatan, fakultas, dan program studi, karena kolom tersebut bersifat deskriptif.

2.6. Data Transformation

Proses pengubahan data mentah menjadi bentuk yang diharapkan sehingga bisa diproses lebih lanjut dikenal sebagai transformasi data. Dalam proses ini, data yang awalnya berupa huruf diubah menjadi bentuk angka [18]. Transformasi fitur juga melibatkan rekonstruksi fitur dari dimensi tinggi menuju yang lebih rendah guna mempermudah proses penganalisisan lanjutan [19]. Di tahapan ini, diselenggarakan proses

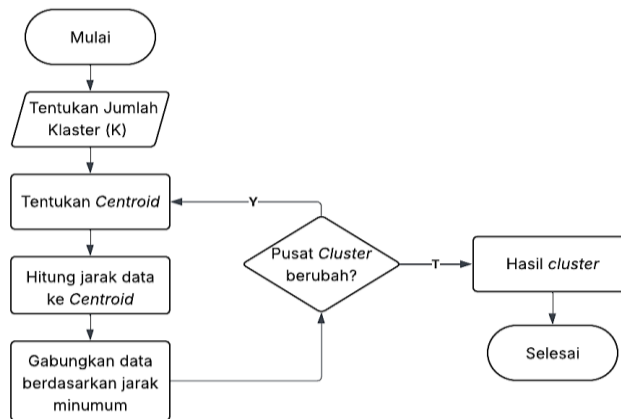
normalisasi data numerik untuk menyamakan skala antar figur agar proses *clustering* lebih akurat. Meskipun semua variabel berada dalam rentang 1-5, normalisasi tetap dilakukan untuk menghindari dominasi fitur tertentu dalam perhitungan jarak. Metode yang dimanfaatkan yaitu *Min-Max Scaling* yaitu dengan melakukan perubahan nilai data ke angka 0 sampai 1 menggunakan rumus. Rumus normalisasi *Min-Max Scaling* yang digunakan:

$$X_{Scaled} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (1)$$

2.7. Analisa Metode

K-Means adalah model *unsupervised learning* untuk mengolah data yang belum pernah diberi label atau kategori. Model ini mendeteksi pola-pola yang kerap dimunculkan oleh data serta memberi respon yang seragam terhadap pola-pola tersebut pada setiap titik data [20]. Algoritma *K-Means* diawali dengan menetapkan total kluster yang diharapkan, lalu memperbarui posisi titik pusat (*centroid*) secara berulang berdasarkan data yang tergabung dalam setiap kluster. Proses ini bertujuan untuk meminimalkan variasi data di dalam kluster sehingga menghasilkan pembagian data yang lebih konsisten dan representatif [21].

Algoritma *K-Means* dipilih karena memiliki kompleksitas komputasi rendah, dan mampu menangani *dataset* numerik besar secara efisien dibandingkan beberapa metode *clustering* lainnya. Selain itu, dengan pendekatan berbasis *centroid*, *K-Means* menghasilkan pusat kluster yang jelas sehingga mempermudah interpretasi karakteristik setiap kelompok dalam konteks indikator yang digunakan dalam studi ini [22]. Algoritma ini dinilai paling sesuai untuk tujuan segmentasi kepuasan alumni secara kuantitatif dan terukur. Proses penerapan metode *K-Means* disajikan secara terstruktur dalam *flowchart* pada Gambar 2.



Gambar 2. Flowchart Metode *K-Means*

Berikut uraian penjelasan dari *flowchart* diatas [23]:

1. Tentukan jumlah kluster (K). Jumlah kluster ditetapkan berdasarkan hasil *Elbow Method* dan analisis data *tracer study*, dengan nilai optimal K = 3 yang mewakili tingkat kepuasan tinggi, sedang, dan rendah
2. Menentukan titik pusat awal (*centroid*) yang dipilih dari *dataset* secara acak.
3. Mengukur jarak data ke *centroid* dengan memanfaatkan *Euclidean Distance* berikut:

$$d(x_i, c_j) = \sqrt{\sum_{k=1}^n (x_{ik} - c_{jk})^2} \quad (2)$$

Dalam rumus tersebut, x_i adalah data ke-i, dan c_j merupakan *centroid* dari kluster ke-j.

4. Mengelompokkan data ke dalam kluster dengan jarak terkecil terhadap *centroid* menggunakan prinsip jarak minimum (*min distance rule*).
5. Menghitung *centroid* sebagai rerata anggota dalam tiap kluster sesudah seluruh data terkelompokkan.

$$c_j = \frac{1}{n_j} \sum_{x_i \in C_j} x_i \quad (3)$$

Proses tersebut dilakukan berulang sampai tidak ada lagi perubahan pada posisi *centroid* (konvergen)

6. Jika *centroid* masih berubah, maka proses perhitungan jarak dan pengelompokkan data akan diulang kembali, namun jika *centroid* sudah konvergen (tidak berubah), maka proses iterasi dihentikan karena hasil pengelompokkan dianggap stabil dan optimal.

7. Setelah proses iterasi selesai, hasil yang diperoleh adalah pembagian data ke dalam klaster (k) yang telah terbentuk.

Sebagai perbandingan, studi yang dilakukan oleh Hasnah Vithon Carelsa, Rio Andika Malik, dan Dwipa Junika Putra pada tahun 2023 yang menerapkan algoritma *K-Means Clustering* untuk menganalisis kepuasan mahasiswa terhadap layanan kantin di kampus menunjukkan bahwa metode ini berhasil dalam mengelompokkan responden berdasarkan variasi pola kepuasan yang berbeda [24]. Penelitian ini mengambil beberapa indikator layanan serta tahapan KDD untuk mengidentifikasi kelompok mahasiswa dengan tingkat kepuasan yang bervariasi.

Hasil temuan ini sejalan dengan hasil dari penelitian ini, di mana alumni juga terkelompok ke dalam beberapa klaster dengan karakteristik kepuasan yang berbeda terhadap fasilitas pembelajaran. Perbedaannya terletak pada objek penelitian, di mana studi penelitian Hasnah Vithon Carelsa, Rio Andika Malik, dan Dwipa Junika Putra lebih berfokus pada layanan operasional sehari-hari di kantin, sedangkan penelitian ini berfokus pada fasilitas pembelajaran sebagai komponen utama dari mutu pendidikan tinggi. Dengan demikian, penerapan *K-Means* dalam kedua penelitian ini menunjukkan konsistensi kesesuaian bahwa algoritma tersebut mampu mengidentifikasi pola kepuasan pengguna layanan pendidikan secara kuantitatif dan mendukung pengambilan keputusan berbasis data.

3. HASIL DAN PEMBAHASAN

Hasil penerapan metode yang telah dijelaskan di bagian sebelumnya serta analisis proses sekaligus penilaian dalam penelitian diterangkan pada bagian ini. Fokus utamanya adalah menggunakan algoritma *K-Means* untuk mengelompokkan tingkat kepuasan para alumni terhadap fasilitas pembelajaran di UIN Suska Riau dengan berlandaskan data hasil *tracer study*. Setiap tahap penelitian dijelaskan secara terstruktur, mulai dari analisis kebutuhan data, proses seleksi dan prapemrosesan, tahap transformasi, penerapan algoritma *K-Means*, hingga evaluasi dan analisis hasil pengelompokan, untuk memperoleh pola kepuasan alumni yang dapat dijadikan landasan pertimbangan strategis bagi pihak universitas dalam mengambil keputusan.

3.1. Analisis Kebutuhan Data

Data yang dimanfaatkan pada studi ini diambil dari hasil *tracer study* terhadap lulusan program S1 UIN Suska Riau. *Dataset* mencakup seluruh fakultas dan program studi. Variabel yang digunakan dalam penelitian meliputi aspek fasilitas pembelajaran, yaitu Perpustakaan (A), Teknologi Informasi dan Komunikasi (B), Modul Belajar (C), Ruang Belajar (D), Laboratorium (E), Variasi Mata Kuliah yang Ditawarkan (F). Data dikumpulkan dalam format Microsoft Excel dan diolah menggunakan bahasa pemrograman *Python* melalui *Google Colab*. Tabel 1 menampilkan data yang peneliti gunakan pada studi ini.

Tabel 1. Data Kepuasan Alumni

No	Nama	Angkatan	Fakultas	Prodi	A	B	C	D	E	F
1	Anindya Nanda Rozana	2015	FST	TIF	4	4	4	4	4	4
2	Ridho Darmawan	2015	FST	TIF	4	4	4	3	3	3
3	Muhammad Irfan	2015	FST	TIF	5	4	4	2	3	4
4	Kiki Fatmala Sari	2014	FST	TIF	4	4	4	4	3	4
5	Rifki Zuhdi.Ar	2014	FST	TIF	1	1	1	1	1	1
...
6504	Manahara Alamsyah Lubis	2018	FUSH	Ilmu Hadis	4	4	4	4	4	5
6505	Sepriadi	2017	FUSH	Ilmu Hadis	5	5	5	4	5	5
6506	Fachruli Isra Rukmana	2018	FUSH	Ilmu Hadis	4	4	4	4	4	4
6507	Husniatul Aulia	2017	FUSH	Ilmu Hadis	5	5	5	5	5	5
6508	Indah Nurrahmi	2017	FUSH	Ilmu Hadis	4	4	4	4	4	4

3.2. Data Selection

Tahapan ini bertujuan memilih atribut yang penting untuk mencapai tujuan analisis. Dari data *tracer study*, hanya digunakan variabel numerik (A-F) yang langsung berkaitan dengan fasilitas pembelajaran. Atribut seperti nama, angkatan, fakultas, dan program studi tidak digunakan karena bersifat deskriptif dan tidak berpengaruh terhadap hasil klaster. Tabel 2 merupakan *dataset* setelah melalui proses seleksi.

3.3. Data Preprocessing

Tahap data *preprocessing* dilakukan untuk memastikan data konsistensi *dataset* sebelum dianalisis. Proses ini meliputi penghapusan data duplikat atau berulang, validasi tipe data agar seluruh variabel indikator berada dalam format numerik, serta pemeriksaan konsistensi nilai untuk memastikan seluruh indikator berada dalam rentang skala Likert 1-5. Tahapan ini menghasilkan data yang bersih, terstruktur, dan siap diproses pada tahap transformasi dan *clustering*.

3.4. Data Transformation

Transformasi data menjadi langkah krusial dalam kerangka KDD untuk memastikan data siap diproses oleh algoritma *clustering*. Dalam penelitian ini, transformasi data dilakukan dengan memanfaatkan metode *Min-Max Scaling* guna proses normalisasinya. Normalisasi ini bertujuan menyamakan skala numerik pada seluruh variabel indikator kepuasan alumni.

Meskipun seluruh variabel fasilitas pembelajaran (A-F) sudah menggunakan skala Likert yang seragam (1-5), normalisasi tetap diperlukan untuk menghilangkan potensi yang bias yang disebabkan oleh perbedaan distribusi *mean* dan variasi antar fitur. *Min-Max Scaling* diterapkan guna mengubah nilai menjadi angka 0 sampai 1, sehingga setiap fitur memiliki bobot yang setara dalam perhitungan jarak *Euclidean*.

Tabel 2. Dataset Setelah Seleksi Data (A-F)

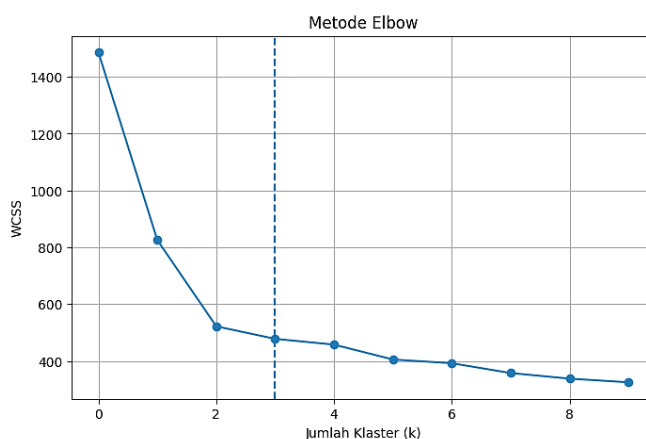
No	A	B	C	D	E	F
1	4	4	4	4	4	4
2	4	4	4	3	3	3
3	5	4	4	2	3	4
4	4	4	4	4	3	4
5	1	1	1	1	1	1
...
6504	4	4	4	4	4	5
6505	5	5	5	4	5	5
6506	4	4	4	4	4	4
6507	5	5	5	5	5	5
6508	4	4	4	4	4	4

3.5. Penerapan Algoritma K-Means

Tahap ini merupakan proses utama penelitian yang bertujuan untuk mengelompokkan Tingkat kepuasan alumni terhadap fasilitas Pembelajaran menggunakan algoritma *K-Means*. Pengelompokan dilakukan berdasarkan enam indikator penilaian utama, yaitu Perpustakaan (A), Teknik Informasi dan Komunikasi (B), Modul Belajar (C), ruang belajar (D), laboratorium (E), dan variasi mata kuliah yang ditawarkan (F), seluruh proses implementasi dilakukan menggunakan *Google Colab* dengan Bahasa pemrograman *Python*.

Sebelum algoritma *K-Means* dijalankan, ada langkah penting yang harus dilakukan yaitu menentukan jumlah kluster yang optimal (K). Dalam penelitian ini, Metode *Elbow method* digunakan sebagai teknik untuk menentukan jumlah kluster yang paling representatif. Cara kerja metode *Elbow* adalah dengan menghitung nilai *Within Cluster Sum of Squares* (WCSS) untuk berbagai variasi jumlah kluster. Nilai WCSS yang semakin kecil mengindikasikan bahwa titik data di dalam kluster semakin dekat dengan *centroid* kluster, yang berarti kualitas *clustering* semakin baik.

Nilai WCSS ini kemudian divisualisasikan dalam sebuah grafik, dimana sumbu X mewakili jumlah kluster (K) dan sumbu Y mewakili nilai WCSS. Titik yang signifikan mengalami penurunan kemudian melandai, yang menyerupai bentuk “siku” atau *elbow*. Berdasarkan analisa grafik *elbow* yang ditunjukkan, penurunan WCSS yang paling signifikan terjadi hingga K=3, setelah itu mulai melandai. Oleh karena itu, jumlah kluster yang ditetapkan dan digunakan dalam penelitian ini adalah tiga kluster. Berikut adalah visualisasi *elbow method* pada Gambar 3 yang menunjukkan titik optimal K untuk pengelompokan data kepuasan alumni.

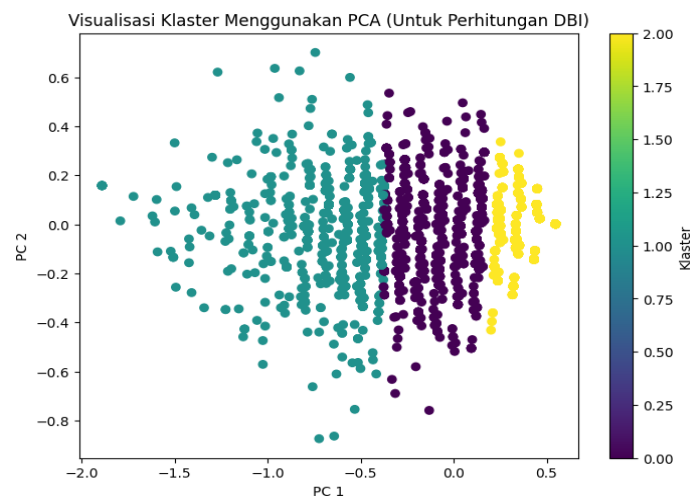


Gambar 3. Grafik *Elbow* Untuk Penentuan Jumlah Kluster

3.6. Pengujian

Setelah penerapan algoritma *K-Means*, langkah selanjutnya adalah melakukan pengujian untuk mengevaluasi seberapa baik hasil *clustering* yang diperoleh. Pengujian ini dilakukan dengan menggunakan metrik evaluasi DBI. DBI merupakan indikator evaluasi dalam analisis *clustering* yang digunakan untuk mengukur tingkat keterpisahan antar kluster, sehingga menunjukkan seberapa jelas batas pemisah yang terbentuk dari masing-masing kluster[25]. DBI mengevaluasi hasil pengelompokan dengan memperhatikan tingkat kekompakan dan jarak setiap kluster dengan kluster terdekatnya[26].

Nilai DBI kurang dari 1 (<1) umumnya dianggap sebagai indikasi kluster yang stabil dan baik. Hasil penelitian menunjukkan nilai DBI sebesar 0,874 menunjukkan kualitas kluster yang baik. Diperkuat oleh visualisasi PCA yang memperlihatkan pemisah antar kluster dan kluster yang terbentuk tidak terlalu saling tumpang tindih. Hal ini semakin menguatkan bahwa *K-Means* berhasil menghasilkan kluster yang representatif untuk data kepuasan alumni. Gambar 4 merupakan hasil visualisasi kluster menggunakan PCA untuk perhitungan DBI:



Gambar 4. Visualisasi Sebaran Kluster Alumni Menggunakan PCA

3.7. Analisa Hasil

Tahapan analisis hasil dilakukan untuk meninjau karakteristik kluster yang terbentuk dan menginterpretasikan pola kepuasan alumni terhadap fasilitas pembelajaran di UIN Suka Riau. Proses ini dilakukan dengan menggabungkan hasil pengelompokan *K-Means*, evaluasi metrik validasi kluster, serta penilaian rata-rata indikator pada setiap kluster. Analisis bertujuan untuk mengidentifikasi segmen alumni berdasarkan Tingkat kepuasan serta merumuskan implikasi strategis bagi pengembang fasilitas Pembelajaran di universitas.

Tabel 3. Jumlah kluster yang terbentuk

Klaster	Jumlah
0	3.175
1	1.112
2	2.221

Berdasarkan Tabel 3 pengelompokan menghasilkan tiga kluster dengan jumlah anggota yaitu, kluster 0=3.175 anggota, Klaster 1=1.112 anggota, Klaster 2=2.221 anggota. Setiap kluster memiliki pola penilaian yang berbeda pada enam indikator fasilitas pembelajaran yaitu, Perpustakaan (A), Teknologi Informasi dan Komunikasi (B), Modul Belajar (C), Ruang Belajar (D), Laboratorium (E), Variasi Mata Kuliah yang Ditawarkan (F). Berdasarkan hasil pengelompokan yang diperoleh, tahap analisis berikutnya dilakukan dengan menyajikan statistik deskriptif tiap indikator kepuasan pada setiap kluster. Penyajian ini dimaksudkan untuk memberikan gambaran kuantitatif mengenai perbedaan karakteristik tingkat kepuasan alumni terhadap fasilitas pembelajaran antar kluster yang terbentuk.

Tabel 4. Statistik Deskriptif Indikator Kepuasan Alumni per Kluster

Indikator	Klaster 0	Klaster 1	Klaster 2
Perpustakaan	4.05	3.09	4.87
Teknologi Informasi dan Komunikasi	3.99	2.99	4.82
Modul Belajar	4.00	3.10	4.86

Indikator	Klaster 0	Klaster 1	Klaster 2
Ruang Belajar	3.89	2.72	4.80
Laboratorium	3.90	2.83	4.80
Variasi Mata Kuliah	4.05	3.39	4.86

Tabel 4 memberikan informasi statistik deskriptif yang menunjukkan nilai rata-rata setiap indikator kepuasan alumni untuk setiap klaster yang dibentuk menggunakan algoritma *K-Means*. Hasil tersebut menunjukkan adanya perbedaan yang signifikan antara kelompok alumni. Klaster 0 menunjukkan tingkat kepuasan yang tinggi dengan nilai rata-rata antara 3.89 hingga 4.05, yang menunjukkan bahwa mayoritas alumni dalam kelompok ini merasa puas dengan fasilitas pembelajaran, terutama terkait dengan aspek perpustakaan, modul belajar, dan variasi mata kuliah. Klaster 1 merupakan kelompok dengan tingkat kepuasan terendah, di mana seluruh indikator memiliki nilai dalam rentang 2.72 hingga 3.39. Nilai terendah terlihat pada indikator teknologi informasi dan komunikasi, ruang belajar, serta laboratorium, yang menandakan adanya kelemahan pada fasilitas fisik dan layanan pendukung pembelajaran. Sementara itu, Klaster 2 menunjukkan tingkat kepuasan yang sangat tinggi dengan nilai rata-rata berada pada kisaran 4.80 hingga 4.87 di semua indikator yang mencerminkan kelompok alumni yang menilai fasilitas pembelajaran sangat baik dan mampu memenuhi ekspektasi akademik mereka. Perbedaan pola nilai di antara klaster ini menegaskan bahwa hasil klusterisasi berhasil membentuk segmentasi kepuasan alumni yang berbeda secara kuantitatif.

Berdasarkan data statistik dalam tabel 4, masing-masing klaster menunjukkan ciri khas yang berbeda dalam hal tingkat terhadap setiap indikator fasilitas pembelajaran. Perbedaan ini akan dijelaskan secara lebih rinci melalui pemaparan profil setiap klaster sebagai berikut ini.

1. Klaster 0

Klaster 0 merupakan kelompok dominan dengan tingkat kepuasan tinggi yang mencakup 3.175 alumni. Nilai rata-rata seluruh indikator berada pada kisaran 3,80–4,05, yang menunjukkan persepsi positif terhadap fasilitas pembelajaran. Aspek perpustakaan, modul belajar, dan variasi mata kuliah memperoleh penilaian tertinggi, menandakan bahwa layanan akademik pada klaster ini telah mampu memenuhi kebutuhan mayoritas alumni.

2. Klaster 1

Klaster 1 terdiri dari 1.112 alumni dengan tingkat kepuasan terendah, ditunjukkan oleh nilai rata-rata indikator pada kisaran 2,70–3,30. Penilaian paling rendah terdapat pada ruang belajar dan laboratorium, yang mengindikasikan keterbatasan pada fasilitas fisik dan sarana pendukung pembelajaran. Kelompok ini mencerminkan segmen alumni yang memerlukan perhatian prioritas dalam peningkatan kualitas fasilitas.

3. Klaster 2

Klaster 2 berjumlah 2.221 alumni dan menunjukkan tingkat kepuasan sangat tinggi, dengan nilai rata-rata seluruh indikator berada pada kisaran 4,80–4,87. Tingginya skor pada semua aspek menggambarkan bahwa fasilitas pembelajaran dinilai sangat baik dan telah memenuhi ekspektasi akademik secara optimal oleh kelompok ini.

Analisa ketiga klaster menunjukkan bahwa setiap kelompok alumni memiliki kebutuhan dan tingkat kepuasan yang berbeda, sehingga universitas perlu menerapkan strategi peningkatan fasilitas yang lebih terarah. Klaster 1 yang berada pada kategori kepuasan rendah menandakan perlu perbaikan pada ruang belajar (D), teknologi informasi dan komunikasi (B), dan laboratorium (E). Klaster 0 memiliki tingkat kepuasan tinggi tetap membutuhkan pemeliharaan dan peningkatan berkelanjutan agar kualitas layanan pembelajaran dapat dipertahankan. Sementara itu, klaster 2 dengan dengan kepuasan sangat tinggi menunjukkan bahwa sebagian fasilitas telah memenuhi ekspektasi alumni, sehingga fokus pengembangan dapat diarahkan pada modernisasi sarana dan inovasi pembelajaran. Secara keseluruhan, hasil *clustering* memberikan dasar yang jelas bagi universitas untuk menetapkan prioritas peningkatan fasilitas berdasarkan kebutuhan masing-masing segmen alumni. Hasil klusterisasi kepuasan alumni dalam penelitian ini konsisten dengan temuan penelitian terdahulu yang menerapkan algoritma *K-Means* pada data *tracer study* dan kepuasan layanan pendidikan, yang umumnya menghasilkan tiga klaster kepuasan alumni [2], [14], sehingga memperkuat validitas metode yang digunakan.

Algoritma *K-Means* telah banyak diterapkan dalam berbagai penelitian sebelumnya, salah satunya pada penelitian yang dilakukan oleh Bagas Parlambang dan Fauziah tahun 2020 [27], yang menilai tingkat kepuasan mahasiswa terhadap kinerja dosen di Universitas Nasional. Hasil dari penelitian ini menunjukkan bahwa klaster dengan nilai rata-rata kepuasan tertinggi adalah klaster 2 yang terdiri dari total 297 titik data, dengan hasil pada aspek reliabilitas itu sebanyak 46 dengan aspek puas. Keunggulan penelitian ini terletak pada fokus analisis yang secara spesifik mengkaji fasilitas pembelajaran serta penggunaan jumlah responden

yang besar, yang memungkinkan pembentukan klaster yang lebih stabil dan representatif. Secara teoritis, temuan ini mendukung konsep kualitas layanan pendidikan, di mana kepuasan alumni mencerminkan kesesuaian antara ekspektasi dan kondisi fasilitas pembelajaran, dengan klaster kepuasan rendah mengindikasikan adanya kesenjangan layanan terutama pada ruang belajar dan laboratorium. Hasil klasterisasi ini dapat digunakan sebagai dasar bagi universitas dalam menentukan prioritas pengembangan fasilitas, di mana kelompok dengan kepuasan rendah menjadi fokus utama perbaikan, sedangkan kelompok dengan kepuasan tinggi perlu dipertahankan melalui pemeliharaan dan peningkatan berkelanjutan. Hasil penelitian ini tidak hanya menggambarkan kondisi kepuasan alumni, tetapi juga memberikan landasan kontekstual dan kebijakan bagi peningkatan kualitas fasilitas pembelajaran.

4. KESIMPULAN

Penelitian ini menyimpulkan bahwa algoritma *K-Means Clustering* berhasil mengidentifikasi pola kepuasan alumni UIN Suska Riau terhadap fasilitas pembelajaran secara efektif dan terstruktur. Implementasi model dilakukan melalui dengan menggunakan kerangka KDD yang terdiri dari tahap normalisasi dengan metode *Min-Max Scaling* untuk menyamakan bobot fitur dan memastikan ketepatan akurasi perhitungan jarak *Euclidean*. Penentuan optimalitas klaster didasarkan pada metode *Elbow* yang menetapkan tiga klaster ($K=3$), sementara kualitas klasterisasi divalidasi dengan nilai DBI sebesar 0,874, yang mengindikasikan bahwa klaster yang terbentuk stabil dan terpisah dengan baik.

Hasil klasterisasi membagi alumni menjadi tiga kelompok berbeda, klaster 0 merupakan kelompok dominan dengan tingkat kepuasan tinggi, menunjukkan keberhasilan fasilitas seperti perpustakaan dan variasi mata kuliah, klaster 1 yang menunjukkan tingkat kepuasan rendah, dengan indikator ruang belajar dan laboratorium sebagai titik kritis yang memerlukan intervensi dan peningkatan prioritas dari pihak universitas, dan klaster 2 yang merepresentasikan tingkat kepuasan sangat tinggi di seluruh indikator sehingga pengembangannya dapat difokuskan pada inovasi fasilitas. Hasil penelitian ini memberikan segmentasi kepuasan yang jelas dan dapat digunakan sebagai dasar perencanaan peningkatan kualitas layanan pembelajaran. Kelemahan penelitian ini terletak pada penggunaan variabel yang terbatas, yaitu hanya enam indikator fasilitas pembelajaran, serta cakupan responden yang dibatasi pada alumni S1. Diharapkan penelitian selanjutnya menambahkan variabel lain yang relevan, memperluas cakupan responden, serta membandingkan *K-Means* dengan metode *clustering* lain atau menggunakan validasi klaster tambahan untuk memperoleh hasil yang lebih komprehensif

REFERENSI

- [1] I. Wayan Supriana, C. Pramarta, and L. Arida Ayu Rahning Putri, "RESISTOR Journal | 1 Aplikasi Pengukur Tingkat Kepuasan Alumni Berdasarkan Data Tracer Study Berbasis Metode Machine Learning", [Online]. Available: <https://s.id/jurnalresistor>
- [2] F. A. Fernaldy, A. A. Arifiyanti, and D. S. Y. Kartika, "Klasterisasi Tracer Study Alumni Universitas Xyz Menggunakan Algoritma K-Means," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 13, no. 1, Jan. 2025, doi: 10.23960/jitet.v13i1.5581.
- [3] S. Asyuti and A. A. Setyawan, "Data Mining Dalam Penggunaan Presensi Karyawan Denga Cluster Means," *Jurnal Ilmiah Sains Teknologi Dan Informasi*, vol. 1, no. 1, pp. 01–10, 2023, [Online]. Available: <https://jurnal.alimspublishing.co.id/index.php/JITI/article/download/6/6>
- [4] I. Virgo, S. Defit, and Y. Yuhandri, "Klasterisasi Tingkat Kehadiran Dosen Menggunakan Algoritma K-Means Clustering," *Jurnal Sistim Informasi dan Teknologi*, pp. 23–28, Mar. 2020, doi: 10.37034/jsisfotek.v2i1.17.
- [5] W. Handayani, Y. Yuneфри, D. Arif, P. Studi Teknik Informatika Fakultas Ilmu Komputer Universitas Lancang Kuning, and J. K. Yos Sudarso, "Optimalisasi K-Means Clustering Berbasis Web dalam Analisis Kepuasan Masyarakat terhadap Korem 031/Wira Bima," 2024.
- [6] M. Indrayani, M. Iqbal, and D. Nasution, "Pemetaan Pilihan Lulusan Smk Panca Budi Medan Menggunakan Algoritma K-Means Dan Visualisasi Data," 2025. [Online]. Available: <http://jurnal.goretanpena.com/index.php/JSSR>
- [7] D. Sartika, M. Iqbal, and Z. Sitorus, "Analisis Algoritma Exponensial Smoothing dan K-Means untuk Optimalisasi Penyerimaan Mahasiswa Baru di Universitas Haji Sumatera Utara-Dewi Sartika et.al Analisis Algoritma Exponensial Smoothing dan K-Means untuk Optimalisasi Penerimaan Mahasiswa Baru di Universitas Haji Sumatera Utara," vol. 07, no. 02, 2025, doi: 10.54209/jatilima.v7i02.1269.
- [8] M. F. Amalia and D. B. Arianto, "Implementasi Algoritma K-Means Clustering Dalam Klasterisasi Kabupaten/Kota Provinsi Jawa Barat Berdasarkan Faktor Pemicu Stunting Pada Balita," *SIMKOM*, vol. 9, no. 1, pp. 36–46, Jan. 2024, doi: 10.51717/simkom.v9i1.356.
- [9] A. Sulistiyawati and E. Supriyanto, "Implementasi Algoritma K-means Clustering dalam Penentuan Siswa Kelas Unggulan," vol. 15, no. 2.

- [10] R. A. Farissa, R. Mayasari, and Y. Umaidah, "Perbandingan Algoritma K-Means dan K-Medoids Untuk Pengelompokkan Data Obat dengan Silhouette Coefficient," 2021. [Online]. Available: <http://jurnal.polibatam.ac.id/index.php/JAIC>
- [11] J. Sutrisno *et al.*, "Klasterisasi Data Hasil Studi Pelacakan Tentang Karir Dan Pekerjaan Lulusan Perguruan Tinggi Menggunakan Algoritma K-Means," *Jurnal Komputer dan Informatika*, vol. 157, no. 2, pp. 157–164, 2023, doi: 10.35508/jicon.v11i.2.12031.
- [12] A. P. Nanda, D. Eko, H. Pramono, and S. Hartati, "Menentukan Tingkat Kepuasan Mahasiswa Terhadap Pelayanan Akademik Menggunakan Metode Algoritma K-Means," vol. 11.
- [13] A. Khaerunnisa, "Analisis Tingkat Kelulusan Mahasiswa di Unisba dengan menggunakan Algoritma K-Means Clustering," *Jurnal Riset Matematika*, pp. 67–76, Jul. 2022, doi: 10.29313/jrm.v2i1.1018.
- [14] J. P. Tanjung, B. A. Wijaya, and M. Ridho, "Implementasi Algoritma K-Means Dalam Mengukur Tingkat Kepuasan Siswa Terhadap Pelayanan Pada SMA Swasta Bani Adam AS," *Data Sciences Indonesia (DSI)*, vol. 3, no. 1, pp. 1–11, Aug. 2023, doi: 10.47709/dsi.v3i1.2775.
- [15] D. Prasetyawan, A. Mulyanto, and R. Gatra, "Pemetaan Lintasan Karir Alumni Berdasarkan Analisis Cluster: Kombinasi K-Means dan Reduksi Dimensi Autoencoder," *Edumatic: Jurnal Pendidikan Informatika*, vol. 9, no. 1, pp. 198–207, Apr. 2025, doi: 10.29408/edumatic.v9i1.29713.
- [16] J. Alejandrino, "Application of Data Mining in Knowledge Management: A Review," *International Journal of Advanced Trends in Computer Science and Engineering*, vol. 10, no. 4, pp. 2690–2696, 2021, doi: 10.30534/ijatcse/2021/061042021.
- [17] I. Lubis and A. Marwan Elhanafi, "Visualisasi Clustering Penderita Stunting Di Kabupaten Langkat Berbasis WebGIS," *Juli*, 2023, [Online]. Available: <https://ojs.trigunadharma.ac.id/index.php/jsk/index>
- [18] L. Diah Wiranti, E. Budianita, A. Nazir, F. Insani, and R. Susanti, "Penerapan Algoritma K-Means Untuk Mengelompokkan Tingkat Stres Akademik Pada Mahasiswa," *Technology and Science (BITS)*, vol. 7, no. 1, 2025, doi: 10.47065/bits.v7i1.7410.
- [19] R. Maulana and F. Fathoni, "Analisis Clustering Rekomendasi Mata Kuliah Peminatan Berdasarkan Karir Alumni Menggunakan Machine Learning," *JOISIE (Journal Of Information Systems And Informatics Engineering)*, vol. 9, no. 1, p. 109, Jun. 2025, doi: 10.35145/joisie.v9i1.4954.
- [20] M. Cui, "Introduction to the K-Means Clustering Algorithm Based on the Elbow Method", doi: 10.23977/accaf.2020.010102.
- [21] I. Gede, I. Sudipa, and M. Darmawiguna, "BUKU AJAR DATA MINING," 2024. [Online]. Available: <https://www.researchgate.net/publication/377415198>
- [22] A. M. Ikotun, A. E. Ezugwu, L. Abualigah, B. Abuhaija, and J. Heming, "K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data," *Inf Sci (N Y)*, vol. 622, pp. 178–210, Apr. 2023, doi: 10.1016/j.ins.2022.11.139.
- [23] F. Amin, D. S. Anggraeni, and Q. Aini, "Penerapan Metode K-Means dalam Penjualan Produk Souq.Com," *Applied Information System and Management (AISM)*, vol. 5, no. 1, pp. 7–14, Apr. 2022, doi: 10.15408/aism.v5i1.22534.
- [24] H. Vithon Carelsa, R. Andika Malik, and D. J. Putra, "Pengukuran Tingkat Kepuasan Mahasiswa Terhadap Pelayanan di Kantin Kampus Menggunakan Algoritma K-means Clustering," vol. 1, no. 2, pp. 6–11, 2023.
- [25] R. Ishak, "Clustering Prestasi Akademik Lulusan Menggunakan Metode K-Means," *Jambura Journal of Electrical and Electronics Engineering*, 2024.
- [26] R. K. Dinata, Bustami, and S. Retno, "Optimizing the Evaluation of K-means Clustering Using the Weight Product," *Revue d'Intelligence Artificielle*, vol. 38, no. 4, pp. 1223–1233, Aug. 2024, doi: 10.18280/ria.380416.
- [27] B. Parlambang and Fauziah, "Implementasi Algoritma K-Means Dalam Proses Penilaian Kuesioner Kepada Dosen Guna Mendukung Kepuasan Mahasiswa Terhadap Dosen," *Jurnal Ilmiah Teknologi dan Rekayasa*, vol. 25, no. 2, pp. 161–173, 2020, doi: 10.35760/tr.2020.v25i2.2719.