

PENGELOMPOKAN KUALITAS DOSEN UNTUK EVALUASI KINERJA AKADEMIK DOSEN MENGGUNAKAN ALGORITMA K-MEDOIDS

Ibnu Alhafizh Urmaini^a, M.Hasmil Adiya^b

^aInstitut Bisnis dan Teknologi Pelita Indonesia, Jl. Jend. Ahmad Yani No. 78-88 Pekanbaru, ibnu.alhafizh@student.pelitaindonesia.ac.id

^bInstitut Bisnis dan Teknologi Pelita Indonesia, Jl. Jend. Ahmad Yani No. 78-88 Pekanbaru, hasmil.adiya@lecturer.pelitaindonesia.ac.id

INFORMASI ARTIKEL

Sejarah Artikel:

Diterima Redaksi: 19 November 2025

Revisi Akhir: 20 April 2026

Diterbitkan Online: 30 Agustus 2026

KATA KUNCI

K-Medoids, Clustering, Kualitas Dosen, Evaluasi Kinerja, Data Mining

KORESPONDENSI

E-mail: ibnu.alhafizh@student.pelitaindonesia.ac.id

ABSTRACT

Kualitas dosen memiliki peran penting dalam peningkatan mutu perguruan tinggi. Oleh karena itu, diperlukan metode evaluasi yang mampu mengelompokkan kualitas dosen secara objektif dan akurat berbasis data. Penelitian ini bertujuan untuk melakukan clustering kualitas dosen di Institut Bisnis dan Teknologi Pelita Indonesia berdasarkan indikator pendidikan dan pengajaran, penelitian, pengabdian kepada masyarakat, serta penunjang tridharma menggunakan algoritma K-Medoids. Data diperoleh dari Laporan Kinerja Dosen (LKD) pada periode 1 tahun, kemudian dilakukan preprocessing sebelum proses clustering dilakukan. Penentuan jumlah cluster optimal menggunakan metode Elbow menghasilkan 3 cluster yang kemudian dikelompokkan sebagai kategori sangat aktif, cukup aktif, dan kurang aktif dalam menggambarkan kualitas kinerja dosen. Evaluasi performa model dilakukan menggunakan Davies-Bouldin Index (DBI) dan Silhouette Score guna memastikan kualitas pengelompokan. Hasil evaluasi menunjukkan nilai DBI sebesar 1.2627 dan Silhouette Score sebesar 0.3121, yang menandakan bahwa pemisahan cluster telah cukup baik namun masih terdapat sebagian data yang memiliki kemiripan antar cluster.

1. PENDAHULUAN

Kualitas dosen merupakan salah satu faktor penentu utama dalam menciptakan lingkungan pendidikan yang berkualitas di perguruan tinggi. Dosen yang berkualitas tidak hanya dituntut untuk memiliki kompetensi akademik yang mumpuni, tetapi juga harus mampu menciptakan proses pembelajaran yang efektif, inspiratif, dan mendukung perkembangan mahasiswa [1]. Penilaian terhadap kualitas dosen menjadi penting untuk memastikan bahwa standar pendidikan yang tinggi dapat terus dipertahankan dan ditingkatkan.

Penilaian kualitas dosen seringkali menghadapi berbagai tantangan. Data penilaian dosen bersifat multidimensi, mencakup aspek seperti pendidikan dan pengajaran, penelitian, serta pengabdian kepada masyarakat. Metode penilaian konvensional seperti menggunakan rata-rata skor, seringkali tidak mampu menangkap kompleksitas dan variasi data secara efektif [2]. Kurangnya alat atau sistem yang dapat mengelompokkan dosen berdasarkan kualitasnya secara objektif dapat menghambat upaya peningkatan mutu pendidikan [3].

Salah satu solusi yang dapat diterapkan adalah menggunakan algoritma K-Medoids, sebuah metode clustering

yang efektif untuk mengelompokkan data berdasarkan kemiripan karakteristiknya [3]. Berbeda dengan K-Means, K-Medoids lebih robust terhadap outlier dan noise, sehingga cocok untuk data penilaian dosen yang mungkin memiliki variasi dan ketidakseragaman. Beberapa penelitian sebelumnya menyimpulkan bahwa algoritma K-Medoids dapat melakukan pengelompokan data dengan baik [4].

2. TINJAUAN PUSTAKA

2.1. Indikator Penilaian Kualitas Dosen

Perkembangan pendidikan tinggi di Indonesia pada era globalisasi menuntut adanya pemberian layanan terbaik. Perguruan tinggi harus dapat memberikan pelayanan terbaik untuk bersaing pada era globalisasi. Peran dosen sangat besar dalam mewujudkan perguruan tinggi yang berkualitas dan memiliki daya saing tinggi. Dosen, dengan tugas utama melaksanakan kegiatan pendidikan dan pengajaran, penelitian dan pengembangan, pengabdian kepada masyarakat, serta penunjang, menempati posisi sebagai kunci utama dalam menjalankan kegiatan di perguruan tinggi [5].

2.1.1 Pendidikan dan Pengajaran

Pendidikan adalah suatu proses sadar dan terencana untuk mengembangkan potensi, keterampilan, pengetahuan, serta pembentukan sikap dan karakter individu. Pendidikan tidak hanya sebatas transfer ilmu pengetahuan, tetapi juga mencakup pembentukan nilai, moral, serta pengembangan kepribadian yang berkelanjutan. Pengajaran merupakan bagian dari pendidikan yang lebih spesifik berfokus pada proses penyampaian ilmu, keterampilan, dan pengalaman belajar dari pendidik kepada peserta didik [6]. Pengajaran menekankan pada interaksi antara guru dan siswa melalui metode, strategi, dan media pembelajaran tertentu agar peserta didik dapat memahami, menguasai, dan menerapkan pengetahuan yang diberikan..

2.1.2 Penelitian dan Pengembangan

Kompetensi penelitian dan publikasi ilmiah mencerminkan kontribusi dosen dalam pengembangan ilmu pengetahuan [7]. Publikasi di jurnal nasional maupun internasional serta partisipasi dalam seminar akademik menunjukkan kualitas keilmuan dan produktivitas akademik seorang dosen. Evaluasi dari mahasiswa juga sering digunakan sebagai salah satu parameter penting, karena mencerminkan kepuasan dan efektivitas pembelajaran dari sudut pandang mahasiswa. Jumlah publikasi dijadikan indikator kualitas dosen dengan menghitung total artikel yang telah dipublikasikan dalam jurnal nasional maupun internasional selama rentang waktu tertentu. Pendekatan ini memberikan gambaran yang lebih menyeluruh terhadap produktivitas dan kontribusi keilmuan dosen dalam jangka panjang..

2.1.3 Pengabdian Kepada Masyarakat

Pengabdian kepada masyarakat (PKM) merupakan suatu konsep yang melibatkan penerapan pengetahuan, keterampilan, dan sumber daya dari institusi pendidikan untuk memecahkan masalah yang dihadapi oleh masyarakat secara langsung [8]. PKM adalah salah satu bentuk pelaksanaan Tri Dharma Perguruan Tinggi yang bertujuan untuk menerapkan ilmu pengetahuan, teknologi, dan keterampilan untuk membantu menyelesaikan permasalahan yang ada di masyarakat. Kegiatan PKM biasanya dilakukan dalam bentuk pelatihan, pendampingan, penyuluhan, atau pemberdayaan masyarakat di berbagai bidang seperti pendidikan, kesehatan, ekonomi, dan lingkungan. Melalui PKM, perguruan tinggi tidak hanya menjadi pusat pengembangan ilmu, tetapi juga berperan aktif dalam memberikan kontribusi nyata bagi kemajuan dan kesejahteraan masyarakat. Aktivitas PKM dihitung berdasarkan jumlah kegiatan yang dilakukan dosen dalam rentang waktu tertentu. Konversi ini memungkinkan setiap indikator direpresentasikan sebagai fitur numerik yang dapat digunakan dalam analisis statistik atau pengelompokan data dosen secara kuantitatif.

2.1.4 Penunjang Tridharma

Penunjang dosen merupakan segala bentuk kegiatan, tugas, maupun kewajiban tambahan di luar tridharma utama perguruan tinggi yang berfungsi mendukung peningkatan kualitas kinerja akademik dosen [9]. Kegiatan penunjang meliputi aktivitas seperti keanggotaan dalam organisasi profesi, keterlibatan dalam kepanitiaan institusi, peran sebagai

narasumber atau pembicara, serta kontribusi dalam kegiatan penjaminan mutu dan pengembangan kelembagaan.

Unsur penunjang menjadi bagian penting dalam evaluasi kinerja karena menunjukkan dedikasi, tanggung jawab, dan keterlibatan dosen secara lebih luas di luar bidang pengajaran, penelitian, dan pengabdian. Melalui kegiatan penunjang, dosen tidak hanya berperan sebagai pengajar dan peneliti, tetapi juga sebagai penggerak dalam pengembangan institusi dan peningkatan kualitas akademik secara menyeluruh.

2.2. Clustering

Clustering adalah metode dalam data mining dan machine learning yang digunakan untuk mengelompokkan sekumpulan data ke dalam beberapa grup atau kluster berdasarkan kemiripan karakteristik di antara data tersebut [10]. Tujuan utama dari clustering adalah untuk mengidentifikasi struktur atau pola tersembunyi dalam data tanpa menggunakan label atau kategori yang telah ditentukan sebelumnya (unsupervised learning). Objek-objek dalam satu kluster memiliki tingkat kesamaan yang tinggi satu sama lain, sementara objek dari kluster yang berbeda memiliki perbedaan yang signifikan. Metode ini sering digunakan dalam berbagai bidang seperti segmentasi pasar, analisis perilaku konsumen, pengelompokan dokumen, pengenalan pola, dan analisis citra.

2.2.1 Metode Elbow

Metode Elbow merupakan salah satu metode untuk menentukan jumlah cluster yang tepat melalui persentase hasil perbandingan antara jumlah cluster yang akan membentuk siku pada suatu titik [11]. Jika nilai cluster pertama dengan nilai cluster kedua membentuk sudut dalam grafik atau nilainya mengalami penurunan paling besar, maka jumlah cluster tersebut dianggap sebagai jumlah yang tepat. Untuk mendapatkan perbandingannya, dilakukan perhitungan Sum of Square Error (SSE) dari masing-masing nilai cluster. Semakin besar jumlah cluster (K), maka nilai SSE akan semakin kecil. Rumus SSE ditunjukkan pada Persamaan 1.

$$SSE = \sum_{k=1}^K \sum_{xi} |X_i - C_k|^2 \quad (1)$$

Keterangan :

K= cluster ke-c

xi= jarak data obyek ke-i

ck= pusat cluster ke-i

2.3. Silhouette Score

Silhouette Score digunakan untuk melihat kualitas dan kekuatan cluster, seberapa baik atau buruknya suatu obyek ditempatkan dalam suatu cluster. Metode ini merupakan gabungan dari metode separasi dan kohesi [12].

$$s_i = \frac{b_i - a_i}{\max(a_i, b_i)} \quad (2)$$

Keterangan :

ai : rata-rata jarak antara titik iii dengan semua titik dalam kluster yang sama.

bi : rata-rata jarak antara titik iii dengan semua titik di kluster terdekat lainnya.

Nilai Silhouette Score berkisar dari -1 hingga 1:

Mendekati 1 → titik sangat cocok dengan klasternya.

Mendekati 0 → titik berada di antara dua klaster.

Negatif → titik mungkin salah penempatan klaster.

2.4. K-Medoids

K-Medoids merupakan kelompok metode partitional clustering yang meminimalkan jarak antara titik dalam cluster dan titik yang ditunjuk sebagai pusat cluster. Berbeda dengan algoritma K-Means yang menggunakan nilai rata-rata sebagai pusat cluster, K-Medoids menggunakan objek sebagai perwakilan (medoid) sebagai pusat cluster untuk setiap cluster [13].

Algoritma K-Medoids memiliki kelebihan untuk mengatasi kelemahan pada algoritma K-Means yang sensitif terhadap noise dan outlier. Kelebihan lainnya yaitu hasil proses clustering tidak bergantung pada urutan masuk dataset. K-Medoids lebih kuat untuk mengatasi kebisingan dan pencilaan karena meminimalkan sejumlah dissimilarities berpasangan, bukan jumlah kuadrat jarak Euclidean..

Langkah-langkah Algoritma K-Medoids :

1. Inisialisasi pusat cluster sebanyak k
2. Alokasikan setiap data ke cluster terdekat menggunakan Euclidean Distance:

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}; i = 1, 2, 3, \dots, n \quad (3)$$

3. Pilih secara acak objek sebagai kandidat medoid baru
4. Hitung jarak setiap objek dengan kandidat medoid baru
5. Hitung total simpangan (S) = total distance baru - total distance lama
6. Jika $S < 0$, tukar objek untuk membentuk k objek baru sebagai medoid
7. Ulangi langkah 3-5 hingga tidak terjadi perubahan medoid.

2.5. Data Mining

Data mining merupakan gabungan sejumlah disiplin ilmu komputer yang didefinisikan sebagai proses penemuan pola-pola baru dari kumpulan data sangat besar, meliputi metode-metode yang merupakan irisan dari artificial intelligence, machine learning, statistics, dan database systems [14].

Data mining adalah proses pencarian data secara otomatis untuk mendapatkan model dari database yang besar, serta metode untuk menemukan informasi tersembunyi dalam database. Data mining merupakan bagian dari proses Knowledge Discovery in Databases (KDD) untuk menemukan informasi dan pola yang berguna dalam data. Proses KDD dapat digambarkan sebagai berikut [15]:

1. Data Selection - pemilihan data dari sekumpulan data operasional sebelum tahap penggalian informasi dimulai
2. Pre-processing/cleaning - proses cleaning pada data yang mencakup membuang duplikasi data, memeriksa data inkonsisten, dan memperbaiki kesalahan
3. Transformation - mengubah data kedalam bentuk yang sesuai untuk ditambang
4. Data Mining - proses mencari pola atau informasi menarik dalam data terpilih menggunakan teknik tertentu
5. Interpretasi/Evaluasi - menampilkan pola informasi dalam bentuk yang mudah dimengerti menggunakan visualisasi.

2.6. Davies Bouldin Index (DBI)

Davies Bouldin Index (DBI) merupakan metode evaluasi clustering yang termasuk dalam internal evaluation. Metode ini mengukur kualitas clustering melalui metric kohesi dan metric separasi. Berbeda dengan external evaluation yang menggunakan informasi terdahulu seperti Rand Measure, f-measure, dan jaccard index, internal evaluation hanya menggunakan data yang ada [16].

DBI bertujuan memaksimalkan jarak intercluster antara satu cluster dengan cluster yang lain. DBI didefinisikan sebagai ukuran rata-rata kemiripan antar cluster yang terbentuk. Nilai DBI yang kecil menunjukkan kualitas clustering yang baik, dimana cluster yang terbentuk memiliki kohesi tinggi dan separasi yang jelas. Dalam konteks deteksi outlier, DBI dapat digunakan untuk mendeteksi outlier pada masing-masing cluster yang terbentuk.. Rumus sederhana davies-bouldin-index :

$$DBI = \frac{1}{k} \sum_{i=1}^k \max_{j \neq i} \left(\frac{S_i + S_j}{d_{ij}} \right) \quad (4)$$

Keterangan :

k = jumlah cluster

S_i =rata-rata jarak antar titik dalam cluster ke-i (mengukur kekompakan)

d_{ij} =jarak antara pusat cluster ke-i dan ke-j (mengukur keterpisahan)

Rasio $(S_i + S_j)/d_{ij}$ makin besar $\rightarrow$ dua cluster makin mirip $\rightarrow$ hasil clustering makin jelek

2.7. Python

Python merupakan bahasa pemrograman yang dapat melakukan eksekusi instruksi multiguna secara langsung dengan orientasi objek menggunakan semantik dinamis. Bahasa ini memiliki kemampuan tinggi yang menggabungkan kapabilitas dan sintaks kode yang sangat baik, dilengkapi dengan pustaka dasar yang komprehensif dan besar.

Python dirancang agar mudah dimengerti dan dipelajari, serta dapat berjalan di berbagai platform seperti Mac, Linux, dan Windows. Sifat open source Python memungkinkan kontribusi pengembangan dengan hak kekayaan intelektual dimiliki oleh PSF.

Python memiliki library yang menyediakan berbagai fungsi analisis data, machine learning, visualisasi data, dan data processing tools. Keberadaan library-library tersebut membuat Python sangat diminati pada bidang analisis dan data science [17].

2.8. Principal Component Analysis

Principal Component Analysis (PCA) merupakan metode reduksi dimensi yang digunakan untuk menyederhanakan data berdimensi tinggi tanpa menghilangkan informasi penting di dalamnya [18]. Teknik ini mengubah variabel awal menjadi principal components yang merupakan kombinasi linear dari variabel asli, tersusun berdasarkan besarnya variasi data yang dapat dijelaskan.

PCA menghasilkan komponen pertama yang memuat variasi terbesar, diikuti komponen kedua, dan seterusnya. Dengan memanfaatkan PCA, proses analisis seperti clustering atau visualisasi menjadi lebih efektif karena data menjadi lebih ringkas, pola lebih mudah dikenali, dan noise dapat diminimalkan.

Prosedur PCA meliputi: pengujian nilai KMO untuk melihat kecukupan sampel secara keseluruhan, Bartlett's test untuk mengetahui korelasi antar variabel, pengujian matriks korelasi anti image, perbaikan model dengan eliminasi variabel jika diperlukan, analisis komponen utama, dan interpretasi hasil PCA.

2.9. Google Colab

Google Colab (Colaboratory) adalah platform berbasis cloud yang dikembangkan oleh Google Research, dirancang untuk memungkinkan pengguna menjalankan kode Python langsung dari browser. Colab sangat berguna untuk berbagai aplikasi seperti machine learning, analisis data, dan pendidikan, tanpa perlu pengaturan atau instalasi tambahan.

Platform ini berbasis pada Jupyter Notebook dan memberikan akses gratis ke resource komputasi seperti GPU dan TPU, yang sangat membantu dalam pemrosesan data dan pelatihan model machine learning. Google Colab merupakan IDE untuk pemrograman Python dimana pemrosesan dilakukan oleh server Google yang memiliki perangkat keras dengan performa tinggi [19]. Dari sisi perangkat lunak, Google Colab telah menyediakan hampir sebagian besar pustaka (library) yang dibutuhkan untuk pemrosesan data dan analisis.

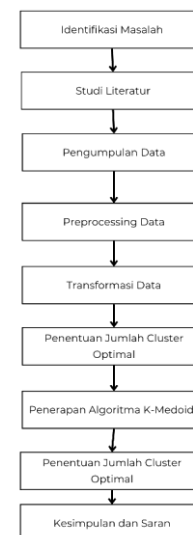
3. METODOLOGI

3.1. Tempat dan Waktu Penelitian

Penelitian ini dilaksanakan di Institut Bisnis dan Teknologi Pelita Indonesia dengan fokus pada data penilaian kualitas dosen selama rentang waktu semester ganjil 2024/2025 dan semester genap 2025. Pengumpulan data dilakukan melalui sumber-sumber yang relevan, seperti hasil pendidikan dan pengajaran, penelitian, pengabdian kepada masyarakat, serta penunjang tridharma. Proses pengumpulan data berlangsung selama periode yang diperlukan untuk memastikan kelengkapan dan akurasi data yang akan digunakan dalam analisis menggunakan algoritma K-Medoids.

3.2. Kerangka Penelitian

Kerangka penelitian merupakan tahapan sistematis yang digunakan sebagai pedoman dalam pelaksanaan penelitian. Setiap tahap dilakukan secara berurutan agar penelitian dapat berjalan terarah dan menghasilkan kesimpulan yang sesuai dengan tujuan yang telah ditetapkan. Adapun kerangka penelitian yang digunakan dalam penelitian ini dapat dilihat pada Gambar 1.



Gambar 1. Kerangka Penelitian

Penelitian ini dilaksanakan melalui beberapa tahapan sistematis yang dimulai dari:

1. Identifikasi masalah mengenai belum adanya evaluasi objektif terhadap kualitas dosen di Institut Bisnis dan Teknologi Pelita Indonesia.
2. Tahap studi literatur dilakukan dengan mengumpulkan dan menelaah berbagai referensi terkait kualitas dosen, algoritma K-Medoids, dan teknik data mining untuk menentukan pendekatan yang tepat.
3. Data penelitian diperoleh dari Laporan Kinerja Dosen (LKD) yang mencakup empat aspek tridharma perguruan tinggi meliputi pendidikan dan pengajaran, penelitian, pengabdian kepada masyarakat, serta kegiatan penunjang.
4. Data yang terkumpul kemudian melalui tahap preprocessing untuk memastikan kebersihan dan kelengkapan data.
5. Dilanjutkan dengan transformasi data melalui normalisasi nilai dan penyamaan skala antar variabel.
6. Penentuan jumlah cluster optimal dilakukan menggunakan metode Elbow untuk mendapatkan pola pengelompokan yang alami.
7. Algoritma K-Medoids diterapkan dengan menentukan medoid awal, menghitung jarak antar objek, dan membentuk cluster berdasarkan kedekatan dengan medoid melalui proses iterasi hingga mencapai kondisi stabil.
8. Evaluasi hasil clustering dilakukan menggunakan Davies-Bouldin Index dan Silhouette Score untuk mengukur kualitas pengelompokan, serta interpretasi cluster berdasarkan kategori "Sangat Aktif", "Cukup Aktif", dan "Kurang Aktif".
9. Penelitian diakhiri dengan penyimpulan hasil analisis dan pemberian saran untuk peningkatan kualitas dosen serta rekomendasi pengembangan penelitian selanjutnya.

3.3. Metode Penelitian

Penelitian ini menggunakan algoritma K-Medoids untuk melakukan clustering kualitas dosen di Institut Bisnis dan Teknologi Pelita Indonesia berdasarkan data penilaian yang

mencakup aspek pendidikan dan pengajaran, penelitian, pengabdian kepada masyarakat, serta penunjang tridharma. Algoritma K-Medoids dipilih karena kemampuannya dalam mengelompokkan data berdasarkan kemiripan karakteristiknya serta ketahanannya terhadap outlier dan noise. Sebagai bagian dari unsupervised machine learning, algoritma ini mengelompokkan data tanpa label kelas yang telah ditentukan sebelumnya dengan memilih objek nyata (medoids) sebagai pusat cluster, sehingga lebih stabil dan akurat dibandingkan algoritma lain seperti K-Means.

Proses clustering diawali dengan memplot data penilaian dosen ke dalam ruang multidimensi dimana setiap dimensi merepresentasikan fitur penilaian. Algoritma K-Medoids kemudian mencari medoids optimal untuk memisahkan data ke dalam cluster sangat aktif, cukup aktif, dan kurang aktif. Tujuan utama algoritma ini adalah meminimalkan jarak antara titik data dalam satu cluster dengan medoids-nya untuk menghasilkan pengelompokan yang akurat dan konsisten. Hasil pengelompokan ini diharapkan dapat menjadi dasar bagi pihak institusi dalam mengambil keputusan manajemen sumber daya manusia, seperti program pelatihan, pembinaan, atau insentif bagi dosen, serta meningkatkan mutu pendidikan di Institut Bisnis dan Teknologi Pelita Indonesia

3.4. Pengumpulan Data

Pengumpulan data kualitas dosen di Institut Bisnis dan Teknologi Pelita Indonesia dilakukan melalui berbagai sumber meliputi hasil pendidikan dan pengajaran, laporan penelitian, pengabdian masyarakat, serta penunjang tridharma. Data mencakup periode semester ganjil 2024/2025 dan semester genap 2025 yang dianggap mampu menggambarkan dinamika dan perkembangan kualitas dosen secara komprehensif. Indikator penilaian utama terdiri dari penelitian, pengabdian kepada masyarakat, pendidikan dan pengajaran, serta penunjang tridharma yang menjadi parameter penting dalam menilai efektivitas pembelajaran.

Tabel 1. Laporan Kinerja Dosen

N o	Nam a	Jumlah Pengajara n	Jumlah Publikas i	Jumla h PKM	Jumlah Penunjan g
1	d1	26	3	5	6
2	d2	23,25	6	4	5
3	d3	54,5	4	3	3
4	d4	45,25	22	15	53
5	d5	65,25	6	5	11
⋮	⋮	⋮	⋮	⋮	⋮
21	d21	48,5	3	3	5
22	d22	30,5	2	5	2
23	d23	54,75	3	4	3
24	d24	31	2	2	3
25	d25	45,25	4	3	10
⋮	⋮	⋮	⋮	⋮	⋮
46	d46	38,5	7	4	8
47	d47	24	7	0	0

48	d48	19,5	6	3	6
49	d49	59	10	6	9
50	d50	31	4	5	18
⋮	⋮	⋮	⋮	⋮	⋮
71	d71	83,95	16	7	28
72	d72	17,65	2	3	7
73	d73	29,5	7	12	7
74	d74	24,25	11	2	4
75	d75	24	0	0	0

Data Laporan Kinerja Dosen (LKD) yang digunakan berjumlah 75 data, dimana setiap data merepresentasikan satu dosen dengan identitas anonim berkode d1 hingga d75. Keempat variabel penelitian meliputi jumlah pengajaran yang menunjukkan aktivitas belajar mengajar, jumlah publikasi yang menggambarkan produktivitas karya ilmiah, jumlah PKM yang mencerminkan keterlibatan dalam pengabdian masyarakat, serta jumlah penunjang yang merepresentasikan kegiatan tambahan di luar kewajiban tridharma.

Data yang terkumpul kemudian diproses menggunakan algoritma K-Medoids untuk mengelompokkan dosen ke dalam kategori sangat aktif, cukup aktif, dan kurang aktif. Pemilihan periode dua semester bertujuan untuk menangkap konsistensi dan perkembangan kualitas dosen secara menyeluruh. Hasil clustering diharapkan dapat menjadi dasar bagi institusi dalam mengambil keputusan strategis seperti program pelatihan, pembinaan, atau pemberian insentif untuk meningkatkan mutu pendidikan dan kualitas dosen secara berkelanjutan.

3.5. Teknik Pengujian

Pengujian merupakan tahap penting dalam penelitian ini untuk mengevaluasi performa model clustering K-Medoids dalam mengelompokkan kualitas dosen di Institut Bisnis dan Teknologi Pelita Indonesia. Evaluasi dilakukan menggunakan Silhouette Score yang mengukur kualitas clustering berdasarkan rasio antara jarak dalam cluster dan jarak antar cluster. Davies-Bouldin Index (DBI) juga digunakan sebagai metrik evaluasi, dimana semakin rendah nilai DBI menunjukkan performa model yang semakin baik, dengan karakteristik cluster yang memiliki jarak internal kecil dan jarak antar cluster yang besar.

Pengujian ini bertujuan untuk menilai kemampuan algoritma K-Medoids dalam menghasilkan pengelompokan yang akurat dan stabil, khususnya dalam menangani data penilaian dosen yang bersifat multidimensi. Hasil pengujian diharapkan dapat memberikan wawasan mendalam tentang distribusi kualitas dosen serta menjadi dasar bagi pihak institusi dalam mengambil keputusan strategis untuk meningkatkan mutu pendidikan dan kualitas dosen secara berkelanjutan.

3.6. Perhitungan Manual

Pada tahap ini dilakukan proses perhitungan menggunakan algoritma K-Medoids untuk melakukan pengelompokan data kualitas dosen. Algoritma K-Medoids dipilih karena keunggulannya dalam mengurangi pengaruh data ekstrem (outlier) dibandingkan dengan algoritma K-Means. Setiap cluster direpresentasikan oleh satu objek nyata (medoid)

yang memiliki total jarak terpendek terhadap seluruh objek lain dalam cluster tersebut.

Sebelum melakukan clustering, dilakukan standarisasi data menggunakan Z-score karena setiap variabel dalam dataset memiliki satuan dan rentang nilai yang berbeda, seperti jumlah pengajaran, publikasi, dan kegiatan pengabdian masyarakat. Perbedaan skala ini dapat menyebabkan variabel dengan nilai besar mendominasi proses perhitungan jarak pada algoritma clustering.

1. Hitung rata-rata dan standar deviasi setiap kolom
 $\text{Mean} = (x_1 + x_2 + \dots + x_n) / n$ atau $(\sum \text{semua nilai variabel}) / n$
 $\text{Standar Deviasi} = \sqrt{((x_1 - \text{mean})^2 + (x_2 - \text{mean})^2 + \dots + (x_n - \text{mean})^2) / (n-1)}$ atau $S = \sqrt{((\sum (x - \text{mean})^2) / (n-1))}$
 Keterangan :
 x = nilai data asli
 n = banyak data
 $\sum$ = notasi penjumlahan
 $\sqrt{}$ = akar kuadrat
 Mean = rata-rata
2. Standarisasi menggunakan rumus Z-Score menggunakan rumus:

$$Z = \frac{X - \bar{X}}{S}$$
 Keterangan :
 x = nilai asli data
 $\bar{x}$ = nilai rata-rata variabel
 S = standar deviasi variabel

3.6.1 Perhitungan Jumlah Cluster

Penentuan jumlah cluster merupakan tahap awal dalam proses clustering untuk mengetahui banyaknya kelompok yang sesuai dengan karakteristik data. Penelitian ini menggunakan metode Elbow untuk menentukan jumlah cluster optimal dengan menghitung nilai Sum of Squared Error (SSE) setiap kemungkinan jumlah cluster (k). Nilai SSE menunjukkan total jarak kuadrat antara setiap data dengan pusat cluster-nya (medoid), dimana semakin kecil nilai SSE menunjukkan tingkat kekompakan data dalam cluster yang semakin baik. Namun, penurunan nilai SSE akan semakin kecil seiring bertambahnya jumlah cluster.

1. Tentukan cluster awal ($K=2$)
 Misal pembagian awal: Cluster 1: {d1, d2, d3} dan Cluster 2: {d4, d5}
2. Hitung Centroid masing-masing cluster
 Centroid Cluster 1 (C_1):

$$Z = \frac{(d_1 + d_2 + d_3)}{\text{jumlah dosen}}$$
 Keterangan :
 Z = centroid cluster
 d_1 = dosen 1
 d_2 = dosen 2
 d_3 = dosen 3
3. Hitung jarak tiap titik ke centroidnya
 Rumus jarak Euclidean:
 $d = \sqrt{(x_1 - \mu_1)^2 + (x_2 - \mu_2)^2 + (x_3 - \mu_3)^2 + (x_4 - \mu_4)^2}$
 Keterangan :
 μ = centroid cluster
 x = nilai data asli

4. Hitung SSE total

$$\text{SSE} = \sum d^2$$
 Keterangan :
 d = jarak tiap titik ke centroid
 Lakukan ulang pada $k=3, k=4, k=5$

3.6.2 Perhitungan K-Medoids

Pada tahap ini dilakukan perhitungan algoritma K-Medoids untuk mengelompokkan data kualitas dosen. K-Medoids dipilih karena menggunakan objek nyata (medoid) sebagai pusat cluster sehingga lebih tahan terhadap outlier. Perhitungan menggunakan data terstandarisasi (Z-score) untuk atribut Jumlah Pengajaran, Publikasi, PKM, dan Penunjang. Langkah implementasi meliputi:

1. Penentuan jumlah cluster sebanyak 3 cluster
2. Pemilihan medoid awal (d_1, d_2, d_3)
3. Penghitungan jarak setiap titik ke medoid awal menggunakan Euclidean distance

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}; i = 1, 2, 3, \dots, n$$
 Keterangan :
 $d(x, y)$ = Jarak Euclidean antara dua titik x dan y
 $\sqrt{}$ = Akar kuadrat (square root)
 $\sum$ = Simbol penjumlahan (summation)
 $i=1$ = Indeks awal penjumlahan
 n = Jumlah dimensi/data fitur
 x_i = Nilai titik x pada dimensi ke- i
 y_i = Nilai titik y pada dimensi ke- i
 $(x_i - y_i)^2$ = Kuadrat selisih antara nilai x dan y pada dimensi ke- i
4. Iterasi medoid dengan memilih medoid baru (d_4, d_5, d_6) secara acak pada iterasi kedua. Proses iterasi dilakukan berulang hingga diperoleh cluster yang optimal dengan medoid terbaik.

3.6.3 Perhitungan Evaluasi Clustering

Evaluasi hasil clustering dilakukan untuk menilai kualitas pengelompokan data yang dihasilkan algoritma K-Medoids menggunakan Davies-Bouldin Index (DBI) dan Silhouette Score. Evaluasi DBI digunakan untuk menilai pemisahan antar cluster dan tingkat kekompakan dalam masing-masing cluster, dimana semakin kecil nilai DBI menunjukkan kualitas clustering yang semakin baik dengan cluster yang kompak dan terpisah jelas. Langkah-langkahnya :

1. Kelompokkan berdasarkan jarak terkecil
2. Hitung S_i (Rata Rata Jarak Setiap Member Ke Medoid)

$$S_i = \frac{\sum_{k=1}^n d(k, \text{medoid})}{n}$$
 Keterangan :
 S_i = Nilai rata-rata jarak untuk cluster i
 n = Jumlah anggota cluster
 $\sum (d(k, \text{medoid}))$ = Penjumlahan jarak seluruh titik k ke medoid cluster
 $d(k, \text{medoid})$ = jarak antara titik ke- k dan medoid
3. Hitung M_{ij} (Jarak antar medoid)
4. Hitung R_{ij} (Rasio antar cluster)

$$R_{ij} = \frac{S_i + S_j}{M_{ij}}$$
 Keterangan :
 R_{ij} = Nilai kemiripan antara cluster i dan cluster j

S_i = Rata-rata jarak semua anggota cluster i ke medoid-nya

S_j = Rata-rata jarak semua anggota cluster j ke medoid-nya

M_{ij} = Jarak antara medoid cluster i dan medoid cluster j .

5. Hitung d_i (nilai davies bouldin untuk tiap cluster)

d_i = nilai R terbesar terhadap cluster lain

6. Hitung DBI Akhir

$$DBI = \frac{1}{k} \sum_{i=1}^k D_i$$

Keterangan :

DBI = Nilai Davies–Bouldin Index

K = Jumlah cluster

$\sum D_i$ = penjumlahan nilai indeks untuk tiap cluster i

Silhouette Score mengevaluasi kualitas pemodelan clustering dengan mengukur seberapa baik suatu data berada dalam clusternya dibandingkan dengan kedekatannya terhadap cluster lain. Nilai Silhouette berada pada rentang -1 hingga 1, dimana nilai mendekati 1 menunjukkan kecocokan data dengan cluster yang ditempati, nilai mendekati 0 mengindikasikan data berada di batas antarcluster, dan nilai negatif menandakan kesalahan penempatan cluster. Langkah-langkahnya :

1. Tentukan cluster tiap dosen
2. Perhitungan Silhouette (3 Data)

Nilai Silhouette

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

Keterangan :

s_i = Nilai silhouette untuk objek/data ke- i

a_i = Rata-rata jarak objek ke- i terhadap semua objek lain dalam cluster yang sama

b_i = Rata-rata jarak objek ke- i terhadap cluster terdekat yang berbeda dari cluster-nya

$\max(a_i, b_i)$ = Nilai maksimum antara jarak dalam cluster dan jarak ke cluster lain

Nilai Silhouette Akhir

$$SS = \frac{\sum_{i=1}^n S_i}{n}$$

Keterangan :

SS = Silhouette Score keseluruhan untuk satu hasil clustering

s_i = Nilai silhouette untuk objek ke- i

n = Jumlah total objek/data yang dikelompokkan

$\sum s(i)$ = Penjumlahan seluruh nilai silhouette dari masing-masing objek

4. HASIL DAN PEMBAHASAN

4.1. Pengumpulan Data

Data dalam penelitian ini diperoleh dari Laporan Kinerja Dosen (LKD) periode semester ganjil 2024/2025 dan semester genap 2025, mencakup aspek pendidikan dan pengajaran, penelitian, pengabdian kepada masyarakat, dan penunjang tridharma. Data terdiri dari 75 dosen dengan empat indikator utama: jumlah pengajaran yang menggambarkan intensitas aktivitas mengajar, jumlah publikasi yang menunjukkan produktivitas karya ilmiah, jumlah PKM yang merepresentasikan

keterlibatan dalam pengabdian masyarakat, serta penunjang tridharma.

Variasi nilai pada keempat indikator menunjukkan perbedaan tingkat aktivitas antar dosen. Beberapa dosen menunjukkan nilai tinggi pada pengajaran dan publikasi yang menandakan kinerja akademik aktif dan produktif, sementara lainnya memiliki nilai relatif rendah di semua indikator yang mengindikasikan aktivitas akademik terbatas. Data tersebut selanjutnya diolah menggunakan algoritma K-Medoids untuk mengelompokkan dosen berdasarkan kesamaan karakteristik kinerja mereka.

4.2. Preprocessing Data

4.2.1 Pengecekan Data Kosong dan Data Duplikat

Langkah awal dalam preprocessing adalah memeriksa keberadaan missing values dan data duplikat. Berdasarkan hasil pemeriksaan menggunakan Python (pandas), tidak ditemukan adanya nilai kosong maupun duplikat. Hasil ini menunjukkan data telah lengkap dan unik, sehingga dapat langsung diproses ke tahap berikutnya tanpa memerlukan pembersihan data tambahan.

```
print("Jumlah data kosong:\n", df.isnull().sum())
print("Jumlah data duplikat:", df.duplicated().sum())
```

```
Jumlah data kosong:
No                0
Nama              0
Jumlah Pengajaran 0
Jumlah Publikasi  0
Jumlah PKM        0
Jumlah Penunjang  0
dtype: int64
Jumlah data duplikat: 0
```

Gambar 2. Pengecekan Data kosong dan duplikat

Berdasarkan hasil pemeriksaan kualitas data, dapat disimpulkan bahwa tidak terdapat data kosong maupun data duplikat pada dataset yang digunakan. Nilai nol (0) pada setiap kolom meliputi No, Nama, Jumlah Pengajaran, Jumlah Publikasi, Jumlah PKM, dan Jumlah Penunjang. Mengkonfirmasi bahwa seluruh data telah terisi dengan lengkap dan unik. Kondisi ini menunjukkan bahwa dataset telah memenuhi kriteria clean data dan siap untuk diproses pada tahap analisis berikutnya tanpa memerlukan data cleaning tambahan.

4.2.2 Data Selection

Dari keseluruhan atribut yang tersedia, dipilih empat variabel untuk proses clustering yaitu Jumlah Pengajaran, Jumlah Publikasi, Jumlah PKM, dan Jumlah Penunjang Tridharma. Keempat variabel ini dinilai paling representatif dalam menilai kualitas dosen secara objektif berdasarkan tridharma perguruan tinggi, mencakup aspek pendidikan, penelitian, pengabdian masyarakat, dan penunjang kegiatan akademik.

```
# Ambil hanya kolom yang diinginkan
X = df[["Jumlah Pengajaran", "Jumlah Publikasi", "Jumlah PKM", "Jumlah Penunjang"]]
```

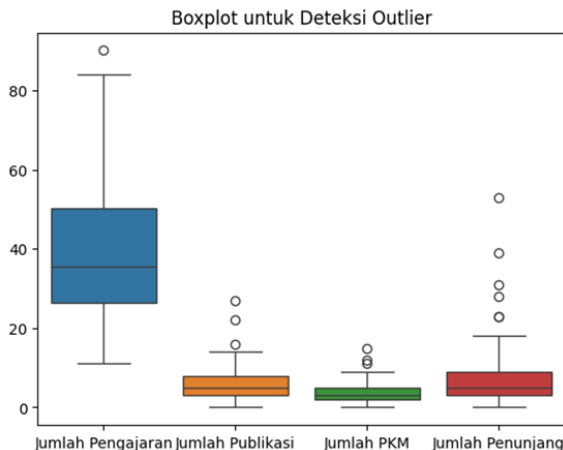
Gambar 3. Pemilihan Variabel

Kode program berfungsi untuk memilih kolom-kolom tertentu dari DataFrame df guna memastikan hanya data relevan yang digunakan dalam analisis. Variabel X menyimpan empat kolom terpilih, yaitu "Jumlah Pengajaran", "Jumlah Publikasi", "Jumlah PKM", dan "Jumlah Penunjang". Pemilihan keempat variabel ini bertujuan memfokuskan analisis pada indikator utama kinerja dosen yang mencakup aspek pengajaran, penelitian

melalui publikasi, pengabdian masyarakat melalui PKM, serta kegiatan penunjang tridharma lainnya.

4.2.3 Pengecekan Outlier

Pengecekan outlier dilakukan menggunakan metode boxplot. Hasil analisis menunjukkan adanya beberapa nilai ekstrem, khususnya pada variabel jumlah publikasi. Mengingat algoritma K-Medoids relatif tahan terhadap outlier, nilai-nilai ekstrem tersebut tetap dipertahankan dalam analisis tanpa dilakukan penghapusan atau transformasi lebih lanjut.



Gambar 4. Deteksi Outlier

Berdasarkan analisis boxplot pada keempat variabel penelitian, teridentifikasi adanya outlier dengan karakteristik yang bervariasi. Variabel Jumlah Pengajaran menunjukkan sebaran data terlebar dengan beberapa nilai ekstrem di atas 80, mengindikasikan adanya outlier pada data pengajaran. Sementara itu, variabel Jumlah Publikasi, Jumlah PKM, dan Jumlah Penunjang memiliki sebaran data lebih sempit meskipun tetap menunjukkan beberapa outlier di bagian atas. Temuan ini mengungkapkan bahwa sebagian besar data terkonsentrasi pada rentang nilai rendah, sementara beberapa dosen menunjukkan aktivitas yang secara signifikan lebih tinggi dibandingkan mayoritas. Secara keseluruhan, hasil boxplot merefleksikan ketidakseimbangan kontribusi di antara dosen dalam keempat kategori kegiatan tridharma perguruan tinggi.

4.3. Data Transformation

Data numerik yang dianalisis memiliki skala yang berbeda-beda antar variabel. Untuk mengatasi perbedaan skala ini, dilakukan transformasi menggunakan metode z-score normalization. Transformasi ini menghasilkan seluruh variabel dengan skala distribusi seragam, yaitu rata-rata 0 dan standar deviasi 1. Proses normalisasi penting untuk memastikan tidak ada variabel yang mendominasi proses pembentukan cluster dalam analisis clustering selanjutnya.

```
scaler = StandardScaler()
X_scaled = scaler.fit_transform(X)

columns = ["Jumlah Pengajaran", "Jumlah Publikasi", "Jumlah PKM", "Jumlah Penunjang"]

# Konversi hasil transformasi ke DataFrame
X_scaled = pd.DataFrame(X_scaled, columns=columns)

# Tampilkan
print(X_scaled.head())
```

	Jumlah Pengajaran	Jumlah Publikasi	Jumlah PKM	Jumlah Penunjang
0	-0.829809	-0.716539	0.363422	-0.252154
1	-0.999449	-0.041973	-0.025959	-0.360800
2	0.928279	-0.491684	-0.415339	-0.593930
3	0.357671	3.555714	4.257228	5.102330
4	1.591417	-0.041973	0.363422	0.317472

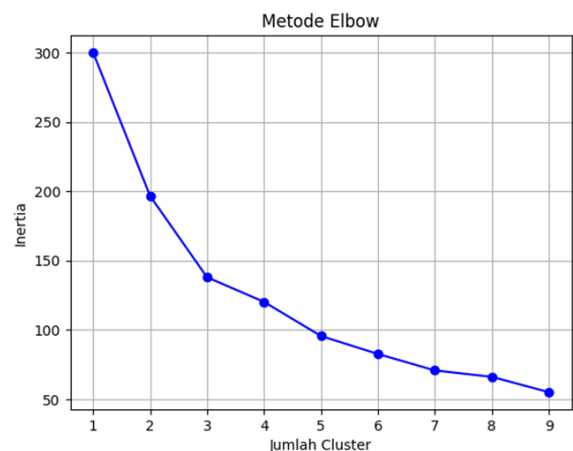
Gambar 5. Data Transformation

Proses standarisasi data dilakukan menggunakan StandardScaler dari pustaka scikit-learn dengan tujuan menyamakan skala antar variabel sehingga memiliki mean 0 dan standar deviasi 1. Standarisasi diterapkan pada keempat variabel: Jumlah Pengajaran, Jumlah Publikasi, Jumlah PKM, dan Jumlah Penunjang. Hasil transformasi dengan scaler.fit_transform(X) dikonversi ke bentuk DataFrame untuk memudahkan pembacaan. Output standarisasi menunjukkan nilai-nilai berupa bilangan positif dan negatif yang merepresentasikan jarak data dari rata-rata dalam satuan standar deviasi, dimana nilai positif besar menunjukkan data jauh di atas rata-rata, sedangkan nilai negatif mengindikasikan data berada di bawah rata-rata.

4.4. Menentukan Jumlah Clustering

4.5.2 Metode Elbow

Penentuan jumlah cluster optimal dilakukan menggunakan metode Elbow dengan memplot nilai inertia terhadap jumlah cluster dari 1 hingga 10. Metode ini diimplementasikan melalui iterasi jumlah cluster dari 1 sampai 9, dimana untuk setiap nilai k, model KMeans dilatih menggunakan data terstandarisasi (X_scaled). Nilai inertia - yang merepresentasikan total jarak kuadrat antara data dengan pusat cluster - disimpan dalam list inertia. Nilai-nilai inertia tersebut kemudian divisualisasikan untuk mengidentifikasi titik "elbow" yang mengindikasikan jumlah cluster optimal. Visualisasi ini memungkinkan pemilihan jumlah cluster yang seimbang, tidak terlalu sedikit maupun berlebihan.



Gambar 6. Hasil Metode Elbow

Hasil penerapan metode Elbow menunjukkan hubungan antara jumlah cluster dan nilai inertia, dengan sumbu X merepresentasikan jumlah cluster (1-9) dan sumbu Y menunjukkan nilai inertia sebagai total jarak kuadrat antara titik data dengan pusat cluster. Teramati penurunan inertia yang signifikan dari cluster 1 ke 3, diikuti pelandaian penurunan setelah cluster ke-3 atau ke-4. Titik pelandaian ini disebut "elbow point" yang menandakan jumlah cluster optimal. Berdasarkan pola grafik, elbow point teridentifikasi pada cluster ke-3, sehingga jumlah cluster ideal untuk pembagian data adalah 3 cluster. Penambahan cluster di atas 3 tidak memberikan penurunan inertia yang signifikan.

4.5. K-Medoids

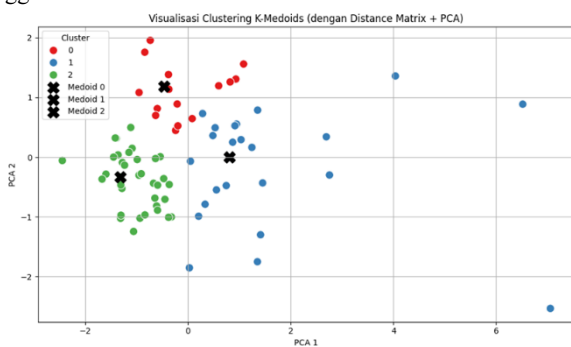
Setelah jumlah cluster ditetapkan, algoritma K-Medoids diaplikasikan untuk mengelompokkan dosen ke dalam 3 cluster. K-Medoids dipilih karena ketahanannya terhadap outlier yang

lebih baik dibanding K-Means. Proses dilakukan pada data ternormalisasi dan hasilnya divisualisasikan menggunakan PCA dalam dua dimensi untuk mempermudah interpretasi.

Implementasi K-Medoids dimulai dengan penghitungan jarak antar data menggunakan metrik Euclidean melalui fungsi `pairwise_distances` pada data terskalakan (`X_scaled`), menghasilkan `distance matrix` sebagai input utama. Setelah menentukan jumlah cluster $k=3$, dipilih titik awal (`initial_medoids`) sebagai medoid awal. Objek `kmedoids_instance` kemudian diproses untuk memperoleh hasil pengelompokan melalui `get_clusters()` dan medoid akhir melalui `get_medoids()`. Hasil clustering disimpan dalam DataFrame baru (`df_clustered`) dengan mengisi kolom "cluster" yang awalnya bernilai -1, kemudian diperbarui berdasarkan pengelompokan akhir, sehingga setiap data dosen memiliki label cluster sesuai kelompoknya.

4.5.1 Hasil Clustering K-Medoids

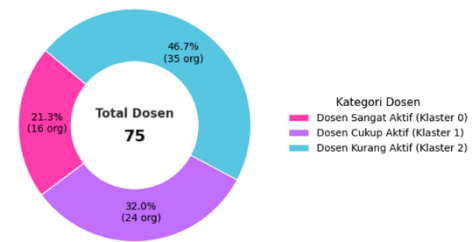
Berdasarkan penerapan algoritma K-Medoids, teridentifikasi tiga cluster dosen yang terbentuk berdasarkan empat variabel kinerja: jumlah pengajaran, jumlah publikasi, jumlah kegiatan PKM, dan jumlah penunjang tridharma. Hasil clustering tersebut berhasil mengelompokkan 75 dosen ke dalam tiga kategori yang merepresentasikan tingkat produktivitas dan kontribusi yang berbeda dalam pelaksanaan tridharma perguruan tinggi.



Gambar 7. Hasil Clustering K-Medoids

Visualisasi hasil clustering menggunakan algoritma K-Medoids dengan teknik Principal Component Analysis (PCA) menunjukkan tiga cluster berbeda yang diwakili oleh warna: merah (cluster 0), biru (cluster 1), dan hijau (cluster 2). Setiap cluster menampilkan titik besar berwarna hitam berbentuk "X" yang merepresentasikan medoid sebagai pusat cluster (Medoid 0, 1, dan 2). Hasil visualisasi menunjukkan data dalam setiap cluster memiliki kedekatan jarak dengan medoid-nya, mengindikasikan pembagian kelompok yang cukup baik. PCA berfungsi mereduksi dimensi data sehingga hasil clustering dapat divisualisasikan dalam dua dimensi berdasarkan komponen utama (PCA 1 dan PCA 2).

Distribusi Dosen Berdasarkan Hasil Clustering K-Medoids



Gambar 8. Distribusi Dosen Berdasarkan Hasil K-Medoids

Distribusi dosen berdasarkan hasil clustering K-Medoids menunjukkan pembagian 75 dosen ke dalam tiga kategori aktivitas. Cluster 2 (dosen kurang aktif) merupakan kelompok terbesar dengan 35 orang (46,7%), diikuti Cluster 1 (dosen cukup aktif) sebanyak 24 orang (32%), dan Cluster 0 (dosen sangat aktif) sebanyak 16 orang (21,3%). Distribusi ini mengindikasikan bahwa sebagian besar dosen (46,7%) masih berada pada tingkat aktivitas rendah, sementara hanya 21,3% yang tergolong sangat aktif dalam menjalankan tridharma perguruan tinggi.

4.5.2 Rata-rata Tiap Variabel per Cluster

Rata-rata tiap variabel per cluster berfungsi memberikan gambaran umum karakteristik setiap kelompok hasil clustering. Melalui perhitungan nilai rata-rata, dapat diidentifikasi profil dominan tiap cluster seperti skor jumlah pengajaran tinggi, jumlah publikasi banyak, keterlibatan PKM besar, atau jumlah penunjang yang banyak. Informasi ini mempermudah interpretasi, membandingkan perbedaan antar cluster, serta memberikan label kategori "sangat aktif", "cukup aktif", dan "kurang aktif" secara lebih objektif.

Tabel 2. Rata-rata Tiap Variabel per Cluster

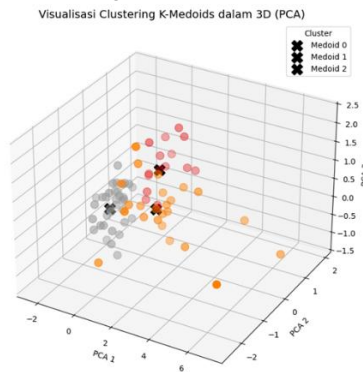
Cluster	Jumlah Pengajaran	Jumlah Publikasi	Jumlah PKM	Jumlah Penunjang
0	55,65	4,3125	3,75	4,8125
1	46,21625	10,20833333	6,333333333	15,70833333
2	27,40857143	4,285714286	2,657142857	4,628571429

Berdasarkan analisis rata-rata variabel per cluster, teridentifikasi karakteristik berbeda pada setiap kelompok. Cluster 1 mencatat nilai rata-rata tertinggi pada publikasi (10,21), PKM (6,33), dan penunjang (15,71), mengkonfirmasi kategori dosen sangat aktif. Cluster 0 menunjukkan rata-rata pengajaran tertinggi (55,65) dengan publikasi dan penunjang lebih rendah, merepresentasikan dosen cukup aktif yang berfokus pada pengajaran. Sementara Cluster 2 memiliki rata-rata terendah pada semua variabel, khususnya pengajaran (27,41) dan PKM (2,66), mengindikasikan dosen kurang aktif. Pola ini mempertegas perbedaan tingkat produktivitas antar cluster dalam pelaksanaan tridharma perguruan tinggi.

4.5.3 Visualisasi clustering K-Medoids

Visualisasi clustering merupakan teknik penyajian hasil pengelompokan data dalam bentuk grafik yang memungkinkan identifikasi pola, distribusi, dan hubungan antar cluster secara visual. Melalui visualisasi, dapat dievaluasi kualitas pembentukan cluster, tingkat pemisahan antar cluster, serta

karakteristik masing-masing cluster untuk mempermudah interpretasi hasil clustering.



Gambar 9. Visualisasi Clustering K-Medoids 3D

Visualisasi tiga dimensi dengan PCA menampilkan hasil pengelompokan K-Medoids dimana setiap titik mewakili satu data dosen dengan warna berbeda untuk setiap cluster: oranye (cluster pertama), merah (cluster kedua), dan abu-hitam (cluster ketiga). Tanda X berukuran besar menunjukkan posisi medoid sebagai pusat cluster yang dipilih langsung dari data asli. Pola sebaran menunjukkan ketiga cluster memiliki area berbeda meskipun terdapat sedikit tumpang tindih, yang merupakan kondisi wajar mengingat karakteristik data tidak terpisah secara sempurna, namun tetap menampilkan struktur pengelompokan yang jelas.

4.6. Evaluasi Cluster

4.6.1 Menggunakan DBI

Evaluasi hasil clustering menggunakan Davies-Bouldin Index (DBI) menghasilkan nilai 1,2627 yang tergolong dalam kategori cukup baik. Nilai DBI mengukur kualitas pemisahan antar cluster berdasarkan rasio antara jarak antar cluster dan penyebaran dalam cluster. Semakin kecil nilai DBI menunjukkan kualitas clustering yang semakin baik, dengan cluster yang homogen dan terpisah jelas.

Perhitungan dilakukan menggunakan fungsi `davies_bouldin_score` dari `scikit-learn` dengan input data ternormalisasi (`X_scaled`) dan label cluster. Hasil ini mengindikasikan bahwa algoritma K-Medoids telah membentuk kelompok yang cukup terpisah, meskipun masih terdapat kemiripan antar cluster yang relatif sedang. Model telah mampu mengelompokkan dosen berdasarkan karakteristik aktivitasnya dengan baik, namun masih terdapat potensi peningkatan kualitas pemisahan cluster.

4.6.2 Menggunakan Silhouette Score

Evaluasi menggunakan Silhouette Score menghasilkan nilai 0,3121 yang mengindikasikan kualitas pemisahan antar cluster dalam kategori cukup baik. Silhouette Score mengukur tingkat kerapatan data dalam cluster yang sama dan jaraknya terhadap cluster terdekat, dengan rentang nilai -1 hingga 1.

Nilai 0,3121 menunjukkan bahwa sebagian besar data telah dikelompokkan ke dalam cluster yang sesuai, meskipun masih terdapat beberapa data yang memiliki kemiripan dengan cluster lain. Hasil ini konsisten dengan evaluasi DBI sebelumnya, yang bersama-sama mengonfirmasi bahwa model K-Medoids telah berhasil mengelompokkan dosen berdasarkan karakteristik

aktivitasnya, walau masih terdapat ruang untuk peningkatan kualitas pemisahan cluster.

5. KESIMPULAN DAN SARAN

5.1. Kesimpulan

Berdasarkan Pengelompokan Kualitas Dosen Untuk Evaluasi Kinerja Akademik Dosen Menggunakan Algoritma K-Medoids terdapat beberapa kesimpulan, antara lain:

1. Algoritma K-Medoids berhasil mengelompokkan kualitas dosen IBTPI berdasarkan indikator penilaian yang digunakan, yaitu pendidikan & pengajaran, penelitian, pengabdian kepada masyarakat, dan penunjang tridharma.
2. Jumlah cluster optimal yang terbentuk adalah 3 berdasarkan metode Elbow, yaitu kategori: Sangat Aktif, Cukup Aktif, dan Kurang Aktif. Hal ini menunjukkan adanya perbedaan yang jelas dalam tingkat aktivitas kinerja dosen di IBTPI.
3. Hasil evaluasi menggunakan Davies-Bouldin Index sebesar 1,2627 dan Silhouette Score sebesar 0,3121 menunjukkan bahwa kualitas pemisahan cluster sudah cukup baik, namun masih terdapat beberapa data yang memiliki kemiripan antar cluster sehingga berpotensi ditingkatkan lagi.

5.2. Saran

Beberapa saran yang dapat dipertimbangkan :

1. Variabel penilaian perlu diperluas dengan menambahkan indikator lain seperti sertifikasi profesi, kepuasan mahasiswa, inovasi pembelajaran, dan keterlibatan organisasi untuk memperoleh hasil yang lebih komprehensif.
2. Frekuensi pengumpulan data sebaiknya dilakukan berkala, minimal setiap semester, agar pola perkembangan kualitas dosen dari waktu ke waktu dapat dianalisis lebih mendalam.
3. Penelitian selanjutnya dapat membandingkan algoritma K-Medoids dengan metode clustering lain, seperti K-Means, DBSCAN, atau Hierarchical Clustering, untuk memperoleh hasil pemodelan yang lebih optimal.

DAFTAR PUSTAKA

- [1] A. Zein, S. Farizy, and E. Suharyanto, "SENTIMEN ANALISIS PADA KOMENTAR PENDEK EVALUASI DOSEN OLEH MAHASISWA (EDOM) PROGRAM STUDI SISTEM INFORMASI," *J. Ilmu Komput.*, vol. V, no. 01, 2022.
- [2] M. Z. Naufal, M. P. Hadi, F. A. Ghifari, R. A. Alhakim, S. M. Pallawagau, and A. D. Kalifia, "Scientica Scientica," *J. Ilm. Sain dan Teknol.*, vol. 3, pp. 406–410, 2025, [Online]. Available: <https://jurnal.pranataindonesia.ac.id/index.php/jik/article/view/113%0Ahttps://jurnal.pranataindonesia.ac.id/index.php/jik/article/download/113/66>
- [3] A. Lisdiyanto, "Sistem Penilaian Kinerja Tridharma Dosen Menggunakan SAW," *J. Teknol. Dan Sist. Inf. Bisnis*, vol. 5, no. 1, pp. 69–72, 2023, doi:

- 10.47233/jteksis.v5i1.760.
- [4] A. A. D. Sulistyawati and M. Sadikin, "SISTEMASI: Jurnal Sistem Informasi Penerapan Algoritma K-Medoids untuk Menentukan Segmentasi Pelanggan," *Sist. J. Sist. Inf.*, vol. 10, no. 3, pp. 516–526, 2021, [Online]. Available: <http://sistemasi.ftik.unisi.ac.id>
- [5] I. Asmadi, G. Gopur, Y. Hendry, M. Ghofur, Y. Deswin, and N. Nurwanto, "Strategi Pengembangan Karir Dosen Sebagai Upaya dan Peningkatan Kualitas Pendidikan Tinggi," *JiIP - J. Ilm. Ilmu Pendidik.*, vol. 7, no. 5, pp. 5125–5129, 2024, doi: 10.54371/jiip.v7i5.4167.
- [6] E. Iip Syaripudin, D. Konkon Furkony, and S. Al-Musaddadiyah Garut, "Kualitas dan Kiprah Dosen PTKIS Sebagai Cendekiawan Ekonomi Islam," *J. NARATAS*, vol. 3, no. 2, pp. 1–8, 2021.
- [7] I. P. D. Suarnatha, "Sistem Pendukung Keputusan Penilaian Kinerja Dosen Menggunakan Metode Hybrid Ahp Dan Topsis," *J. Teknol. Dan Ilmu Komput. Prima*, vol. 5, no. 1, pp. 11–18, 2022, doi: 10.34012/jutikomp.v5i1.2579.
- [8] D. Apriadi, Nurul Hidayat, Nizhamuddin AB, Ahmatang, and Sudarto, "Kuliah Kerja Nyata: Pengabdian Kepada Masyarakat Melalui Kegiatan Pendampingan Pendidikan," *J. Pengabd. Masy. Paguntaka*, vol. 1, no. 1, pp. 25–30, 2022, doi: 10.61457/jumpa.v1i1.2.
- [9] I. Amal and I. Syah, "Peningkatan Kompetensi Profesional Dosen Melalui Pendidikan Dan Pelatihan Pada Akademik Keperawatan (AKPER) YAPENAS 21 Kabupaten Maros," vol. 2, no. 2, pp. 147–161, 2025.
- [10] T. Alfina and B. Santosa, "Analisa Perbandingan Metode Hierarchical Clustering, K-Means dan Gabungan Keduanya dalam Membentuk Cluster Data (Studi Kasus : Problem Kerja Praktek Jurusan Teknik Industri ITS)," *Anal. PerbandinganMetode Hierarchical Clust. K-means dan Gabungan Keduanya dalam Clust. Data*, vol. 1, no. 1, pp. 1–5, 2012.
- [11] Merliana, N. P. Eka., Ernawati., Santoso, and A. Joko, "Prosiding Seminar Nasional Multi Disiplin Ilmu & Call for Papers Unisbank (Sendi_U)," *Anal. Penentuan Jumlah Clust. Terbaik Pada Metod. K-Means Clust.*, pp. 978–979, 2015, [Online]. Available: <https://www.unisbank.ac.id/ojs/index.php/sendu/article/view/3333>
- [12] T. M. Kodinariya and P. R. Makwana, "Review on determining number of Cluster in K-Means Clustering. International Journal," *Int. J.*, vol. 1, no. 6, pp. 90–95, 2013.
- [13] S. Velamparambil, S. Mackinnon-Cormier, J. Perry, R. Lemos, M. Okoniewski, and J. Leon, "GPU accelerated Krylov subspace methods for computational electromagnetics," *Proc. 38th Eur. Microw. Conf. EuMC 2008*, vol. 1, no. 9, pp. 1312–1314, 2008, doi: 10.1109/EUMC.2008.4751704.
- [14] S. P. Dewi and E. Rahayu, "Penerapan Data Mining Untuk Prediksi Penjualan Produk Terlaris Menggunakan Metode K-Nearest Neighbor," *Build. Informatics, Technol. Sci.*, vol. 3, no. 4, pp. 639–648, 2022, doi: 10.47065/bits.v3i4.1408.
- [15] N. Ayuningtyas and F. M. Basysyar, "Penerapan Data Mining pada Penjualan Produk MS Glow Menggunakan Metode Naive Bayes untuk Strategi Pemasaran," vol. 5, no. 2, 2022.
- [16] S. Ramadhani, D. Azzahra, and T. Z., "Perbandingan Algoritma K-Means Dan K-Medoids Pada Text Mining untuk Klasifikasi Tesis Mahasiswa Berdasarkan Pengujian Davies Bouldin Index," *Digit. Zo. J. Teknol. Inf. dan Komun.*, vol. 13, no. 1, pp. 24–33, 2022.
- [17] S. F. Pane and Y. A. Saputra, *Big Data: Classification Behavior Menggunakan Python*, no. 54. in Data Science. Kreatif, 2020.
- [18] T. Saepurohman and B. E. Putro, "Analisis Principal Component Analysis (PCA) Untuk Mereduksi Faktor-Faktor yang Mempengaruhi Kualitas Kulit Kikil Sapi," *Semin. dan Konf. Nas. IDEC*, no. 2008, pp. 2–3, 2019.
- [19] R. Gelar Guntara, "Pemanfaatan Google Colab Untuk Aplikasi Pendeteksian Masker Wajah Menggunakan Algoritma Deep Learning YOLOv7," *J. Teknol. Dan Sist. Inf. Bisnis*, vol. 5, no. 1, pp. 55–60, 2023, doi: 10.47233/jteksis.v5i1.750.