Skema Pendanaan: PEMULA

LAPORAN PENELITIAN

Analisis Klaster Dosen Perguruan Tinggi Menggunakan Algoritma K-Means dan K-Medoids Berdasarkan Indikator Kinerja Akademik

TIM PENELITI

Ketua : Bayu Satria Pratama, S.Kom., M.Kom. (230001)

Anggota : Ir. Gatot Purwanto, M.M. (890004)

FAKULTAS/PUSAT STUDI

UNIVERSITAS BUDI LUHUR

Maret 2026

RINGKASAN

Perguruan tinggi merupakan institusi strategis dalam meningkatkan kualitas sumber daya manusia melalui Tri Dharma Perguruan Tinggi yang mencakup pengajaran, penelitian, dan pengabdian kepada masyarakat. Namun, evaluasi kinerja dosen menghadapi tantangan karena data yang beragam dan multidimensi, sehingga sulit dianalisis secara manual dan berpotensi menimbulkan bias. Penelitian ini bertujuan untuk melakukan pengelompokan dosen secara objektif menggunakan pendekatan *data mining* guna mendukung pengambilan keputusan strategis institusi.

Metodologi yang digunakan mengacu pada kerangka kerja CRISP-DM (*Cross Industry Standard Process for Data Mining*) yang terdiri dari enam tahapan, mulai dari *business understanding* hingga *deployment*. Penelitian ini membandingkan dua algoritma *clustering*, yaitu K-Means dan K-Medoids, dengan memanfaatkan indikator kinerja komprehensif seperti pengajaran, publikasi, pengabdian, Jabatan Akademik (JAD), Beban Kerja Dosen (BKD), dan skor SINTA. Data dinormalisasi menggunakan metode *Min-Max Scaling* dan kualitas klaster dievaluasi menggunakan *Davies-Bouldin Index* (DBI).

Hasil penelitian menunjukkan bahwa algoritma K-Means secara konsisten menghasilkan nilai DBI yang lebih rendah dibandingkan K-Medoids pada setiap variasi jumlah klaster. Jumlah klaster paling optimal ditemukan pada K=5 dengan nilai DBI sebesar 0,141 untuk K-Means. Klaster yang terbentuk berhasil memetakan profil dosen ke dalam lima kategori: *High Performers, Overburdened Achievers, Balanced Moderate, Critical Development,* dan *Underutilized*.

Kesimpulannya, K-Means merupakan algoritma terbaik untuk mengelompokkan data kinerja dosen dalam studi ini. Hasil klasterisasi ini dapat digunakan sebagai instrumen evaluasi internal untuk menyusun program pembinaan, pemberian insentif, serta optimalisasi pengembangan karier dosen demi memperkuat daya saing perguruan tinggi.

Kata Kunci: K-Means, K-Medoids, Kinerja Dosen, *Clustering*, CRISP-DM.

PRAKATA

Puji syukur dipanjatkan kehadiran Allah SWT, yang telah memberikan berkah dan hidayah-Nya sehingga kami memiliki kesempatan untuk melakukan penelitian dengan judul "Analisis Klaster Dosen Perguruan Tinggi Menggunakan Algoritma *K-Means* dan *K-Medoids* Berdasarkan Indikator Kinerja Akademik.".

Pada kesempatan ini tim penulis menyampaikan ucapan terima kasih kepada Universitas Budi Luhur terutama Direktorat Penelitian dan Pengabdian Masyarakat, pihak Fakultas Teknologi Informasi Universitas Budi Luhur yang telah mendukung penelitian ini. Tanpa bantuan dan dukungan ini sangat sulit bagi kami untuk dapat melakukan penelitian ini. Tim penulis telah berusaha untuk menyempurnakan tulisan ini, namun sebagai manusia kamipun menyadari akan keterbatasan maupun kekhilafan dan kesalahan yang tanpa kami sadari. Oleh karena itu, saran dan kritik untuk perbaikan laporan akhir ini akan sangat dinantikan.

Jakarta, Maret 2026

Penulis

DAFTAR ISI

DAFTAR TABEL

DAFTAR GAMBAR

DAFTAR LAMPIRAN

BAB I

PENDAHULUAN

1.1 Latar Belakang dan Rumusan Masalah

Perguruan tinggi merupakan institusi strategis dalam meningkatkan kualitas sumber daya manusia (SDM). Keberhasilan perguruan tinggi tidak terlepas dari kualitas dosen sebagai pelaksana Tri Dharma Perguruan Tinggi yang mencakup pengajaran, penelitian, dan pengabdian kepada Masyarakat[1]. Kinerja dosen menentukan mutu lulusan, produktivitas penelitian, serta kontribusi perguruan tinggi terhadap masyarakat dan daya saing bangsa[2].

Di era globalisasi dan perkembangan teknologi informasi, perguruan tinggi dituntut untuk bersaing di tingkat nasional maupun internasional. Hal ini menuntut peningkatan mutu tridharma secara berkelanjutan, yang diukur melalui indikator kinerja dosen seperti pengajaran, publikasi, pengabdian masyarakat, jabatan akademik (JAD), beban kerja dosen (BKD), dan skor SINTA[1]. Indikator-indikator ini mencerminkan peran dosen dalam menghasilkan inovasi, kontribusi ilmiah, serta akuntabilitas kinerja akademik.

Namun, evaluasi kinerja dosen menghadapi tantangan. Data yang digunakan sangat beragam, berasal dari berbagai sumber, dan memiliki dimensi berbeda. Misalnya, data pengajaran meliputi jumlah mata kuliah, sedangkan publikasi memperhitungkan kualitas jurnal dan indeksasi. Data pengabdian masyarakat bersifat variatif, sementara JAD dan BKD menunjukkan beban struktural, serta skor SINTA merepresentasikan produktivitas penelitian. Kompleksitas data ini sulit dianalisis secara manual dan berpotensi menimbulkan bias, sehingga diperlukan pendekatan berbasis data untuk pengelompokan dosen secara objektif[3].

Salah satu pendekatan yang relevan adalah analisis klaster (clustering), yang bertujuan mengelompokkan data berdasarkan kemiripan karakteristik[3,4]. Penelitian ini menerapkan dua algoritma populer, yaitu *K-Means* dan *K-Medoids*. *K-Means* membagi data berdasarkan *centroid*, sedangkan *K-Medoids* memilih medoid yang lebih tahan terhadap outlier [5–9]. Untuk mendukung proses penelitian, digunakan kerangka kerja CRISP-DM (*Cross Industry Standard Process for Data Mining*) yang mencakup tahapan sistematis mulai dari pemahaman masalah hingga evaluasi model[7].

Penelitian terdahulu yang dilakukan oleh Dahnial (2024) menggunakan metode *clustering K-Means* untuk mengelompokkan dosen berdasarkan publikasi jurnal. Sementara itu, Hendrawan (2023) bertujuan meningkatkan mutu dosen dengan memanfaatkan data evaluasi dosen sebagai dasar dalam proses clustering. Selanjutnya, Purwayoga (2021) melakukan pengelompokan dosen menggunakan metode *K-Means* berdasarkan penilaian mahasiswa terhadap kinerja dosen. Adapun Virgo (2020) mengelompokkan dosen menggunakan *K-Means* berdasarkan tingkat kedisiplinan dosen dalam kehadiran mengajar.

Namun, penelitian-penelitian tersebut masih memiliki keterbatasan karena hanya menggunakan satu algoritma klastering dan indikator kinerja yang bersifat parsial. Belum banyak penelitian yang melakukan analisis klaster dosen dengan mengombinasikan beberapa indikator kinerja akademik secara komprehensif serta membandingkan kinerja dua algoritma klastering yang berbeda, khususnya *K-Means* dan *K-Medoids*. Oleh karena itu, penelitian ini dilakukan untuk mengisi celah penelitian tersebut dengan menganalisis pengelompokan dosen perguruan tinggi menggunakan algoritma *K-Means* dan *K-Medoids* berdasarkan indikator kinerja akademik yang lebih menyeluruh, sehingga diharapkan dapat menghasilkan klaster dosen yang lebih akurat dan representatif. Hasil klasterisasi dapat menjadi instrumen evaluasi internal maupun eksternal untuk mendukung akreditasi dan daya saing perguruan tinggi.

Secara akademik, penelitian ini memperkaya literatur mengenai penerapan data mining dalam pengelolaan pendidikan tinggi, khususnya dalam membandingkan keunggulan *K-Means* dan *K-Medoids* pada data kinerja dosen. Evaluasi kualitas klaster menggunakan *Davies-Bouldin Index* (DBI) untuk memastikan validitas hasil[7,8].

Dengan demikian, penelitian ini tidak hanya menawarkan kontribusi teoretis berupa model klasterisasi kinerja dosen, tetapi juga kontribusi praktis dalam mendukung pengambilan keputusan strategis berbasis data. Hasilnya diharapkan membantu perguruan tinggi dalam meningkatkan kualitas tridharma, mengoptimalkan peran dosen, serta memperkuat daya saing di era global.

1.2 Pendekatan Pemecahan Masalah

1.2.1 Identifikasi Masalah dan Tujuan Penelitian

Penelitian diawali dengan mengidentifikasi permasalahan utama, yaitu kompleksitas dalam mengevaluasi kinerja dosen yang mencakup berbagai indikator: pengajaran, publikasi, pengabdian, jabatan akademik (JAD) dan beban kerja dosen (BKD). Tujuan yang ditetapkan adalah membentuk kelompok dosen yang memiliki karakteristik kinerja serupa untuk mendukung perencanaan SDM di perguruan tinggi.

1.2.2 Pengumpulan Data

Data yang digunakan diperoleh dari sumber resmi perguruan tinggi, meliputi data pengajaran, publikasi ilmiah, kegiatan pengabdian, JAD dan BKD. Data ini mencerminkan kinerja dosen secara komprehensif, baik dalam aspek tridharma maupun indikator administratif.

1.2.3 Praproses Data

Tahap ini dilakukan untuk memastikan kualitas data yang akan dianalisis.
Praproses meliputi:

- Pembersihan data (data cleaning) untuk mengatasi data kosong, duplikasi, atau inkonsistensi.
- Transformasi data agar seluruh variabel memiliki format yang seragam.
- Normalisasi data dengan metode Min-Max Scaling sehingga setiap indikator berada pada rentang nilai yang sama, guna menghindari dominasi satu variabel terhadap hasil klasterisasi.

1.2.4 Penerapan Algoritma Clustering

Dua algoritma clustering digunakan dalam penelitian ini.

1.2.5 Evaluasi Hasil Klasterisasi

Hasil pengelompokan dari kedua algoritma dievaluasi menggunakan DBI, evaluasi ini dilakukan untuk membandingkan performa K-Means dan K-Medoids serta menentukan metode yang lebih sesuai untuk data kinerja dosen.

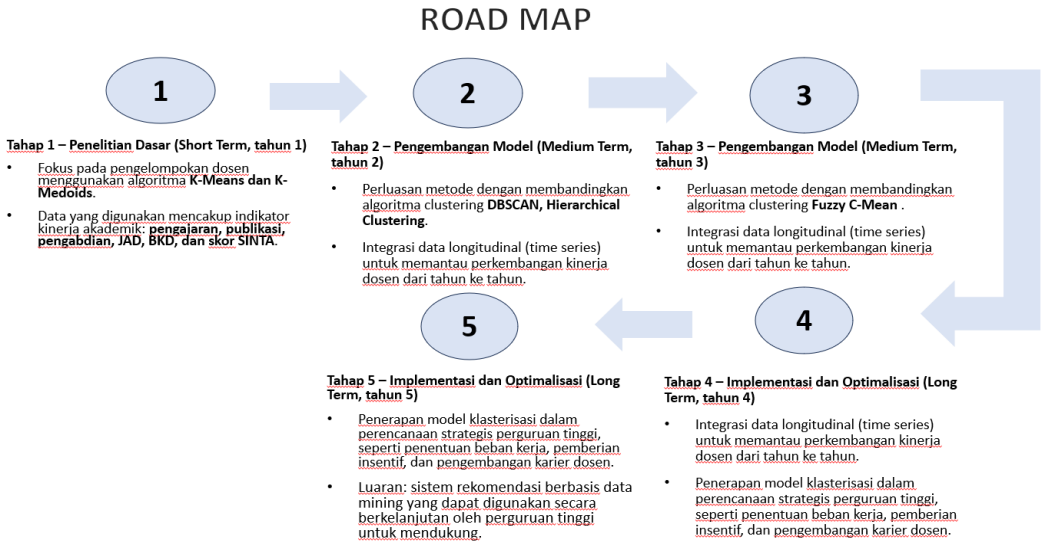
1.2.6 Interpretasi Hasil Klasterisasi

Setiap klaster yang terbentuk dianalisis untuk memahami karakteristik dosen yang termasuk di dalamnya.

1.3 *State Of The Art* dan Kebaruan

Penelitian sebelumnya masih terbatas karena umumnya hanya menggunakan satu algoritma klastering dan indikator kinerja yang parsial. Penelitian ini bertujuan mengisi celah tersebut dengan mengombinasikan indikator kinerja akademik yang lebih komprehensif serta membandingkan algoritma K-Means dan K-Medoids untuk menghasilkan klaster dosen yang lebih akurat dan representatif, sehingga dapat mendukung evaluasi kinerja, akreditasi, dan daya saing perguruan tinggi.

1.4 Peta Jalan Penelitian

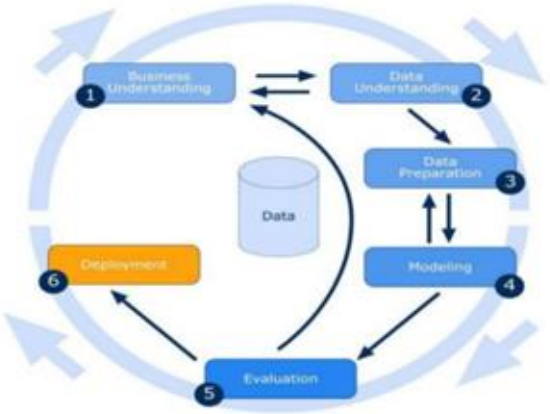


BAB II

METODE

2.1 *Crips-DM*

Metode penelitian ini dirancang untuk menjawab rumusan masalah dengan mengelompokkan dosen berdasarkan indikator kinerja akademik menggunakan algoritma *K-Means* dan *K-Medoids*. Pendekatan yang digunakan mengacu pada kerangka *CRISP-DM* (*Cross Industry Standard Process for Data Mining*), yang terdiri atas enam tahapan utama:



Gambar 1 CRISP-DM *(Cross Industry Standard Process for Data Mining)*

2.1.1 *Business Understanding*

Tahap awal adalah memahami permasalahan, yaitu kebutuhan evaluasi kinerja dosen secara komprehensif untuk mendukung perencanaan SDM perguruan tinggi. Indikator yang digunakan mencakup pengajaran, publikasi, pengabdian, Jabatan Akademik (JAD), Beban Kerja Dosen (BKD).

2.1.2 *Data Understanding*

Data diperoleh dari sistem informasi perguruan tinggi, laporan BKD. Aktivitas meliputi identifikasi sumber data, analisis kualitas (kelengkapan, konsistensi), serta eksplorasi deskriptif awal. Luaran tahap ini adalah dataset awal yang merepresentasikan kinerja dosen.

2.1.3 *Data Preparation*

Dataset kemudian diproses agar siap digunakan. Tahap ini meliputi data cleaning (menghapus duplikat, memperbaiki data kosong, mengatasi *outlier*), data integration dari berbagai sumber, transformasi data ke format numerik, serta normalisasi dengan *Min-Max Scaling* agar setiap variabel setara. Luaran berupa dataset final yang bersih dan terstandarisasi.

2.1.4 *Modeling*

Proses inti adalah penerapan algoritma clustering:

- *K-Means* : mengelompokkan data ke dalam k klaster dengan menghitung jarak Euclidean ke centroid.
- *K-Medoids* : mirip K-Means tetapi menggunakan medoid (data aktual) sebagai pusat, lebih tahan terhadap outlier.

Jumlah klaster ditentukan dengan metode Elbow dan hasil dievaluasi menggunakan *Davies-Bouldin Index* (DBI). Luaran adalah model clustering yang mampu mengelompokkan dosen sesuai indikator akademik.

2.1.5 *Evaluation*

Model dievaluasi dengan menilai nilai DBI dan interpretasi profil tiap klaster. Hasil analisis diharapkan menghasilkan kelompok dosen dengan ciri khas, misalnya dosen berfokus pengajaran, penelitian, atau kinerja seimbang. Luaran berupa evaluasi model dan deskripsi klaster dosen.

2.1.6 *Deployment*

Tahap akhir adalah implementasi hasil penelitian untuk mendukung kebijakan perguruan tinggi. Hasil klasterisasi digunakan sebagai dasar:

- Menyusun program pembinaan untuk dosen berkinerja rendah.
- Memberikan penghargaan atau insentif bagi dosen berprestasi.
- Mengatur strategi pengembangan SDM dan beban kerja dosen.

2.2 *K-Means*

Algoritma *K-Means* merupakan salah satu metode yang banyak digunakan dalam proses *clustering* untuk mengelompokkan data berdasarkan kemiripan karakteristiknya. Metode ini termasuk dalam kategori clustering non-hierarki, yang membagi dataset ke dalam beberapa klaster berbeda sehingga data dengan karakteristik serupa berada dalam satu klaster, sedangkan data dengan karakteristik berbeda ditempatkan di klaster lain[10–12]. Proses *K-Means* dimulai dengan menentukan jumlah klaster (k) yang diinginkan, kemudian memilih centroid awal sebagai pusat klaster. Selanjutnya, jarak antara setiap data dengan centroid dihitung menggunakan metode *Euclidean Distance*, dan data dikelompokkan ke dalam klaster dengan centroid terdekat.

Setelah pengelompokan awal, posisi *centroid* diperbarui dengan menghitung rata-rata seluruh anggota dalam masing-masing klaster, dan langkah ini dilakukan secara berulang (iteratif) hingga tidak ada perubahan signifikan pada centroid atau keanggotaan klaster telah stabil (konvergen). Rumus perhitungan jarak *Euclidean* ditunjukkan pada persamaan (1), perhitungan centroid baru sebagai rata-rata anggota klaster terdapat pada persamaan (2), dan proses penentuan keanggotaan klaster dijelaskan pada persamaan (3). Dengan demikian, algoritma *K-Means* mampu mengelompokkan data secara sistematis dan efektif berdasarkan kemiripan karakteristik setiap objek[13,14].

$$d(x+y) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}$$

$$C_k = \frac{1}{n_k}\sum d_i$$

$$c_j = (\frac{1}{n_j}) \sum_{x\ \epsilon\ Culster_j} x$$

2.3 K-Medoids

Algoritma *K-Medoids* merupakan metode *clustering* yang bekerja dengan prinsip serupa K-Means, namun menggunakan data aktual sebagai pusat klaster (*medoid*) sehingga lebih tahan terhadap nilai outlier. Proses dimulai dengan pemilihan medoid awal secara acak, kemudian setiap data dikelompokkan ke dalam klaster berdasarkan jarak *Manhattan* terdekat terhadap medoid yang telah ditentukan . Metode ini memungkinkan pengelompokan data dilakukan dengan memperhatikan kedekatan antar objek dalam bentuk jarak absolut, sehingga karakteristik klaster menjadi lebih representatif dibandingkan penggunaan rata-rata seperti *pada K-Means*[7,15].

Setelah pengelompokan awal, posisi *medoid* diperbarui dengan memilih kandidat medoid yang meminimalkan total jarak antar anggota dalam klaster. Langkah ini dilakukan secara iteratif hingga *medoid* mencapai kondisi stabil, di mana tidak ada perubahan signifikan pada posisi pusat klaster atau keanggotaan data dalam klaster . Perhitungan jarak Manhattan yang digunakan untuk menentukan kedekatan data dengan medoid ditunjukkan pada persamaan . Dengan demikian, algoritma *K-Medoids* mampu menghasilkan klaster yang lebih robust terhadap data yang ekstrem dan memberikan representasi yang lebih akurat dari struktur kelompok dalam dataset[7,15].

$$d(a_x, b_y) = \sum_{z=1}^{n} |a_z - b_z|$$

2.4 *Min-Max*

Metode *Min-Max* merupakan salah satu teknik normalisasi data yang digunakan untuk mengubah nilai asli ke dalam rentang tertentu, umumnya antara 0 hingga 1, melalui proses transformasi linier. Teknik ini bekerja dengan menyesuaikan setiap nilai berdasarkan nilai minimum (batas bawah) dan maksimum (batas atas) dari masing-masing atribut. Nilai hasil normalisasi diperoleh dengan mengurangkan nilai awal dengan nilai minimum, kemudian membaginya dengan selisih antara nilai maksimum dan minimum pada atribut tersebut. Rumus perhitungan normalisasi *Min-Max* ditunjukkan pada persamaan [16].

$$x = \frac{x_i - \min(x)}{\max(x) - \min(x)}$$

2.5. *Davies-Bouldin Index* (DBI)

Kualitas hasil klasterisasi dalam penelitian ini dievaluasi menggunakan metode *Davis-Bouldin Index* (DBI), yang merupakan salah satu ukuran untuk menilai performa clustering. Nilai DBI yang lebih rendah menunjukkan bahwa klaster yang terbentuk memiliki tingkat pemisahan yang baik antar klaster serta tingkat kemiripan yang tinggi di dalam klaster. Penentuan jumlah klaster yang optimal dilakukan dengan membandingkan nilai DBI dari beberapa percobaan, di mana jumlah klaster terbaik adalah yang menghasilkan nilai DBI paling kecil. Formula perhitungan Davis-Bouldin Index ditunjukkan pada persamaan [11,14].

$$DB = (\frac{1}{n}) \sum_{i=1}^{N} max_{j \neq i} (R_{ij})$$

BAB III

HASIL PELAKSANAAN PENELITIAN

3.1 *Business Understanding*

Perguruan tinggi berperan penting dalam meningkatkan kualitas sumber daya manusia melalui kinerja dosen dalam pelaksanaan Tri Dharma Perguruan Tinggi. Untuk menghadapi persaingan nasional dan internasional, peningkatan mutu dosen harus dilakukan secara berkelanjutan dengan mengacu pada berbagai indikator kinerja akademik yang kompleks dan multidimensi.

Namun, keberagaman data kinerja dosen dari berbagai sumber menyulitkan evaluasi secara manual dan beresiko bias. Oleh karena itu, diperlukan pendekatan analisis data yang objektif dan sistematis untuk mengelompokkan dosen berdasarkan kinerja akademik mereka. Dengan demikian, perguruan tinggi dapat mengambil keputusan strategis yang tepat untuk pengembangan kualitas dosen, peningkatan mutu pendidikan, dan pengelolaan sumber daya akademik secara optimal.

3.2 *Data Understanding*

Data diperoleh dari sistem informasi perguruan tinggi, khususnya laporan Beban Kerja Dosen (BKD) yang diunduh dari situs SISTER Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi (Kemendikbudristek). Aktivitas yang dilakukan meliputi identifikasi sumber data, analisis kualitas data (meliputi kelengkapan dan konsistensi), serta eksplorasi deskriptif awal. Hasil dari tahap ini adalah dataset awal yang merepresentasikan kinerja dosen homebase secara komprehensif.

3.3 *Data Preparation*

Dataset diproses terlebih dahulu agar siap digunakan, meliputi tahap pembersihan data seperti penghapusan duplikat, perbaikan data yang kosong, serta penanganan *outlier*. Selain itu, data dari berbagai sumber digabungkan, kemudian diubah ke dalam format numerik dan dinormalisasi menggunakan metode *Min-Max Scaling* agar semua variabel memiliki skala yang seimbang. Hasil proses ini berupa dataset akhir yang bersih dan terstandarisasi. Data yang diperoleh adalah sebagai berikut:

Tabel 1 Laporan BKD

NO.	JAD	A/B	Lebih A/B	C	Lebih C	D	Lebih D	E	Lebih E	Σ Kinerja	Σ Beban Lebih	Simpulan
1	Lektor Kepala	6.5	8	6.25	6.875	1.25	0	1.75	0	15.75	14.875	M
2	Lektor	6.3125	18.9375	7.5	0	2	0	0.125	0	15.9375	18.9375	M
3	Lektor	7	0	6.25	0	0.33	0	1	0	14.58	0	M
***	***	***	***	***	***	***	***	***	***	***	***	***
***	***	***	***	***	***	***	***	***	***	***	***	***

***	***	***	***	***	***	***	***	***	***	***	***	***
342	Asisten Ahli	0	0	0	0	0	0	0	0	0	0	TM
343	Asisten Ahli	6.8125	7.25	7.125	4	0.6417	0	0.5	0	15.0792	11.25	M
344	Asisten Ahli	10.75	4.75	3.75	0	1	0	0.375	0.5	15.875	5.25	M

Setelah dilakukan proses pembersihan data (data cleaning), yaitu dengan menghapus data dosen yang sedang dibebastugaskan, tahap selanjutnya adalah melakukan proses normalisasi data menggunakan metode *Min-Max*. Selain itu, variabel kategorikal seperti JAD dosen dan simpulan terlebih dahulu dikonversi menjadi nilai numerik. Hal ini dilakukan karena algoritma klastering seperti *K-Means* dan *K-Medoids* hanya dapat memproses data dalam bentuk angka.

Proses normalisasi *Min-Max* bertujuan untuk mengubah rentang nilai setiap atribut ke dalam skala yang sama, umumnya berada pada interval 0 hingga 1. Dengan penyamaan skala ini, setiap variabel memiliki kontribusi yang seimbang dalam proses perhitungan jarak, sehingga tidak ada atribut tertentu yang mendominasi hasil analisis klastering.

Setelah tahap normalisasi selesai, dataset yang telah diolah kemudian disiapkan sebagai input untuk proses pemodelan klastering. Pada tahap ini, data telah berada dalam kondisi yang lebih bersih, terstruktur, serta memiliki distribusi nilai yang seragam. Kondisi tersebut diharapkan dapat meningkatkan akurasi, stabilitas, dan kualitas hasil pengelompokan yang dihasilkan oleh algoritma klastering. Adapun hasil dari proses normalisasi data tersebut dapat dilihat pada Tabel 2.

Tabel 2 Tabel Dataset Normalisasi

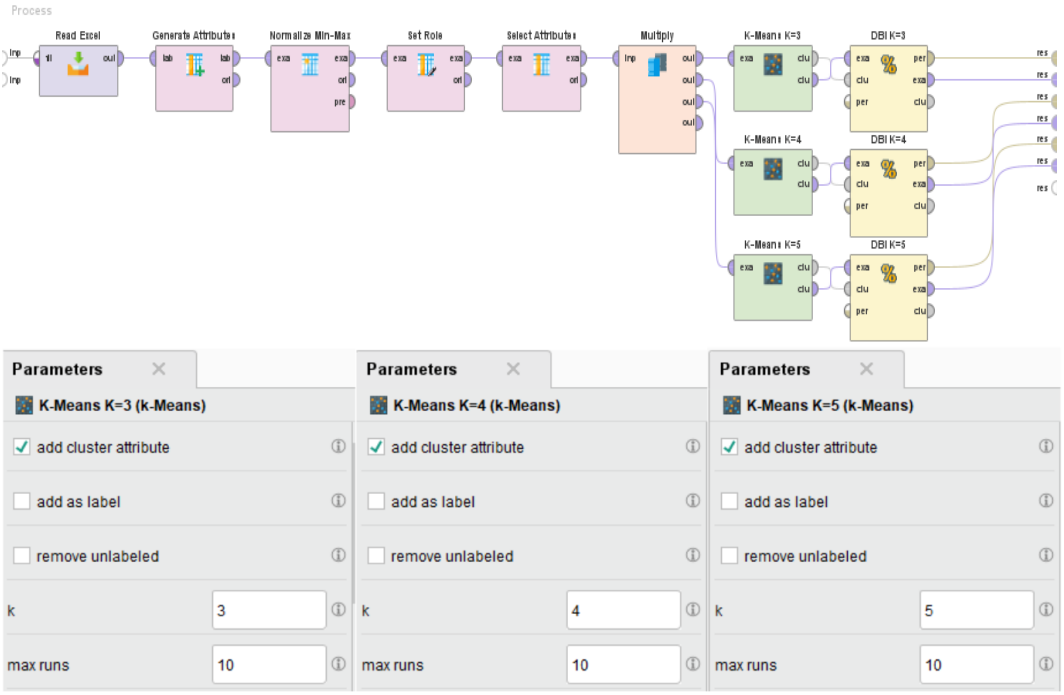
NO.	JAD	A/B	Lebih A/B	C	Lebih C	D	Lebih D	E	Lebih E	Σ Kinerja	Σ Beban Lebih	Simpulan
1	0.47	0.44	0.14	0.52	0.19	0.20	0.00	0.47	0.00	0.98	0.23	1.00
2	0.24	0.43	0.34	0.63	0.00	0.32	0.00	0.03	0.00	1.00	0.29	1.00
3	0.24	0.47	0.00	0.52	0.00	0.05	0.00	0.27	0.00	0.91	0.00	1.00
***	***	***	***	***	***	***	***	***	***	***	***	***
***	***	***	***	***	***	***	***	***	***	***	***	***
***	***	***	***	***	***	***	***	***	***	***	***	***
342	0.18	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
343	0.18	0.46	0.13	0.59	0.11	0.10	0.00	0.13	0.00	0.94	0.17	1.00
344	0.18	0.73	0.09	0.31	0.00	0.16	0.00	0.10	0.04	0.99	0.08	1.00

4.2 *Modeling*

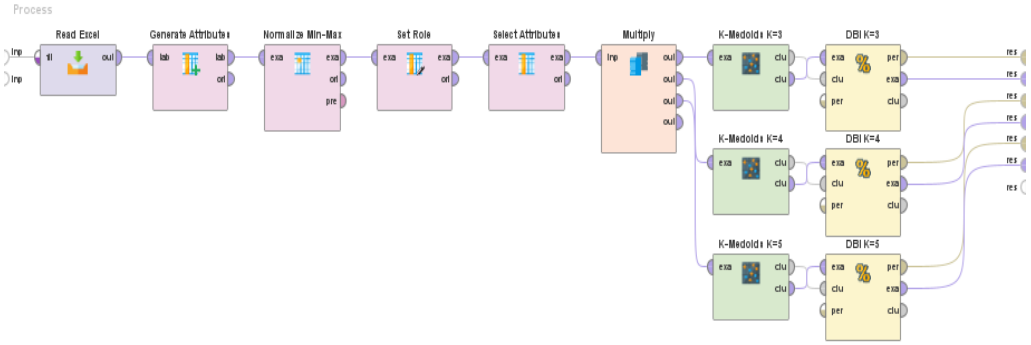
Pada tahap ini, dataset yang telah melalui proses pengumpulan, pembersihan (data *cleaning*), serta normalisasi, selanjutnya digunakan sebagai input dalam proses pemodelan (*modeling*). Tahap ini bertujuan untuk melakukan pengelompokan data

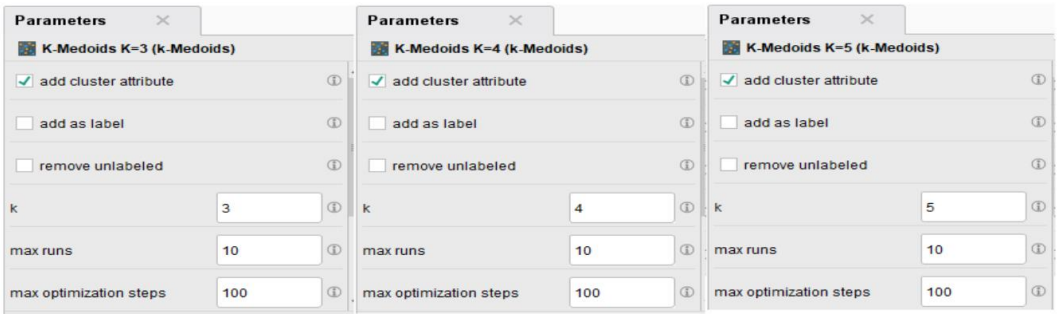
(*clustering*) berdasarkan kemiripan karakteristik antar dosen, khususnya yang berkaitan dengan indikator kinerja akademik.

Model yang digunakan dalam penelitian ini adalah algoritma *K-Means* dan *K-Medoids.* Kedua algoritma tersebut dipilih karena memiliki pendekatan yang berbeda dalam menentukan pusat klaster (*centroid/medoid*), sehingga memungkinkan dilakukan analisis perbandingan untuk memperoleh hasil klasterisasi yang optimal. Algoritma *K-Means* bekerja dengan menentukan titik pusat klaster berdasarkan rata-rata (mean) dari anggota klaster, sedangkan *K-Medoids* menggunakan objek data sebenarnya sebagai pusat klaster (*medoid*), sehingga cenderung lebih tahan terhadap *noise* dan *outlier*.



Gambar 2 Model *K-Means*





Gambar 3 *Model K-Medoids*

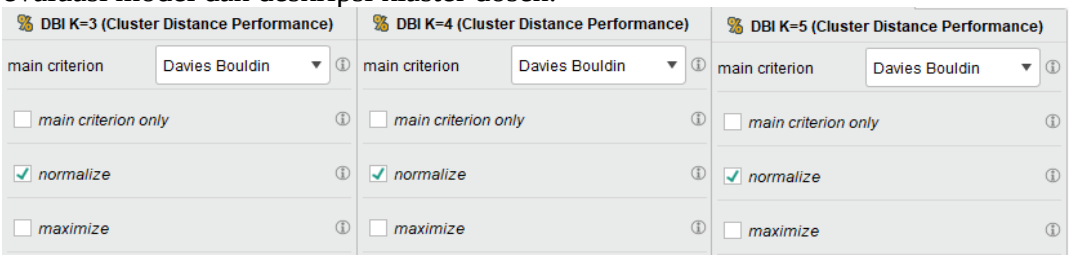
Pada proses ini, dataset akan diproses oleh masing-masing algoritma untuk membentuk sejumlah klaster yang telah ditentukan sebelumnya. Pengelompokan dilakukan dengan menghitung tingkat kedekatan atau jarak antar data berdasarkan atribut yang digunakan, sehingga data dosen dengan karakteristik yang serupa akan berada dalam satu klaster yang sama.

Hasil dari kedua metode tersebut kemudian akan dianalisis dan dibandingkan untuk mengetahui metode mana yang menghasilkan pengelompokan yang lebih akurat, stabil, dan representatif terhadap kondisi sebenarnya. Evaluasi ini dapat didasarkan pada kualitas klaster yang terbentuk, seperti tingkat homogenitas dalam klaster dan heterogenitas antar klaster.

Dengan demikian, tahap modeling ini menjadi bagian yang sangat penting dalam penelitian, karena berperan dalam mengungkap pola, kecenderungan, serta karakteristik kinerja akademik dosen secara lebih mendalam dan sistematis, yang selanjutnya dapat dijadikan dasar dalam pengambilan keputusan atau rekomendasi.

3.5 *Evaluation*

Model dievaluasi dengan menilai nilai DBI dan interpretasi profil tiap klaster. Hasil analisis diharapkan menghasilkan kelompok dosen dengan ciri khas, misalnya dosen berfokus pengajaran, penelitian, atau kinerja seimbang. Luaran berupa evaluasi model dan deskripsi klaster dosen.



Gambar 4 Evaluasi DBI

Setelah jumlah klaster ditentukan dalam tiga variasi, yaitu 2, 3, dan 5 klaster, proses klasterisasi kemudian dilakukan dengan menerapkan algoritma *K-Means* dan *K-Medoids* pada masing-masing variasi tersebut. Setiap skenario jumlah klaster diuji untuk melihat bagaimana pola pengelompokan data dosen terbentuk berdasarkan indikator yang digunakan. Hasil dari proses klasterisasi pada masing-masing algoritma selanjutnya dievaluasi menggunakan metode *Davis-Bouldin Index* (DBI) untuk mengukur kualitas klaster yang dihasilkan.

Evaluasi dilakukan dengan membandingkan nilai DBI dari setiap kombinasi algoritma dan jumlah klaster, di mana nilai DBI yang lebih kecil menunjukkan kualitas klaster yang lebih baik, baik dari segi homogenitas dalam klaster maupun separasi antar klaster. Dengan demikian, melalui proses ini dapat ditentukan jumlah klaster yang paling optimal serta algoritma yang memberikan performa terbaik dalam mengelompokkan data. Adapun hasil perhitungan nilai DBI dari seluruh percobaan tersebut disajikan secara rinci pada tabel.

Tabel 3 Hasil DBI

Jumlah Klaster	DBI K-Means	DBI K-Medoids
K=2	0.247	0.286
K=3	0.194	0.237
K=5	0.141	0.179

Berdasarkan hasil evaluasi menggunakan *Davis-Bouldin Index* (DBI) pada variasi jumlah klaster (K=2, K=3, dan K=5), dapat disimpulkan bahwa algoritma *K-Means* secara konsisten menghasilkan nilai DBI yang lebih kecil dibandingkan dengan *K-Medoids* pada setiap jumlah klaster. Hal ini menunjukkan bahwa *K-Means* memiliki kualitas pengelompokan yang lebih baik, ditinjau dari tingkat homogenitas dalam klaster dan separasi antar klaster.

Selain itu, nilai DBI terkecil diperoleh pada jumlah klaster K=5, yaitu sebesar 0,141 untuk *K-Means* dan 0,179 untuk *K-Medoids*. Dengan demikian, dapat disimpulkan bahwa jumlah klaster yang paling optimal dalam penelitian ini adalah 5 klaster, dengan algoritma terbaik adalah *K-Means* karena menghasilkan performa clustering yang paling optimal berdasarkan nilai DBI terendah.

Hasil pengelompokan data menggunakan algoritma K-Means dengan jumlah 3 klaster serta hasil pengelompokan menggunakan K-Medoids dengan jumlah klaster yang sama disajikan pada Tabel 4.

Tabel 4 Hasil Klaster K-Means & K-Medoids

No	Cluster K-Means	Cluster K-Medoids	JAD	***	***	***	Σ kinerja	ΣBeban Lebih	Simpulan
1	cluster_4	cluster_1	0.47	***	***	***	0.98	0.23	1.00
2	cluster_4	cluster_1	0.24	***	***	***	1.00	0.29	1.00
3	cluster_4	cluster_1	0.24	***	***	***	0.91	0.00	1.00
***	***	***	***	***	***	***	***	***	***
***	***	***	***	***	***	***	***	***	***
***	***	***	***	***	***	***	***	***	***
342	cluster_1	cluster_0	0.18	***	***	***	0.00	0.00	0.00
343	cluster_4	cluster_1	0.18	***	***	***	0.94	0.17	1.00
344	cluster_0	cluster_1	0.18	***	***	***	0.99	0.08	1.00

3.6 *Deployment*

Pada tahap ini, hasil pengelompokan dosen disampaikan kepada pihak yang berwenang, seperti pimpinan fakultas atau bagian akademik, sebagai dasar dalam merumuskan strategi peningkatan kualitas dan pengembangan dosen secara lebih terarah dan efektif. Informasi klaster dosen ini dapat digunakan untuk evaluasi kinerja, perencanaan pelatihan, serta pengambilan keputusan terkait pengelolaan sumber daya manusia di perguruan tinggi. Berikut penamaan klaster yang dapat dilihat pada Tabel 5.

Tabel 5 Penamaan Klaster

Klaster	**Nama Klaster**	**ΣKinerja**	**ΣBeban Lebih**	**Interpretasi**
Klaster 4	High Performers	Sangat Tinggi	Rendah	Efisien & Optimal
Klaster 3	Overburdened Achievers	Tinggi	Tinggi	Produktif tapi Berisiko
Klaster 2	Balanced Moderate	Sedang	Sedang	Stabil
Klaster 1	Critical Development	Rendah	Bervariasi	Perlu Intervensi
Klaster 0	Underutilized	Rendah/Sedang	Sangat Rendah	Potensi Terpendam

BAB IV KESIMPULAN

KESIMPULAN

4.1 Kesimpulan

Penelitian ini menyimpulkan bahwa penerapan algoritma *clustering* melalui kerangka kerja CRISP-DM mampu memberikan solusi objektif dalam mengevaluasi kinerja dosen yang kompleks dan multidimensi. Berdasarkan hasil pengujian, algoritma *K-Means* terbukti memiliki performa yang lebih unggul dan konsisten dibandingkan algoritma *K-Medoids* dalam mengelompokkan data kinerja akademik. Hal ini divalidasi melalui nilai *Davies-Bouldin Index* (DBI) yang secara konsisten lebih rendah pada setiap skenario jumlah klaster, yang menunjukkan tingkat homogenitas internal dan separasi antar kelompok yang lebih baik.

Akurasi pengelompokan yang paling optimal ditemukan pada formasi 5 klaster (K=5) dengan nilai DBI terendah sebesar 0,141. Melalui model ini, dosen berhasil dipetakan ke dalam lima kategori profil kinerja yang sangat spesifik, yaitu: *High Performers* (sangat tinggi dan efisien), *Overburdened Achievers* (produktif namun berisiko karena beban kerja tinggi), *Balanced Moderate* (stabil), *Critical Development* (perlu intervensi), serta *Underutilized* (potensi terpendam dengan beban kerja sangat rendah).

Hasil klasterisasi ini tidak hanya memberikan kontribusi teoretis pada literatur *data mining* di pendidikan tinggi, tetapi juga menjadi instrumen praktis bagi pimpinan perguruan tinggi. Dengan adanya pengelompokan yang akurat, institusi dapat mengambil keputusan strategis berbasis data, seperti pemberian insentif bagi dosen berprestasi, penyusunan program pembinaan yang tepat sasaran, hingga optimalisasi beban kerja untuk meningkatkan daya saing institusi di tingkat global.

4.2 Saran

Penelitian selanjutnya disarankan membandingkan algoritma clustering lain seperti DBSCAN, Hierarchical, atau Fuzzy C-Means untuk meningkatkan akurasi. Integrasi data longitudinal juga penting agar tren kinerja dosen dapat dipantau dari waktu ke waktu. Secara praktis, pengembangan sistem rekomendasi berbasis data mining yang terhubung dengan sistem informasi perguruan tinggi dapat mendukung keputusan strategis, seperti penentuan beban kerja, insentif, dan rencana karier dosen, sehingga optimalkan SDM dan daya saing institusi.

LAMPIRAN

Lampiran 1 Realisasi Penggunaan Anggaran

Dana Disetujui: Rp 5.000.000

Jenis Pembelajaan	Komponen	Item	Kuantitas	Biaya Satuan	Total
Belanja Bahan	ATK	Pulpen, Kertas, Map, Dll	1 Paket	Rp 250.000	Rp 250.000
Belanja Bahan	Bahan penelitian (habis pakai)	Bahan habis pakai	1 Paket	Rp 400.000	Rp 400.000
Pengumpulan Data	Honor pembantu peneliti	1 orang x 5 har	5 hari	Rp 80.000	Rp 400.000
Pengumpulan Data	FGD	Sewa tempat + konsumsi	1 kali	Rp 600.000	Rp 600.000
Pengumpulan Data	Konsumsi	Snack & makan saat survei	5 kali	Rp 80.000	Rp 400.000
Pengumpulan Data	Transport	Transportasi lapangan	5 kali	Rp 100.000	Rp 500.000
Pengumpulan Data	Penginapan	1 orang x 2 malam	2 malam	Rp 200.000	Rp 400.000
Analisis Data	Honor pengolah data	Data entry & analisis	1 orang	Rp 600.000	Rp 600.000
Analisis Data	Honor narasumber	Expert input / validasi data	1 orang	Rp 350.000	Rp 350.000
Lainnya	Biaya pendaftaran HKI & Publikasi	Pengetikan & layout laporan	1 HKI	Rp 1.100.000	Rp 1.100.000

Lampiran 2 Biodata dan Anggota Tim Peneliti

A. Identitas Diri

1.	Nama Lengkap (dengan gelar)	:	Bayu Satria Pratama, S.Kom., M.Kom.
2.	Jenis Kelamin	:	Laki-Laki
3.	Jabatan Fungsional	:	Tenaga Pengajar
4.	NIP/NIDN/ID-SINTA	:	230001/ 9846770671130352/ 6923113
5.	Tempat, Tanggal Lahir	:	Jakarta, 14 Mei 1992
6.	E-mail	:	bayupratama@budiluhur.ac.id
7.	Nomor Handphone	:	087771315750
8.	Alamat	:	Jl. Karyawan 2 RT. 003/009, Kec. Karang Tengah, Kota Tangerang, Banten.

B. Riwayat Pendidikan

Nama Perguruan Tinggi	S1	S2
	Universitas Budi Luhur	Universitas Budi Luhur
Bidang Ilmu	Ilmu Komputer	Ilmu Komputer
Tahun Masuk-Lulus	2010-2014	2022-2024

C. Pengalaman Penelitian (5 Tahun Terakhir)

No.	Tahun	Judul Penelitian	Pendanaan	
			Sumber	Jumlah (Rp.)
1.	2023	Algoritme K-Means Untuk Klasterisasi Lulusan Perguruan Tinggi Berbasis Hasil Studi Pelacakan	Universitas Budi Luhur	15.000.000
2.	2025	Perbandingan K-Means dan K-Medoids dalam Pengelompokkan Tingkat Kejahatan pada Provinsi Jawa Tengah	Mandiri	1.000.000

D. **Publikasi Artikel Ilmiah dalam Jurnal (5 Tahun Terakhir)**

No.	Judul	Nama Jurnal	Volume/Nomor/Tahun
1.	Algoritme K-Means Untuk Klasterisasi Lulusan Perguruan Tinggi Berbasis Hasil Studi	J-ICON: Jurnal Komputer dan Informatika	Volume 11 Nomor 2 Tahun 2023

	Pelacakan		
2.	Perbandingan K-Means dan K-Medoids dalam Pengelompokkan Tingkat Kejahatan pada Provinsi Jawa Tengah	Idealis: Indonesia Journal Information System	Volume 8 Nomor 2 Tahun 2025

E. **Pemakalah Seminar Ilmiah (5 Tahun Terakhir)**

No.	Nama Temu Ilmiah/Seminar	Judul Artikel Ilmiah*	Waktu dan Tempat
1.			
2.			

F. **Perolehan HKI (5 Tahun Terakhir)**

No.	Judul/Tema HKI*	Tahun	Jenis	Nomor P/ID
1.	Algoritme K-Means Untuk Klasterisasi Lulusan Perguruan Tinggi Berbasis Hasil Studi Pelacakan	2023	Program Komputer	EC00202365045
2.	Pengembangan Pembelajaran Visual pada Smartphone untuk Disabilitas Tuna Rungu	2025	Hak cipta nasional	EC002026024397

Jakarta, 24 Maret 2026

Pengusul

(Bayu Satria Pratama, S.Kom., M.Kom.)

Lampiran 3 Catatan Harian

No	Tanggal	Kegiatan
1	Sep-25	**Catatan: Business Understanding & Literasi**
		Melakukan studi literatur mengenai indikator kinerja dosen (Tri Dharma, JAD, SINTA) dan koordinasi awal tim peneliti untuk penyamaan persepsi.
2	Sep-25	**Catatan: Belanja Bahan & ATK**
		Pembelian bahan habis pakai dan ATK (pulpen, kertas, map) untuk menunjang administrasi dan dokumentasi penelitian.
3	Oktober 2025	**Catatan: Data Understanding & Pengumpulan Data**
		Melakukan survei lapangan, pengumpulan data BKD dan skor SINTA, serta koordinasi dengan pembantu peneliti untuk entri data awal.
4	Oktober 2025	**Catatan: Focus Group Discussion (FGD)**
		Melaksanakan FGD bersama pihak terkait di tempat sewa yang telah dianggarkan untuk validasi awal variabel penelitian.
5	Nov-25	**Catatan: Data Preparation**
		Melakukan pembersihan data (*cleaning*), transformasi, dan normalisasi menggunakan *Min-Max Scaling* agar data siap diolah oleh algoritma.
6	Desember 2025	**Catatan: Modeling (Proses Clustering)**
		Menjalankan algoritma K-Means dan K-Medoids menggunakan perangkat lunak pengolah data untuk mencari perbandingan nilai DBI terbaik.
7	Desember 2025	**Catatan: Evaluation bersama Narasumber**
		Pertemuan dengan narasumber ahli untuk validasi hasil klaster (K=5) dan memastikan pengelompokan sesuai dengan kondisi riil di lapangan.
8	Januari 2026	**Catatan: Analisis Data & Penyusunan Laporan**

		Finalisasi analisis data dan penyusunan draf laporan akhir penelitian berdasarkan hasil klasterisasi yang telah divalidasi.
9	Februari 2026	**Catatan: Deployment**
		Penyusunan artikel publikasi ilmiah dan pendaftaran Hak Kekayaan Intelektual (HKI) atas model/sistem pengelompokan yang dihasilkan.
10	Februari 2026	**Catatan: Pelaporan Keuangan & Administrasi**
		Finalisasi seluruh laporan keuangan (SPJ) dan penyerahan laporan akhir kepada pihak institusi.