Teknika 20 (2): 311-322

JURNAL TEKNIKA
ISSN: 0854-3143
e-ISSN: 2622-3481
Journal homepage: http://jurnal.polsri.ac.id/index.php/teknika
Journal Email: teknika@polsri.ac.id

Analisis Deskriptif Data SKP Pegawai Kominfo Menggunakan Algoritma *K-Means Clustering* untuk Identifikasi Pola Kinerja

Alan Marta Dinata[1], Melda Agarina*[2]

[*1,2] Program Studi Sistem Informasi, Institut informatika dan Bisnis Darmajaya, Bandar Lampung, Indonesia
*Email Penulis Korespondensi: **meldaagarina@darmajaya.ac.id**

Abstrak

Perkembangan teknologi informasi mendorong peningkatan tata kelola pemerintahan, termasuk pengelolaan kinerja ASN. Kominfo telah memiliki data Sasaran Kinerja Pegawai (SKP), namun pemanfaatannya masih terbatas pada penilaian deskriptif tanpa analisis mendalam. Penelitian ini bertujuan mengidentifikasi pola kinerja pegawai Kominfo menggunakan algoritma K-Means Clustering dalam kerangka CRISP-DM. Data yang digunakan terdiri atas 404 pegawai dengan 10 variabel kinerja seperti disiplin, produktivitas, kualitas kerja, orientasi pelayanan, komitmen, kerja sama, dan nilai SKP. Analisis dilakukan menggunakan Python pada Google Colaboratory. Penentuan jumlah klaster optimal menggunakan Elbow Method menunjukkan titik siku yang jelas pada k=5, ditandai oleh penurunan Within-Cluster Sum of Squares (WCSS) yang mulai stabil setelah nilai tersebut. Evaluasi kualitas klaster melalui Silhouette Coefficient juga mendukung pemilihan lima klaster. Hasil penelitian mengungkap bahwa kelima klaster tersebut merepresentasikan variasi tingkat kinerja mulai dari tinggi hingga rendah, dengan mayoritas pegawai berada pada kategori sedang hingga baik. Disiplin dan kualitas kerja menjadi indikator paling konsisten, sedangkan komitmen dan kerja sama muncul sebagai variabel pembeda utama antar klaster. Penelitian ini membuktikan bahwa algoritma K-Means efektif dalam memetakan kinerja pegawai secara objektif dan berbasis data. Temuan ini dapat digunakan sebagai dasar penyusunan strategi pembinaan serta pengembangan SDM untuk meningkatkan produktivitas organisasi.

Kata kunci— *Clustering, Data Mining, K-Means, Kinerja Pegawai, Kominfo.*

Abstract

Advances in information technology have driven improvements in governance, including the management of civil servant performance. The Ministry of Communication and Information Technology already has Employee Performance Target (SKP) data, but its use is still limited to descriptive assessments without in-depth analysis. This study aims to identify patterns of Kominfo employee performance using the K-Means Clustering algorithm within the CRISP-DM framework. The data used consists of 404 employees with 10 performance variables such as discipline, productivity, work quality, service orientation, commitment, cooperation, and SKP

Received: November 14, 2025; Online Available: January 4, 2026

values. The analysis was performed using Python on Google Colaboratory. Determining the optimal number of clusters using the Elbow Method shows a clear elbow point at k=5, marked by a decrease in the Within-Cluster Sum of Squares (WCSS) that begins to stabilise after that value. Evaluating cluster quality through the Silhouette Coefficient also supports the selection of five clusters. The results of the study reveal that the five clusters represent variations in performance levels from high to low, with the majority of employees falling into the moderate to good category. Discipline and work quality are the most consistent indicators, while commitment and cooperation emerge as the main differentiating variables between clusters. This study proves that the K-Means algorithm is effective in mapping employee performance objectively and based on data. These findings can be used as a basis for formulating coaching strategies and human resource development to improve organizational productivity.

Keywords— *Clustering, Data Mining, K-Means, Employee Performance, Ministry of Communication and Information Technology.*

1. PENDAHULUAN

Perkembangan teknologi informasi di era digital telah mendorong transformasi besar dalam tata kelola pemerintahan, termasuk dalam bidang manajemen sumber daya manusia. Pemerintah Indonesia berupaya memperkuat efektivitas dan efisiensi kinerja aparatur sipil negara (ASN) melalui penerapan sistem pemerintahan berbasis elektronik (SPBE) sebagaimana tertuang dalam Peraturan Presiden Nomor 95 Tahun 2018. Salah satu komponen penting dalam pelaksanaan SPBE adalah pengelolaan data kinerja pegawai yang lebih terintegrasi, transparan, dan berbasis data [1], [2].

Penilaian kinerja aparatur negara selama ini dilakukan melalui Sasaran Kinerja Pegawai (SKP) yang menjadi instrumen utama dalam mengukur capaian kinerja individu berdasarkan tugas dan fungsi jabatan [3], [4]. Namun, pengelolaan dan analisis data SKP masih sering dilakukan secara manual dan administratif, tanpa dukungan sistem analisis data yang optimal. Kondisi ini mengakibatkan hasil evaluasi kinerja seringkali belum menggambarkan kondisi riil pegawai, serta belum dimanfaatkan secara maksimal dalam pengambilan keputusan terkait pembinaan dan pengembangan sumber daya manusia di instansi pemerintah [1], [5], [6].

Dalam konteks inilah, penerapan data mining menjadi relevan sebagai pendekatan analitis untuk menemukan pola tersembunyi dalam data kinerja pegawai. Salah satu metode yang banyak digunakan adalah algoritma *K-Means Clustering*, yang termasuk ke dalam pendekatan *unsupervised learning* dan mampu mengelompokkan data berdasarkan kemiripan karakteristik [7], [8], [9]. Pemilihan algoritma *K-Means* dalam penelitian ini didasarkan pada beberapa pertimbangan akademis. Pertama, *K-Means* memiliki kemampuan komputasi yang efisien untuk data berukuran menengah hingga besar, menjadikannya sesuai untuk memproses data SKP pegawai dalam jumlah ratusan. Kedua, algoritma ini menghasilkan klaster yang mudah diinterpretasikan karena berbasis pada *centroid*, sehingga memudahkan manajemen dalam membaca pola kinerja yang terbentuk. Ketiga, *K-Means* bekerja optimal pada data numerik terstandarisasi seperti indikator kinerja pegawai, yang memang memiliki skala penilaian yang seragam. Selain itu, dibandingkan metode lain seperti *K-Medoids*, DBSCAN, atau *Fuzzy C-Means*, *K-Means* umumnya memberikan hasil yang stabil dan lebih sederhana untuk diterapkan pada konteks organisasi pemerintah yang membutuhkan analisis cepat dan implementatif [10],[11].

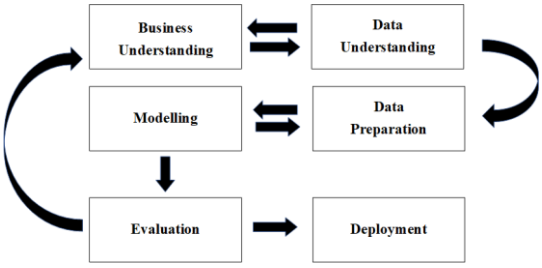
Penelitian sebelumnya menunjukkan efektivitas *K-Means* dalam mengelompokkan data kepegawaian pada berbagai instansi pemerintah, termasuk untuk analisis kepadatan pegawai dan pengelompokan tingkat kinerja. Meskipun demikian, penelitian yang secara spesifik menerapkan algoritma ini pada data SKP pegawai di lingkungan Kementerian Komunikasi dan Informatika (Kominfo) masih sangat terbatas. Padahal, analisis *clustering* pada data SKP dapat memberikan manfaat signifikan, di antaranya mengidentifikasi pola kinerja pegawai, menentukan kelompok berkinerja tinggi atau rendah, dan mendukung kebijakan pembinaan yang lebih tepat sasaran [12],

[13].

Berdasarkan kebutuhan tersebut, penelitian ini berjudul "Analisis Deskriptif Data SKP Pegawai Kominfo Menggunakan Algoritma *K-Means Clustering* untuk Identifikasi Pola Kinerja". Tujuan penelitian ini adalah menganalisis data SKP pegawai Kominfo menggunakan algoritma *K-Means* untuk mengidentifikasi pola kinerja, sehingga dapat menjadi dasar pengambilan keputusan yang lebih objektif dan berbasis data dalam pengelolaan sumber daya manusia.

2. METODE PENELITIAN

Penelitian ini menerapkan metode Data *Mining* dengan mengacu pada tahapan CRISP-DM (*Cross-Industry Standard Process for Data Mining*) seperti terlihat pada Gambar 1, yang terdiri atas enam langkah utama dan digambarkan melalui rancangan penelitian pada Gambar 1[14].



Gambar 1 Tahapan CRISP DM

2.1 Business Understanding

Tahap ini bertujuan untuk memahami permasalahan dan tujuan bisnis yang akan dicapai melalui analisis data SKP pegawai di Kementerian Komunikasi dan Informatika (Kominfo). Permasalahan utama yang diangkat yaitu belum optimalnya pemanfaatan data SKP dalam mengidentifikasi pola kinerja pegawai. Analisis ini dilakukan untuk memberikan gambaran objektif terkait tingkat kinerja pegawai berdasarkan indikator seperti disiplin, produktivitas, kualitas kerja, orientasi pelayanan, komitmen, dan kerja sama. Metode analisis yang digunakan adalah data *mining* dengan pendekatan *unsupervised learning*, menggunakan algoritma *K-Means Clustering* untuk mengelompokkan pegawai berdasarkan pola kinerjanya. Lingkungan kerja yang digunakan adalah *Google Colaboratory*, karena mendukung integrasi *Python* dan pustaka analisis data seperti Pandas, *Matplotlib*, dan *Scikit-learn*.

2. 2 Data Understanding

Tahap ini mencakup pemahaman struktur data 404 pegawai dengan 10 variabel melalui *Exploratory Data Analysis* (EDA), yang meliputi visualisasi distribusi variabel dengan *histogram* dan boxplot, identifikasi *outlier*, analisis korelasi antar fitur, serta pemeriksaan kualitas data seperti *missing* values, duplikasi, dan konsistensi nilai. Langkah ini memastikan bahwa data berada dalam kondisi layak untuk tahap pra-pemrosesan dan rekayasa fitur berikutnya.

2. 3 Data Preparation

Pada tahap ini dilakukan pembersihan data serta *feature engineering* untuk memastikan variabel yang digunakan akurat dan informatif. Proses meliputi penanganan *missing values*, penghapusan duplikasi, serta pemeriksaan dan pengelolaan *outlier*. Selanjutnya dilakukan pemilihan fitur berdasarkan EDA dan korelasi, penghapusan fitur yang redundan, serta transformasi atau normalisasi variabel. Tahap ini diakhiri dengan scaling menggunakan *StandardScaler* atau *MinMaxScaler* untuk menyeragamkan skala fitur sebelum proses

klasterisasi. Seluruh langkah ini menjadi fondasi penting agar algoritma *K-Means* menghasilkan klaster yang optimal dan valid.

2. 4 Modelling

Pada tahap ini, sebelum proses pemodelan dilakukan, perlu ditentukan terlebih dahulu jumlah kluster yang paling optimal. Penentuan jumlah kluster dapat dilakukan menggunakan metode *Elbow*, yaitu teknik yang membantu menentukan jumlah kluster terbaik dengan melihat titik "siku" pada grafik antara jumlah klaster (k) dan nilai *within-cluster sum of squares* (WCSS). Titik siku tersebut menunjukkan jumlah klaster yang efisien, di mana penambahan klaster berikutnya tidak memberikan peningkatan signifikan terhadap hasil [15], [16]. Setelah jumlah klaster optimal diperoleh, tahap selanjutnya adalah melakukan pemodelan klasterisasi menggunakan algoritma *K-Means*. Proses K-Means berupaya meminimalkan *Within-Cluster Sum of Squares (WCSS)* dengan rumus (1).

$$WCSS = \sum_{k=1}^{K} \sum_{i \in C_k} |\, x_i - \mu_k|^2 \quad (1)$$

di mana:

- K = jumlah klaster,
- C_i = klaster ke-i,
- x = data pada klaster,
- μ_i = centroid klaster ke-i,
- $||x - \mu_k||$ = jarak *Euclidean* antara data dan *centroid*.

Algoritma *K-Means* terdiri atas beberapa tahapan sistematis berikut:

1. Inisialisasi, pilih jumlah klaster kkk dan tentukan centroid awal (secara acak atau metode *k-means++*).
2. *Assignment Step, s*etiap data dihitung jaraknya ke semua centroid menggunakan jarak Euclidean, lalu ditempatkan pada klaster dengan jarak terdekat seperti tersaji pada Rumus (2).

$$C_i = \{x: |x - \mu_i|^2 \leq |x - \mu_j|^2,\ j \neq i\} \quad (2)$$

3. *Update Step*, hitung ulang posisi centroid dengan menghitung rata-rata semua data dalam klaster seperti tersaji pada Rumus (3).

$$\mu_i = \frac{1}{|C_i|} \sum_{x \in C_i} x \quad (3)$$

4. Iterasi, ulangi langkah *Assignment–Update* hingga centroid stabil atau perubahan WCSS sangat kecil.
5. Konvergensi, algoritma berhenti ketika klaster tidak berubah atau perubahan nilai WCSS sudah minimal.

2.5 Evaluation

Pada tahap ini dilakukan evaluasi kualitas model menggunakan metrik *Silhouette Coefficient*. Metode ini digunakan untuk menentukan jumlah klaster yang paling optimal pada data berskala rasio. Saat diterapkan, algoritma *Silhouette* menghitung rata-rata jarak suatu data terhadap data lain dalam klaster yang sama dan antar klaster berbeda. Nilai indeks yang mendekati 1 menunjukkan hasil klasterisasi yang kuat atau baik, sedangkan nilai yang mendekati 0 menandakan bahwa proses klasterisasi kurang optimal [14].

Tabel 1 Interpretasi *Silhouette Coefficient*

Silhouette Coefficient	Interpretasi
0.71 – 1.00	*Cluster* Kuat
0.51 – 0.70	*Cluster* Layak/Sesuai

0.26 – 0.50	*Cluster* Lemah
<0.25	Tidak Dapat Disebut *Cluster*

2.6 Deployment

Menyusun laporan akhir yang memaparkan seluruh proses dan hasil penelitian, mencakup tahapan mulai dari pemilihan data, metode yang diterapkan, hingga pembahasan hasil serta kesimpulan akhir penelitian.

3. HASIL DAN PEMBAHASAN

Hasil penelitian ini berupa proses *clustering* yang dilakukan berdasarkan tahapan CRISP-DM dengan menerapkan algoritma *K-Means*, menggunakan *Google Colab* sebagai platform dan *Python* sebagai bahasa pemrograman.

3.1 Business Understanding

Penelitian ini bertujuan meningkatkan efektivitas manajemen kinerja pegawai di Kementerian Komunikasi dan Informatika (Kominfo) melalui pemanfaatan data Sasaran Kinerja Pegawai (SKP) yang selama ini belum dianalisis secara mendalam. Penilaian kinerja yang masih bersifat deskriptif dan subjektif menyulitkan manajemen dalam mengidentifikasi kategori kinerja pegawai secara objektif. Metode *Clustering* dengan algoritma *K-Means* digunakan untuk mengelompokkan pegawai berdasarkan indikator kinerja seperti disiplin, produktivitas, kualitas kerja, orientasi pelayanan, komitmen, kerja sama, dan nilai SKP. Hasil pengelompokan diharapkan dapat membantu manajemen memahami pola kinerja pegawai serta mendukung pengambilan keputusan dalam pembinaan dan pengembangan sumber daya manusia secara lebih tepat dan berbasis data.

3.2 Data Understanding

Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari Kementerian Komunikasi dan Informatika (Kominfo) berupa data Sasaran Kinerja Pegawai (SKP). Data tersebut berisi informasi mengenai kinerja pegawai berdasarkan sejumlah indikator penilaian. Variabel-variabel yang digunakan dalam penelitian ini ditunjukkan pada Tabel 2.

Tabel 2 Variabel Penelitian

Kode Variabel	**Keterangan**
NIP	Nomor Induk Pegawai
Nama_Pegawai	Nama Pegawai
Disiplin	Tingkat kedisiplinan dalam bekerja
Produktivitas	Kemampuan menghasilkan output kerja
Kualitas_Kerja	Ketepatan dan ketelitian dalam menyelesaikan pekerjaan
Orientasi_Pelayanan	Kemampuan memberikan pelayanan terbaik
Komitmen	Tingkat tanggung jawab dan dedikasi terhadap pekerjaan
Kerja_Sama	Kemampuan bekerja sama dengan rekan kerja
Nilai_SKP	Nilai akhir hasil penilaian kinerja pegawai
Label_Kinerja	Kategori tingkat kinerja pegawai (Tinggi, Sedang, Rendah)

Berdasarkan Tabel 2, data penelitian menggunakan format .xlsx (*Excel*) dengan total 10 atribut. Atribut-atribut tersebut merepresentasikan indikator kinerja pegawai yang digunakan sebagai dasar dalam proses *clustering* menggunakan algoritma *K-Means* untuk mengidentifikasi pola dan kelompok kinerja pegawai di lingkungan Kominfo.

3.3 Data Preparation

Pada tahap ini, langkah pertama yang dilakukan adalah memilih atribut yang relevan

untuk proses *modeling* dengan menghapus data yang tidak diperlukan. Atribut yang dihapus adalah NIP dan Nama_Pegawai, karena atribut tersebut tidak memiliki pengaruh langsung terhadap proses analisis dan tidak relevan dalam pengelompokan kinerja pegawai. Setelah pemilihan atribut, dilakukan pemeriksaan *missing value* untuk memastikan tidak ada data yang kosong atau hilang pada setiap kolom. Pemeriksaan ini dilakukan agar proses analisis dapat berjalan dengan baik tanpa adanya nilai yang tidak valid.

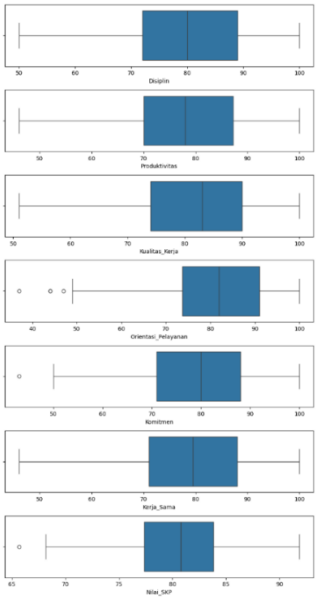
```
missing_values = data_dropped.isnull().sum()
missing_values
```

	0
Disiplin	0
Produktivitas	0
Kualitas_Kerja	0
Orientasi_Pelayanan	0
Komitmen	0
Kerja_Sama	0
Nilai_SKP	0
Label_Kinerja	0

dtype: int64

Gambar 2 *Missing Value*

Berdasarkan hasil pemeriksaan pada Gambar 2, tidak ditemukan adanya *missing value* pada seluruh atribut yang digunakan. Hal ini menunjukkan bahwa data bersih dan siap digunakan untuk tahap analisis selanjutnya. Langkah berikutnya adalah pemeriksaan *outlier* untuk mendeteksi adanya data ekstrem yang dapat memengaruhi hasil pengelompokan. Deteksi *outlier* dilakukan dengan menggunakan metode visualisasi seperti *boxplot* serta analisis statistik pada masing-masing variabel numerik.

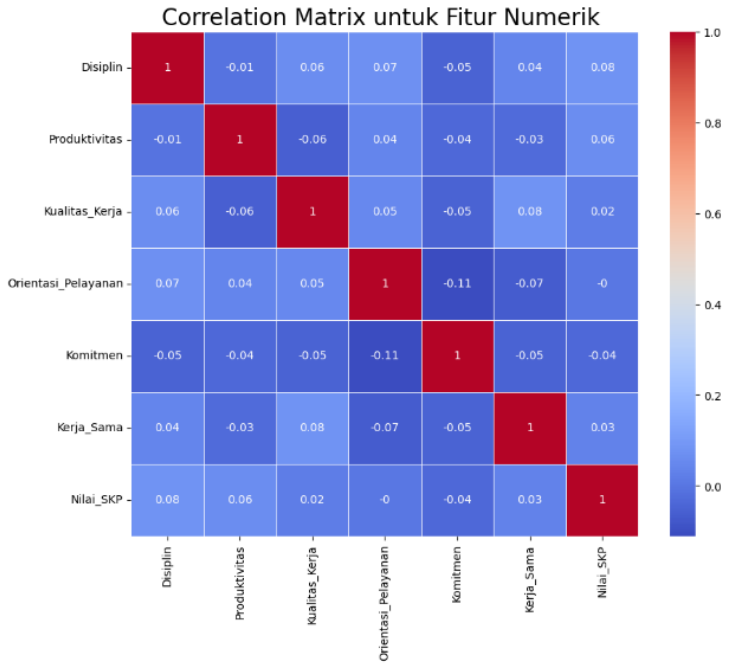


Gambar 3 *outlier*

Berdasarkan Gambar 3, ditemukan beberapa fitur menunjukkan keberadaan *outlier*, terutama pada Orientasi_Pelayanan, Komitmen, dan Nilai_SKP. *Outlier* pada Orientasi_Pelayanan berada di rentang nilai rendah (40–50), kemungkinan disebabkan oleh tugas pegawai yang tidak berhubungan langsung dengan layanan publik atau penilaian yang lebih ketat. Pada Komitmen, satu nilai ekstrem muncul akibat rendahnya konsistensi kerja, kehadiran, atau motivasi pegawai. Sementara itu, outlier pada Nilai_*SKP* dapat terjadi karena pegawai baru, target kinerja yang tidak tercapai, atau evaluasi atasan yang ketat. Secara umum, outlier tersebut masih

mencerminkan variasi nyata dalam performa pegawai, sehingga tetap dipertahankan untuk menjaga keaslian data.

Setelah mengidentifikasi keberadaan *outlier*, langkah selanjutnya adalah melakukan analisis lanjutan terhadap kualitas dan relevansi setiap fitur melalui evaluasi korelasi. Tahap ini penting untuk memastikan bahwa fitur yang digunakan dalam proses pemodelan benar-benar memberikan kontribusi informasi yang signifikan serta tidak menimbulkan redundansi data.



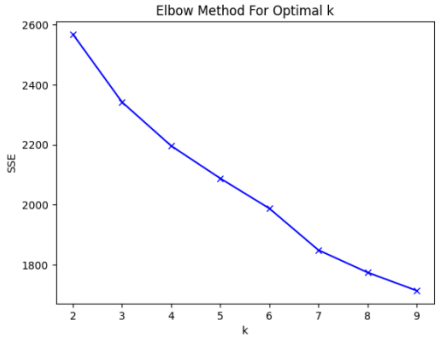
Gambar 4 *Correlation Matrix*

Mengacu pada Gambar 4, hasil analisis korelasi menunjukkan bahwa seluruh fitur kinerja memiliki hubungan yang sangat rendah satu sama lain (-0,11 hingga 0,08). Nilai korelasi yang kecil ini menandakan tidak adanya multikolinearitas, sehingga masing-masing fitur dianggap memberikan informasi yang independen. Berdasarkan temuan tersebut, seluruh fitur dipertahankan karena relevan secara konseptual terhadap penilaian kinerja ASN dan tidak menunjukkan redundansi data. Proses *feature engineering* hanya memerlukan normalisasi untuk menyetarakan skala antar variabel, tanpa perlu eliminasi atau penggabungan fitur. Dengan demikian, semua fitur dinilai layak digunakan dalam pemodelan *K-Means* karena memberikan kontribusi informasi yang unik dalam pembentukan *cluster*.

Tahap berikutnya adalah transformasi dan standarisasi data, yaitu mengubah data numerik ke dalam skala yang sama agar setiap variabel memiliki kontribusi yang seimbang dalam proses *clustering*. Proses ini penting untuk memastikan bahwa variabel dengan rentang nilai besar tidak mendominasi hasil pengelompokan pada algoritma *K-Means*.

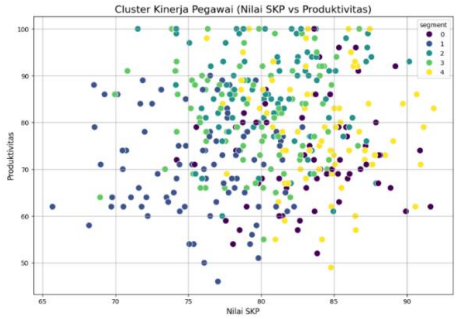
3.4 Modelling

Agar proses pengelompokan data dengan algoritma *K-Means* dapat dilakukan secara optimal, diperlukan penentuan jumlah *cluster* yang tepat. Penentuan jumlah cluster ini dapat dioptimalkan menggunakan metode ***Elbow***, di mana titik optimal akan terlihat pada grafik sebagai bentuk siku (*elbow point*) yang menunjukkan jumlah *cluster* terbaik.



Gambar 5 Hasil K Optimal dengan metode *Elbow*

Berdasarkan Gambar 5 yang menampilkan hasil pencarian nilai K optimal, diketahui bahwa jumlah klaster yang paling sesuai dalam analisis ini adalah K = 5. Titik siku (*elbow point*) terlihat berada pada kisaran K = 5 atau K = 6, namun K = 5 dipilih karena setelah titik tersebut, penurunan nilai SSE mulai melambat secara signifikan.

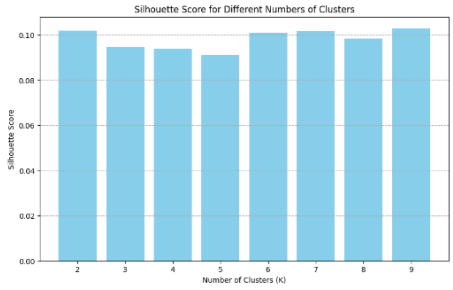


Gambar 6 *Clustering* K = 5

Gambar 6 menunjukkan hasil penerapan algoritma *K-Means Clustering* pada data kinerja pegawai Kominfo yang menghasilkan 5 *cluster* dengan karakteristik berbeda. *Cluster* (Kinerja tertinggi, disiplin dan kualitas kerja sangat baik), *Cluster* 1 (Kinerja baik, komitmen dan kerja sama tinggi, produktivitas sedikit lebih rendah), *Cluster* 2 (Kinerja cukup baik, disiplin dan produktivitas baik namun komitmen menurun), *Cluster* 3 (Kinerja sedang, produktivitas cukup tinggi tetapi kerja sama rendah) dan *Cluster* 4 (Kinerja terendah, seluruh indikator berada pada level sedang hingga rendah). Hasil pengelompokan ini membantu manajemen memahami pola kinerja pegawai dan mendukung pengambilan keputusan dalam pembinaan serta pengembangan SDM.

3.5 Evaluation

Tahap berikutnya adalah melakukan evaluasi untuk menilai kualitas dan hasil dari *cluster* serta model yang telah terbentuk. Metode yang digunakan dalam proses ini adalah *Silhouette Coefficient.*



Gambar 7 *Silhouette Coefficients*

Pada gambar 7 hasil *Silhouette Plot* untuk K = 8 clusters, diperoleh nilai rata-rata *Silhouette Coefficient* sebesar 0,10. Nilai ini menunjukkan bahwa kualitas pemisahan antar *cluster* masih tergolong lemah.



Gambar 8 *Clustering* K=8

Pada gambar 8 hasil *Silhouette Score* untuk K=8 yang rendah (0,10) menandakan bahwa pemisahan antar *cluster* kurang jelas. Dengan kata lain, nilai K=8 tidak menghasilkan pengelompokan yang baik, karena banyak data masih tumpang tindih antar *cluster*.

3.6 Deployment

Berdasarkan hasil pengelompokan yang terbentuk pada nilai **K=5**, karakteristik setiap *cluster* dapat diidentifikasi melalui variabel penciri yang menunjukkan perbedaan pola kinerja antar kelompok pegawai, sehingga setiap *cluster* memiliki ciri khas yang membedakannya dari *cluster* lainnya seperti tersaji pada Gambar 9 sampai dengan Gambar 13.

a. *Cluster* 0

	NIP	Nama_Pegawai	Disiplin	Produktivitas	Kualitas_Kerja	Orientasi_Pelayanan	Komitmen	Kerja_Sama	Nilai_SKP	Label_Kinerja	segment
6	19755325316860	Rizky Maulana	87	75	97	96	81	87	87.17	Tinggi	0
8	19755495394815	Andi Saputra	98	68	79	88	81	89	83.83	Sedang	0
17	19757120997885	Anton Siregar	86	62	94	77	63	92	79.00	Sedang	0
19	19757733221635	Dwi Lestari	96	79	95	90	91	96	91.17	Tinggi	0
36	19752644658402	Budi Prasetyo	96	75	97	85	60	87	83.33	Sedang	0

Gambar 9 *cluster* 0

Cluster 0 merepresentasikan pegawai berkinerja unggul, dengan skor tinggi pada disiplin, produktivitas, kualitas kerja, komitmen, dan kerja sama. Nilai SKP berada pada kisaran 85–91, menunjukkan konsistensi dan performa yang stabil memiliki Karakteristik yaitu Disiplin tinggi (>85), Produktivitas dan kualitas kerja tinggi, Komitmen dan kerja sama sangat baik dan Nilai SKP: 85–91. Dengan demikian, *Cluster* 0 mencerminkan kelompok pegawai berprestasi tinggi.

b. *Cluster* 1

	NIP	Nama_Pegawai	Disiplin	Produktivitas	Kualitas_Kerja	Orientasi_Pelayanan	Komitmen	Kerja_Sama	Nilai_SKP	Label_Kinerja	segment
0	19751787846414	Siti Rahma	65	83	83	76	74	85	77.67	Sedang	1
3	19758395928407	Lia Handayani	79	65	70	88	61	81	74.00	Sedang	1
10	19752934879560	Nanda Sari	60	75	84	94	65	71	74.83	Sedang	1
12	19756950318585	Fina Rahmi	72	62	84	80	96	85	79.83	Sedang	1
13	19751068574553	Citra Dewi	62	79	88	91	92	85	82.83	Sedang	1

Gambar 10 *cluster* 1

Cluster 1 menggambarkan pegawai dengan kinerja menengah, dengan skor disiplin, kualitas kerja, dan komitmen pada kisaran 65–80. Nilai SKP berada pada rentang 74–83, menunjukkan performa yang cukup stabil namun masih memerlukan peningkatan, khususnya pada aspek kedisiplinan dan produktivitas yang memiliki karakteristik utama yaitu Disiplin sedang (65–79), Produktivitas dan kualitas kerja cukup baik, Komitmen bervariasi dan Nilai SKP: 74–83. *Cluster* ini mencerminkan kelompok pegawai berperforma menengah yang masih memiliki ruang pengembangan.

c. *Cluster* 2

	NIP	Nama_Pegawai	Disiplin	Produktivitas	Kualitas_Kerja	Orientasi_Pelayanan	Komitmen	Kerja_Sama	Nilai_SKP	Label_Kinerja	segment
2	19755537253172	Fina Rahmi	87	93	83	89	93	96	90.17	Tinggi	2
4	19755298312065	Vina Kartika	89	81	67	65	91	88	80.17	Sedang	2
14	19751991681409	Tono Rahman	98	95	77	82	67	80	83.17	Sedang	2
15	19754578544903	Maya Hapsari	65	78	77	92	64	98	79.00	Sedang	2
16	19756974559824	Citra Dewi	67	85	61	62	98	95	78.00	Sedang	2

Gambar 11 *cluster* 2

Cluster 2 menggambarkan pegawai dengan kinerja baik hingga tinggi, dengan variasi performa antar individu. Nilai SKP berada pada rentang 78–90, menunjukkan stabilitas pada aspek disiplin, kualitas kerja, dan kerja sama, meskipun konsistensi sebagian pegawai masih dapat ditingkatkan yang memiliki karakteristik utama yaitu Disiplin sedang–tinggi (62–89), Kualitas kerja dan orientasi pelayanan stabil (sedang–tinggi), Komitmen dan kerja sama relatif tinggi (>85 pada sebagian besar pegawai) dan Nilai SKP: 78–90. *Cluster* ini mencerminkan pegawai berkinerja baik dengan potensi untuk mencapai kategori unggul.

d. *Cluster* 3

	NIP	Nama_Pegawai	Disiplin	Produktivitas	Kualitas_Kerja	Orientasi_Pelayanan	Komitmen	Kerja_Sama	Nilai_SKP	Label_Kinerja	segment
1	19753563451924	Mega Putri	87	64	92	87	87	75	82.00	Sedang	3
5	19753652062880	Mega Putri	70	70	95	94	82	73	80.67	Sedang	3
7	19753253890010	Dwi Lestari	84	92	99	83	84	64	84.33	Sedang	3
9	19755684118973	Yudi Pratama	92	65	94	90	81	64	81.00	Sedang	3
11	19759646907153	Tono Rahman	86	88	94	92	74	75	84.83	Sedang	3

Gambar 12 *cluster* 3

Cluster 3 menggambarkan pegawai dengan kinerja sedang, ditandai dengan nilai disiplin, produktivitas, dan kualitas kerja pada kisaran 70–87, serta nilai SKP 80–84. Indikator komitmen dan kerja sama cenderung fluktuatif (64–75), sehingga kinerja kelompok ini belum mencapai kategori tinggi yang memiliki karakteristik utama yaitu Disiplin dan produktivitas sedang hingga baik, Kualitas kerja sedang, Komitmen dan kerja sama cukup baik namun tidak konsisten dan Nilai SKP: 80–84. Secara keseluruhan, *cluster* ini mencerminkan pegawai berkinerja rata-rata yang memerlukan peningkatan konsistensi dan motivasi kerja.

e. *Cluster* 4

	NIP	Nama_Pegawai	Disiplin	Produktivitas	Kualitas_Kerja	Orientasi_Pelayanan	Komitmen	Kerja_Sama	Nilai_SKP	Label_Kinerja	segment
50	19751924418326	Dedi Kurniawan	70	55	97	73	75	58	80.98	Sedang	4
61	19758570405288	Dewi Anggraini	74	84	93	75	55	63	77.36	Sedang	4
62	19756459271371	Siti Rahma	78	89	81	79	67	75	80.75	Sedang	4
64	19755687332650	Fadli Akbar	79	86	64	87	77	57	83.58	Sedang	4
68	19755175401317	Farhan Maulana	89	76	82	76	93	67	89.07	Tinggi	4

Gambar 13 *cluster* 4

Cluster 4 menunjukkan pegawai dengan kinerja sedang hingga tinggi, namun belum seimbang antar indikator. Disiplin, kualitas kerja, dan komitmen relatif baik (70–97), sementara produktivitas dan kerja sama masih bervariasi, dengan beberapa nilai berada di bawah rata-rata. Nilai SKP berkisar 77–89, dengan mayoritas pada kategori Sedang yang mrmiliki karakteristik utama yaitu Disiplin dan kualitas kerja cukup tinggi (70–97), produktivitas dan kerja sama bervariasi, beberapa rendah, Komitmen dan orientasi pelayanan stabil pada tingkat sedang–baik dan Nilai SKP: 77–89. *Cluster* ini mencerminkan pegawai berpotensi baik tetapi belum optimal karena ketidakkonsistenan pada produktivitas dan kerja sama.

Temuan utama penelitian menunjukkan bahwa data kinerja pegawai dapat dikelompokkan secara jelas ke dalam beberapa cluster berdasarkan pola skor pada indikator-indikator SKP. Pembentukan *cluster* ini mengungkap perbedaan tingkat kinerja antar kelompok

pegawai, seperti kelompok dengan kinerja tinggi, sedang, dan perlu pembinaan.

Dampak model terhadap kinerja organisasi adalah memberikan dasar analitis yang objektif bagi manajemen dalam mengevaluasi kinerja pegawai. Model *clustering* membantu organisasi mengenali area kekuatan, kelemahan, serta identifikasi kelompok pegawai yang membutuhkan intervensi khusus, sehingga pengambilan keputusan menjadi lebih tepat dan berbasis data. Tindak lanjut yang dapat dilakukan organisasi meliputi:

1. Penyusunan strategi pengembangan SDM berbeda per *cluster*, seperti pelatihan tambahan untuk *cluster* berkinerja rendah dan program penguatan untuk cluster berkinerja tinggi.
2. Penetapan kebijakan *reward* and *punishment* yang lebih akurat, karena evaluasi kinerja didasarkan pada pola data yang terukur.
3. Perbaikan perencanaan beban kerja dan distribusi tugas, dengan memanfaatkan informasi karakteristik tiap *cluster*.

Dengan demikian, temuan jumlah *cluster* tidak hanya menggambarkan kondisi kinerja saat ini, tetapi juga menjadi dasar penting bagi organisasi dalam meningkatkan efektivitas manajemen kinerja.

4. KESIMPULAN

Penelitian ini berhasil menerapkan algoritma *K-Means Clustering* dalam menganalisis data Sasaran Kinerja Pegawai (SKP) di Kementerian Komunikasi dan Informatika (Kominfo) dengan mengikuti tahapan CRISP-DM. Berdasarkan hasil evaluasi menggunakan *Elbow Method*, jumlah klaster optimal yang diperoleh adalah lima (K=5). Validitas klaster kemudian diperkuat melalui pengujian *Silhouette Coefficient*, yang menunjukkan nilai pada kategori layak sehingga struktur klaster dapat diterima. Pembentukan lima klaster tersebut menggambarkan variasi tingkat kinerja pegawai dari kategori tertinggi hingga terendah. Sebagian besar pegawai berada pada klaster dengan kategori sedang hingga baik, menunjukkan bahwa kinerja pegawai Kominfo secara umum berada pada tingkat yang cukup stabil. Indikator disiplin dan kualitas kerja muncul sebagai variabel yang paling konsisten di seluruh klaster, sedangkan komitmen dan kerja sama menjadi pembeda utama antar klaster. Secara keseluruhan, penerapan *K-Means* terbukti efektif dalam memetakan pola kinerja pegawai secara objektif dan berbasis data. Hasil *clustering* ini dapat digunakan oleh manajemen Kominfo untuk mendukung pengambilan keputusan, khususnya dalam evaluasi kinerja, pembinaan, serta pengembangan kompetensi pegawai.

5. SARAN

Penelitian ini masih memiliki ruang untuk pengembangan lebih lanjut. Beberapa saran yang dapat menjadi acuan bagi penelitian selanjutnya antara lain:

1. Membandingkan algoritma *K-Means* dengan metode *clustering* lain seperti *K-Medoids*, *Fuzzy C-Means*, atau DBSCAN untuk memperoleh algoritma dengan kualitas pengelompokan terbaik.
2. Menambahkan variabel pendukung kinerja, misalnya masa kerja, jabatan, atau tingkat pendidikan, agar pola kinerja yang terbentuk lebih komprehensif.
3. Mengembangkan model prediksi kinerja pegawai menggunakan algoritma *supervised learning* untuk menghasilkan sistem pendukung keputusan yang lebih proaktif.

DAFTAR PUSTAKA

[1] D. Prameswara And E. Yulianingsih, “Sistem Informasi Kinerja Pegawai Pada Kantor Camat Sungai Lilin Berbasis Website,” Vol. 19, No. 3.

[2] A. Rifqi And R. T. Aldisa, “Sistem Pendukung Keputusan Penilaian Kinerja Karyawan Menerapkan Metode Topsis”.

[3] A. F. Kalyana, D. Hamidin, And S. Nirwan, “Implementasi Algoritma Decision Tree Pada Kinerja Karyawan Menggunakan Kerangka Objective And Key Result (Okr),” *J.*

Kecerdasan Buatan Dan Teknol. Inf., Vol. 4, No. 1, Pp. 01–12, Jan. 2025, Doi: 10.69916/Jkbti.V4i1.160.

[4] M. A. Abdullah And R. T. Aldisa, "Perbandingan Metode Preference Selection Index Dan Kombinasi Preference Selection Index Dan Topsis Dalam Penilaian Kinerja Karyawan Hotel".

[5] S. Sahibu, R. Bambang, I. Taufik, And A. Agusriandi, "Penerapan Data Mining Dalam Analisis Penilaian Kinerja Pegawai Menerapkan Metode K-Means," *J. Media Inform. Budidarma*, Vol. 7, No. 1, P. 22, Jan. 2023, Doi: 10.30865/Mib.V7i1.5100.

[6] A. N. Z. Hidayah And A. F. Rozi, "Penerapan Data Mining Dalam Menentukan Kinerja Karyawan Terbaik Dengan Menggunakan Metode Algoritma C4.5 (Studi Kasus : Universitas Mercu Buana Yogyakarta)," *Artif. Intell.*, No. 2, 2021.

[7] W. Alfian, K. -, And T. Hidayat, "Analisis Clustering Pegawai Berdasarkan Tingkat Kedisiplinan Menggunakan Algoritma K-Means Dan Davies-Bouldin Index," *J. Electr. Eng. Comput. Jeecom*, Vol. 6, No. 2, Pp. 437–448, Oct. 2024, Doi: 10.33650/Jeecom.V6i2.9556.

[8] H. L. Sari And I. Y. Beti, "Penerapan Data Mining Dalam Pengelompokkan Buku Yang Dipinjam Menggunakan Algoritma K-Means".

[9] P. Marpaung, I. Pebrian, And W. Putri, "Penerapan Data Mining Untuk Pengelompokan Kepadatan Penduduk Kabupaten Deli Serdang Menggunakan Algoritma K-Means," Vol. 6, 2023.

[10] W. Cesar, R. R. Saputra, And A. Wibowo, "Klasterisasi Kepadatan Pegawai Dengan Metode K-Means Untuk Prediksi Kebutuhan Casn Instansi Pemerintah," Vol. 10, No. 2, 2023.

[11] G. B. Kaligis And S. Yulianto, "Analisa Perbandingan Algoritma K-Means, K-Medoids, Dan X-Means Untuk Pengelompokkan Kinerja Pegawai," *It-Explore J. Penerapan Teknol. Inf. Dan Komun.*, Vol. 1, No. 3, Pp. 179–193, Oct. 2022, Doi: 10.24246/Itexplore.V1i3.2022.Pp179-193.

[12] M. E. A. Rivan And R. A. Sonaru, "Perbandingan Metode K-Means Dan Ga K-Means Untuk Clustering Dataset Heart Disease Patients," Vol. 9, No. 3, 2022.

[13] F. Pratama And M. F. Azima, "Penerapan Algoritma K-Means Clustering Dalam Mendeteksi Kerusakan Perangkat Laboratorium Komputer Berbasis Android," Vol. 19, No. 1.

[14] G. F. Ibanez And G. W. Wiriasto, "Kombinasi Principal Component Analysis Dengan Algoritma K-Means Untuk Klasterisasi Data Stunting".

[15] A. Natalis And Y. Nataliani, "Pemanfaatan K-Means Clustering Dan Analytic Hierarchy Process Terhadap Penilaian Prestasi Kerja Pegawai," *J. Tek. Inform. Dan Sist. Inf.*, Vol. 8, No. 1, Apr. 2022, Doi: 10.28932/Jutisi.V8i1.4243.

[16] J. Faran And A. Triayudi, "Penerapan Algoritma K-Means Data Mining Untuk Clustering Kinerja Karyawan Koperasi".