

Implementasi K-Means dan Elbow dalam Sistem Informasi Kepegawaian untuk Pengelompokan Data Karyawan PTPN Medan

Windi Aninda Fitrah^{1*}, Aninda Muliani Harahap²

^{1,2} *Fakultas Sains dan Teknologi, Universitas Islam Negeri Sumatera Utara,*

Jl. Lap. Golf No.120, Medan, Indonesia

E-mail: windianindafrh@gmail.com, anindamh@uinsu.ac.id

**Corresponding Author*

Abstrak— Pengelolaan data kepegawaian secara manual di unit kebun PTPN IV Regional I Medan sering mengalami kendala dalam hal efisiensi dan objektivitas evaluasi kinerja karyawan. Penelitian ini bertujuan untuk mengembangkan sistem informasi kepegawaian berbasis web yang mampu mengelompokkan karyawan secara otomatis menggunakan algoritma K-Means Clustering dengan jumlah kluster optimal yang ditentukan oleh Metode Elbow. Penelitian menggunakan metode kuantitatif dengan data primer berupa atribut karyawan seperti absensi, masa kerja, usia, pendidikan, kinerja, dan divisi. Sistem dirancang menggunakan framework Laravel dan diuji menggunakan metode Black Box Testing. Hasil penelitian menunjukkan bahwa integrasi K-Means dan Elbow dapat menghasilkan pengelompokan karyawan ke dalam tiga kategori kinerja (tinggi, sedang, rendah) secara akurat dan efisien. Sistem ini dapat digunakan oleh HRD sebagai dasar dalam pengambilan keputusan manajerial secara objektif.

Kata Kunci — Data Mining; Elbow Method; K-Means Clustering; Laravel; Sistem Informasi Kepegawaian

Abstract— Manual management of employee data at PTPN IV Regional I Medan often faces inefficiencies and lacks objectivity in performance evaluation. This study aims to develop a web-based employee information system that automatically clusters employees using the K-Means Clustering algorithm, with the optimal number of clusters determined by the Elbow Method. A quantitative method was applied using primary data consisting of employee attributes such as attendance, tenure, age, education, performance, and division. The system was developed using the Laravel framework and tested with Black Box Testing. The results showed that the integration of K-Means and Elbow effectively grouped employees into three performance categories (high, medium, low) accurately and efficiently. This system can assist HR departments in making objective managerial decisions.

Keywords — Data Mining; Elbow Method; K-Means Clustering; Laravel; Personnel Information System

I. PENDAHULUAN

Perkembangan teknologi informasi [1] yang semakin pesat di era digitalisasi ini telah mendorong berbagai perusahaan, termasuk di sektor perkebunan, untuk mengintegrasikan sistem digital yang canggih ke dalam manajemen sumber daya manusia (SDM) guna meningkatkan efektivitas dan produktivitas organisasi. Sistem informasi kepegawaian[2] berbasis web kini menjadi solusi inovatif yang tidak hanya meningkatkan efisiensi proses administrasi, seperti pengolahan data absensi, rekam jejak karyawan, dan penggajian, tetapi juga berperan penting dalam mendukung pengambilan keputusan strategis berbasis data (*data-driven decision making*)[3] yang lebih akurat dan tepat waktu. Namun demikian, salah satu tantangan utama yang masih dihadapi oleh banyak perusahaan, termasuk PTPN IV Regional I Medan, adalah proses evaluasi kinerja karyawan yang masih sering dilakukan secara manual, bersifat subjektif, dan rentan terhadap bias, sehingga memerlukan pendekatan yang lebih sistematis dan terukur.

Namun demikian, salah satu tantangan utama yang masih dihadapi oleh banyak perusahaan, termasuk PTPN IV Regional I Medan, adalah proses evaluasi kinerja karyawan yang masih sering dilakukan secara manual, bersifat subjektif, dan rentan terhadap bias, sehingga memerlukan pendekatan yang lebih sistematis dan terukur. Kondisi ini semakin diperparah oleh ketidakkonsistenan dalam penerapan standar evaluasi antar divisi, keterbatasan waktu manajer dalam melakukan assessment yang

mendalam, serta kurangnya data historis yang terintegrasi untuk mendukung penilaian kinerja yang objektif.

Sebagai perusahaan Badan Usaha Milik Negara (BUMN) yang bergerak di bidang perkebunan kelapa sawit dan karet, PTPN IV Regional I Medan menghadapi kompleksitas dalam pengelolaan data kepegawaian akibat jumlah karyawan yang besar (mencapai ribuan orang) dan tersebar di berbagai unit kebun yang lokasinya terpencil. Kondisi geografis ini menambah tantangan dalam hal konektivitas dan real-time data processing. Karakteristik industri perkebunan yang memiliki siklus kerja musiman, variasi cuaca yang mempengaruhi produktivitas, serta ketergantungan pada tenaga kerja lapangan yang intensif menambah kompleksitas dalam pengelolaan SDM. Meskipun data dasar seperti absensi dan masa kerja telah dicatat secara digital melalui sistem yang ada, perusahaan belum memiliki sistem yang mampu melakukan pengelompokan (*clustering*) data karyawan secara analitis untuk mengevaluasi produktivitas, potensi, serta kebutuhan pengembangan SDM secara lebih mendalam.

Pemilihan algoritma K-Means dalam penelitian ini didasarkan pada beberapa pertimbangan strategis dan teknis yang relevan dengan karakteristik data kepegawaian PTPN IV Regional I Medan. Pertama, algoritma K-Means memiliki kemampuan yang sangat baik dalam menangani dataset berukuran besar dengan kompleksitas komputasi $O(nk)$, dimana n adalah jumlah data, k adalah jumlah cluster, dan i adalah jumlah iterasi, sehingga sangat sesuai untuk mengelola ribuan data karyawan yang tersebar di berbagai unit kebun. Kedua, K-Means memberikan hasil clustering yang mudah diinterpretasikan oleh manajemen karena menghasilkan centroid yang dapat dijelaskan secara intuitif sebagai "profil rata-rata" dari setiap kelompok kinerja karyawan. Ketiga, algoritma ini relatif robust terhadap outlier dengan implementasi yang tepat dan dapat memberikan hasil yang konsisten, hal ini penting mengingat variabilitas data karyawan perkebunan yang dipengaruhi faktor musiman dan geografis.

Berdasarkan penelitian sebelumnya, algoritma K-Means telah terbukti efektif untuk melakukan clustering data kepegawaian dengan mengelompokkan karyawan berdasarkan karakteristik tertentu. Beberapa varian K-Means seperti K-Means++ juga telah menunjukkan peningkatan signifikan dalam hal akurasi dan kecepatan konvergensi. Selain itu, penelitian oleh trivusi [4] menyatakan bahwa penerapan metode elbow (*elbow method*) dapat membantu menentukan jumlah klaster optimal dengan menganalisis titik penurunan signifikan dalam *within-cluster sum of squares* (WCSS), sehingga meningkatkan akurasi pengelompokan. Validasi tambahan dengan silhouette analysis juga direkomendasikan untuk memastikan kualitas cluster yang dihasilkan. Lebih lanjut, A. Rajsya and A. Rachman Manga [5] menunjukkan bahwa sistem informasi berbasis web yang mengintegrasikan algoritma K-Means dan metode elbow mampu meningkatkan efisiensi pengelolaan SDM di lingkungan perkebunan dengan menyediakan insights yang lebih terstruktur dan dapat ditindaklanjuti melalui visualisasi data interaktif.

Namun demikian, penelitian terdahulu seperti yang dilakukan oleh A. Y. Fitrianna, N. Priatna, and J. A. Dahlan [6] dan A. Yahya and R. Kurniawan [7] masih memiliki keterbatasan karena belum mengintegrasikan hasil clustering ke dalam sistem informasi kepegawaian berbasis web yang dapat beroperasi secara otomatis dan *real-time*. Sistem yang ada juga belum sepenuhnya memanfaatkan multi-atribut karyawan seperti tingkat kehadiran (absensi), usia, masa kerja, latar belakang pendidikan, penilaian kinerja, divisi kerja, riwayat pelatihan, dan faktor-faktor kontekstual lainnya untuk menghasilkan segmentasi yang lebih komprehensif dan relevan dengan kebutuhan manajerial. Selain itu, aspek scalability dan *maintainability* sistem juga belum mendapat perhatian yang memadai dalam implementasi jangka panjang.

Tujuan utama dari penelitian ini adalah membangun sebuah sistem informasi kepegawaian berbasis web yang tidak hanya mampu mengklasifikasikan karyawan ke dalam tiga kelompok kinerja (tinggi, sedang, rendah) secara otomatis dan objektif menggunakan pendekatan K-Means dan metode elbow, Sistem ini juga menyajikan visualisasi data interaktif, laporan lengkap, dan rekomendasi strategis berbasis machine learning untuk manajemen. Selain itu, dilengkapi dengan fitur pemantauan (*monitoring*) dan peringatan (*alerting*) untuk mendeteksi tren kinerja secara real-time, serta kemampuan prediksi (*forecasting*) kebutuhan SDM di masa depan [8]. Keberadaan fitur-fitur ini, sistem diharapkan dapat menjadi alat pendukung keputusan (*decision support system*) yang handal untuk perencanaan karier, pelatihan, rotasi jabatan, succession planning, dan penguatan SDM—semuanya berbasis analisis data yang akurat dan dapat dipertanggung jawabkan.

II. METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan metode pengembangan sistem yang terstruktur[9]. Penelitian dilakukan untuk merancang dan mengimplementasikan sistem kepegawaian berbasis web yang mengintegrasikan algoritma K-Means Clustering[10] dan Metode informasi Elbow dalam proses pengelompokan karyawan. Berikut adalah proses tahapan penelitian yang dilakukan dalam perancangan dan implementasi [11] sistem kepegawaian berbasis web. Setiap tahapan disusun secara terstruktur, mulai dari identifikasi masalah, pengumpulan data, pengolahan data dengan metode KDD, penerapan algoritma K-Means Clustering serta Metode Elbow, hingga tahap analisis, perancangan, dan implementasi sistem. Alur tahapan penelitian dapat dilihat pada Gambar 1.

Identifikasi masalah

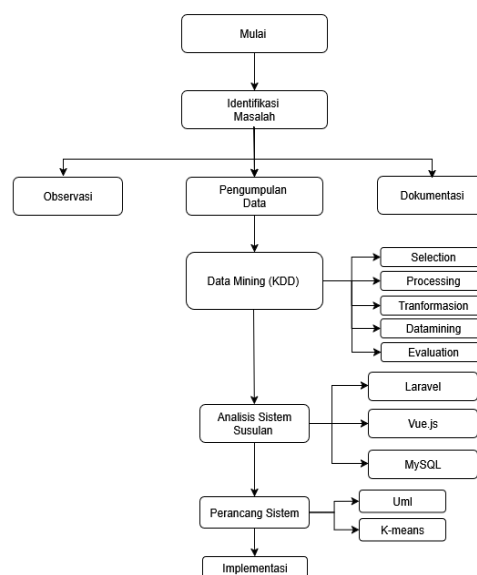
Tahap identifikasi masalah dilakukan untuk melihat kesenjangan dalam pengelolaan data kepegawaian pada unit kebun PTPN IV Regional I Medan. Oleh karena itu, diperlukan sistem kepegawaian berbasis web yang mampu mengelompokkan karyawan secara objektif dengan memanfaatkan algoritma K-Means Clustering untuk membentuk klaster berdasarkan atribut karyawan, serta metode Elbow untuk menentukan jumlah klaster optimal. Penelitian ini dilaksanakan pada periode Januari hingga Juni 2025 dengan tujuan menghasilkan sistem yang dapat mendukung pengelolaan data kepegawaian secara lebih efektif, akurat, dan terstruktur.

Pengumpulan data

Data penelitian diperoleh melalui observasi dan dokumentasi. Data yang digunakan mencakup absensi karyawan, masa kerja, usia, tingkat pendidikan, nilai kinerja, dan divisi kerja. Seluruh data dianonimkan untuk menjaga kerahasiaan. Akses sistem dirancang terbatas, di mana admin memiliki kendali penuh, sedangkan karyawan hanya dapat melihat data pribadinya.

Data Mining (KDD)

Proses pengolahan data dilakukan dengan pendekatan Knowledge Discovery in Database (KDD). Tahapan ini terdiri atas pemilihan data (selection), pembersihan dan normalisasi (processing), transformasi ke bentuk numerik, penerapan algoritma K-Means Clustering, serta evaluasi hasil. Algoritma K-Means digunakan untuk mengelompokkan karyawan berdasarkan kesamaan atribut, sedangkan metode Elbow digunakan untuk menentukan jumlah klaster optimal dengan meminimalkan nilai Within-Cluster Sum of Squares (WCSS) [5].



Gambar 1. Tahapan Penelitian

Metode analisis data dilakukan melalui penerapan dua algoritma utama:

1. Pada setiap cluster terdapat titik pusat (*centroid*) yang mempresentasikan cluster tersebut. Menetapkan jumlah kelompok (k) yang diinginkan dalam kumpulan data.
2. Menentukan posisi awal titik pusat (*centroid*) secara acak sebagai titik referensi awal.
3. K-Means Clustering digunakan untuk mengelompokkan data karyawan berdasarkan kesamaan nilai dari enam atribut [12], [13]. Fungsi objektif algoritma ini ditunjukkan pada Persamaan (1), yang bertujuan untuk meminimalkan total jarak kuadrat antara setiap data dengan pusat klaster (*centroid*).

$$J = \sum_{i=1}^k \sum_{x \in C_i} \|x - \mu_i\|^2 \quad (1)$$

Deskripsi dari Persamaan (1), J merupakan fungsi objektif atau total jarak kuadrat dalam klaster, k adalah jumlah klaster yang ditentukan, C_i merupakan himpunan data pada klaster ke- i , x adalah data, dan μ_i merupakan centroid atau titik pusat klaster ke- i . Algoritma K-Means akan terus memperbarui posisi centroid hingga nilai J minimum tercapai, sehingga data dalam setiap klaster memiliki kemiripan yang tinggi sementara perbedaan antar klaster semakin jelas.

4. Metode Elbow digunakan untuk menentukan jumlah klaster optimal dengan meminimalkan nilai WCSS (*Within-Cluster Sum of Squares*). Nilai WCSS dihitung sebagaimana ditunjukkan pada Persamaan (2).

$$WCSS = \sum_{i=1}^k \sum_{x \in C_i} \|x - \mu_i\|^2 \quad (2)$$

Deskripsi dari Persamaan (2), WCSS merepresentasikan jumlah kuadrat jarak dalam setiap klaster. Variabel k adalah jumlah klaster yang ditentukan, C_i adalah himpunan data dalam klaster ke- i , x merupakan data anggota klaster, dan μ_i adalah centroid dari klaster ke- i . Nilai $\|x - \mu_i\|^2$ menunjukkan jarak Euclidean kuadrat antara data dengan centroid. Semakin kecil nilai WCSS, semakin homogen anggota data dalam klaster.

5. Melakukan klasifikasi objek berdasarkan kedekatan jarak dengan centroid.
Grafik WCSS (Jumlah Kuadrat Dalam Klaster) terhadap jumlah klaster k titik 'siku' (siku) menunjukkan jumlah klaster optimal [14] yang digunakan dalam proses clustering.

Analisis Sistem Usulan

Hasil data mining kemudian dianalisis untuk membangun sistem usulan. Analisis dilakukan dengan menentukan kebutuhan fungsional sistem serta pemilihan teknologi yang digunakan, yaitu framework Laravel, Vue.js untuk antarmuka, serta MySQL sebagai basis data.

Perancangan Sistem

Perancangan sistem menggunakan Unified Modeling Language (UML) untuk memodelkan alur proses, struktur data, dan interaksi pengguna [9]. Sistem dibangun dengan framework Laravel berbasis PHP, basis data MySQL, dan dijalankan melalui XAMPP [10][15]. Antarmuka dikembangkan dengan pendekatan Model-View-Controller (MVC) [16] agar logika bisnis terpisah dari tampilan. Algoritma K-Means Clustering diintegrasikan untuk mendukung pengelompokan data karyawan, dengan proses clustering hanya dapat dijalankan oleh admin melalui sistem panel.

Implementasi

Tahap implementasi dilakukan dengan membangun sistem informasi kepegawaian berbasis web menggunakan Laravel, Vue.js, dan MySQL. Sistem menyediakan panel admin untuk menjalankan proses clustering, sedangkan karyawan hanya dapat mengakses informasi pribadinya melalui dashboard interaktif. Setelah implementasi, dilakukan pengujian menggunakan metode Black Box Testing untuk memastikan fungsionalitas sistem tanpa melihat kode sumber [17]. Pengujian mencakup login, input data karyawan, proses clustering, serta tampilan hasil klaster. Hasil pengujian menunjukkan bahwa sistem berjalan sesuai kebutuhan dan mampu mendukung pengambilan keputusan manajerial di PTPN IV.

III. HASIL DAN PEMBAHASAN

Tahapan Clustering dan Penentuan Jumlah Kluster

Penelitian ini menggunakan data karyawan yang terdiri dari atribut: masa kerja, absensi, nilai kinerja, pendidikan, usia, dan divisi kerja. Data tersebut menjadi dasar dalam proses clustering menggunakan metode K-Means dan Elbow, dapat dilihat dari Tabel 1.

Penentuan Jumlah Kluster (Metode Elbow)

Perhitungan WCSS dilakukan untuk berbagai nilai k (jumlah kluster) mulai dari k=1 hingga k=10, kemudian diamati pola penurunan nilai WCSS. Hasil perhitungan metode Elbow berdasarkan nilai WCSS untuk setiap jumlah kluster ditunjukkan pada Tabel 2. Pada tabel 2 titik tekuk terlihat pada k = 3, karena setelah itu penurunan WCSS mulai melambat, artinya penambahan kluster tidak memberikan peningkatan signifikan.

Tabel 1. Data Karyawan

No	Nama	Masa Kerja	Absensi	Nilai	Pendidikan	Usia	Divisi
1	Rosnawati Situmorang, S.Ag	25	70	79	6	53	12
2	Asrial	25	100	93	3	53	9
3	Irwansyah	25	90	84,33	3	46	11
4	Rusli Hadinoto	10	90	89	3	46	12
5	Ahmad Rusin	25	70	95	2	57	12
6	Rina Yuliana	15	90	89	3	36	12
7	Marianto, S.Pd	31	100	94,3	6	57	12
8	Roni Sahpin	26	100	90,75	3	53	3
9	Sri Budi	17	100	95,17	3	40	10
10	Dony Heriyanto Sihombing, St	12	70	88	6	39	9
11	Basuki	11	90	84,33	3	36	12
12	Muslim Maang, S.Pd	23	90	90	6	56	12
13	Nasip	25	80	90,43	1	57	6
14	Suhendri	34	90	92,67	3	53	6
15	Fenky Irawan	25	70	79	6	53	12
...			...	...	...	...	...
491	Indah Maulina Safitri Sitorus	6	80	92,4	6	36	5

Tabel 2. Hasil Perhitungan WCSS untuk Setiap Nilai k

Jumlah Kluster (k)	WCSS
1	2946
2	2200,236
3	1892,637
4	1659,595
5	1510,639
6	1372,821
7	1269,069
8	1186,855
9	1090,512
10	1036,879

Tabel 3. Centroid Awal

Centroid	Masa Kerja	Absensi	Nilai	Pendidikan	Usia	Divisi	Cluster
1	25	70	79	6	53	12	C1 (Rendah)
2	25	100	93	3	53	9	C2 (Sedang)
3	25	100	84,33	3	46	11	C3 (Tinggi)

Menentukan Jumlah Cluster

Pada tahap penentuan jumlah cluster, Metode Elbow digunakan dengan menghitung nilai WCSS dari $k=1$ hingga $k=10$. Titik siku terlihat pada $k=3$, sehingga jumlah cluster optimal ditetapkan tiga kelompok. Proses clustering kemudian dilakukan dengan algoritma K-Means, diawali dengan pemilihan centroid secara acak, lalu menghitung jarak setiap data menggunakan Euclidean Distance dan menempatkannya ke cluster terdekat. Iterasi berlanjut dengan pembaruan centroid berdasarkan rata-rata atribut hingga stabil. Hasil akhir menunjukkan: C1 (absensi rendah, masa kerja lama) kinerja rendah, C2 (absensi tinggi, masa kerja panjang, kinerja konsisten) kinerja sedang, dan C3 (usia lebih muda, absensi tinggi, produktivitas baik) kinerja tinggi. Dengan demikian, karyawan terbagi dalam tiga kelompok yang merepresentasikan variasi kinerja berdasarkan atribut penelitian.

Menentukan pusat Cluster

Pada tahap ini akan menentukan nilai pusat Cluster (*Centroid*) dari data yang digunakan [18]. Nilai Centroid Awal ditentukan secara random dari tabel Data Awal. Pada tabel 3. Berdasarkan tabel centroid awal yang ditampilkan, pemilihan centroid ini tampaknya menggunakan strategi untuk merepresentasikan tiga kategori performa karyawan yang berbeda. Centroid pertama dengan nilai 79 dirancang untuk mewakili kelompok dengan performa rendah, ditandai dengan absensi yang lebih rendah (70) dibandingkan dua centroid lainnya. Centroid kedua dengan nilai 93 merepresentasikan kelompok performa sedang dengan absensi penuh (100), sementara centroid ketiga dengan nilai 84.33 mewakili kategori tinggi meski secara numerik nilainya berada di antara kategori rendah dan sedang. Pemilihan ini kemungkinan didasarkan pada kombinasi seluruh atribut, bukan hanya fokus pada satu variabel nilai saja. Masa kerja yang seragam (25 tahun) untuk semua centroid menunjukkan bahwa atribut ini mungkin tidak menjadi faktor pembeda utama dalam pengelompokan, atau nilai 25 merepresentasikan karakteristik umum dari dataset.

Menghitung Jarak Clustering

Perhitungan jarak antara setiap data karyawan dengan pusat cluster dilakukan menggunakan rumus Euclidean Distance [19]. Rumus ini mengukur kedekatan data (a_i) terhadap centroid (b_j) pada cluster ke- j . Hasil perhitungan menentukan jarak terkecil antara data dengan centroid, sehingga setiap data secara otomatis ditempatkan pada cluster yang paling dekat. Dengan demikian, Euclidean Distance berperan penting dalam proses K-Means karena memastikan setiap karyawan dikelompokkan berdasarkan kemiripan atributnya terhadap centroid masing-masing cluster. Hasil perhitungan jarak dapat dilihat pada tabel 4. Pada tabel 4. Hasil pengelompokan data karyawan menggunakan algoritma K-Means dengan tiga cluster (C1, C2, C3). Nilai pada kolom C1, C2, dan C3 merupakan jarak (euclidean distance) setiap data terhadap centroid masing-masing cluster. Karyawan ditampilkan dalam cluster dengan jarak terkecil, sebagaimana ditampilkan pada kolom *Cluster*

Setelah mengetahui hasil clustering pada iterasi pertama, langkah selanjutnya adalah menghitung rata-rata data berdasarkan cluster yang sama. Rata-rata data dari cluster tersebut akan menjadi nilai centroid kedua dapat dilihat pada tabel 5. Pada tabel 5. menggambarkan rata-rata karakteristik tiap cluster. Cluster pertama (C1) memiliki masa kerja panjang namun tingkat absensi rendah sehingga dinilai sebagai kinerja rendah. Cluster kedua (C2) didominasi karyawan dengan masa kerja paling lama, absensi tinggi, dan nilai kinerja stabil sehingga digolongkan kinerja sedang. Sementara itu, cluster ketiga (C3) berisi karyawan dengan usia lebih muda, absensi tinggi, serta produktivitas yang baik sehingga dipertahankan sebagai kinerja tinggi.

Setelah dilakukan iterasi kedua dan diperoleh rata-rata karakteristik tiap cluster pada Tabel 5, proses perhitungan dilanjutkan dengan pembaruan centroid pada iterasi berikutnya. Iterasi ketiga ini bertujuan untuk memastikan apakah posisi centroid sudah stabil atau masih mengalami perubahan signifikan. Hasil pembaruan centroid pada iterasi ketiga dapat dilihat pada Tabel 6.

Tabel 4. Hasil Perhitungan Jarak Data Dengan Centroid Awal Iterasi 1

No	Karyawan	C1	C2	C3	Cluster
1	Rosnawati Situmorang, S.Ag	0,00	33,38	31,42	C1
2	Asrial	33,38	0,00	11,32	C2
3	Irwansyah	31,42	11,32	0,00	C3
4	Rusli Hadinoto	19,57	34,63	33,88	C1
5	Ahmad Rusin	26,23	11,40	18,35	C2
6	Rina Yuliana	37,39	20,35	14,93	C3
7	Mariato, S.Pd	16,91	31,17	34,15	C1
8	Roni Sahpin	33,60	6,49	12,46	C2
9	Sri Budi	22,53	33,74	33,44	C1
10	Dony Heriyanto Sihombing, St	21,33	36,04	33,83	C1
11	Basuki	22,86	38,33	34,60	C1
12	Muslim Maang, S.Pd	23,11	11,83	15,69	C2
13	Nasip	17,54	20,87	24,23	C1
14	Suhendri	17,69	31,47	33,53	C1
15	Fenky Irawan	33,42	23,77	22,18	C3
...	...	...	...	...	...
491	Indah Maulina Safitri Sitorus	42,17	25,99	23,90	C3

Tabel 5. Pusat Centroid Baru Iterasi 2

Centroid	Masa Kerja	Absensi	Nilai	Pendidikan	Usia	Divisi	Cluster
1	22,95	74,68	89,89	3,54	48,17	9,23	C1 (Rendah)
2	28,55	93,81	92,65	3,62	53,86	8,42	C2 (Sedang)
3	16,24	92,33	89,32	3,68	39,83	9,98	C3 (Tinggi)

Tabel 6. Pusat Centroid Baru Iterasi 3

Centroid	Masa Kerja	Absensi	Nilai	Pendidikan	Usia	Divisi	Cluster
1	23,46	74,29	90,42	3,46	48,74	9,33	C1 (Rendah)
2	29,73	93,81	91,40	3,54	54,28	8,66	C2 (Sedang)
3	13,42	90,55	90,17	3,94	38,21	9,43	C3 (Tinggi)

Tabel 7. Hasil Perhitungan Jarak Data Dengan Centroid Awal Iterasi 3

No	Karyawan	C1	C2	C3	Cluster
1	Rosnawati Situmorang, S.Ag	13,52	27,60	30,18	C1
2	Asrial	26,24	8,08	21,25	C2
3	Irwansyah	26,66	13,60	17,94	C2
4	Rusli Hadinoto	14,71	32,28	22,44	C1
5	Ahmad Rusin	18,64	8,41	22,83	C2
6	Rina Yuliana	30,07	24,63	10,28	C3
7	Mariato, S.Pd	13,12	24,52	33,35	C1
8	Roni Sahpin	26,94	9,31	22,56	C2
9	Sri Budi	12,64	30,80	21,55	C1
10	Dony Heriyanto Sihombing, St	16,03	33,65	20,83	C1
11	Basuki	19,50	36,24	21,78	C1
12	Muslim Maang, S.Pd	17,70	9,05	20,48	C2
13	Nasip	10,97	15,32	24,88	C1
14	Suhendri	12,81	24,41	32,91	C1
15	Fenky Irawan	23,84	26,23	6,20	C3
...	...	...	...	...	...
491	Indah Maulina Safitri Sitorus	34,02	30,92	13,35	C3

Pada tabel 6. Centroid tiap cluster sudah stabil dan menggambarkan profil kinerja karyawan. Cluster pertama (C1) berisi karyawan dengan masa kerja panjang namun absensi rendah sehingga dikategorikan kinerja rendah. Cluster kedua (C2) mencakup karyawan dengan masa kerja paling lama, absensi tinggi, dan kinerja konsisten yang dipetakan sebagai kinerja sedang. Sedangkan cluster ketiga (C3) didominasi karyawan lebih muda dengan absensi tinggi dan produktivitas baik sehingga

digolongkan kinerja tinggi. Berdasarkan hasil iterasi ketiga, posisi centroid tiap cluster sudah stabil dan tidak mengalami perubahan signifikan. Untuk memastikan hasil tersebut, dilakukan kembali perhitungan jarak setiap data karyawan terhadap centroid baru. Hasil perhitungan jarak pada iterasi ketiga ditampilkan pada Tabel 7. Pada tabel 7. Posisi centroid tidak lagi berubah secara signifikan dalam jumlah iterasi yang relatif sedikit, terutama jika data sudah cukup terstruktur dan pemilihan awal centroid dilakukan dengan baik. Algoritma *K-Means* akan berhenti sebelum mencapai iterasi maksimum yang ditentukan, karena posisi centroid tidak lagi berubah setelah beberapa iterasi awal. Dari data tersebut dapat ditarik kesimpulan bahwa posisi cluster pada iterasi ke - 2 tidak mengalami perubahan dengan iterasi ke - 3.

Hasil Clustering

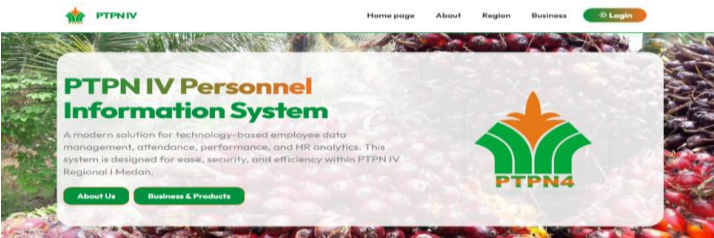
Setelah melalui beberapa iterasi perhitungan dengan algoritma *K-Means* dan menentukan jumlah kluster optimal menggunakan metode *Elbow*, sistem berhasil mengelompokkan seluruh data karyawan menjadi tiga kluster utama. dapat dilihat pada tabel 8. Pada tabel 8. Hasil perhitungan *K-Means* dan penentuan jumlah kluster optimal dengan metode *Elbow*, data karyawan PTPN IV terbagi menjadi tiga kelompok, yaitu 169 karyawan kategori rendah, 194 kategori sedang, dan 128 kategori tinggi. Proses iterasi menunjukkan bahwa nilai centroid pada tiap kluster menjadi stabil setelah beberapa kali perhitungan, sehingga hasil pengelompokan dapat dianggap konsisten. Temuan ini memberikan gambaran distribusi kinerja karyawan yang dapat dimanfaatkan sebagai dasar evaluasi dan pengambilan keputusan manajerial secara lebih objektif.

Implementasi Sistem

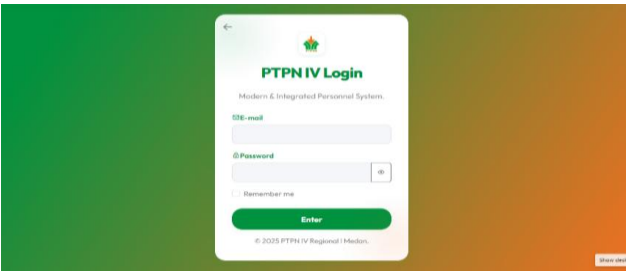
Proses membangun dan mengimplementasikan aplikasi secara keseluruhan, baik dari segi perangkat lunak maupun perangkat keras, dikenal dengan istilah implementasi[20]. Pada tahap implementasi, sistem informasi kepegawaian yang telah dirancang kemudian diwujudkan dalam bentuk aplikasi berbasis web. Implementasi ini mencakup pembuatan antarmuka pengguna serta integrasi fungsi-fungsi utama sistem. Adapun hasil implementasi dapat dilihat pada beberapa tampilan berikut, dimulai dari halaman utama sistem. Pada gambar 2 menampilkan nama sistem "Sistem Informasi Kepegawaian PTPN IV" dengan deskripsi modernisasi pengelolaan data kepegawaian berbasis teknologi. Tersedia tombol navigasi seperti About Us dan Business & Products , serta tombol Login untuk masuk ke dashboard sesuai level akses pengguna. Berikut merupakan Tampilan halaman login dapat dilihat pada gambar 3.

Tabel 8. Hasil Clustering

Klaster	Jumlah Karyawan	Kategori
C1	169	Rendah
C2	194	Sedang
C3	128	Tinggi



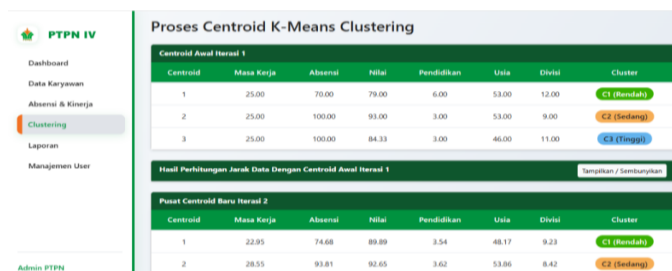
Gambar 2. Halaman Utama



Gambar 3. Halaman Login



Gambar 4. Dashboard Utama



Gambar 5. Proses Iterasi dan Perhitungan Centroid pada Clustering

Pada gambar 3 Pengguna sistem wajib memasukkan alamat email dan kata sandi untuk mengakses dashboard. Fitur “Remember me” disediakan untuk memudahkan akses berulang. Kemudian akan masuk pada Tampilan Dashboard Utama dapat dilihat pada gambar 4. Pada gambar 4 Menampilkan statistik ringkas seperti jumlah karyawan (491 orang), jumlah divisi (12), jumlah pengguna (admin dan karyawan), jumlah data absensi, serta laporan yang dihasilkan. Untuk melihat proses iterasi dan perhitungan Centroid pada Clustering dapat dilihat pada Tampilan Halaman Clustering seperti pada gambar 5.

Pada gambar 5 Sistem ini menampilkan proses pengelompokan karyawan menggunakan algoritma K-Means. Tahap awal memperlihatkan centroid iterasi pertama yang dipilih secara acak. Sistem kemudian menghitung jarak setiap data karyawan terhadap centroid tersebut untuk menentukan klaster sementara. Pada iterasi kedua, posisi centroid diperbarui dengan mengambil rata-rata dari data yang masuk dalam masing-masing klaster, sehingga pembagian kelompok menjadi lebih presisi. Memudahkan menjelaskan, setiap klaster diberi kode warna khusus: hijau (C1) merepresentasikan kelompok kinerja rendah, oranye (C2) menunjukkan kinerja sedang, dan biru (C3) menggambarkan kinerja tinggi. Proses iterasi ini terus berlanjut hingga posisi centroid tidak lagi mengalami perubahan signifikan atau mencapai batas iterasi yang ditentukan, sehingga menghasilkan pengelompokan karyawan yang optimal berdasarkan parameter kinerja yang dianalisis.

Perbedaan utama penelitian ini dengan penelitian sebelumnya terletak pada integrasi hasil clustering ke dalam sistem informasi kepegawaian berbasis web yang beroperasi secara otomatis dan real-time. Jika penelitian terdahulu masih terbatas pada penggunaan algoritma K-Means dan metode Elbow tanpa menghubungkannya langsung ke sistem informasi kepegawaian, penelitian ini mengembangkan sistem berbasis Laravel dengan dukungan database MySQL yang mampu mengelompokkan karyawan berdasarkan multi-atribut seperti absensi, usia, masa kerja, pendidikan, kinerja, dan divisi kerja. Selain itu, sistem yang dirancang juga menyediakan dashboard interaktif, laporan kinerja yang lebih komprehensif, serta kemudahan akses bagi admin dan karyawan. Dengan demikian, penelitian ini tidak hanya menghasilkan model clustering yang akurat, tetapi juga menghadirkan solusi praktis yang dapat langsung diterapkan dalam mendukung pengambilan keputusan manajerial di PTPN IV secara lebih efisien dan berbasis data.

IV. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa sistem informasi kepegawaian berbasis web yang dikembangkan berhasil memenuhi tujuan penelitian, yaitu mengelompokkan data karyawan secara otomatis dan tujuan menggunakan algoritma K-Means Clustering dengan bantuan Metode Elbow. Sistem ini mampu mengklasifikasikan karyawan ke dalam

tiga kategori berdasarkan kinerja (tinggi, sedang, dan rendah), dengan memanfaatkan atribut seperti absensi, masa kerja, usia, pendidikan, nilai kinerja, dan divisi kerja. Proses clustering berjalan secara optimal setelah dilakukan normalisasi data dan penentuan jumlah kluster menggunakan Elbow Method. Hasil yang diperoleh menunjukkan bahwa sistem dapat mempermudah bagian SDM dalam melakukan analisis dan evaluasi kinerja karyawan, serta mendukung pengambilan keputusan manajerial secara lebih efisien dan berbasis data. Sebagai pengembangan lebih lanjut, penelitian ini dapat dilanjutkan dengan menambahkan fitur visualisasi hasil clustering secara interaktif untuk memudahkan pemahaman pengguna. Selain itu, sistem integrasi dengan database kepegawaian real-time yang sudah ada di lingkungan PTPN IV dapat meningkatkan efisiensi dan validitas data yang digunakan. Penelitian lanjutan juga disarankan untuk mengukur akurasi hasil clustering menggunakan metode evaluasi seperti skor siluet atau Indeks Davies-Bouldin, serta menerapkan metode clustering lainnya seperti DBSCAN atau hierarchical clustering untuk perhitungan kinerja dan hasil. Dengan demikian, sistem yang dikembangkan dapat terus disempurnakan dan digunakan secara berkelanjutan dalam mendukung manajemen sumber daya manusia di lingkungan perusahaan.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Universitas Islam Negeri Sumatera Utara, khususnya Program Studi Sistem Informasi Fakultas Sains dan Teknologi, serta semua pihak yang telah memberikan dukungan dan bantuan sehingga penelitian ini dapat diselesaikan dengan baik.

DAFTAR PUSTAKA

- [1] R. Hefifiya Siregar and A. M. Harahap, "Sistem Informasi Perpustakaan Berbasis Web pada Perpustakaan Fakultas Saintek UINSU Web-Based Library Information System at the UINSU Faculty of Saintek Library," 2024.
- [2] N. Apipah, "Bab 2 ," 2025.
- [3] W. Powell Niken Retnowati, M. Windarti, K. Kunci, K. Cluster, and W. Powell, "Penentuan Jumlah Cluster pada K-Means Cluster Menggunakan Metode," *Jurnal Matematika*, vol. 18, no. 1, 2024.
- [4] Trivusi, "K-Means Clustering: Pengertian, Cara Kerja, Kelebihan, dan Kekurangannya."
- [5] A. Rajsya and A. Rachman Manga, "Rancang Bangun Penerapan Metode Elbow Pada K-Means Untuk Clustering Data Persediaan Barang," *Literatur Informatika & Komputer*, vol. 1, no. 4, pp. 395–403, 2024, doi: 10.33096/linier.vxix.xxxx.
- [6] A. Y. Fitrianna, N. Priatna, and J. A. Dahlan, "Pengembangan Model E-Book Interaktif Berbasis Pembelajaran Induktif untuk Melatihkan Kemampuan Penalaran Aljabar Siswa SMP," *Jurnal Cendekia: Jurnal Pendidikan*, vol. 05, no. 02, pp. 1562–1577, 2021, [Online]. Available: <https://j-cup.org/index.php/cendekia/article/view/653>
- [7] A. Yahya and R. Kurniawan, "Implementasi Algoritma K-Means untuk Pengelompokan Data Penjualan Berdasarkan Pola Penjualan," *Malcom : Indonesian Journal of Machine Learning and Computer Science*, vol. 5, no. 1, pp. 350–358, Jan. 2025, doi: 10.57152/malcom.v5i1.1773.
- [8] N. A. Maori, "Metode Elbow dalam Optimasi Jumlah Cluster pada K-Means Clustering," *Jurnal Simetris*, vol. 14, 2023.
- [9] M. K. Suhardi, "2. Suhardi-Laporan Akademik BOPTN 2022," 2022.
- [10] B. Setiawan and T. Triase, "Implementasi Desain UI/UX Aplikasi Ourticle kedalam Aplikasi Berbasis Android ," *Sibatik Journal: Jurnal Ilmiah Bidang Sosial, Ekonomi, Budaya, Teknologi, dan Pendidikan*, vol. 2, no. 3, pp. 805–818, Feb. 2023, doi: 10.54443/sibatik.v2i3.665.
- [11] U. Aplikasi Kerangka Berpikir *et al.*, "Computational Thinking Pada Pemebelajaran Faraid Di Era Digital," 2023. [Online]. Available: <https://ojs.literasiedu.org/index.php/jotter>
- [12] A. A. Haryada, T. Triase, M. Jannah, and H. Amri, "Implementasi Repository E- papers PT. Harian Waspada," *Jurnal Ilmiah Sistem Informasi dan Teknik Informatika (JISTI)*, vol. 7, no. 1, pp. 120–132, Apr. 2024, doi: 10.57093/jisti.v7i1.197.
- [13] D. Anggraeni and A. Muliani Harahap, "Aplikasi GIS Pencarian Agen Bus dan Travel di Kota Medan dengan Metode Dijkstra dan Topsis Berbasis Mobile," 2024. [Online]. Available: <http://jurnal.goretanpena.com/index.php/JSSR>

- [14] M. F. Siemensma, A. E. van der Windt, E. M. van Es, J. W. Colaris, and D. Eygendaal, "Management of the stiff elbow: a literature review," *EFORT Open Rev*, vol. 8, no. 5, pp. 351–360, 2023, doi: 10.1530/EOR-23-0039.
- [15] A. M. Harahap and A. Ikhwan, "Implementation of Information Technology Governance in Man 1 Medan Using the Cobit 5 Framework," *Sinkron*, vol. 8, no. 1, pp. 241–246, Jan. 2023, doi: 10.33395/sinkron.v8i1.11936.
- [16] A. Ikhwan, M. Salsabila Nasution, and N. Fahimah, "Perancangan Sistem Informasi Manajemen Penentuan Kamar Asrama Taruna Berbasis Website pada Poltekbang Medan," *Jurnal Ilmiah Ilmu dan Teknologi Rekayasa*, vol. 6, no. 1, p. 2024, 2024, [Online]. Available: <https://ejournal.polsub.ac.id/index.php/jiitr>
- [17] R. Kusumastuti, E. Bayunanda, A. Muhammad Rifa, M. Ryandy Ghonim Asgar, and F. Inti Ilmawati, "Clustering Titik Panas Menggunakan Algoritma Agglomerative Hierarchical Clustering (AHC) Hot Spot Clustering Using Agglomerative Hierarchical Clustering (AHC) Algorithm," *Cogito Smart Journal* |, vol. 8, no. 2.
- [18] M. Badri *et al.*, "Some rights reserved BY-NC-SA 4.0 International License Implementasi Augmented Reality pada Media Pengenalan Prodi Sistem Informasi FST Uinsu Medan," vol. 7, no. 2, pp. 109–121, 2022, doi: 10.36341/rabit.vxix.xxx.
- [19] F. N. Cahya, Yudi Mahatma, and Siti Rohmah Rohimah, "Perbandingan Metode Perhitungan Jarak Euclidean dengan Perhitungan Jarak Manhattan pada K-Means Clustering Dalam Menentukan Penyebaran Covid di Kota Bekasi," *JMT : Jurnal Matematika dan Terapan*, vol. 5, no. 1, pp. 43–55, Feb. 2023, doi: 10.21009/jmt.5.1.5.
- [20] D. M. Novela, K. Nasution, and Z. Siambaton, "Penerapan Algoritma K-Means Clustering dalam Aplikasi Penilaian Kinerja Pegawai di Kantor Kepala Desa Cinta Makmur Berbasis Web," *Hello World Jurnal Ilmu Komputer*, vol. 4, no. 1, pp. 49–65, Apr. 2025, doi: 10.56211/helloworld.v4i1.790.