

EVALUASI KINERJA PEGAWAI NEGERI SIPIL PEMERINTAH DAERAH KABUPATEN GROBOGAN MENGGUNAKAN ALGORITMA K-MEANS

Herra Profiliana¹, Jati Sasongko Wibowo²

^{1,2}Teknik Informatika, Fakultas Teknologi Informasi dan Industri, Universitas Stikubank
e-mail: proviliana@gmail.com, jatisw@edu.unisbank.ac.id

ABSTRAK

Profesi Pegawai Negeri Sipil (PNS) semakin diminati oleh banyak individu karena beberapa faktor menarik. Salah satunya adalah stabilitas pendapatan yang menawarkan keamanan finansial bagi PNS dan keluarganya. PNS juga menawarkan jaminan pensiun yang menarik, memungkinkan mereka pensiun dengan martabat dan keuangan yang terjamin. Namun, daya tarik utama PNS juga mencakup tanggung jawab mereka sebagai abdi masyarakat. Mereka harus memberikan pelayanan yang adil kepada warga masyarakat dengan efektif, efisien, produktif, dan profesional sesuai dengan peran mereka. Untuk mencapai tujuan ini, organisasi PNS memerlukan sumber daya manusia (SDM) berkualitas. Di Kabupaten Grobogan, Badan Kepegawaian, Pendidikan, dan Pelatihan Daerah (BKPPD) memiliki peran penting dalam menjalankan evaluasi kinerja PNS secara rutin setiap tahun. Evaluasi ini melibatkan berbagai aspek, termasuk Surat Keputusan Penilaian (SKP), tingkat pendidikan, kompetensi, dan disiplin pegawai. Hasil evaluasi mengklasifikasikan PNS menjadi tiga kluster: berkualitas baik (24%), cukup (30%), dan perlu perbaikan (16%). Pendekatan evaluasi komprehensif ini membantu BKPPD memastikan PNS di Kabupaten Grobogan tetap menjaga dan meningkatkan tingkat kualitas pelayanan kepada masyarakat.

Kata Kunci: PNS, kluster, rapidminer, penilaian kinerja

1. PENDAHULUAN

Semakin meningkatnya minat individu terhadap karier sebagai Pegawai Negeri Sipil (PNS) dapat dijelaskan oleh berbagai faktor yang menarik. Beberapa alasan utama yang mendorong minat ini adalah stabilitas pendapatan, jaminan pensiun, serta risiko kecil untuk diberhentikan dari pekerjaan. Bukti konkret dari minat ini dapat ditemukan dalam jumlah pelamar yang membanjiri pendaftaran CPNS pada tahun 2021. Sebanyak 3.482.989 individu bersaing untuk merebutkan 128.016 formasi yang tersedia, menegaskan daya tarik yang dimiliki oleh profesi PNS.

PNS memiliki peran krusial sebagai pelayan masyarakat yang harus memberikan layanan yang adil dan berkualitas. Oleh karena itu, mutlak diperlukan sumber daya manusia (SDM) yang memiliki kapabilitas untuk menjalankan tugasnya dengan efektif, efisien, produktif, dan profesional sesuai dengan peran dan tanggung jawab yang diemban. Dalam konteks ini, Badan Kepegawaian, Pendidikan, dan Pelatihan Daerah (BKPPD) memiliki peran penting dalam membentuk dan mengembangkan SDM berkualitas. Salah satu cara yang diambil adalah melalui pelaksanaan evaluasi kinerja PNS di Pemerintah Kabupaten Grobogan setiap tahunnya [1].

Pentingnya evaluasi kinerja bagi BKPPD terletak pada kemampuannya untuk memberikan penilaian yang akurat terhadap kinerja pegawai saat ini [2]. Selain itu, evaluasi ini juga berfungsi sebagai sumber umpan balik yang berharga dalam proses pembinaan PNS, seperti pengangkatan, kenaikan pangkat, pengangkatan dalam jabatan, serta program pendidikan dan pelatihan. Lebih lanjut, evaluasi kinerja ini tidak terbatas pada penilaian sasaran kinerja pegawai semata. Aspek-aspek pendukung lainnya, seperti tingkat pendidikan, kompetensi (termasuk pelatihan struktural, fungsional, teknis, serta partisipasi dalam seminar atau workshop), dan kedisiplinan pegawai, juga menjadi bahan pertimbangan dalam evaluasi [7].

Untuk membantu proses evaluasi kinerja pegawai negeri sipil di Kabupaten Grobogan, dilakukan pendekatan dengan menggunakan metode k-means untuk mengelompokkan data. K-means adalah sebuah algoritma clustering yang digunakan untuk mengklasifikasikan data dengan memisahkan atau mengelompokkan data ke dalam kategori yang jelas dan berbeda satu sama lain. Teknik clustering ini memanfaatkan Algoritma K-means karena algoritma ini berfokus pada jarak antara data. Keunggulan utama dari Algoritma K-means adalah tingkat ketelitian yang tinggi dalam mengukur objek, sehingga algoritma ini menjadi pilihan yang efisien dan dapat diandalkan dalam proses evaluasi kinerja [3].

2. METODE PENELITIAN

Metode penelitian yang digunakan dalam penulisan penelitian ini adalah sebagai berikut :

2.1. Data mining

Data mining [4], yang juga sering disebut sebagai Knowledge Discovery in Databases (KDD), merujuk kepada suatu metode otomatis untuk menggali informasi dari dalam basis data yang sangat besar. Tujuannya adalah untuk mengidentifikasi pola dengan menerapkan pengelompokan (clustering) atau pencarian hubungan (association) [8]. Secara umum, proses data mining mencakup tugas seperti menjelaskan, meramalkan, mengestimasi, mengklasifikasikan, mengelompokkan, dan mengidentifikasi asosiasi dalam data [14].

2.2. Clustering

Clustering data [5], merupakan salah satu teknik dalam Data Mining yang bersifat tanpa supervisi (unsupervised). Terdapat dua jenis clustering data yang sering digunakan dalam upaya pengelompokan data, yakni hierarchical (hierarki) dan non-hierarchical (non hirarki). Tujuan utama dari teknik clustering data ini adalah untuk mengoptimalkan fungsi tujuan yang telah ditetapkan dalam proses pengelompokan [13]. Fungsi ini umumnya bertujuan untuk mengurangi variasi di dalam setiap kelompok (cluster) serta meningkatkan variasi antara kelompok-kelompok tersebut [9].

2.3. Algoritma K-means

K-Mean merupakan teknik analisis kelompok yang bertujuan untuk mengelompokkan objek pengamatan ke dalam sejumlah K kelompok (cluster) [15]. Dalam metode ini, tiap objek akan ditempatkan dalam satu kelompok yang memiliki nilai rata-rata (mean) yang mewakili kelompok tersebut. K-Mean berupaya untuk menemukan pusat (centroid) dari setiap kelompok dengan melakukan iterasi untuk memperbaiki hasilnya. Tujuan dari metode ini adalah untuk mengorganisir data ke dalam kelompok-kelompok sehingga objek-objek dengan karakteristik yang serupa akan ditempatkan dalam satu kelompok, sementara objek-objek dengan karakteristik yang berbeda akan ditempatkan dalam kelompok yang berbeda pula.

Berikut adalah langkah-langkah teknik Clustering dengan menggunakan algoritma k-means, seperti yang dijelaskan oleh Siska [11]:

- a. Mulailah dengan menentukan jumlah cluster (k) yang akan dibentuk.
- b. Lakukan inisialisasi pusat cluster dengan cara acak atau random. Pusat-pusat ini akan digunakan sebagai pusat awal untuk setiap cluster.
- c. Hitung jarak antara setiap data dengan masing-masing centroid menggunakan rumus Euclidean Distance, seperti yang dinyatakan dalam persamaan berikut:

$$D_{(x,y)} = \sqrt{(X_{1x} - X_{1y})^2 + (X_{2x} - X_{2y})^2 + \dots + (X_{3x} - X_{3y})^2}$$

Keterangan:

$D_{(x,y)}$ = Jarak data ke ke pusat cluster

X_{kx} = Data ke i pada attribute ke k

X_{ky} = Titik pusat ke j pada attribute ke k

- d. Kelompokkan data berdasarkan jarak terdekat antara data dan centroid.
- e. Tentukan pusat cluster baru dengan menghitung rata-rata dari data yang ada dalam cluster tersebut menggunakan rumus yang sesuai.

$$V_{ij} = \frac{1}{N_i} \sum_{k=0}^{N_i} X_{kj}$$

Keterangan :

V_{ij} = rata-rata dari variable ke-j di dalam cluster ke-i

N_i = mewakili jumlah data yang menjadi anggota dari cluster ke-i

i, k = i dan k adalah indeks untuk cluster

j = j adalah indeks untuk variable

X_{kj} = dihitung dengan mengambil nilai rata-rata dari semua data $X_{k,j}$ yang termasuk ke dalam cluster tersebut

- f. Posisi suatu data akan ditentukan oleh jarak terpendek antara pusat cluster dan data tersebut.

- g. Ulangi langkah 3 hingga 5 hingga tidak ada perubahan anggota dalam setiap cluster. Jika dalam beberapa iterasi masih terdapat perubahan, maka ulangi proses clustering. Jika tidak ada perubahan, maka proses clustering dianggap selesai.

2.4. RapidMiner

RapidMiner merupakan perangkat lunak yang sering digunakan untuk melakukan ekstraksi data dengan berbagai metode dalam data mining [10]. Software ini menyediakan kemudahan bagi penggunaannya untuk mengolah data dalam jumlah besar dengan bantuan berbagai operator yang dapat digunakan untuk mengubah data. Hasil analisis yang dihasilkan oleh RapidMiner juga dapat disajikan secara visual melalui grafik. Beberapa fitur unggulan yang dimiliki oleh RapidMiner mencakup [12]:

- Menyediakan beragam algoritma data mining, termasuk decision tree dan k-means, serta algoritma lainnya.
- Menampilkan grafis yang canggih, seperti diagram histogram yang tumpang tindih, diagram pohon (tree chart), dan plot 3D Scatter.
- Plugin-plugin beragam tersedia di RapidMiner, termasuk plugin teks yang memungkinkan analisis teks.
- Merupakan platform yang komprehensif, mencakup prosedur data mining dan machine learning, seperti ETL (extraction, transformation, loading), data preprocessing, visualisasi, modeling, dan evaluasi.
- Proses data mining di RapidMiner didasarkan pada operator-operator yang dapat disusun secara bertingkat, dijelaskan melalui format XML, dan dapat dibuat melalui antarmuka grafis.
- Mengintegrasikan proyek data mining dengan Weka dan statistik R.

3. HASIL DAN PEMBAHASAN

3.1. Persiapan Data

Pada penelitian ini, tahapan awal yang dilakukan adalah pengumpulan data Pegawai Negeri Sipil (PNS) pada kabupaten Grobogan yang diambil pada data pegawai kondisi saat ini [6].

3.2. Preprocessing Data

Setelah data disiapkan, langkah selanjutnya yang dilakukan yaitu *preprocessing data* dari data asli untuk dapat diolah dengan *k-means* pada aplikasi rapidminer. Didalam *preprocessing data* dilakukan proses *cleaning* dan seleksi data. Setelah dilakukan seleksi data, atribut yang diperoleh yaitu jenis jabatan, pendidikan, diklat struktural, diklat fungsional, diklat teknis, seminar, skp, dan hukuman disiplin.

3.3. Transformasi Data

Pada transformasi data dilakukan proses pembobotan data dengan mengubah data string kedalam bentuk numerik agar dapat diolah. Berikut ini merupakan tabel hasil transformasi data dengan menghasilkan 4 atribut dengan di inisialkan A1 yaitu pendidikan, A2 kompetensi, A3 SKP, dan A4 hukuman disiplin.

Tabel 1. Hasil transformasi data dengan empat atribut

Data ke	NIP	A1	A2	A3	A4
1	19790309xxxxxxx009	15	25	25	5
2	19710314xxxxxxx002	15	25	25	5
3	19650530xxxxxxx001	20	25	25	5
4	19720114xxxxxxx006	15	15	30	5
5	19710720xxxxxxx005	15	40	25	5
6	19720904xxxxxxx002	20	25	25	5
7	19650130xxxxxxx002	15	25	30	5
8	19810516xxxxxxx011	15	40	25	5
9	19780308xxxxxxx002	20	25	30	5
10	19750426xxxxxxx010	15	25	25	5
....
66	19701209xxxxxxx011	15	17,5	25	5

67	19730228xxxxxxxx009	15	17,5	25	5
68	19690201xxxxxxxx019	10	17,5	25	5
69	19650412xxxxxxxx001	5	17,5	25	5
70	19720712xxxxxxxx018	5	17,5	25	5

Hasil transformasi data diatas didapat setelah melalui proses pembobotan pada setiap kriteria yang dijelaskan pada tabel dibawah ini :

Tabel 2. Hasil transformasi data pembobotan setiap kriteria

Kriteria	Indikator	Sub Bobot	Bobot		
Pendidikan	S-3	25	25		
	S-2	20			
	S-1/D-IV	15			
	D-III	10			
	D-II/D-I/SLTA	5			
	Dibawah SLTA	1			
Kompetensi	Dikstru (bagi Jabatan Struktural)	Sudah ikut	15	40	
		Belum ikut	0		
	Dikfung (bagi Jabatan Fungsional)	Sudah ikut	15		
		Belum ikut	0		
	Diktekn (bagi jabatan pelaksana nilai max 22,5)	Sudah ikut	15		22,5
		Belum ikut	0		0
Seminar (bagi jabatan pelaksana nilai max 17,5)	Sudah ikut	10	17,5		
	Belum ikut	0	0		
SKP	91-100	30	30		
	76-90	25			
	61-75	15			
	51-60	5			
	50 ke bawah	1			
Hukdis	Tidak Pernah Dikenai Hukdis	5	5		
	Ringan	3			
	Sedang	2			
	Berat	1			
Total Bobot			100		

3.4. Penggunaan K-Means

Berikut adalah cara menerapkan Algoritma K-Means pada data Pegawai Negeri Sipil Kabupaten Grobogan dengan langkah-langkah yang dapat dijelaskan secara berbeda:

- Mulailah dengan menentukan bahwa akan ada 3 kluster yang akan dibentuk.
- Dalam tahap pengujian, pilih secara acak 3 data pegawai yang akan menjadi pusat awal dari kluster-klusternya.

$$C1 = 10, 15, 25, 5$$

$$C2 = 15, 17,5, 25, 5$$

$$C3 = 20, 25, 30, 5$$

Kluster 1 diambil dari data PNS dengan NIP 19840218xxxxxxxx004,

Kluster 2 diambil dari data PNS dengan NIP 19890101xxxxxxxx001,

Kluster 3 diambil dari data PNS dengan NIP 19780308xxxxxxxx002.

- Selanjutnya, lakukan perhitungan jarak terpendek untuk setiap kluster. Setelah selesai menghitung jarak-jarak ini, langkah berikutnya adalah mengelompokkan data berdasarkan jarak terpendek menggunakan rumus Euclidean:

$$D_{(x,y)} = \sqrt{(X_{1x} - X_{1y})^2 + (X_{2x} - X_{2y})^2 + \dots + (X_{nx} - X_{ny})^2}$$

$$D_{(x,y)} = \text{Jarak data ke } x \text{ ke pusat cluster } y$$

X_{nx} = Data ke i pada attribute ke n
 X_{ny} = Titik pusat ke j pada attribute ke n

Jarak data dengan kluster 1

$$D_{(x1,c1)} = \sqrt{(15 - 10)^2 + (25 - 15)^2 + (25 - 25)^2 + (5 - 5)^2}$$

$$D_{(x2,c1)} = \sqrt{(15 - 10)^2 + (25 - 15)^2 + (25 - 25)^2 + (5 - 5)^2}$$

$$D_{(x1,c1)} = \sqrt{(20 - 10)^2 + (25 - 15)^2 + (25 - 25)^2 + (5 - 5)^2}$$

Lanjutkan perhitungan jarak kluster 1 sampai pada pegawai terakhir.

$$D_{(x70,c1)} = \sqrt{(5 - 10)^2 + (17,5 - 15)^2 + (25 - 25)^2 + (5 - 5)^2}$$

Jarak data dengan kluster 2

$$D_{(x1,c2)} = \sqrt{(1 - 1)^2 + (2 - 1)^2 + (2 - 2)^2 + (5 - 5)^2}$$

$$D_{(x2,c2)} = \sqrt{(15 - 10)^2 + (25 - 15)^2 + (25 - 25)^2 + (5 - 5)^2}$$

$$D_{(x3,c2)} = \sqrt{(20 - 10)^2 + (25 - 15)^2 + (25 - 25)^2 + (5 - 5)^2}$$

Lanjutkan perhitungan jarak kluster 2 sampai pada pegawai terakhir.

$$D_{(x70,c2)} = \sqrt{(5 - 10)^2 + (-15)^2 + (25 - 25)^2 + (5 - 5)^2}$$

Jarak data dengan kluster 3

$$D_{(x1,c2)} = \sqrt{(1 - 1)^2 + (2 - 1)^2 + (2 - 2)^2 + (5 - 5)^2}$$

$$D_{(x2,c2)} = \sqrt{(15 - 10)^2 + (25 - 15)^2 + (25 - 25)^2 + (5 - 5)^2}$$

$$D_{(x3,c2)} = \sqrt{(20 - 10)^2 + (25 - 15)^2 + (25 - 25)^2 + (5 - 5)^2}$$

Lanjutkan perhitungan jarak kluster 2 sampai pada pegawai terakhir.

$$D_{(x70,c2)} = \sqrt{(5 - 10)^2 + (17,5 - 15)^2 + (25 - 25)^2 + (5 - 5)^2}$$

d. Berikut ini adalah tabel yang menampilkan hasil perhitungan jarak pada Iterasi 1.

Tabel 3. Hasil perhitungan jarak Iterasi Pertama.

No	NIP	Jarak Ke Centroid			Jarak Terdekat
		C1	C2	C3	
1	19790309xxxxxxxx009	11,18034	7,5	7,071068	C3
2	19710314xxxxxxxx002	11,18034	7,5	7,071068	C3
3	19650530xxxxxxxx001	14,14214	9,013878	5	C3
4	19720114xxxxxxxx006	7,071068	5,59017	11,18034	C2
5	19710720xxxxxxxx005	25,4951	22,5	16,58312	C3
6	19720904xxxxxxxx002	14,14214	9,013878	5	C3
7	19650130xxxxxxxx002	12,24745	9,013878	5	C3
8	19810516xxxxxxxx011	25,4951	22,5	16,58312	C3
9	19780308xxxxxxxx002	15	10,30776	0	C3

10	19750426xxxxxxxx010	11,18034	7,5	7,071068	C3
....
66	19701209xxxxxxxx011	5,59017	0	10,30776	C2
67	19730228xxxxxxxx009	5,59017	0	10,30776	C2
68	19690201xxxxxxxx019	2,5	5	13,46291	C1
69	19650412xxxxxxxx001	5,59017	10	17,5	C1
70	19720712xxxxxxxx018	5,59017	10	17,5	C1

- e. Hitung rata-rata dari data pegawai yang termasuk dalam kluster tersebut untuk menentukan pusat kluster yang baru.
- f. Setelah memperoleh pembaruan untuk centroid, langkah berikutnya adalah memulai kembali iterasi dengan menggunakan pusat centroid yang baru. Ini melibatkan melakukan perhitungan yang serupa seperti yang dilakukan pada iterasi pertama, terus menerus, hingga tidak ada data yang mengalami perpindahan ke kluster lainnya.
- g. Setelah tidak ada data yang berpindah, jumlahkan masing-masing titik kluster terakhir untuk mengetahui seberapa baik kluster tersebut terbentuk. Hasil perhitungan titik pusat centroid akhir dapat ditemukan dalam tabel berikut:

Tabel 4. Hasil perhitungan titik pusat centroid akhir

	A1	A2	A3	A4
C1	7,1875	17,1875	25	5
C2	16,83333	15,58333	25,5	4,8
C3	15,625	28,54167	25,41667	5

$$\begin{aligned} \text{Cluster 1} &= 7,1875 + 17,1875 + 11,95833333 + 25 + 5 \\ &= 54,375 \end{aligned}$$

$$\begin{aligned} \text{Cluster 2} &= 16,83333 + 15,58333 + 25,5 + 4,8 \\ &= 62,71666667 \end{aligned}$$

$$\begin{aligned} \text{Cluster 3} &= 15,625 + 28,54167 + 25,41667 + 5 \\ &= 74,58333333 \end{aligned}$$

- h. Pada tahap akhir, lakukan perhitungan rasio dengan menghitung perbandingan antara BCV (Between-Cluster Variance) dan WCV (Within-Cluster Variance) atau SSE (Sum of Squared Errors) pada iterasi terakhir.

$$BCV = \frac{1}{nk} \sum_i^k d(m_i, m_i)$$

$$BCV = d(m_1, m_2) + d(m_1, m_3) + d(m_2, m_3)$$

$$BCV = \sqrt{\begin{aligned} &(7,1875 - 16,83333)^2 + (17,1875 - 15,58333)^2 + \\ &(25 - 25,5)^2 + (5 - 4,8)^2 \end{aligned}}$$

$$+ \sqrt{\begin{aligned} &(7,1875 - 15,625)^2 + (17,1875 - 28,54167)^2 + \\ &(25 - 25,41667)^2 + (5 - 5)^2 \end{aligned}}$$

$$+ \sqrt{\begin{aligned} &(16,83333 - 15,625)^2 + (15,58333 - 25,54167)^2 + \\ &(25,5 - 25,41667)^2 + (4,8 - 5)^2 \end{aligned}}$$

$$BCV = 36,96159125$$

Tabel 5. Jarak centroid terkecil pada iterasi 4

NIP	Jarak terdekat	NIP	Jarak terdekat
19790309xxxxxxxx009	3,620447	19770420xxxxxxxx003	1,997846
19710314xxxxxxxx002	3,620447	19870126xxxxxxxx002	1,997846
19650530xxxxxxxx001	5,644257	19660227xxxxxxxx001	3,264668
19720114xxxxxxxx006	4,898101	19660511xxxxxxxx001	1,997846
19710720xxxxxxxx005	11,48293	19680930xxxxxxxx001	3,264668
19720904xxxxxxxx002	5,644257	19670912xxxxxxxx001	3,264668
19650130xxxxxxxx002	5,825888	19660319xxxxxxxx010	3,264668
19810516xxxxxxxx011	11,48293	19700507xxxxxxxx016	1,997846
19780308xxxxxxxx002	7,258855	19840802xxxxxxxx002	3,093592
19750426xxxxxxxx010	3,620447	19760112xxxxxxxx004	3,620447
19660906xxxxxxxx007	3,264668	19890611xxxxxxxx001	3,620447
19680525xxxxxxxx001	5,818768	19670330xxxxxxxx014	1,997846
19670514xxxxxxxx003	5,005802	19760126xxxxxxxx012	1,997846
19720923xxxxxxxx007	1,997846	19790901xxxxxxxx006	2,829808
19751022xxxxxxxx002	5,644257	19890101xxxxxxxx001	2,706422
19660703xxxxxxxx007	5,536972	19760910xxxxxxxx011	2,209709
19660912xxxxxxxx009	1,997846	19680125xxxxxxxx004	2,209709
19850331xxxxxxxx004	3,620447	19811128xxxxxxxx001	2,706422
19800220xxxxxxxx014	6,660153	19800607xxxxxxxx002	2,829808
19840218xxxxxxxx004	3,563048	19780704xxxxxxxx016	2,829808
19710410xxxxxxxx007	1,997846	19760410xxxxxxxx003	2,706422
19780314xxxxxxxx010	1,997846	19750105xxxxxxxx003	2,706422
19850226xxxxxxxx006	3,264668	19671005xxxxxxxx024	2,209709
19680309xxxxxxxx003	1,640418	19700905xxxxxxxx012	2,209709
19831018xxxxxxxx008	11,48293	19870602xxxxxxxx003	3,620447
19760308xxxxxxxx009	11,48293	19680227xxxxxxxx004	2,209709
19770808xxxxxxxx004	15,63194	19840422xxxxxxxx022	2,706422
19700125xxxxxxxx001	4,63044	19660504xxxxxxxx001	2,209709
19780120xxxxxxxx005	3,620447	19961216xxxxxxxx003	2,829808
19741028xxxxxxxx004	3,620447	19890724xxxxxxxx008	2,829808
19700525xxxxxxxx004	3,620447	19701209xxxxxxxx011	2,706422
19691125xxxxxxxx003	5,644257	19730228xxxxxxxx009	2,706422
19761106xxxxxxxx005	1,997846	19690201xxxxxxxx019	2,829808
19850218xxxxxxxx009	3,264668	19650412xxxxxxxx001	2,209709
19680828xxxxxxxx007	3,264668	19720712xxxxxxxx018	2,209709

$$\begin{aligned}
WCV(SSE) &= \sum_{j=i}^n \sum_{p \in c_i} d(p, m_i)^2 \\
&= 3,620447^2 + 3,620447^2 + 5,644257^2 + 4,898101^2 + 11,48293^2 + 5,644257^2 + \\
&5,825888^2 + 11,48293^2 + 7,258855^2 + 3,620447^2 + 3,264668^2 + 5,818768^2 + \\
&5,005802^2 + 1,997846^2 + 5,644257^2 + 5,536972^2 + 1,997846^2 + 3,620447^2 + \\
&6,660153^2 + 3,563048^2 + 1,997846^2 + 1,997846^2 + 3,264668^2 + 1,640418^2 + \\
&11,48293^2 + 11,48293^2 + 15,63194^2 + 4,630444^2 + 3,620447^2 + 3,620447^2 + \\
&3,620447^2 + 5,644257^2 + 1,997846^2 + 3,264668^2 + 3,264668^2 + 1,997846^2 + \\
&1,997846^2 + 3,264668^2 + 1,997846^2 + 3,264668^2 + 3,264668^2 + 3,264668^2 + \\
&1,997846^2 + 3,093592^2 + 3,620447^2 + 3,620447^2 + 1,997846^2 + 1,997846^2 +
\end{aligned}$$

$$2,829808^2 + 2,706422^2 + 2,209709^2 + 2,209709^2 + 2,706422^2 + 2,829808^2 + 2,829808^2 + 2,706422^2 + 2,706422^2 + 2,209709^2 + 3,620447^2 + 2,209709^2 + 2,706422^2 + 2,209709^2 + 2,829808^2 + 2,829808^2 + 2,706422^2 + 2,706422^2 + 2,829808^2 + 2,209709^2 + 2,209709^2$$

$$WCV = 1588,8$$

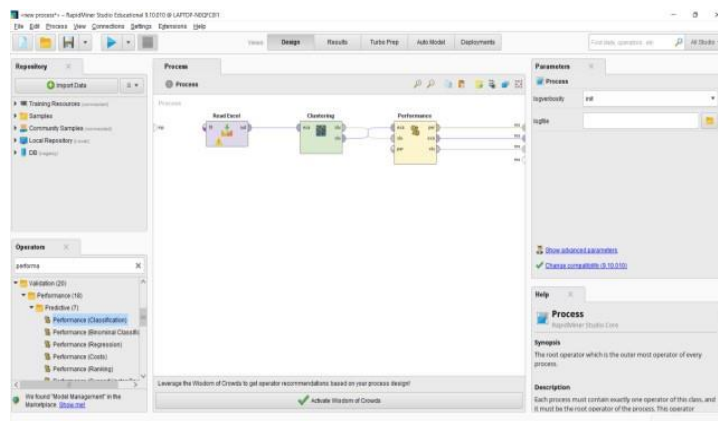
Dalam pengujian ini, yang menggunakan perbandingan Variasi Antara Kelas (Between-Class Variation - BCV) dan Variasi Dalam Kelas (Within-Class Variation - WCV).

$$Rasio = \frac{BCV}{WCV} = \frac{36,96159125}{1588,8} = 0,023263841$$

Ditemukan bahwa nilai rasio 0,023263841 memenuhi kriteria penilaian yang sangat baik ketika diperiksa dalam tabel penilaian rasio.

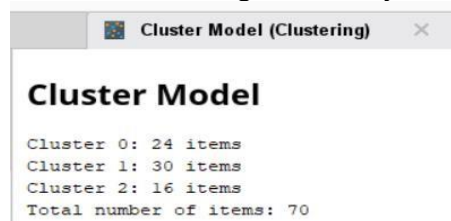
3.5. Implementasi RapidMiner

Berikut ini merupakan langkah penggunaan implementasi *k-means clustering* pada tools Rapidminer. Langkah awal yaitu data PNS dibuat dalam data set baru dengan format .xlsx (excel) dan sudah melalui proses preprocessing untuk bisa dianalisis menggunakan aplikasi Rapidminer Studio. Tentukan pemodelan desain clustering menggunakan algoritma K-Means kemudian lakukan drag and drop operators yang akan digunakan dan hubungkan setiap operatornya serta atur paramater disetiap operatornya seperti gambar dibawah ini :



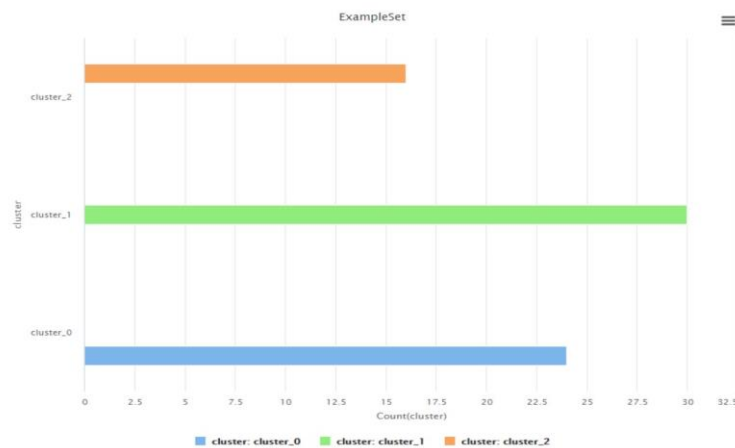
Gambar 1. Desain Clustering Menggunakan Algoritma K-Means

Setelah menghubungkan setiap operator dan mengatur setiap operator selanjutnya, proses selanjutnya adalah menjalankan proses secara lokal *Run Proses Locally*. Dari hasil clustering data sebanyak 70 data PNS di Pemerintah Kabupaten Grobogan yang dipilih secara acak dengan implementasi pada tools RapidMiner diperoleh tiga cluster baik sebanyak 24 (34%) pegawai, cluster cukup sebanyak 30 (43%) pegawai dan cluster kurang baik sebanyak 16 (23%) pegawai.



Gambar 2. Hasil Cluster

Inilah visualisasi model cluster K-Means dalam bentuk diagram batang yang dapat ditemukan pada gambar 3.



Gambar 3. Tampilan Chart Bar Algoritma K-Means Clustering

3.6. Perbandingan Hasil Perhitungan

Berdasarkan hasil penelitian yang dilakukan pada data PNS di Pemerintah Daerah Kabupaten Grobogan, selain menggunakan perhitungan manual dengan aplikasi Microsoft Excel, penelitian ini juga memanfaatkan aplikasi data mining RapidMiner Studio. Berikut adalah perbandingan hasil perhitungan yang ditemukan:

Tabel 6. Perbandingan Hasil Perhitungan

Perbandingan	Perhitungan Manual			Perhitungan RapidMiner		
Jumlah Literasi	4			-		
Jumlah Cluster	C0	C1	C2	C0	C1	C2
	16	30	24	24	30	16
Jumlah Centroid	3			3		
Jumlah Nilai Rasio (%)	22,85714	42,85714	34,28571	34,28571	42,85714	22,85714

Pada tabel 6 dapat dilihat perhitungan dalam menggunakan Microsoft Excel dan RapidMiner menghasilkan jumlah data yang sama setiap keanggotaan clusternya namun dengan urutan cluster yang berbeda. Perbedaan urutan cluster tersebut terjadi pada cluster 0 dan cluster 2 dimana pada perhitungan menggunakan Microsoft Excel masing-masing berjumlah 16 data dan 24 data. Sedangkan pada perhitungan menggunakan RapidMiner masing-masing berjumlah 24 data dan 16 data. Ini disebabkan oleh perbedaan dalam menetapkan centroid awal. Dalam menentukan centroid awal dengan Microsoft Excel, peneliti melakukan penentuan secara acak. Sementara, dalam aplikasi RapidMiner, proses penentuan centroid awal dilakukan secara otomatis oleh aplikasi.

4. KESIMPULAN

Dari hasil penelitian dan pengujian yang dilakukan oleh peneliti dapat disimpulkan bahwa MetodeDK-Means ClusteringA dapat diimplementasikan kedalam evaluasi kinerja pegawai negeri sipil di Lingkungan Pemerintah Kabupaten Grobogan. Berdasarkan pengolahan data yang telah dilakukan menghasilkan 3 kelompok PNS yang mempunyai tingkat kinerja baik berjumlah 24 (34%) pegawai, cluster cukup sebanyak 30 (43%) pegawai dan cluster kurang baik sebanyak 16 (23%) pegawai dari 70 data sampel pegawai yang dipilih secara random. Masih banyak pegawai negeri sipil pada Pemerintah Kabupaten Grobogan dengan kinerja yang kurang baik. Dengan rapidminer dapat membantu mengetahui kualitas kinerja PNS kabupaten Grobogan dengan perhitungan kurang dari satu detik sehingga tidak membutuhkan waktu lama untuk menilai kinerja pegawai. Hasil perhitungan menggunakan tools RapidMiner tidak banyak perubahan secara signifikan dengan perhitungan manual karena menghasilkan jumlah data yang sama setiap keanggotaan clusternya namun dengan urutan cluster yang berbeda. Didalam proses pengujian atau evaluasi hasil perhitungan clustering K-Means antara tools RapidMiner dan manual sama-sama memiliki kualitas perhitungan yang baik dimana masing-masing nilai evaluasi yang didapat < 1 dan mendekati angka 0.

DAFTAR PUSTAKA

- [1] Perbup Grobogan Nomor 94 Tahun 2021, Kedudukan, Susunan Organisasi, Tugas, Fungsi, Uraian Tugas Jabatan Dan Tata Kerja Badan Kepegawaian, Pendidikan Dan Pelatihan Daerah Kabupaten Grobogan, Purwodadi, 2021.
- [2] Peraturan Pemerintah Nomor 46 Tahun 2011, Penilaian Prestasi Kerja Pegawai Negeri Sipil, Jakarta, 2011
- [3] Regina, Sandra dkk, "Clustering Kualitas Kinerja Karyawan Pada Perusahaan Bahan Kimia Menggunakan Algoritma K-Means," Jakarta: STMIK Nusa Mandiri, 2021.
- [4] Triandin, Melissa dkk, "Data Mining dalam Mengukur Tingkat Keaktifan Siswa dalam Mengikuti Proses Belajar pada SMP IT Andalas Cendekia dengan Menggunakan Metode K-Means Clustering," Padang: Universitas Putra Indonesia, 2021
- [5] F. T. Rantika, I. F. Astuti and A. H. Kridalaksana, "Clustering Data PNS Status Tugas Belajar Dan Ijin Belajar Menggunakan Metode Fuzzy C-Means (Studi Kasus : PNS Di Lingkungan Pemerintah Provinsi Kalimantan Timur)," Prosiding Seminar Sains dan Teknologi FMIPA Unmul, vol. 1, pp. 1-5, 2015
- [6] E. F. Pranasta, Penerapan Data Mining Untuk Penilaian Kinerja Pns Pada Bkpp Kabupaten Gayo Lues Menggunakan Metode Clustering, Bandung: E Fauzi Pranasta, 2015.
- [7] S. Henrei, "Sistem Informasi Manajemen Kepegawaian (SIMPEG)," 14 Desember 2014. [Online]. Available: <http://shinzotech.com/product/38/sistem-informasi-manajemen-kepegawaian-simpeg>. [Accessed 7 April 2018].
- [8] F. N. R. F. J. Aziz, B. d. Setiawan and A. Issa, "Implementasi Algoritma K-Means untuk Klasterisasi Kinerja Akademik Mahasiswa," Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer, vol. 2, pp. 2243-2251, 2018.
- [9] L. Parlina, A. P. Widarto, A. Wanto and M. R. Lubis, "Memanfaatkan Algoritma K-Means Dalam Menentukan Pegawai Yang Layak Mengikuti Asessment Center Untuk Clustering Program Sdp," CESS (Journal of Computer Engineering System and Science), vol. 3, pp. 87-93, 2018.
- [10] CTI, B. R., Gafar, A. A., Fajriani, N., Ramdani, U., Uyun, F. R., Ransi, N., dkk, "Implementasi k-means clustering pada rapidminer untuk analisis daerah rawan kecelakaan. Dalam Prosiding seminar nasional riset kuantitatif terapan 2017 (Vol. 1), 2017.
- [11] S. T. Siska, "Analisa dan Penerapan Data Mining untuk menentukan Kubikasi Air Terjual berdasarkan Pengelompokan Pelanggan menggunakan Algoritma K-Means Clustering," J. Teknol. Inf. Pendidik., vol. 9, no. 1, pp. 153–162, 2016.
- [12] Aprilla, d., C Dennis, "Belajar data mining dengan rapidminer," Jakarta, Indonesia, 2013.
- [13] Suprayogi, "Data Mining Clustering," Erlin Kusuma, 2017.
- [14] Turban, E, "Decision Support Systems and Intelligent Systems Edisi Bahasa Indonesia Jilid 1," Andi: Yogyakarta, 2005.
- [15] L. Kaufman, and P. J. Rousseuw, "Finding Groups in Data: An Introduction to Cluster Analysis.," Biometrics, 1990.