



# Penerapan Algoritma *K-Means* dalam Proses *Clustering* Penilaian Kinerja Aparatur Sipil Negara di Sekretariat DPRD Pematangsiantar

Della Aulia<sup>1</sup>, M Safii<sup>2</sup>, Dedi Suhendro<sup>3</sup>

<sup>1</sup>Mahasiswa Program Studi Sistem Informasi, STIKOM Tunas Bangsa, Pematangsiantar

<sup>2,3</sup>AMIK Tunas Bangsa, Pematangsiantar

<sup>1</sup>dellaaulia854@gmail.com, <sup>2</sup>m.safii@amiktunasbangsa.ac.id, <sup>3</sup>dedi.su@amiktunasbangsa.ac.id

## Abstract

*This study discusses the assessment of the performance of the ASN (State Civil Apparatus) in the Pematangsiantar DPRD Secretariat based on the quality of its work. In carrying out their performance, employees are still often truant and have a poor work ethic. In this case the research aims to improve employee welfare. During this time the amount of employee income is only based on the group and position they have. This study uses the K-Means method to classify or classify employee performance quality assessments based on SKP (employee work objectives) with additional income based on work quality assessment and employee behavior. After conducting this research there is a result that in carrying out the performance of employees include Quality, Quantity, Time, Cost of each task activity so as to produce a system capable of assisting the appraisal officer in evaluating the quality of employee performance using the K-Means algorithm as information to find out the employee including very good, enough, and less.*

**Keywords:** Data Mining, Performance of the State Civil Apparatus, K-Means Clustering Algorithm, RapidMiner

## Abstrak

*Penelitian ini membahas tentang penilaian kinerja ASN (Aparatur Sipil Negara) di Sekretariat DPRD Pematangsiantar berdasarkan kualitas kerjanya. Dalam melaksanakan kerjanya pegawai masih sering bolos dan memiliki etika kerja yang kurang baik. Dalam hal ini dilakukannya penelitian bertujuan untuk meningkatkan kesejahteraan pegawai. Selama ini besaran penghasilan pegawai hanya didasarkan pada golongan dan jabatan yang dimilikinya. Penelitian ini menggunakan metode K-Means untuk mengelompokkan atau mengklusterisasi penilaian kualitas kinerja pegawai berdasarkan SKP (Sasaran Kerja Pegawai) dengan adanya tambahan penghasilan berdasarkan penilaian kualitas kerja dan perilaku pegawai. Setelah melakukan penelitian ini terdapat hasil bahwa dalam melaksanakan kerjanya pegawai meliputi Quality, Quantity, Time, Cost dari setiap kegiatan tugasnya sehingga menghasilkan sistem yang mampu membantu petugas penilai dalam melakukan penilaian kualitas kinerja pegawai dengan menggunakan algoritma K-Means sebagai informasi untuk mengetahui pegawai tersebut termasuk sangat baik, cukup, dan kurang.*

**Kata kunci:** Data Mining, Kinerja Aparatur Sipil Negara, Algoritma K-Means Clustering, RapidMiner

## 1. PENDAHULUAN

Penilaian kinerja (*Performance Appraisal*) merupakan suatu proses yang digunakan untuk menentukan apakah seorang pegawai melakukan pekerjaannya sesuai dengan tugas dan tanggung jawabnya. Penilaian kinerja atau *Performance Appraisal* ini juga bisa sebagai dasar untuk kenaikan gaji, penurunan jabatan maupun bonus atau tambahan bagi pegawai. Disisi lain penilaian kinerja yang dilakukan dengan baik dan profesional dapat meningkatkan loyalitas dan motivasi pegawai sehingga tujuan suatu kantor dapat tercapai sesuai dengan yang diharapkan.

Di lingkup instansi, kegiatan penilaian kinerja dapat diterapkan untuk menilai kinerja pegawai. Salah satu aspek yang dapat digunakan sebagai acuan dalam penilaian kinerja pegawai adalah penilaian yang berdasarkan SKP (Sasaran Kerja Pegawai) yang sudah ditetapkan. Untuk dapat lebih membantu proses penilaian tersebut, perlu adanya pengelolaan dan pengolahan data yang baik. Pengolahan data yang dimaksud adalah data-data yang dimiliki hasil penilaian berdasarkan SKP dapat diproses dengan baik agar data-data tersebut dapat memberikan suatu informasi yang lebih jelas dan akurat. Pimpinan kantor perlu mengetahui hasil kinerja pegawai, dari hasil tersebut akan diketahui kualitas kinerja pegawai dalam melaksanakan tugas dan tanggung jawabnya, sehingga pegawai bisa mendapatkan tambahan upah atau bonus dari kualitas kerjanya dan tidak berdasarkan jabatan atau golongannya.

Penelitian yang dilakukan oleh [1] bahwa hasil pengelompokan kinerja dosen menggunakan *K-Means* terbentuk data kelompok dosen sangat baik terdiri dari 12 anggota dengan total nilai *centroid* 48.550, data kelompok dosen sangat baik terdiri 29 anggota dengan total nilai *centroid* 40.340, data kelompok dosen cukup baik 10 anggota dengan total nilai *centroid* 37.963 dan kelompok dosen kurang baik terdiri dari 9 anggota dengan total nilai *centroid* 37.033.

Penelitian lainnya dilakukan oleh [2] bahwa pengelompokan kualitas kerja pegawai menggunakan algoritma *K-Means* dalam pemeliharaan kesehatan pegawai terdiri atas 5 kelompok kerja dengan pegawai yang berbeda dan penyakit yang berbeda, dimana penyakit dan jumlah penyakit dalam kelompok tersebut akan digunakan sebagai bahan pertimbangan Bu Fatma untuk merencanakan program kesehatan pegawai di semester berikutnya.

Dalam penelitian ini diharapkan adanya peningkatan kinerja pegawai. Melalui pelaksanaan penilaian kinerja pegawai yang dilakukan oleh peneliti akan menghasilkan klasterisasi pegawai dimana data-datanya diperoleh dari penilaian pejabat penilai terhadap pegawai secara objektif dan transparan.

## 2. METODOLOGI PENELITIAN

### 2.1. Data Mining

Data mining merupakan serangkaian proses untuk menggali nilai tambah berupa informasi yang selama ini tidak diketahui secara manual dari suatu basis data. Data mining mulai ada sejak 1990-an sebagai cara yang benar dan tepat untuk mengambil pola dan informasi yang digunakan untuk menemukan hubungan antara data untuk melakukan pengelompokan ke dalam satu atau lebih cluster sehingga objek - objek yang berada dalam satu cluster akan mempunyai kesamaan yang tinggi antara satu dengan lainnya. Data mining merupakan bagian dari proses penemuan pengetahuan dari basis data Knowledge Discovery in [3].

### 2.2. Algoritma *K-Means*

*K-Means* merupakan salah satu metode *clustering* non-hirarki yang berusaha mempartisi data yang ada ke dalam bentuk satu atau lebih cluster. Algoritma ini mempartisi data ke dalam *cluster* sehingga data yang memiliki karakteristik yang sama dikelompokkan ke dalam satu *cluster* yang sama dan data yang mempunyai



karateristik yang berbeda di kelompokkan ke dalam cluster yang lain. Algoritma ini sederhana untuk diterapkan dan dijalankan, relatif cepat, mudah diadaptasi, dan umum digunakan dalam praktek [4].

Menurut (Khomaruddin, 2016) *K-Means* merupakan salah satu metode *cluster* analisis non hirarki yang berusaha untuk mempartisi objek yang ada ke dalam satu atau lebih cluster atau kelompok objek yang mempunyai karakteristik yang berbeda dikelompokkan ke dalam *cluster* lain.

Menurut (witten, 2018) adapun langkah-langkah metode *K-Means* antara lain sebagai berikut [4]:

- a) Pilih secara acak k buah data sebagai pusat *cluster*
- b) Jarak antara data dan pusat cluster dihitung menggunakan *Euclidian Distance*. Untuk menghitung jarak semua data ke setiap titik pusat *cluster* dapat menggunakan teori jarak *Euclidean* yang dirumuskan sebagai berikut :

$$d_{ij} = \sqrt{\sum_{k=1}^p \{X_{ik} - X_{jk}\}^2} \quad (1)$$

Keterangan :

- $d_{ij}$  = Jarak objek antara objek i dan j
  - $p$  = Dimensi data
  - $X_{ik}$  = Koordinat dari objek i pada dimensi k
  - $X_{jk}$  = Koordinat dari objek j pada dimensi k
- c) Data ditempatkan dalam *cluster* yang terdekat, dihitung dari tengah *cluster*.
  - d) Pusat *cluster* baru akan ditentukan bila semua data telah ditetapkan dalam *cluster* terdekat.
  - e) Proses penentuan pusat *cluster* dan penempatan data dalam *cluster* diulangi sampai nilai *cendroid* tidak berubah lagi.

### 2.3. RapidMiner

RapidMiner merupakan perangkat lunak yang bersifat terbuka (*open source*). *RapidMiner* adalah sebuah solusi untuk melakukan analisis terhadap data mining, text mining dan analisis pengelompokan. *RapidMiner* menggunakan berbagai teknik *deskriptif* dan prediksi dalam memberikan wawasan kepada pengguna sehingga dapat membuat keputusan yang paling baik. *RapidMiner* memiliki kurang lebih 500 operator *datamining*, termasuk operator untuk input, output, data *preprocessing* dan visualisasi. *RapidMiner* merupakan software yang berdiri sendiri untuk analisis data dan sebagai mesin data mining yang dapat diintegrasikan pada produknya sendiri. *RapidMiner* ditulis dengan menggunakan bahasa java sehingga dapat bekerja di semua sistem operasi. [4].

## 3. HASIL DAN PEMBAHASAN

Dalam penelitian ini data yang digunakan adalah data SKP (Sasaran Kerja Pegawai) tahun 2017 yang diperoleh dari Kantor DPRD Kota Pematangsiantar. Data yang ada dalam penelitian ini dikelompokkan menjadi 3 *cluster*, data yang diperoleh akan dihitung terlebih dahulu. Pada penelitian ini aplikasi yang digunakan yaitu *RapidMiner* sebagai pengujian data yang dianalisa.

### 3.1. Pengolahan Data

Berikut uraian perhitungan manual proses Algoritma *K-Means clustering*.

#### a) Menentukan Data Yang Akan Di *Cluster*

Sampel data yang akan digunakan dalam proses *clustering* kinerja pegawai berdasarkan SKP (Sasaran Kerja Pegawai). Berikut daftar tabel penggunaan data dapat dilihat pada tabel 1 :

**Tabel 1.** Data Sasaran Kerja Pegawai

No	Nama	SKP	Perilaku
1	Evi	86.22	84
2	Berliana	85.87	83.65
3	Nurul	84.52	81.35
4	Ratna	86.19	82.91
5	Isum	85.33	83.2
6	Sudirson	84.8	82.24
7	Edy	84.19	81.98
8	Jekson	85.55	83.17
9	Tri	84.13	81.12
10	Sutiani	91.58	82.92
11	Susi	84.49	82.82
12	Suyani	86.19	84.03
13	Wanden	85.52	86.25
14	Wulandani	84.73	81.96
15	Rumiris	85.86	83.59
16	Lisma	85.08	82.68
17	Sumiah	85.03	82.86
18	Oloan	84.05	82.27
19	Nazmah	85.03	82.86
20	Hajjah	85.13	82.28
21	Haposan	85.55	83.17
22	Binahar	84.67	82.22
23	Setia	85.03	82.86

#### b) Menentukan Nilai *k* Jumlah *Cluster*

Menentukan jumlah *cluster* yang diinginkan. Dalam penelitian ini data-data yang ada akan dikelompokkan menjadi tiga *cluster*. Yaitu *cluster* dengan nilai tertinggi (C1), *cluster* dengan nilai tengah (C2), dan *cluster* dengan nilai terendah (C3).

#### c) Menentukan Nilai *Centroid* (*Centroid* Pusat)

Penentuan pusat *cluster* awal ditentukan secara random yang diambil dari data yang ada dalam *range* dan didapat titik pusat dari setiap *cluster* dapat dilihat pada tabel 2

**Tabel 2.** *Centroid* Data Awal

	SKP	PERILAKU
<i>Cluster 1</i>	91.58	82.92
<i>Cluster 2</i>	85.08	82.68

	SKP	PERILAKU
<i>Cluster 3</i>	84.05	82.270

d) Menghitung Jarak Setiap Data Terhadap *Centroid* (Pusat *Cluster*)

Setelah data nilai pusat *cluster* awal ditentukan, maka selanjutnya adalah menghitung jarak masing-masing data terhadap pusat *cluster*. Proses pencarian jarak terpendek pada iterasi 1 dapat dilihat pada perhitungan dan tabel dibawah ini.

$$D_{Evi,C1} = \sqrt{(86,22 - 91,58)^2 + (84 - 82,92)^2} = 5,467723$$

$$D_{Evi,C2} = \sqrt{(86,22 - 85,08)^2 + (84 - 82,68)^2} = 1,744133$$

$$D_{Evi,C3} = \sqrt{(86,22 - 84,05)^2 + (84 - 82,27)^2} = 2,775212$$

Hasil dari keseluruhan perhitungan dapat dilihat pada tabel 3 sebagai berikut :

**Tabel 3.** Hasil Perhitungan Jarak Pusat *Cluster* Iterasi 1

No	Nama	SKP	Perilaku	C1	C2	C3	Jarak Terpendek
1	Evi	86.22	84.00	5.467723	1.744133	2.775212	1.744133022
2	Berliana	85.87	83.65	5.756475	1.251	2.284032	1.2509996
3	Nurul	84.52	81.35	7.232462	1.443087	1.033102	1.033102125
4	Ratna	86.19	82.91	5.390009	1.133578	2.233652	1.133578405
5	Isum	85.33	83.2	6.256269	0.576975	1.582182	0.576974869
6	Sudirson	84.80	82.24	6.814015	0.521536	0.7506	0.521536192
7	Edy	84.19	81.98	7.449544	1.132299	0.322025	0.322024844
8	Jekson	85.55	83.17	6.03518	0.67897	1.749286	0.678969808
9	Tri	84.13	81.12	7.664366	1.826499	1.152779	1.15277925
...	...	...	...	...	...	...	...
23	Setia	85.03	82.86	6.550275	0.186815	1.143897	0.186815417

e) Menentukan Posisi *Cluster* atau Pengelompokan

Dalam menentukan posisi *cluster* masing-masing data hasil SKP berdasarkan perilaku.

**Tabel 4.** Hasil Cluster

No	Nama	Kelompok
1	Evi	2
2	Berliana	2
3	Nurul	3
4	Ratna	2
5	Isum	2
6	Sudirson	2
7	Edy	3
8	Jekson	2
9	Tri	3
...	...	...
23	Setia	2



Dapat dilihat pada tabel di atas bahwa pengelompokan antar *cluster* mendapat hasil C1 dengan 1 data, C2 dengan 17 data dan C3 sebanyak 5 data. Proses *K-Means* akan terus beriterasi sampai pengelompokan data sama dengan pengelompokan data iterasi sebelumnya. Dengan kata lain proses akan terus melakukan iterasi sampai data pada iterasi terakhir sama dengan iterasi sebelumnya.

f) Menghitung *centroid* baru menggunakan hasil dari setiap anggota pada masing-masing *cluster* setelah didapatkan hasil jarak dari setiap objek pada iterasi ke-1 maka lanjut ke iterasi ke-2 pada perhitungan dan tabel dibawah ini.

C1 :

$$D_{SKP,C1} \frac{91,58}{1} = 91,58$$

$$D_{Perilaku,C1} \frac{82,92}{1} = 82,92$$

C2 :

$$D_{SKP,C2} \frac{86,22+85,87+86,19+85,33+84,80+85,55+84,49+86,19+85,52+85,86+85,08+85,03+85,03+85,13+85,55+84,67+85,03}{17} = 85,38471$$

$$D_{Perilaku,C2} \frac{84,00+83,65+82,91+83,2+82,24+83,17+82,82+84,03+86,25+83,59+82,68+82,86+82,86+82,28+83,17+82,22+82,86}{17} = 83,22294$$

C3 :

$$D_{SKP,C3} \frac{84,52+84,19+84,13+84,73+84,05}{5} = 84,324$$

$$D_{Perilaku,C3} \frac{81,35+81,98+81,12+81,96+82,27}{5} = 81,736$$

**Tabel 5. Centroid Baru Iterasi 1**

	SKP	PERILAKU
C1	91,58	82,92
C2	85,38471	83,22294

g) Selanjutnya dilakukan kembali langkah ke 4 sampai 6. Jika *centroid* hasil iterasi dengan *centroid* sebelumnya bernilai sama seperti posisi *cluster* data hasil tidak mengalami perubahan maka proses iterasi berhenti. Namun jika nilai *centroid* tidak sama serta posisi data masih berubah maka proses iterasi berlanjut pada iterasi berikutnya. Berikut tabel hasil *cluster* iterasi ke 2 :

**Tabel 6. Centroid Baru**

Cluster 1	91.58	82.92
Cluster 2	85,38471	83,22294
Cluster 3	84,324	81,736

Langkah selanjutnya adalah melakukan perhitungan jarak terhadap data SKP perilaku dengan menggunakan data *centroid* baru.

$$D_{Evi,C1} = \sqrt{(86,22 - 91,58)^2 + (84 - 82,92)^2} = 5,467723$$

$$D_{Evi,C2} = \sqrt{(86,22 - 85,38471)^2 + (84 - 83,22294)^2} = 1,140849$$

$$D_{Evi,C3} = \sqrt{(86,22 - 84,324)^2 + (84 - 81,736)^2} = 2,953051$$

Hasil dari keseluruhan perhitungan dapat dilihat pada tabel 7 sebagai berikut :

**Tabel 7. Hasil Perhitungan Jarak Pusat Cluster Iterasi 2**

No	Nama	SKP	Perilaku	C1	C2	C3	Jarak Terpendek
1	Evi	86.22	84.00	5.467723	1.140849	2.953051	1.140849104
2	Berliana	85.87	83.65	5.756475	0.646444	2.460389	0.646443825
3	Nurul	84.52	81.35	7.232462	2.062917	0.432911	0.432911076
4	Ratna	86.19	82.91	5.390009	0.863962	2.204593	0.863962265
5	Isum	85.33	83.2	6.256269	0.059321	1.776325	0.059321422
6	Sudirson	84.80	82.24	6.814015	1.143702	0.693247	0.693247431
7	Edy	84.19	81.98	7.449544	1.724014	0.278374	0.278373849
8	Jekson	85.55	83.17	6.03518	0.173565	1.886646	0.1735653
9	Tri	84.13	81.12	7.664366	2.448806	0.645827	0.645826602
...	...	...	...	...	...	...	...
23	Setia	85.03	82.86	6.550275	0.507487	1.327333	0.507486513

**Tabel 8. Hasil Cluster Iterasi 2**

No	Nama	Kelompok
1	Evi	2
2	Berliana	2
3	Nurul	3
4	Ratna	2
5	Isum	2
6	Sudirson	3
7	Edy	3
8	Jekson	2
9	Tri	3
...	...	...
23	Setia	2

Dapat dilihat pada hasil perhitungan manual pada data di atas mendapatkan hasil akhir yang berbeda pada iterasi 1 dan iterasi 2. Dimana hasil untuk iterasi 2 bernilai C1 = 1, C2 = 12, dan C3 = 8. Untuk itu proses *K-Means* akan dilanjutkan ke iterasi berikutnya. Berikut perhitungan dan tabel iterasi ke-3 dibawah ini :

Menghitung *Centroid* Baru :

C1 :

$$D_{SKP,C1} = \frac{91,58}{1} = 91,58$$

$$D_{Perilaku,C1} = \frac{82,92}{1} = 82,92$$

C2 :

$$D_{SKP,C2} = \frac{86,22+85,87+86,19+85,33+85,55+84,49+86,19+85,52+85,86+85,08+85,03+85,03+85,55+85,03}{14} = 85,49571$$

$$D_{Perilaku,C2} = \frac{84,00+83,65+82,91+83,2+83,17+82,82+84,03+86,25+83,59+82,68+82,86+82,86+83,17+82,86}{14} = 83,43214$$

C3 :

$$D_{SKP,C3} = \frac{84,52+84,80+84,19+84,13+84,73+84,05+85,13+84,67}{8} = 84,5275MVH$$

$$D_{Perilaku,C3} = \frac{81,35+82,24+81,98+81,12+81,96+82,27+82,28+82,22}{8} = 81,9275$$

**Tabel 9.** Hasil Perhitungan *Centroid* Baru Pada Iterasi 3

	SKP	PERILAKU
C1	91,58	82,92
C2	85,49571	83,43214

Langkah selanjutnya adalah melakukan perhitungan jarak terhadap data SKP perilaku dengan menggunakan data *centroid* baru.

$$D_{Evi,C1} = \sqrt{(86,22 - 91,58)^2 + (84 - 82,92)^2} = 5,467723$$

$$D_{Evi,C2} = \sqrt{(86,22 - 85,49571)^2 + (84 - 83,43214)^2} = 0,920354$$

$$D_{Evi,C3} = \sqrt{(86,22 - 84,5275)^2 + (84 - 81,9275)^2} = 2,675783$$

Hasil dari keseluruhan perhitungan dapat dilihat pada Tabel 5.0 sebagai berikut :

**Tabel 10.** Hasil Perhitungan Jarak Pusat Cluster Iterasi 3

No	Nama	SKP	Perilaku	C1	C2	C3	Jarak Terpendek
1	EVI	86.22	84.00	5.467723	0.920354	2.675783	0.920354
2	BERLIANA	85.87	83.65	5.756475	0.433072	2.183876	0.433072
3	NURUL	84.52	81.35	7.232462	2.299421	0.577549	0.577549
4	RATNA	86.19	82.91	5.390009	0.868715	1.931117	0.868715
5	ISUM	85.33	83.2	6.256269	0.285222	1.504414	0.285222
6	SUDIRSON	84.80	82.24	6.814015	1.380298	0.414623	0.414623
7	EDY	84.19	81.98	7.449544	1.952846	0.341559	0.341559
8	JEKSON	85.55	83.17	6.03518	0.267705	1.609134	0.267705
9	TRI	84.13	81.12	7.664366	2.685364	0.900035	0.900035
...	...	....	...	...	...	...	...
23	SETIA	85.03	82.86	6.550275	0.737724	1.059275	0.737724

**Tabel 11.** Hasil *Cluster*

No	Nama	Kelompok
1	Evi	2
2	Berliana	2
3	Nurul	3
4	Ratna	2
5	Isum	2
6	Sudirson	3
7	Edy	3
8	Jekson	2
9	Tri	3
...	...	...
23	Setia	2



Proses *K-Means* masih terus berlanjut ke iterasi berikutnya dikarenakan masih ada perubahan dan hasilnya belum stabil. Berikut perhitungan dan tabel iterasi ke-4 dibawah ini:

Menghitung Centroid Baru :

C1 :

$$D_{SKP,C1} \frac{91,58}{1} = 91,58$$

$$D_{Perilaku,C1} \frac{82,92}{1} = 82,92$$

C2 :

$$D_{SKP,C2} \frac{86,22+85,87+86,19+85,33+85,55+86,19+85,52+85,86+85,08+85,03+85,03+85,55+85,03}{13} = 85,57308$$

$$D_{Perilaku,C2} \frac{84,00+83,65+82,91+83,2+83,17+84,03+86,25+83,59+82,68+82,86+82,86+83,17+82,86}{13} = 83,47923$$

C3 :

$$D_{SKP,C3} \frac{84,52+84,80+84,19+84,13+84,49+84,73+84,05+85,13+84,67}{9} = 84,52333$$

$$D_{Perilaku,C3} \frac{81,35+82,24+81,98+81,12+82,82+81,96+82,27+82,28+82,22}{9} = 82,02667$$

**Tabel 12.** Hasil Perhitungan *Centroid* Baru Pada Iterasi 4

	SKP	PERILAKU
C1	91,58	82,92
C2	85,57308	83,47923
C3	84,52333	82,02677

Langkah selanjutnya adalah melakukan perhitungan jarak terhadap data SKP perilaku dengan menggunakan data *centroid* baru.

$$D_{Evi,C1} = \sqrt{(86,22 - 91,58)^2 + (84 - 82,92)^2} = 5,467723$$

$$D_{Evi,C2} = \sqrt{(86,22 - 85,57308)^2 + (84 - 83,47923)^2} = 0,830488$$

$$D_{Evi,C3} = \sqrt{(86,22 - 84,52333)^2 + (84 - 82,02667)^2} = 2,602445$$

Hasil dari keseluruhan perhitungan dapat dilihat pada tabel berikut :

**Tabel 13.** Hasil Perhitungan Jarak Pusat *Cluster* Iterasi 4

No	Nama	SKP	Perilaku	C1	C2	C3	Jarak Terpendek
1	EVI	86.22	84.00	5.467723	0.830488	2.602445	0.830487844
2	BERLIANA	85.87	83.65	5.756475	0.342528	2.109199	0.342528019
3	NURUL	84.52	81.35	7.232462	2.375415	0.676675	0.676674877
4	RATNA	86.19	82.91	5.390009	0.839415	1.886281	0.839415125
5	ISUM	85.33	83.2	6.256269	0.370211	1.423876	0.370211038
6	SUDIRSON	84.80	82.24	6.814015	1.460596	0.349365	0.349364502
7	EDY	84.19	81.98	7.449544	2.039754	0.336584	0.336584148
8	JEKSON	85.55	83.17	6.03518	0.310091	1.536638	0.310090653
9	TRI	84.13	81.12	7.664366	2.765581	0.988309	0.988309443
...	...	...	...	...	...	...	...
23	SETIA	85.03	82.86	6.550275	0.823638	0.975272	0.823637839



**Tabel 14.** Hasil *Cluster*

No	Nama	Kelompok
1	Evi	2
2	Berliana	2
3	Nurul	3
4	Ratna	2
5	Isum	2
6	Sudirson	3
7	Edy	3
8	Jekson	2
9	Tri	3
...	...	...
23	Setia	2

Proses *K-Means* masih terus berlanjut ke iterasi berikutnya dikarenakan masih ada perubahan dan hasilnya belum stabil. Berikut perhitungan dan tabel iterasi ke-5 dibawah ini :

Menghitung *Centroid* Baru :

C1 :

$$D_{SKP,C1} \frac{91,58}{1} = 91,58$$

$$D_{Perilaku,C1} \frac{82,92}{1} = 82,92$$

C2 :

$$D_{SKP,C2} \frac{86,22+85,87+86,19+85,33+85,55+86,19+85,52+85,86+85,03+85,03+85,55+85,03}{12} = 85,61417$$

$$D_{Perilaku,C2} \frac{84,00+83,65+82,91+83,2+83,17+84,03+86,25+83,59+82,86+82,86+83,17+82,86}{12} = 83,54583$$

C3 :

$$D_{SKP,C3} \frac{84,52+84,80+84,19+84,13+84,49+84,73+85,08+84,05+85,13+84,67}{10} = 84,5793$$

$$D_{Perilaku,C3} \frac{81,35+82,24+81,98+81,12+82,82+81,96+82,68+82,27+82,28+82,22}{10} = 82,092$$

**Tabel 15.** Hasil Perhitungan *Centroid* Baru Pada Iterasi 5

	SKP	PERILAKU
C1	91,58	82,92
C2	85,61417	83,54583
C3	84,579	82,092

Langkah selanjutnya adalah melakukan perhitungan jarak terhadap data SKP dan Perilaku dengan menggunakan data *centroid* baru.

$$D_{Evi,C1} = \sqrt{(86,22 - 91,58)^2 + (84 - 82,92)^2} = 5,467723$$

$$D_{Evi,C2} = \sqrt{(86,22 - 85,61417)^2 + (84 - 83,54583)^2} = 0,757167$$

$$D_{Evi,C3} = \sqrt{(86,22 - 84,579)^2 + (84 - 82,092)^2} = 2,516614$$

Hasil dari keseluruhan perhitungan dapat dilihat pada tabel 16 sebagai berikut :

**Tabel 16.** Hasil Perhitungan Jarak Pusat *Cluster* Iterasi 5

N o	Nama	SKP	Perilaku	C1	C2	C3	Jarak Terpendek
1	Evi	86.22	84.00	5.467723	0.757167	2.516614	0.757167
2	Berliana	85.87	83.65	5.756475	0.276227	2.023375	0.276227
3	Nurul	84.52	81.35	7.232462	2.453342	0.744342	0.744342
4	Ratna	86.19	82.91	5.390009	0.857828	1.806778	0.857828
5	Isum	85.33	83.2	6.256269	0.447606	1.338531	0.447606
6	Sudirson	84.80	82.24	6.814015	1.538853	0.265979	0.265979
7	Edy	84.19	81.98	7.449544	2.116621	0.404802	0.404802
8	Jekson	85.55	83.17	6.03518	0.381272	1.450836	0.381272
9	Tri	84.13	81.12	7.664366	2.843839	1.070694	1.070694
...	...	...	...	...	...	...	...
23	Setia	85.03	82.86	6.550275	0.900898	0.890632	0.890632

**Tabel 17.** Hasil *Cluster*

No	Nama	Kelompok
1	Evi	2
2	Berliana	2
3	Nurul	3
4	Ratna	2
5	Isum	2
6	Sudirson	3
7	Edy	3
8	Jekson	2
9	Tri	3
...	...	...
23	Setia	3

Proses *K-Means* masih terus berlanjut ke iterasi berikutnya dikarenakan masih ada perubahan dan hasilnya belum stabil. Berikut perhitungan dan tabel iterasi ke-6 dibawah ini :

Menghitung *Centroid* Baru :

C1 :

$$D_{SKP,C1} \frac{91,58}{1} = 91,58$$

$$D_{Perilaku,C1} \frac{82,92}{1} = 82,92$$

C2 :

$$D_{SKP,C2} \frac{86,22+85,87+86,19+85,33+85,55+86,19+85,52+85,86+85,55}{9} = 85,80889$$

$$D_{Perilaku,C2} \frac{84,00+83,65+82,91+83,2+83,17+84,03+86,25+83,59+83,17}{9} = 83,77444$$

C3 :



$$D_{SKP,C3} = \frac{84,52+84,80+84,19+84,13+84,49+84,73+85,08+85,03+84,05+85,03+85,13+84,67+85,03}{13} = 84,68308$$

$$D_{Perilaku,C3} = \frac{81,35+82,24+81,98+81,12+82,82+81,96+82,68+82,86+82,27+82,86+82,28+82,22+82,86}{13} = 82,26923$$

**Tabel 18.** Hasil Perhitungan *Centroid* Baru Pada Iterasi 6

	SKP	PERILAKU
C1	91,58	82,92
C2	85,80889	83,77444
C3	84,68308	82,26923

Langkah selanjutnya adalah melakukan perhitungan jarak terhadap data SKP perilaku dengan menggunakan data baru.

$$D_{Evi,C1} = \sqrt{(86,22 - 91,58)^2 + (84 - 82,92)^2} = 5,467723$$

$$D_{Evi,C2} = \sqrt{(86,22 - 85,80889)^2 + (84 - 83,77444)^2} = 0,468922$$

$$D_{Evi,C3} = \sqrt{(86,22 - 84,68308)^2 + (84 - 82,26923)^2} = 2,314669$$

Hasil dari keseluruhan perhitungan dapat dilihat pada tabel berikut :

**Tabel 19.** Hasil Perhitungan Jarak Pusat *Cluster* Iterasi 6

No	Nama	SKP	Perilaku	C1	C2	C3	Jarak Terpendek
1	EVI	86.22	84.00	5.467723	0.468922	2.314669	0.468922
2	BERLIANA	85.87	83.65	5.756475	0.13864	1.820799	0.13864
3	NURUL	84.52	81.35	7.232462	2.745754	0.933584	0.933584
4	RATNA	86.19	82.91	5.390009	0.944727	1.637499	0.944727
5	ISUM	85.33	83.2	6.256269	0.747878	1.133508	0.747878
6	SUDIRSON	84.80	82.24	6.814015	1.836403	0.120522	0.120522
7	EDY	84.19	81.98	7.449544	2.416781	0.571646	0.571646
8	JEKSON	85.55	83.17	6.03518	0.657553	1.250176	0.657553
9	TRI	84.13	81.12	7.664366	3.140819	1.275392	1.275392
...	...	...	...	...	...	...	...
23	SETIA	85.03	82.86	6.550275	1.201198	0.685101	0.685101

**Tabel 20.** Hasil *Cluster*

No	Nama	Kelompok
1	Evi	2
2	Berliana	2
3	Nurul	3
4	Ratna	2
5	Isum	2
6	Sudirson	3
7	Edy	3
8	Jekson	2
9	Tri	3
10	Sutiani	1

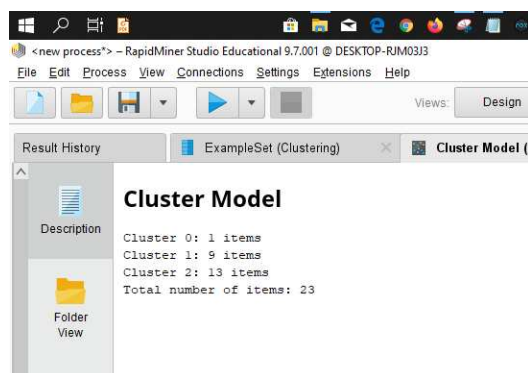


11	Susi	3
12	Suyani	2
13	Wanden	2
14	Wulandani	3
15	Rumiris	2
16	Lisma	3
17	Sumiah	3
18	Oloan	3
19	Nazmah	3
20	Hajjah	3
21	Haposan	2
22	Binahar	3
23	Setia	3

Pada perhitungan ini iterasi berhenti pada iterasi ke-6 karena pada iterasi 5 dan iterasi 6 pengelompokan data yang dilakukan terhadap 3 *cluster* didapatkan hasil yang sama, dan telah mencapai stabil dan *konvergen*. Berdasarkan perhitungan manual pada data SKP (Sasaran Kerja Pegawai) yang telah dilakukan diatas mendapatkan hasil yaitu C1= 1, C2= 9, dan C3= 13 pada posisi data tiap *cluster*.

### 3.2. Hasil Percobaan

Pada pengujian ini kita menggunakan *software RapidMiner Studio Versi 9.7*, dengan menggunakan *software* tersebut kita akan membandingkan hasil dari pengolahan data secara manual dengan hasil pada pengolahan data menggunakan *software*. Untuk *Cluster Model* dapat melihat dalam *Description* yang merupakan tampilan hasil pengelompokan *cluster* dan jumlah anggota dengan *Format Text* seperti terlihat pada gambar 1.



**Gambar 1.** Tampilan *Cluster Model Description*

Keterangan :

1. Jumlah *cluster* 0 (tinggi) berjumlah 1 item.
2. Jumlah *cluster* 1 (sedang) berjumlah 9 item.
3. Jumlah *cluster* 2 (rendah) berjumlah 13 item.
4. Jumlah keseluruhan yaitu 23 data.



#### 4. SIMPULAN

Berdasarkan pembahasan dan hasil *implementasi software Rapidminer* pada pengelompokan menggunakan Data Mining metode *Algoritma K-Means*, maka penulis dapat menyimpulkan sebagai berikut:

- a) Hasil analisis menunjukkan bahwa penerapan Data Mining dalam analisis penilaian kinerja ASN untuk menetapkan besaran TPP di Sekretariat DPRD Pematangsiantar dapat membantu dalam memaksimalkan kinerja pegawai.
- b) Berdasarkan hasil yang diperoleh dari metode *K-Means clustering* yang di *implementasikan* kedalam *Rapidminer* memiliki nilai yang sama yaitu menghasilkan 3 *cluster* yaitu cluster tinggi sebanyak 1 yaitu Sutiani. Sedangkan *cluster* sedang sebanyak 9 yaitu Evi, Berliana, Ratna, Isum, Jekson, Suyani, Wanden, Rumiris, dan Haposan . Dan *cluster* rendah sebanyak 13 yaitu Nurul, Sudirson, Edy, Tri, Susi, Wulandani, Lisma, Sumiah, Oloan, Nazmah, Hajijah, Binahar, Setia. Hasil yang didapat dari penelitian dapat menjadi masukan bagi instansi terkait, pegawai yang menjadi perhatian lebih pada kantor DPRD berdasarkan *cluster* yang telah dilakukan

#### DAFTAR PUSTAKA

- [1] D. Sartika and J. Jumadi, "Seminar Nasional Teknologi Komputer & Sains (SAINTEKS) Clustering Penilaian Kinerja Dosen Menggunakan Algoritma K-Means (Studi Kasus: Universitas Dehasen Bengkulu)," *Semin. Nas. Teknol. Komput. Sains*, pp. 703–709, 2019.
- [2] C. M. Fikri, F. E. M. Agustin, and F. Mintarsih, "Pengelompokan Kualitas Kerja Pegawai Menggunakan Algoritma K-Means++ Dan Cop-Kmeans Untuk Merencanakan Program Pemeliharaan Kesehatan Pegawai Di Pt. Pln P2B Jb Depok," *Pseudocode*, vol. 4, no. 1, pp. 9–17, 2017, doi: 10.33369/pseudocode.4.1.9-17.
- [3] S. R. Ningsih, I. S. Damanik, A. P. Windarto, H. S. Tambunan, J. Jalaluddin, and A. Wanto, "Analisis K-Medoids Dalam Pengelompokkan Penduduk Buta Huruf Menurut Provinsi," *Pros. Semin. Nas. Ris. Inf. Sci.*, vol. 1, no. September, p. 721, 2019, doi: 10.30645/senaris.v1i0.78.
- [4] A. T. R. Saragih, A. S. Sembiring, and M. Sayuthi, "Penerapan Metode Clustering K-Means untuk Proses Seleksi Calon Peserta Lomba MTQ," *Pelita Inform.*, vol. 17, no. April, pp. 117–122, 2018.