



ANALISA PERBANDINGAN ALGORITMA K-MEANS, K-MEDOIDS, DAN X-MEANS UNTUK PENGELOMPOKKAN KINERJA PEGAWAI

(Studi Kasus: Sekretariat DPRD Provinsi Sulawesi Utara)

Gideon Bartolomeus Kaligis¹, Sri Yulianto²

^{1,2} Program Studi Teknik Informatika, FTI UKSW, Salatiga
Jl. Diponegoro 52-60, Salatiga 50711, Indonesia

Email : 672018312@student.uksw.edu¹, sri.yulianto@uksw.edu²

Submitted: 06-10-2022

Riwayat artikel:
Revised: 20-10-2022

Published: 29-10-2022

Abstrak – Kinerja pegawai menjadi rangkuman dalam hal kualitas, kuantitas, jam kerja dan juga kerja sama untuk mencapai suatu tujuan yang telah ditetapkan oleh instansi atau perusahaan, namun dalam Sekretariat DPRD Provinsi Sulawesi Utara belum adanya metode untuk menentukan pengelompokan *kinerja pegawai*. Untuk mengatasi permasalahan ini, diperlukan adanya pengelompokan *kinerja pegawai* di Sekretariat DPRD Provinsi Sulawesi Utara, sehingga bisa menentukan *kinerja pegawai* yang berkualitas. Tujuan dari penelitian ini dengan melakukan perbandingan metode-metode *clustering* untuk mendapatkan metode yang lebih baik dalam pengelompokan cluster terhadap *kinerja pegawai* di Sekretariat DPRD Provinsi Sulawesi Utara. Metode pengelompokan data *kinerja pegawai* yang dibuat menggunakan metode *clustering k-means, k-medoids, x-means* dengan menggunakan lima atribut, yaitu: orientasi pelayanan, integritas, komitmen, disiplin, dan kerjasama, kemudian diolah dengan bantuan rapidminer, sehingga membagi data menjadi dua cluster yang dikategorikan sebagai nilai tinggi (C_1) dan rendah (C_0). Pada tahap berikutnya mencari nilai *davies bouldin index* memakai bantuan rapidminer pada setiap metode yang dipakai untuk melakukan perbandingan serta menentukan metode yang lebih optimal dalam *clustering*. Hasil nilai yang diperoleh dari metode *davies bouldin index* di setiap algoritma, yaitu: *k-means* sebesar -0.377, *k-medoids* sebesar -0.930, dan *x-means* sebesar -0.497, maka algoritma terbaik untuk pengelompokan data *kinerja pegawai* dalam penelitian ini adalah algoritma *k-means*, karena memiliki nilai DBI yang terkecil.

Kata Kunci: Kinerja Pegawai, K-Means, K-Medoids, X-Means, Clustering, Davies Bouldin Index.

Abstract – Employee performance is summarized in terms of quality, quantity, working hours and also cooperation to achieve a goal that has been set by the agency or company, but in the DPRD secretariat of north sulawesi province there is no method to determine the grouping of employee performance. To overcome this problem, it is necessary to group employee performance in the DPRD secretariat of north sulawesi province, so that it can determine the performance of quality employees. The purpose of this study is to compare clustering methods to get a better method of grouping clusters on employee performance at the DPRD secretariat of north sulawesi province. The method of grouping employee performance data made using the *k-means, k-medoids, x-means* clustering method using

five attributes, namely: service orientation, integrity, commitment, discipline, and cooperation, then processed with the help of rapidminer, thus dividing the data. Into two clusters that are categorized as high (C_1) and low values (C_0). In the next stage, look for the davies bouldin index value using the help of rapidminer on each method used to make comparisons and determine the more optimal method in clustering. The results obtained from the davies bouldin index method in each algorithm, namely: *k-means* of -0.377, *k-medoids* of -0.930, and *x-means* of -0.497, then the best algorithm for grouping employee performance data in this study is the algorithm *k-means*, because it has the smallest DBI value.

Keywords: Employee performance, K-Means, K-Medoids, X-Means, Clustering, and Davies Bouldin Index.

I. PENDAHULUAN

Kinerja pegawai merupakan pekerjaan individu di dalam lembaga yang cocok pada otoritas serta tanggung jawab dari setiap pegawai yang berusaha keras untuk mencapai tujuan utama instansi atau institusi sehingga tidak melawan aturan, etis, dan moral [1]. Kebutuhan mengenai informasi dan komunikasi adalah bagian yang penting untuk manusia dalam hal pekerjaan kantoran, apalagi dalam menangani data yang yang begitu besar, oleh karena itu dibutuhkan data mining sebagai teknik komputasi pengolahan data [2]. Dalam data mining bisa diterapkan ke dalam pekerjaan kita untuk menggali kumpulan data berupa pengetahuan yang tidak diketahui secara manual, dalam data mining terdapat metode *clustering*, diketahui *clustering* itu sendiri merupakan salah satu sub beberapa kategori data mining, yang fungsinya akan menentukan proses di mana sampel yang sama dibagi menjadi kelompok pada setiap cluster [3].

Sekretariat DPRD merupakan pelayanan administrasi dan pemberian bantuan dalam tugas DPRD. Dalam praktek atau agenda, Sekretariat DPRD dibutuhkan penataan kerja untuk meningkatkan kapasitas sumber daya manusia dalam pelayanan DPRD, agar bisa membantu untuk memaksimalkan kerja pelayanan yang ada di Sekretariat DPRD Provinsi Sulawesi Utara [4]. Di setiap tahun dilakukan penilaian *kinerja pegawai* untuk mengambil suatu keputusan dan bahan pertimbangan dalam mengetahui kinerja yang diperoleh pegawai dalam pelayanan yang sudah ditetapkan setiap tahunnya melalui form. Terdapat beberapa standar penilaian kinerja pada pegawai dalam pelayanan pada Sekretariat DPRD Provinsi Sulawesi Utara, antara lain berupa orientasi pelayanan, integritas, komitmen, disiplin, dan kerjasama. Pendataan kriteria pada *kinerja pegawai* akan dilakukan proses *clustering*, sehingga data akan dikelompokkan pada setiap cluster yang ditentukan, serta algoritma yang lebih baik akan terlihat dengan menentukan perbandingan nilai pada *davies bouldin index*.

Dengan latar belakang di atas, maka dilakukan kajian terhadap data penilaian *kinerja pegawai* dengan menggunakan bantuan rapidminer, untuk menemukan *clustering* yang diperoleh di setiap algoritma yang digunakan dalam penelitian ini, dimana hasil akhir akan dianalisa dengan melakukan perbandingan *davies bouldin index* pada tiga algoritma diantaranya *k-means*, *k-medoids*, dan *x-means*. Tujuan

dari penelitian ini untuk melakukan perbandingan metode-metode dalam bentuk mendapatkan metode algoritma yang lebih optimal dalam pengelompokan data *kinerja pegawai*.

II. KAJIAN PUSTAKA

A. Peneliti Terdahulu

Penelitian yang membahas *Pengelompokan Data Persediaan Obat Menggunakan Perbandingan Metode K-Means Dengan Hierarchical Clustering Single Linkage*, dengan permasalahan mengantisipasi persediaan obat di Indonesia agar tersalurkan secara merata dan berkelanjutan. Penelitian ini membahas tentang algoritma *k-means* dan algoritma *hierarchical clustering* untuk membandingkan nilai validitasnya yaitu *silhouette index*. Hasil menunjukkan algoritma *k-means* dengan hasil validitas *silhouette index* sebesar 0.8414, sedangkan hasil validitas *silhouette index* pada algoritma *hierarchical clustering* sebesar 0.8629 [5].

Penelitian yang membahas *Analisa Perbandingan Algoritma Clustering untuk Pemetaan Status Gizi Balita di Puskesmas Pasir Jaya*, dengan permasalahan status gizi pada balita harus diperhatikan seperti gizi buruk. Dalam penelitian melakukan perbandingan terhadap algoritma *k-means* dan algoritma *fuzzy c-means* serta memakai 4 parameter yaitu, berat badan, umur, jenis kelamin, tinggi badan dan menggunakan *silhouette index*. Hasil menunjukkan algoritma *k-means* menghasilkan validasi sebesar 0.79, sedangkan algoritma *fuzzy c-means* menghasilkan sebesar 0.78 [6].

Penelitian yang membahas *Analisa Tingkat Kepuasan Mahasiswa Terhadap Layanan Pembelajaran Menggunakan K-Means dan Algoritma Genetika*, dengan permasalahan membandingkan algoritma *k-means* dengan optimasi *k-means* serta algoritma *genetika* untuk melihat tingkat kepuasan mahasiswa dalam melaksanakan proses belajar. Penelitian ini menggunakan 127 sampel yang diperoleh dari pengisian angket dari kepuasan mahasiswa terhadap pembelajaran yang diikuti, dan memakai metode *davies bouldin index* untuk menentukan algoritma mana yang terbaik. Hasil menunjukkan algoritma *k-means* memperoleh nilai DBI 1.990, sedangkan optimasi *k-means* dengan algoritma *genetika* memperoleh nilai DBI 1.593 [7].

B. K-Means

K-Means merupakan suatu teknik penganalisa data atau dikenal dengan teknik data mining untuk menerapkan proses pemodelan data tanpa supervisi (unsupervised) serta menjadi salah satu teknik untuk melakukan pengelompokan di setiap data-data dengan bentuk partisi [6].

Adapun berupa langkah langkah metode K-Means yang akan digunakan untuk jumlah cluster dan penempatan data dalam cluster, antara lain : [8]

- 1) Tahap awal yang dilakukan dengan memilih secara acak k data sebagai pusat cluster.

- 2) Jarak yang akan digunakan untuk menentukan data dengan pusat cluster memakai Euclidean Distance. Dalam perhitungan untuk menentukan jarak data ke masing-masing titik pusat cluster bisa memakai euclidean distance, sebagai berikut; [1]

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

Dimana: d_{ij} merupakan antara objek dengan objek i dan objek j , p merupakan ukuran data, x_{ik} merupakan suatu titik dari objek i pada ukuran k , x_{jk} merupakan suatu titik dari objek j dalam ukuran k .

- 3) Data yang masuk kedalam cluster terdekat akan melakukan perhitungan dari tengah cluster, sehingga data yang menghasilkan cluster baru akan ditentukan bila semua data telah ditetapkan kedalam cluster terdekat.
- 4) Pada tahap akhir dalam proses penentuan pusat cluster untuk menentukan data-data yang masuk dalam cluster, akan di ulangi sampai jumlah centroid tidak bisa berubah lagi.

C. K-Medoids

Algoritma *k-medoids* merupakan teknik partisi *clustering* untuk membuat pengelompokkan data sejumlah n objek yang akan menghasilkan sebuah cluster. Metode ini memakai sebuah bahan di sekumpulan objek untuk mewakili sebuah cluster dari data untuk digunakan [7]. Pada algoritma ini dalam teknik partisi yang dapat diterapkan berdasarkan prinsip yang meminimalisir banyaknya dari ketidaksamaan pada masing-masing objek serta titik acuan yang sesuai di dalam metode ini.

Adapun berupa langkah-langkah yang dilakukan untuk menentukan metode ini, sebagai berikut : [7]

- 1) Siapkan sampel data untuk menyiapkan pusat cluster dalam jumlah cluster yang diperlukan.
- 2) Menentukan nilai secara acak/ nilai centroid.
- 3) Pada tahap selanjutnya menghitung jarak euclidean distance, dengan persamaan: [7]

$$d(x, y) = \sqrt{\sum_{i=1}^n \{X_i - Y_j\}^2} \quad (2)$$

Dimana: $d(x, y)$ merupakan antara objek dengan objek i dan objek j , x_i merupakan nilai atribut pertama dari objek ke i , y_j merupakan nilai atribut pertama dari objek ke j , n merupakan jumlah atribut yang dipakai. Memilih secara acak objek data setiap masing-masing cluster sebagai kandidat medoid baru.

- 4) Memilih secara acak objek data pada setiap cluster yang akan memperoleh medoid baru, sehingga menghasilkan total cost yang dihasilkan dalam

perhitungan jarak yang ditentukan dengan mengambil nilai paling sedikit pada setiap cluster.

- 5) Menghitung jumlah simpanan untuk menentukan nilai jarak baru dan dikurangi jumlah nilai jarak lama. Jika simpanan < 0 , maka akan membentuk sebuah ukuran objek dengan data cluster, sehingga membentuk masing-masing k atau objek baru pada medoid, dan mengulangi langkah 2 – 4 hingga tidak akan terjadi perubahan medoid yang dihasilkan, agar mendapatkan cluster beserta anggota pada setiap cluster.

D. X-Means

Algoritma *x-means* adalah teknik metode yang melengkapi kelemahan pada metode *k-means*, pada metode *k-means* terdapat kekurangan untuk melakukan perhitungan yang relatif lama didalam nilai cluster k yang harus ditentukan pengguna [2]. *X-means* sendiri termasuk dalam pembelajaran tak terarah (unsupervised learning), dimana proses data yang dilakukan akan dikelompokkan sendiri pada setiap data yang menjadi masukannya tanpa diketahui terlebih dulu sasaran kelas yang akan dijadikan cluster.

Adapun berupa langkah-langkah dalam menentukan persamaan algoritma *x-means*, sebagai berikut : [2]

- 1) Pada langkah pertama inialisasi jumlah k .
- 2) Menghitung jarak pada setiap data yang ada dalam pusat cluster dengan memakai persamaan euclidean distance, sebagai berikut: [2]

$$d_{ij} = \sqrt{\sum_{k=1}^n (X_{ik} - X_{jk})^2} \quad (3)$$

Dimana: d_{ij} adalah jarak objek dengan objek i dan j , x_{ik} adalah koordinat dari objek i pada dimensi k , x_{jk} adalah koordinat dari objek j pada dimensi k , n adalah jumlah atribut yang telah digunakan.

- 3) Langkah selanjutnya menentukan pada setiap k yang menghasilkan pusat cluster baru dengan menghitung nilai rata-rata karakter data yang masuk kedalam cluster.
- 4) Mengoptimal nilai dari suatu cluster, dengan menghitung setiap masing-masing cluster yang diperoleh.
- 5) Pada tahap terakhir kembali pada langkah kedua sampai tidak bisa bertukar dengan cluster lain atau mencapai batas maximum iterasi yang di dapat.

E. Data Mining

Data mining merupakan proses pengumpulan data atau informasi yang penting dari data utamanya. Proses pengumpulan data mining terkadang didapatkan dengan memakai metode statistika, matematika, dan juga menggunakan teknologi artificial intelligence [3]. Data mining adalah knowledge discovery in database (KDD) atau pattern recognition, yang dikenal dengan Knowledge discovery (mining) dalam basis data, pengalihan pengetahuan, data, analisis pola, arkeologi

data, pengerukan data, pengumpulan informasi, intelijen bisnis, dan lain-lain [6]. Ada cara dan teknik lain yang dapat dipakai untuk pengumpulan data mining, namun cara tersebut membutuhkan langkah-langkah, agar data yang diinginkan bisa terkumpul. Adapun fungsi dari data mining, yaitu fungsi deskriptif dan fungsi prediktif, deskriptif berfungsi supaya bisa mencari tahu lebih detail data yang sedang diamati, sedangkan prediktif berfungsi untuk mencari model dari data yang sedang diamati.

F. Kinerja

Kinerja merupakan arti dari Kinetika energi Kerja dalam bahasa Inggris memiliki arti yaitu “performance”, istilah ini biasanya menyatakan arti “*kinerja* pekerjaan” atau “*kinerja* sebenarnya” memiliki arti prestasi kerja dan prestasi sebenarnya yang dikerjakan oleh seseorang dengan melakukan tanggung jawabnya. Contohnya pekerjaan di dalam organisasi yang membutuhkan perhatian dan usaha pengoptimalan *kinerja* dari sumber daya manusia atau disingkat dengan SDM [9].

G. Clustering

Clustering merupakan metode analisis data, dan seringkali menjadi salah satu teknik dari data mining, dengan maksud untuk membagi data menggunakan ciri yang serupa ke suatu ‘daerah’ sehingga data dan karakteristik daerah tidak sama dengan yang lain [2][6]. *Clustering* adalah salah satu bagian dari metode data mining dan bersifat tanpa arahan atau disebut dengan “unsupervised” dan juga suatu cara untuk mencari dan membagi data yang mempunyai kesamaan ciri dengan satu data dan data lain” [6].

H. Davies Bouldin Index

Davies-Bouldin Index (DBI) ialah teknik suatu ukuran yang akan digunakan untuk menentukan jumlah cluster yang terbaik setelah proses pengklasteran terselesaikan [8]. Pada metode DBI ini bermaksud untuk mengoptimalkan perbedaan antara klaster yang satu dengan klaster yang lain serta mencoba meminimalkan jarak antar objek pada suatu klaster, dan juga dalam metode DBI ini bisa menjadi parameter setiap algoritma *clustering* untuk membandingkan proses cluster algoritma mana yang lebih optimal [10]. Dalam metode ini semakin kecil nilai yang mendekati 0, semakin optimal *clustering* yang dihasilkan [8].

Untuk perhitungan dalam DBI ini, menghitung rata-rata pada setiap titik berupa jumlah nilai kekompakan yang dibagi dengan jarak antara kedua titik pusat cluster sebagai pemisah. Adapun rumus untuk menghitung *davies bouldin index*, sebagai berikut : [10]

$$DBI = \frac{1}{k} \sum_{i=1}^k R_i, \quad (4)$$

$$R_i = \max R_{ij}, \quad (5)$$

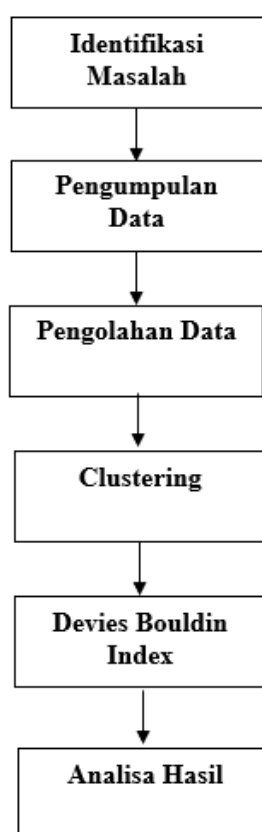
$$R_i = \frac{\text{var } C_i + \text{var } C_j}{C_i - C_j}, \quad (6)$$

Dimana :

- k = merupakan buah cluster
- R_{ij} = jarak antara titik pusat cluster i dan cluster j
- C_i = merupakan pusat dari cluster i
- C_j = merupakan pusat dari cluster j

III. METODE PENELITIAN

Adapun tahapan-tahapan dalam menyelesaikan penelitian ini yang disajikan dalam Gambar 1 di bawah ini.



Gambar 1 Tahapan penelitian

A. Identifikasi Masalah

Belum adanya pengelompokkan penilaian *kinerja pegawai* secara umum dan informasi secara lebih jelasnya mengenai pengelompokkan data *penilaian kinerja* pada Sekretariat DPRD Sulawesi Utara, serta juga penilaian *kinerja pegawai* masih teracak secara random, oleh karena itu sebagai tempat pelayanan perlu adanya penilaian agar mampu menjaga kinerja yang berkualitas. Untuk mengatasi kinerja pegawai yang masih teracak secara random, maka dilakukan penelitian ini untuk mengelompokkan *kinerja pegawai* sesuai kinerja individu setiap pegawai, nantinya akan menggunakan tiga algoritma, serta membandingkan algoritma yang lebih optimal dalam proses pengelompokkan data.

B. Pengumpulan Data

Teknik pengumpulan data dilakukan untuk mendapatkan data yang bisa membantu memecahkan permasalahan dalam penelitian ini. Teknik pengumpulan data yang dilakukan dalam penelitian ini, diantaranya:

- 1) Observasi dilakukan dengan melihat secara langsung sistem penilaian *kinerja pegawai* ketika ini di Sekretariat DPRD Sulawesi Utara.
- 2) Wawancara dilakukan untuk memberikan beberapa pertanyaan yang berkaitan dalam permasalahan yang ada, pada salah satu pegawai yang menjabat sebagai kasubag di Sekretariat DPRD Sulawesi Utara.

C. Pengolahan Data

Menggunakan SKP pada tahun 2017-2021 berupa data kuantitatif, dan mengambil data dalam SKP dengan atribut orientasi pelayanan, integritas, komitmen, disiplin, dan kerjasama. Kemudian diseleksi SKP yang lengkap pada tahun 2017-2021 dalam aplikasi spreadsheet.

D. Clustering

Setelah mendapatkan data lengkap dalam aplikasi spreadsheet, di import ke aplikasi rapidminer dengan menggunakan tiga metode algoritma *k-means*, *k-medoids*, dan *x-means* dan membuat dua kriteria *clustering* yaitu nilai tinggi dan nilai rendah, sehingga data otomatis akan masuk ke dalam cluster yang ditentukan.

E. Davies Bouldin Index

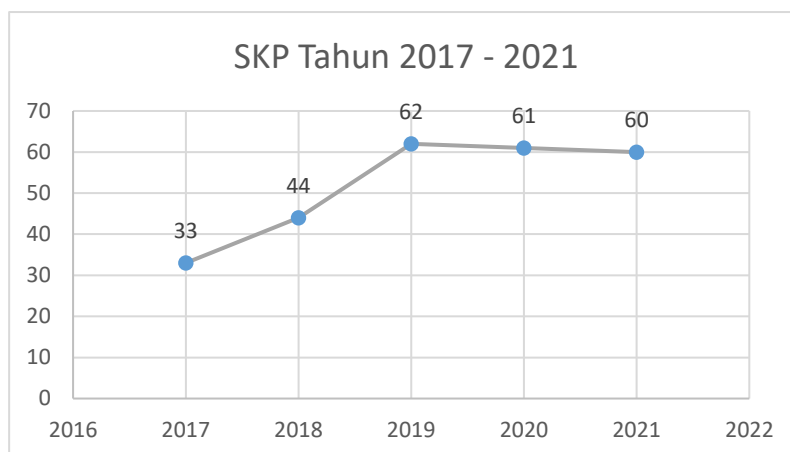
Pada tahap *davies bouldin index*, setiap algoritma akan diuji menggunakan metode ini dalam aplikasi rapidminer untuk menentukan proses *clustering* yang lebih optimal, semakin kecil nilai DBI yang dihasilkan atau nilai yang mendekati 0, maka proses *clustering* yang dihasilkan pada algoritma tersebut lebih optimal.

F. Analisa Hasil

Setelah mendapatkan hasil, akan dilakukan perbandingan nilai DBI yang dihasilkan, untuk menentukan metode *clustering* yang lebih optimal.

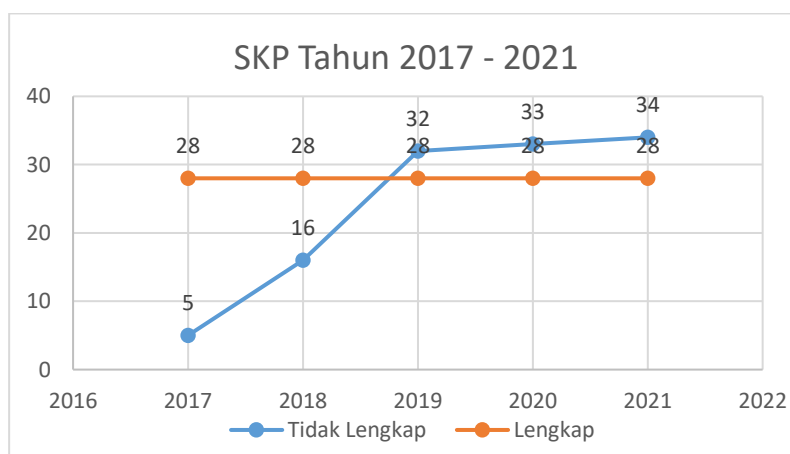
IV. ANALISA HASIL DAN PEMBAHASAN

Dalam melakukan penelitian ini data yang akan digunakan merupakan SKP(Sasaran Kerja Pegawai) pada tahun 2017-2021 yang diperoleh dari instansi Sekretariat DPRD Provinsi Sulawesi Utara. Setiap 1 SKP berjumlah 5 data/atribut berupa orientasi pelayanan, integritas, komitmen, disiplin, dan kerjasama.



Gambar 2. Grafik Line SKP Tahun 2017 - 2021

Pada Gambar 2 di atas adalah data SKP yang diambil dari Sekretariat DPRD Sulawesi Utara, dengan SKP 2017 berjumlah 33 pegawai, 2018 berjumlah 44 pegawai, 2019 berjumlah 62 pegawai, 2020 berjumlah 61 pegawai, dan yang terakhir pada tahun 2021 berjumlah 60 pegawai, sehingga data SKP yang diperoleh dari 2017-2021 berjumlah 260 SKP dari tahun ke tahun. Kemudian seleksi SKP yang lengkap dari 2017-2021, untuk digunakan pada tahap percobaan dalam aplikasi *rapidminer*.

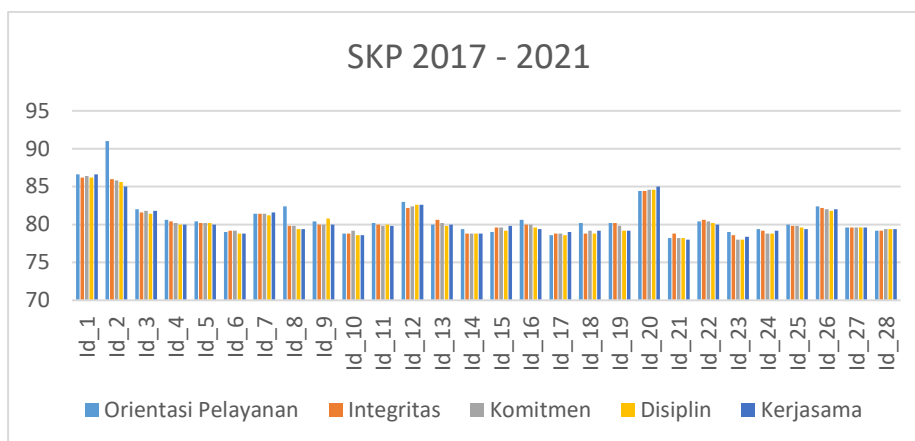


Gambar 3. Grafik Line SKP Tidak Lengkap dan Lengkap Tahun 2017- 2021

Pada Gambar 3 di atas menyatakan bahwa pada tahun 2017 – 2021 SKP yang lengkap berjumlah 28. SKP yang akan dianalisa berupa 1 SKP yang ada di setiap

tahun 2017-2021, dan untuk SKP yang tidak digunakan berupa SKP yang tidak ada setiap tahun 2017-2021. Kemudian SKP yang lengkap pada setiap tahun 2017-2021 akan dijadikan 1 data dengan menghitung nilai rata-ratanya. SKP yang lengkap akan dimasukkan ke dalam data proses.

A. Data Proses



Gambar 4. Grafik Clustered Column SKP Lengkap dengan Atribut

Pada Gambar 4 di atas adalah jumlah SKP yang sudah di seleksi dengan mengambil rata-rata setiap SKP yang lengkap, dan setiap 1 SKP memiliki 5 data/atribut *kinerja pegawai*. Setelah SKP yang di seleksi telah ditentukan, melakukan tahap percobaan pada algoritma *k-means*, *k-medoids*, dan *x-means*, untuk menentukan algoritma yang lebih optimal dalam pengelompokkan *kinerja pegawai* dengan bantuan aplikasi rapidminer.

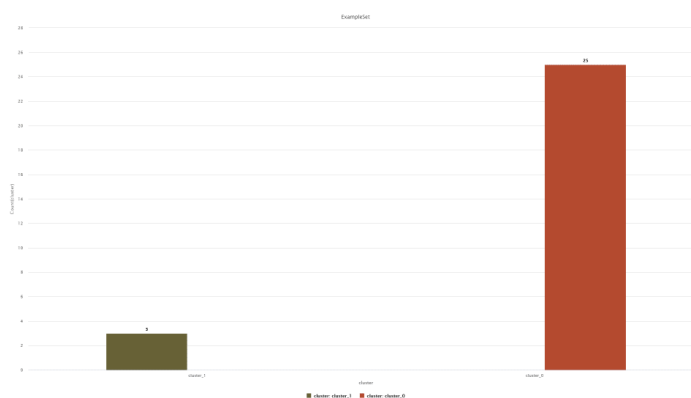
B. Centroid Data

Pada saat sebelum proses *clustering* dengan memakai algoritma *k-means*, *k-medoids*, *x-means*, yang harus dilakukan adalah menentukan nilai titik tengah atau centroid data. Pada proses menentukan nilai titik tengah, secara otomatis nilai-nilai titik tengah akan ditentukan oleh rapidminer.

C. Pengujian Data Menggunakan RapidMiner

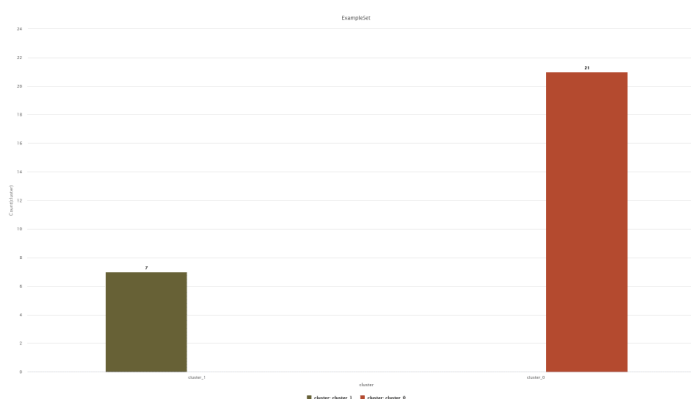
1. Visualisasi Bar(Column) Cluster *k-means*, *k-medoids*, dan *x-means*

Terdapat 2 cluster yang memakai algoritma *k-means*, *k-medoids*, dan *x-means*, dengan memakai 28 SKP yang telah di seleksi dari tahun 2017–2021, *clustering* yang dipakai memiliki atribut nilai tinggi dan nilai rendah. Dalam percobaan ini dilakukan proses *clustering*, untuk melihat cluster yang dihasilkan pada setiap algoritma dalam pengelompokkan *kinerja pegawai*.



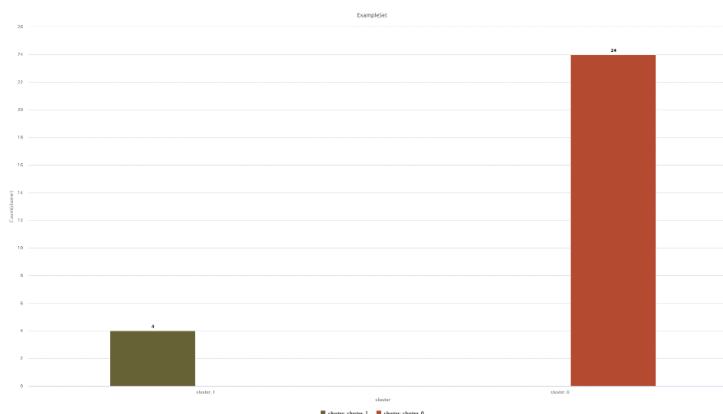
Gambar 5. Grafik Bar(Column) Cluster *k-means*

Pada Gambar 5 di atas menunjukkan bahwa pegawai yang termasuk di kategori nilai tinggi (C_1) berjumlah 3 pegawai, sedangkan di kategori nilai rendah (C_0) berjumlah 25 pegawai.



Gambar 6. Grafik Bar(Column) Cluster *k-medoids*

Pada Gambar 6 di atas menunjukkan adanya perbedaan dengan algoritma sebelumnya pegawai yang termasuk di kategori dengan nilai tinggi (C_1) berjumlah 7 pegawai, sedangkan di kategori nilai rendah (C_0) berjumlah 21 pegawai.

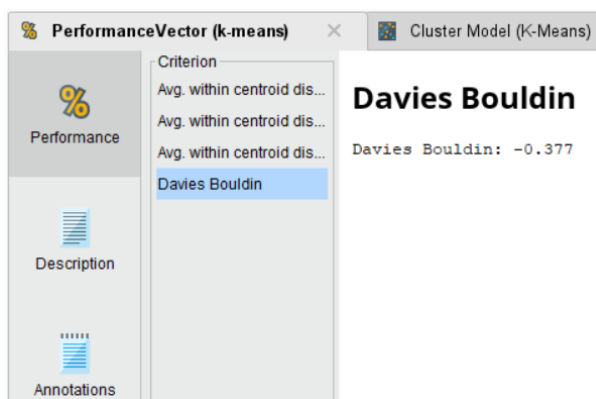


Gambar 7. Grafik Bar(Column) Cluster *x-means*

Pada Gambar 7 di atas menunjukkan terdapat perbedaan jumlah data yang ada di dalam cluster, sehingga berbeda dengan dua algoritma sebelumnya, di kategori nilai tinggi (C_1) berjumlah 4 pegawai, sedangkan di kategori nilai rendah (C_0) berjumlah 24 pegawai.

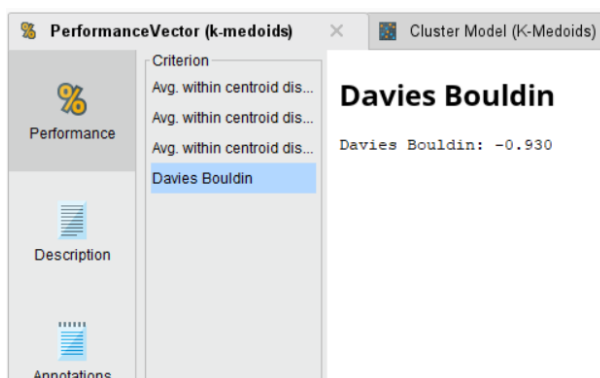
2. Visualisasi nilai *davies bouldin index* algoritma *k-means*, *k-medoids*, dan *x-means*

Setelah percobaan *clustering* yang dihasilkan setiap algoritma, memiliki perbedaan jumlah data yang masuk kedalam masing-masing cluster. Selanjutnya melakukan percobaan dengan menggunakan metode *davies bouldin index* sebagai pembandingan di algoritma *k-means*, *k-medoids*, dan *x-means*. Dalam metode ini dilakukan, untuk menentukan proses algoritma *clustering* yang lebih optimal.



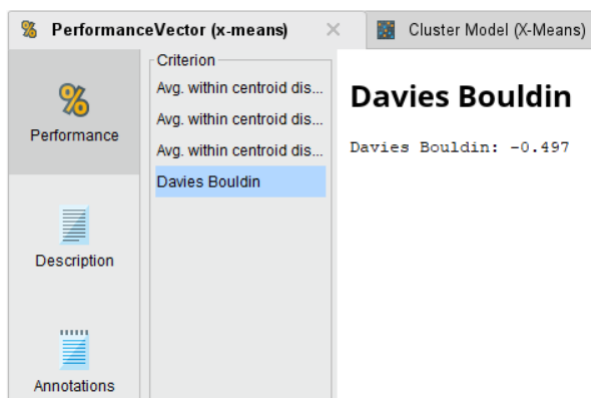
Gambar 8. Nilai *Davies Bouldin Index k-means*

Pada Gambar 8 di atas menunjukkan bahwa *davies bouldin index* di performance vektor *k-means* sebesar -0.377.



Gambar 9. Nilai *Davies Bouldin Index k-medoids*

Pada gambar 9 di atas menunjukkan bahwa *davies bouldin index* di performance vektor *k-medoids* sebesar -0.930, nilai DBI yang dihasilkan lebih tinggi dari nilai DBI dalam percobaan algoritma sebelumnya.



Gambar 10. Nilai *Davies Bouldin Index X-Means*

Pada Gambar 10 di atas menunjukkan bahwa *davies bouldin index* di performance vektor *x-means* sebesar -0.497, nilai DBI yang dihasilkan merupakan nilai kedua tertinggi dari hasil percobaan pada algoritma *k-means* dan algoritma *k-medoids*.

D. Analisa Hasil

Tahap terakhir akan di lakukan analisa terhadap *davies bouldin index* untuk setiap algoritma, agar bisa menentukan metode *clustering* algoritma apa yang lebih optimal.

Tabel 1. Nilai *Davies Bouldin Index*

Algoritma	Davies Bouldin Index
K-Means	-0.377
K-Medoids	-0.930
X-Means	-0.497

Pada tabel 1 terdapat hasil percobaan tiga algoritma yang dilakukan dalam metode DBI, pertama algoritma *k-means* memiliki nilai DBI -0.377, kedua algoritma *k-medoids* memiliki nilai DBI 0.930, dan terakhir algoritma *x-means* memiliki nilai DBI 0.497. Algoritma *k-means* merupakan algoritma yang lebih optimal dari antara algoritma *k-medoids* dan algoritma *x-means*, karena memiliki nilai DBI yang terkecil, semakin optimal algoritma yang di pilih, maka data *kinerja pegawai* yang dikelompokkan mendapatkan hasil yang lebih optimal.

V. KESIMPULAN DAN SARAN

Berdasarkan hasil yang diuji dalam performance *davies bouldin index* pada setiap algoritma yang diterapkan pada penelitian ini, algoritma *k-means* menjadi metode yang terbaik dalam menentukan data *kinerja pegawai* ke dalam kriteria cluster yang dibutuhkan, karena memiliki nilai DBI paling terkecil, di mana cluster yang dihasilkan dalam metode *k-means* kriteria nilai tinggi (C_1) berjumlah 3 pegawai, sedangkan nilai rendah (C_0) berjumlah 25 pegawai.

Agar penelitian ini dapat lebih berkembang perlu adanya pengujian data dengan memakai data skp yang lebih banyak dengan mengambil sampel data 10 tahun terakhir, serta menambahkan jumlah cluster yang akan dipakai, sehingga pengelompokan data bisa lebih luas dalam proses cluster tersebut. Ada juga saran untuk penelitian ini, perlu adanya perbandingan dengan memakai algoritma selain metode *k-means*, *k-medoids*, dan *x-means*, seperti algoritma *k-means (H20)*, algoritma *agglomerative clustering*, algoritma *top down clustering*, algoritma *flatten clustering*, dan lain-lain.

DAFTAR PUSTAKA

- [1] Rosmaini, R., & Tanjung, H. 2019. "Pengaruh Kompetensi, Motivasi Dan Kepuasan Kerja Terhadap Kinerja Pegawai. Maneggio: Jurnal Ilmiah Magister Manajemen." *Keuangan dan Bisnis 2* (1): 1-15.
- [2] Aulia, D., Safii, M., & Suhendro, D. 2021. "Penerapan Algoritma K-Means dalam Proses Clustering Penilaian Kinerja Aparatur Sipil Negera di Sekretariat DPRD Pematangsiantar." *Transformatika 6* (1): 47-60.
- [3] NOYA, A., TULUSAN, F., & TAMPONGANGOY, D. 2021. "Kinerja Sekretariat DPRD Provinsi Sulawesi Utara." *JURNAL ADMINISTRASI PUBLIK 7* (107): 39-51.
- [4] Fatonah, N. S., & Pancarani, T. K. 2022. "Analisa Perbandingan Algoritma Clustering Untuk Pemetaan Status Gizi Balita Di Puskesmas Pasir Jaya." *Konvergensi Teknologi Informasi & Komunikasi 18* (1): 1-9.
- [5] Firdaus, R. D., Laksana, T. G., & Ramadhani, R. D. 2019. "Pengelompokan Data Persediaan Obat Menggunakan Perbandingan Metode K-Means Dengan Hierarchical Clustering Single Linkage." *INISTA: Journal of Informatics, Information System, Software Engineering and Applications 2* (1): 33-48.
- [6] Fatimah, S., Suriati, S., & Usman, A. 2022. "Pengelompokan Tingkat Pemahaman Guru PAUD Terhadap Pembelajaran Berbasis STEAM Menggunakan Metode X-Means Clustering." *Explorer 2* (1): 24-31.
- [7] Rinaldi, A. R., Surlanto, L., Sudrajat, D., & Kurnia, D. A. 2019. "Analisa Tingkat Kepuasan Mahasiswa Terhadap Layanan Pembelajaran Menggunakan K-Means dan Algoritma Genetika." *Jurnal ICT: Information Communication & Technology 18* (1): 60-64.
- [8] Parlina, I., Windarto, A. P., Wanto, A., & Lubis, M. R. 2018. "Memanfaatkan Algoritma K-Means dalam Menentukan Pegawai yang Layak Mengikuti Assessment Center untuk Clustering Program SDP." *CESS (Journal of Computer Engineering, System and Science) 3* (1): 87-93.

- [9] Prاتمanto, D., Wati, F. F., Rousyati, R., & Crisna, A. 2019. "Pengelompokan Siswa Penyandang Disabilitas Berdasarkan Tingkat Tunagrahita Menggunakan Algoritma K-Medoids." *Indonesian Journal on Software Engineering (IJSE)* 5 (1): 134-142.
- [10] Zuhail, N. K. 2022. "Study Comparison K-Means Clustering dengan Algoritma." *In STAINS (SEMINAR NASIONAL TEKNOLOGI & SAINS)* 1 (1): 200-205.