

Clustering Disiplin Pegawai Kecamatan Danau Sipin Kota Jambi Menggunakan Algoritma K-Means

Hetty Rohayani^{*1}, Irmanelly², Rico³, Noneng Marthiawati⁴, Rendi Efdiansyah⁵

Fakultas Sains dan Teknologi, Universitas Muhammadiyah Jambi, Indonesia^{1,4,5}

Fakultas Ekonomi dan Pembangunan, Universitas Muhammadiyah Jambi, Indonesia²,

Fakultas Teknik dan Ilmu Komputer, Universitas Adiwangsa Jambi³.

Email :hettyrohayani@gmail.com¹, 73irmanelly@gmail.com²,

reecho86@gmail.com³, marthiawati93@gmail.com⁴,

rendiefdiansyah7@gmail.com⁵

ABSTRAK

Disiplin merupakan kunci keberhasilan suatu organisasi dalam mencapai tujuannya, karena kedisiplinan merupakan fungsi operasional dari manajemen sumber daya manusia yang sangat penting dan akan menciptakan karyawan yang berkualitas. Disiplin karyawan dapat dilihat melalui kehadiran karyawan. Pentingnya mengevaluasi tingkat disiplin pegawai untuk meningkatkan pelayanan kepada masyarakat dan memudahkan pimpinan untuk mengetahui tingkat kedisiplinan pegawai kantor Camat Danau Sipin Kota Jambi saat ini belum mengelompokkan tingkat disiplin pegawai karena tidak ada system yang dapat membantu dalam prosesnya. Pengelompokan tingkat disiplin karyawan digunakan metode K-Means *Clustering* dalam penelitian ini. Penerapan metode K-Means *Clustering* diimplementasikan dalam aplikasi yang dibuat dengan 3 cluster dari 14 sampel data, dan tidak absen dari rumah. Data yang digunakan adalah data kehadiran untuk 14 karyawan dari tahun 2018 hingga 2021. Hasil penelitian ini adalah algoritma k-means mampu mengklasifikasikan data dengan tingkat kedisiplinan tertinggi. Sedang dan rendah dengan adanya system aplikasi ini, diharapkan pimpinan dapat dengan mudah mengetahui tingkat kedisiplinan pegawai kantor Camat Danau Sipin Kota Jambi sehingga dapat memberikan bonus bagi karyawan dengan tingkat kedisiplinan yang tinggi, dan sanksi bagi karyawan dengan tingkat disiplin yang rendah.

Kata Kunci: Kehadiran, Pegawai, *Clustering*, Algoritma K-Means

ABSTRACT

Discipline is the key to the success of an organization in achieving its goals, because discipline is an operational function of human resource management which is very important and will create quality employees. Employee discipline can be seen through the presence of employees. The importance of evaluating the level of employee discipline to improve service to the community and make it easier for leaders to find out the level of discipline for employees of the Danau Sipin Sub-District Head Office, Jambi City, currently does not classify employee discipline levels because there is no system that can assist in the process. The grouping of employee discipline levels used the K-Means Clustering method in this study. The application of the K-Means Clustering method is implemented in an application made with 3 clusters of 14 data samples, and is not absent from the house. The data used is attendance data for 14 employees from 2018 to 2021. The results of this study are that the k-means algorithm is able to classify data with the highest level of discipline. Medium and low with this application system, it is hoped that leaders can easily find out the level of discipline of

employees of the Danau Sipin sub-district office, Jambi City so that they can provide bonuses for employees with a high level of discipline, and sanctions for employees with a low level of discipline.

Keywords: Attendance, Employees, Clustering, K-Means Algorithm

PENDAHULUAN

Kedisiplinan merupakan kunci keberhasilan suatu organisasi dalam mencapai tujuannya, karena kedisiplinan merupakan fungsi operatif manajemen sumber daya manusia yang sangat penting dan akan menciptakan pegawai berkualitas[1]. salah satu aspek disiplin yang dapat digunakan untuk mengevaluasi kinerja pegawai adalah dengan kehadiran yang dapat dilihat melalui absensi kehadiran pegawai[2].

Menurut Kamus Besar Bahasa Indonesia (KBBI), kehadiran adalah soal kehadiran, berada di tempat (orang, kelompok orang). Kehadiran karyawan biasanya dicatat menggunakan mesin absensi berbasis sidik jari, seperti mesin absensi jam.Cuti, biasanya dikeluarkan dengan sertifikat/formulir khusus. Oleh karena itu, teknik pengelompokan data, atau teknik clustering, digunakan untuk memudahkan pegawai dalam proses penilaian keberadaan seluruh pegawai dalam suatu instansi.

Clustering adalah proses pembentukan sekelompok data dari record-record dalam kelompok yang tidak diketahui berdasarkan kesamaan[3]. Data dengan karakteristik yang sama dikelompokkan ke dalam satu cluster, dan data dengan karakteristik yang berbeda dikelompokkan ke dalam cluster yang berbeda[4]. Teknik clustering juga biasa digunakan dalam berbagai bidang seperti biologi, psikologi, dan ekonomi[5].

Algoritma K-Means

merupakan teknik non-hierarchical clustering yang membagi data ke

dalam cluster-cluster, dimana data dengan karakteristik yang sama dikelompokkan menjadi satu cluster dan data dengan karakteristik yang berbeda dikelompokkan ke dalam cluster lainnya[6]. K-Means mencoba untuk membagi data yang ada menjadi dua atau lebih kelompok[7]. Sehingga hasil dari penerapan metode ini adalah kelompok data absensi pegawai Kecamatan Danau Sipin Kota Jambi manakah yang memiliki kedisiplinan tertinggi, sedang, dan rendah.

Melalui penerapan metode K-Means, penulis akan mengelompokkan data absensi pegawai 2 tahun sebelumnya, yaitu tahun 2020 sampai dengan tahun 2021, data absensi ini kemudian digabungkan dalam 1 tabel dan kemudian akan dibentuk kedalam beberapa kelompok, sehingga menjadi bentuk cluster-cluster, dimana data dalam satu kelompok mempunyai karakteristik yang sama satu sama lain dan mempunyai karakteristik yang berbeda dengan data yang ada dalam kelompok lain[8]. Sehingga nantinya akan menghasilkan sebuah informasi tentang bagaimana pengelompokan pegawai yang memiliki tingkat kedisiplinan tertinggi, sedang, dan rendah di Kecamatan Danau Sipin.

Saat ini Kecamatan Danau Sipin belum melakukan pengelompokan absensi kehadiran pegawai, sehingga tidak dapat mengetahui berapa jumlah pegawai yang tingkat kedisiplinannya tinggi, sedang, dan rendah.pengelompokan ini juga

bertujuan agar pimpinan dapat mengetahui pegawai mana yang memiliki tingkat kedisiplinan tinggi sehingga dapat memberikan apresiasi atau penghargaan berupa bonus kepada pegawai tersebut dan pegawai lainnya agar lebih meningkatkan lagi kedisiplinannya. Sedangkan untuk pegawai yang tingkat kedisiplinannya rendah, pimpinan dapat memberikan sanksi kepada pegawai tersebut agar kedepannya bisa memperbaiki dan meningkatkan kedisiplinannya sehingga visi, misi, dan tujuan desa dapat tercapai.

LANDASAN TEORI

A. Algoritma-Algorithm K-Means

K-Means merupakan salah satu algoritma data mining yang dapat digunakan untuk melakukan pengelompokan atau clustering data. K-Means merupakan metode pengelompokan parsial yang juga dapat digunakan untuk membagi data menjadi kelompok/cluster[9]. Data mining sendiri merupakan proses menganalisis data untuk menemukan pola-pola yang ada pada dataset. Data mining juga memiliki kemampuan untuk menganalisis data dalam jumlah yang sangat besar menjadi informasi berupa pola-pola yang bermakna untuk digunakan oleh pengambil keputusan [10]. Data mining sangat berguna bagi instansi pemerintah dan bisnis yang ingin menganalisis data dalam jumlah yang sangat besar.[11] . Ada banyak pendekatan untuk membuat sebuah cluster, salah satunya adalah dengan membuat aturan yang mendikte keanggotaan dalam kelompok yang sama berdasarkan kesamaan antar anggota. Pendekatan lain adalah dengan membuat satu set fungsi yang mengukur beberapa properti clustering sebagai fungsi dari

beberapa parameter clustering. Metode K-Means merupakan bagian dari algoritma distance clustering yang membagi data menjadi serangkaian cluster, dan algoritma ini hanya bekerja untuk atribut numerik.

Ada beberapa cara untuk mengelompokkan k-means:

- Tentukan jumlah kelompok.
- Secara acak menetapkan data ke grup.
- Hitung pusat kelompok (centroid/rata-rata) dari data-data yang ada dimasing-masing kelompok.
- Alokasikan masing-masing data ke dalam centroid/rata-rata terdekat.

$$D(x_2, x_1) = ||x_2 - x_1 ||_2 =$$

$$\sqrt{\sum_{j=1}^p |X_2 - X_1_j|^2}$$

- Jika masih ada kelompok pergeseran data, perubahan nilai centroid melebihi ambang batas yang ditentukan, atau perubahan nilai fungsi tujuan yang digunakan masih melebihi ambang batas yang ditentukan. Jika demikian, kembali ke langkah 3.

B. Data Mining

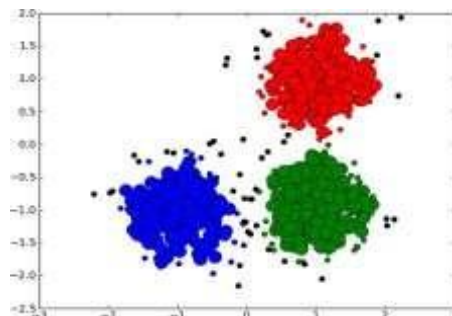
Data mining adalah analisis data untuk menemukan hubungan yang berbeda dan menyimpulkan bahwa mereka sebelumnya tidak diketahui dengan cara yang sekarang dipahami dan berguna bagi pemilik data[12].

Data mining dapat digambarkan sebagai proses penggalian pengetahuan dari sejumlah besar data yang tersedia. Pengetahuan yang dihasilkan dari proses data mining harus baru, dapat dimengerti, dan berguna. Dalam data mining, data disimpan secara elektronik dan

diproses secara otomatis oleh komputer menggunakan teknik dan perhitungan tertentu[13].

C. Clustering

Clustering atau klasifikasi adalah suatu cara membagi seri data menjadi kelompok-kelompok berdasarkan kesamaan tertentu[10].



Gambar 1 Clustering

D. Algoritma K-Means

Algoritma K-Means adalah algoritma clustering iteratif yang pertama-tama membagi dataset menjadi sejumlah K cluster yang ditentukan. Algoritme K-Means mudah diimplementasikan dan dijalankan, relatif cepat, dapat disesuaikan, dan banyak digunakan dalam praktik. Secara historis, K-Means adalah salah satu algoritma terpenting dalam data mining[14].

Algoritma K-Means dimulai dengan memilih K secara acak. dimana K adalah jumlah cluster yang terbentuk. Kemudian secara acak menetapkan nilai K. Untuk sementara nilai adalah pusat cluster, atau biasanya centroid acak, rata-rata, atau "rata-rata". Hitung jarak dari semua data yang ada ke setiap centroid menggunakan rumus Euclidean sampai ditemukan jarak terdekat ke semua data centroid. Mengklasifikasikan semua data berdasarkan kedekatandengan centroid. Ikuti langkah-langkah ini sampai nilai centroid tetap sama[15].

JarakEuclidean yang dirumuskan sebagai berikut:

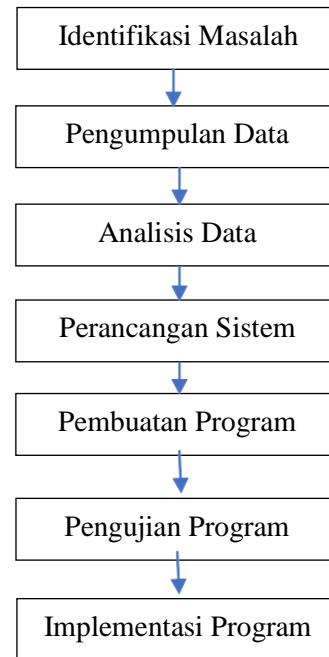
$$D(i, j) = \sqrt{(X_{1j} - X_{1i})^2 + (X_{2j} - X_{2i})^2 + \dots + (X_{kj} - X_{ki})^2} \dots$$

$D(i, j)$ = Jarak data ke i ke pusat cluster j
 X_{ki} = Data ke i pada atribut dat ke k
 X_{kj} = Data ke j pada atribut dat ke k

METODE PENELITIAN

A. Metode Desain Sistem

Penelitian ini menggunakan metode K-means clustering. Alur survei ditunjukkan pada gambar berikut:



Gambar 2. Keraangka Kerja Penelitian

1. Mengidentifikasi masalah

Fase ini merupakan langkah awal dalam penyelidikan, dan identifikasi masalah merupakan upaya untuk menemukan dan mengumpulkan masalah yang diterima selama penyelidikan. Masalah yang saya terima adalah Kecamatan Danau Sipin Kota Jambi tidak mengukur tingkat kedisiplinan pegawai melalui absensi berkelompok. Cari tahu pegawai mana yang memiliki tingkat kedisiplinan tinggi, sedang, dan rendah.

2. Pengumpulan Data

Pada tahap ini dilakukan proses pengumpulan data wawancara dengan Rizalul Fikri, SE, M.Ap, Camat Danau Sipin Kota Jambi data menggunakan metode, dan observasi untuk mengamati dan menganalisis proses

pengelompokan data tingkat kedisiplinan pegawai Kecamatan Danau Sipin Kota Jambi untuk mendapatkan data dan informasi yang dibutuhkan peneliti.

3. Analisis Data

Pada tahap ini, data yang terkumpul dianalisis. Oleh karena itu kendala dan permasalahan yang peneliti temui saat mengelompokkan tingkat kedisiplinan pegawai di Kecamatan Danau Sipin agar dapat dicarikan solusi dari permasalahan tersebut.

4. Perencanaan sistem

Pada tahap ini penulis membuat rancangan sistem yang diusulkan. Perancangan sistem ini meliputi perancangan ASI dan perancangan UML yang meliputi perancangan user interface dan perancangan database sistem untuk lebih mengatur pengelolaan file yang ada.

5. Pembuatan Program

Pada fase ini, penulis menulis program dalam bentuk desain program aplikasi tingkat lapangan rakyat dan menggunakan VisualBasicNet.2010 dalam database Microsoft Access 2010 untuk menulis pengkodean program.

6. Pengujian Program

Pada saat ini, penulis melakukan uji coba program untuk melihat apakah ada masalah dengan program tindakan disiplin pegawai.

7. Implementasi program

Implementasi adalah fase penerapan program aplikasi tingkat disiplin karyawan di organisasi penelitian, memastikan bahwa aplikasi yang dikembangkan bebas bug dan kemudian menjalankan uji coba aplikasi. Pada tahap ini, hasil survei yang dilakukan juga akan dievaluasi. Evaluasi melibatkan evaluasi hasil dari program aplikasi yang diimplementasikan, terlepas dari apakah itu memenuhi persyaratan pengguna.

HASIL DAN PEMBAHASAN

A. Analisis Data

Tahapan analisis data merupakan tahap yang paling penting dalam pengembangan sebuah sistem, karena pada tahap inilah nantinya dilakukan evaluasi kinerja, identifikasi terhadap masalah yang ada, rancangan sistem dan langkah-langkah yang dibutuhkan untuk perancangan yang diinginkan [16].

Tabel 1. Data Absensi Pegawai Kecamatan Danau Sipin Kota Jambi

Sumber: kecamatan Danau Sipin.

No	Nama	Tidak Absen			Tidak Absen	
		Masuk	Sakit	Izin	Alpa	Pulang
1	RIZALUL FIKRI	6	2	15	0	10
2	HAVID DWIRIL	10	10	6	1	9
3	MAHANI	4	8	11	0	0
4	MUHAMMAD HUSNI THAMRIN DEBBY MUTIARA	7	6	8	1	2
5	DESY ANGRAINI	9	10	7	5	16
6	GATOT SETIAWAN RAHMAT HIDAYAT	11	4	6	1	15
7	MUHAMMAD FAUZI EMA SURYANI	26	6	4	2	27
8	ZURKARNAIN RIZAL EFENDI	16	3	11	7	26
9	JUMIATI M.TAMRIN EDI SURYADI	11	4	2	11	40
10	ANDALINA BR BARUS	12	7	10	18	38
11	DEDI KUSWORO AZRIANI LATIF	13	4	6	12	34
12	ENDANG PURNAMA YANTI	21	11	8	11	55
13	ADI JUNAIDI	17	3	10	9	31
14	MUSLAMAH ELYAWATI	16	1	4	9	42

B. Analisis Sistem

Analisis adalah kegiatan yang bertujuan untuk menyelidiki dan menilai bentuk masalah yang ada pada sistem. Agar sistem yang dirancang dapat bekerja dengan baik, maka perlu dilakukan analisis kinerja sistem yang berfokus pada pengembangan sistem. Analisis sistem adalah kegiatan yang

menggambarkan sistem informasi dunia nyata yang lengkap dalam suatu komponen dengan tujuan mengidentifikasi dan menilai masalah yang muncul. Ini mengarah pada solusi perbaikan dan pengembangan yang lebih baik dan berbasis kebutuhan.

C. Alat analisis dan desain sistem

Antarmuka dirancang dengan mengacu pada diagram urutan yang dibuat. Perancangan antarmuka membantu menjelaskan hasil akhir dari sistem dalam hal antarmuka sistem yang berinteraksi langsung dengan pengguna.

1. Bahasa Pemodelan Terpadu (UML)

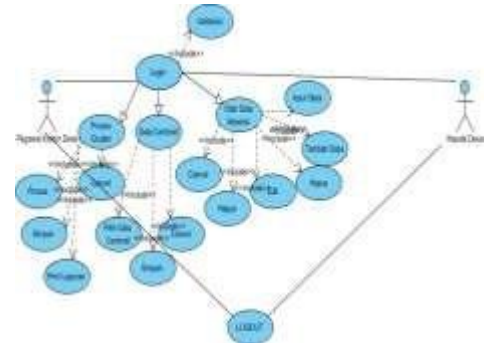
Unified Modeling Language (UML) adalah alat untuk pengembangan perangkat lunak *berorientasi objek*. UML sendiri sangat handal dalam dunia pengembangan sistem berorientasi objek. Ini menyediakan bahasa pemodelan visual yang memungkinkan UML untuk menggunakan semua elemen dan diagram berdasarkan paradigma berorientasi objek untuk membuat desain yang mudah dipahami oleh pengembang sistem dan memiliki mekanisme yang efektif untuk berbagi. UML menyediakan standar untuk merancang model untuk sistem lain, mendefinisikan notasi dan sintaks/semantik.

Secara umum, UML hanya menggunakan empat diagram: *diagram use case*, *diagram kelas*, *diagram aktivitas*, dan *diagram urutan*.

2. Use case Diagram

Use case diagram adalah diagram yang menggambarkan interaksi antara sistem dengan aktor-aktor yang berperan dalam sistem

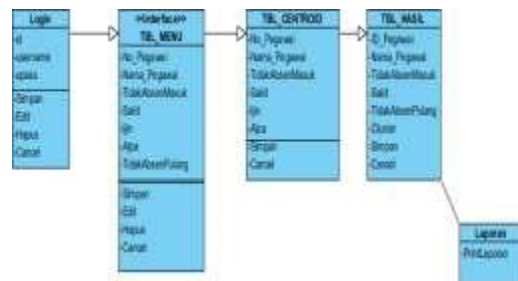
[17].Berikut contohhuse case diagram login:



Gambar 3. Use Case Diagram

3. Class Diagram

Class Diagram adalah diagram yang menggambarkan struktur suatu sistem dalam hal mendefinisikan kelas-kelas yang dibuat untuk membuat sistem.

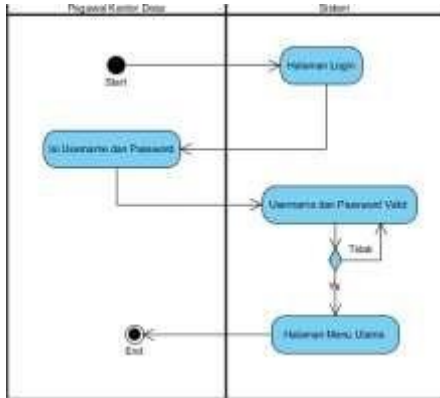


Gambar 4. Class Diagram

4. Activitydiagram

Activity diagram merupakan diagram yang menggambarkan urutan aktivitas dalam sistem yang dapat dilakukan oleh *user* dari awal hingga akhir sehingga dapat memberikan pemahaman keseluruhan bagi *user*[18]. *Activity diagram* juga merupakan tipe diagram khusus dari *statechart* yang berfokus pada aktivitas atau potongan-potongan proses yang sesuai dengan metode atau fungsi

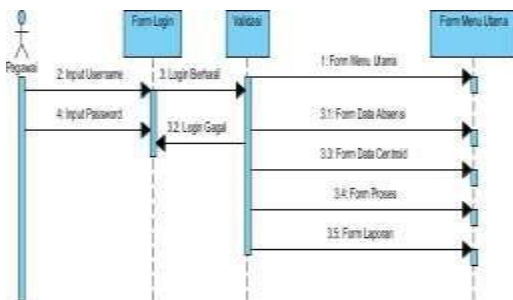
angora dan aktivitas yang telah terjadi [19]. Berikut contoh activity diagram login:



Gambar 5. Activity Diagram

5. Sequence Diagram

Sequencediagrammerupakan diagram yang menunjukkan alur dari sebuah sistem untuk fungsionalitas [20]. Berikut contoh *sequence diagram* login:



Gambar 6. Sequence Diagram

C. Analisis Masalah

Kajian disiplin pegawai di Kecamatan Danau Sipin Kota Jambi terus menimbulkan permasalahan seiring dengan berjalannya proses klasifikasi setiap tahunnya. Dilihat dari alur sistem informasi yang saat ini beroperasi, absensi masih manual dan belum ada sistem pengklasifikasian tingkat kedisiplinan pegawai. Hal ini dapat menyebabkan manajemen kehilangan jejak karyawan mana yang memiliki tingkat kedisiplinan tinggi, sedang, dan rendah. Anda dapat menilai tingkat

kedisiplinan seorang pegawai dengan memberikan bonus kepada pegawai dengan tingkat kedisiplinan yang tinggi dan yang diberi wewenang oleh tingkat kedisiplinan yang rendah. Tentang analisis masalah metode PIECES. Ketika menganalisis suatu sistem, beberapa aspek biasanya dipertimbangkan, seperti kinerja, informasi, ekonomi, keamanan aplikasi, efisiensi, dan layanan pelanggan. Analisis ini disebut analisis PIECES (kinerja, informasi, ekonomi, kontrol, efisiensi, dan layanan).

D. Analisis kebutuhan sistem

Analisis kebutuhan sistem merupakan tahap perancangan sistem yang bertujuan untuk mengidentifikasi masalah dan kebutuhan sistem sehingga pengembang dapat memahami gambaran umum sistem sesuai dengan kebutuhan. Untuk mengetahui tingkat kedisiplinan pegawai di Kecamatan Danau Sipin Kota Jambi, Kecamatan Danau Sipin Kota Jambi membutuhkan suatu sistem yang dapat digunakan untuk mengklasifikasikan tingkat kedisiplinan pegawai. Oleh karena itu, diperlukan suatu cara untuk menyederhanakan proses pengelompokan disiplin pegawai dengan hasil yang baik. Perlunya suatu sistem yang dapat memberikan informasi yang didambakan. Oleh karena itu pada Kecamatan Danau Sipin Kota Jambi perlu membuat aplikasi clustering data mining menggunakan metode K-Means untuk mengelompokkan tingkat disiplin pegawai. Sistem ini diharapkan dapat memberikan solusi yang mendekati solusi dunia nyata yang efektif dan efisien. Selain itu, didukung oleh ketersediaan perangkat keras dan perangkat lunak yang ada, serta ketersediaan orang yang dapat menggunakan komputer, memungkinkan Anda untuk mengelola sistem berbasis komputer.

E. Analisis Pengguna Sistem

Dalam penelitian ini Kota Jambi yang

terlibat dalam penggunaan sistem atau pengguna ketika mengelompokkan tingkat kedisiplinan pegawai-pegawai, Kecamatan, atau pengelola yang secara khusus mengolah data tingkat disiplin pegawai, yaitu kantor pusat. Pengguna ini mampu atau diharapkan dapat menggunakan dan mengoperasikan perangkat keras dengan baik dan menjalankan sistem ini semaksimal mungkin.

F. Tujuan Pengguna Sistem

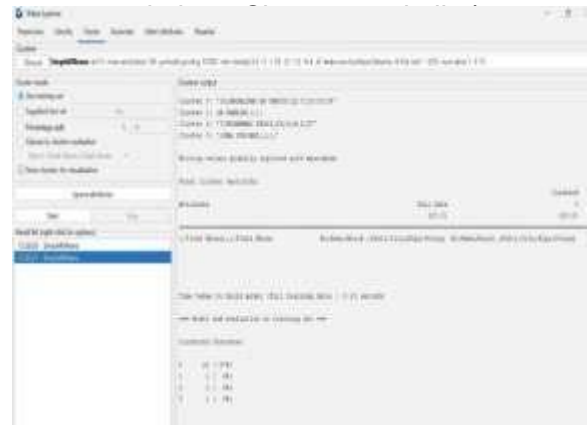
Tujuan utama penggunaan sistem ini adalah untuk memastikan keakuratan data pada tingkat kedisiplinan pegawai yang jelas dan relevan dengan keadaan sebenarnya. Hasilnya, manajemen dan orang-orang terdekatnya nantinya dapat mengetahui kelompok karyawan mana yang memiliki tingkat kedisiplinan tinggi, sedang, dan rendah. Hal ini memungkinkan pemimpin selanjutnya untuk menawarkan bonus dan penghargaan kepada karyawan dengan tingkat disiplin yang lebih tinggi dan membuat keputusan untuk menjatuhkan sanksi kepada karyawan dengan tingkat disiplin yang lebih rendah. Akibatnya, karyawan diharapkan bersaing untuk meningkatkan kedisiplinan.

G. Implementasi Clustering K-Means melalui WEKA

- Masuk ke metode Weka Open file csv import ke Weka
- Pilih Cluster Choose dengan metode SimpleKMeans

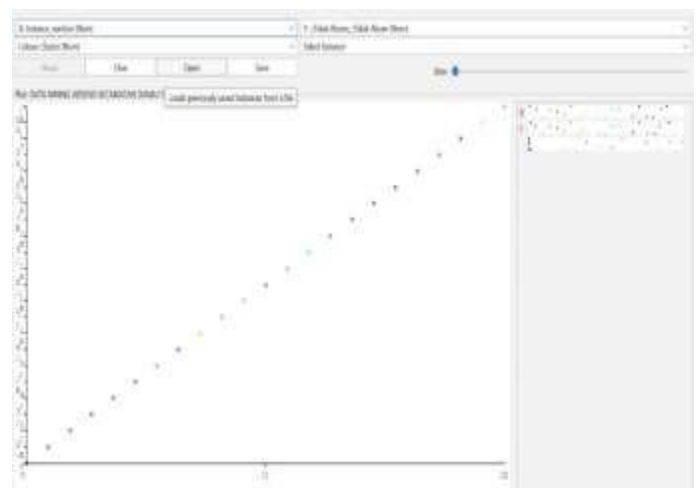


Gambar 7. Proses Import ke Weka



Gambar 8. Hasil Cluster Output

Hasil output grafik dapat dilihat pada gambar 9.



Gambar 9. Output Cluster

KESIMPULAN

Berdasarkan hasil analisis dan pembahasan data mining menggunakan metode non-hierarki, implementasi metode K-Means clustering untuk pengelompokan tingkat disiplin pegawai di Kecamatan Danau Sipin Kota Jambi merupakan aplikasi absensi pegawai yang berhasil diimplementasikan, dan mudah untuk menemukan sekelompok pegawai yang pegawainya memiliki

merupakan bagian dari tahapan proses Knowledge Discovery in Database (KDD) . Jurnal Edik Informatika,” *J. Edik Inform.*, vol. 2, no. 2, pp. 213-219, 2019.

- [13] Patel, “濟無No Title No Title No Title,” pp. 9-25, 2019.
- [14] N. Wakhidah, “Clustering Menggunakan K-Means Algorithm,” *J. Transform.*, vol. 8, no. 1, p. 33, 2010.
- [15] P. Ulil, F. Aulia, and S. Saepudin, “Penerapan Data Mining K-Means Clustering Untuk Mengelompokkan Berbagai Jenis Merk Laptop,” pp. 241-249, 2021.
- [16] Pujiyanto, “ANALISIS SISTEM Pendahuluan Langkah-langkah Analisis Sistem,” pp. 1-13, 2009.
- [17] S. Kurniawan, T. Bayu, “Perancangan Sistem Aplikasi Pemesanan Makanan dan Minuman Pada Cafeteria NO Caffe di TAnjung Balai Karimun Menggunakan Bahasa Pemrograman PHP dan My.SQL,” *J. Chem. Inf. Model.*, vol. 53, no. 9, pp. 1689-1699, 2020.
- [18] 7) Marakas dan O’Brien (2017, “Bab li Landasan Teori,” *J. Chem. Inf. Model.*, vol. 53, no. 9, pp. 8-24, 2018.
- [19] S. m A.S Rosa, “Activity Diagram,” pp. 8-26, 2014.
- [20] Y. Zamrodah, “濟無No Title No Title No Title,” vol. 15, no. 2, pp. 1-23, 2016.