

# **BAB I**

## **PENDAHULUAN**

### **1.1 Latar Belakang**

Manajemen sumber daya manusia merupakan bagian dari manajemen organisasi pada perusahaan dengan fokus terhadap pengelolaan sumber daya manusia, dengan tujuan meningkatkan kontribusi produktif perindividu yang terdapat didalam suatu perusahaan melalui sejumlah cara yang dapat dipertanggung jawabkan. Pengelolaan dan pendayagunaan sumber daya manusia yang baik dan berkisnambungan terhadap sumber daya manusia yang ada sehingga mereka dapat bekerja secara optimal, efektif dan produktif dalam mencapai tujuan perusahaan (Nawawi, 2001).

Sumber daya manusia merupakan individu produktif yang merupakan penggerak dalam suatu organisasi, baik di institusi maupun perusahaan. sumber daya manusia merupakan fungsi asset penting dalam perusahaan sehingga harus terus dilatih dan dikembangkan kemampuan yang dimiliki. terdapat dua pengertian SDM secara makro, yaitu SDM makro merupakan suatu wilayah yang penduduknya memiliki usia produktif, baik yang sudah bekerja maupun yang tidak bekerja. Sedangkan SDM mikro memiliki artian sempit yaitu individu yang bekerja pada sebuah perusahaan maupun institusi tertentu serta menjadi anggota suatu perusahaan dan biasanya disebut dengan pegawai, buruh, pekerja, karyawan, tenaga kerja dan lain – lain (Susan, 2018).

Perusahaan XYZ merupakan perusahaan yang mengelola bisnis penjualan yang secara langsung kepada konsumen. Pada umumnya suatu perusahaan memiliki tujuannya masing masing, dan bagaimana suatu perusahaan harus tetap mempertahankan kualitas dan pelayanan yang baik dari setiap produk yang dihasilkan. Namun pencapaian suatu perusahaan didukung juga oleh sumber daya manusianya, jika sumber daya manusianya baik maka diharapkan kinerja dari perusahaan akan baik. Untuk mendapatkan kinerja yang baik maka perlu dilakukannya evaluasi kinerja.

Evaluasi kinerja dilakukan secara periodik yang telah ditetapkan oleh perusahaan terhadap kinerja karyawan. Evaluasi kinerja mengacu pada suatu sistem formal dan terstruktur yang digunakan untuk dapat mengukur dan memberikan

penilaian terhadap faktor – faktor yang berkaitan dengan pekerjaan, perilaku, hasil, dan tingkat kehadiran (absensi). Absensi merupakan bagian dari aktivitas pelaporan maupun pendataan kehadiran karyawan yang berkenaan dengan tanggung jawab saat bekerja dan kehadiran tepat waktu bisa dikatakan karyawan memiliki sifat disiplin. Kedisiplinan dari seorang karyawan merupakan tolak ukur dalam melakukan penilaian terhadap kinerja karyawan di perusahaan serta perusahaan dapat memberlakukan sanksi apabila kewajiban seorang karyawan dilanggar atau tidak ditaati (Arya et al., 2020). Dengan menggunakan data karyawan maka akan dilakukan pengolahan data dengan memanfaatkan teknik data mining.

Data merupakan aspek penting dalam suatu perusahaan, data dikelola untuk mendapatkan informasi yang berguna bagi perusahaan. Informasi yang didapat dari data tersebut dapat digunakan untuk melakukan strategi bisnis, maupun pengambilan suatu kebijakan atau keputusan. Untuk mendapatkan informasi dari data yang besar, perlu memanfaatkan teknik data mining untuk dapat menggali informasi (Ermatita, 2014). Data mining merupakan proses dalam pencarian pola atau informasi menarik dalam data terpilih dengan menggunakan teknik maupun metode tertentu. Data mining merupakan gabungan dari beberapa disiplin ilmu yang menggabungkan teknik pembelajaran mesin, statistik, pengenalan pola, visualisasi, database untuk dapat menangani permasalahan pengambilan informasi dari *database*.

Dalam teknik data mining dibagi menjadi beberapa kelompok berdasarkan tugas yang dapat dilakukan meliputi, Deskripsi (merupakan pencarian cara untuk dapat menggambarkan pola serta kecenderungan yang ada di dalam data), Estimasi (sama dengan klasifikasi namun variabel target estimasi terarah ke numerik daripada ke arah kategori), Prediksi (sama dengan estimasi dan klasifikasi, namun prediksi memberikan nilai hasil di masa yang akan datang), Asosiasi (menemukan atribut yang muncul dalam waktu yang sama), dan yang terakhir adalah pengklusteran (pengelompokan record terhadap pengamatan dalam memperhatikan dan membentuk kelas objek yang memiliki kemiripan (Mardi Yuli, 2014). Data mining disebut juga dengan Knowledge Discovery in Database (KDD) merupakan sebuah proses pencarian data secara otomatis pada sebuah memori atau database

yang besar. Untuk itu data mining dapat digunakan untuk mengevaluasi kinerja karyawan.

*Clustering* merupakan teknik pembelajaran tanpa pengawasan (*unsupervised*) yang tidak memerlukan dataset berlabel, *clustering* merupakan pengelompokan sekumpulan objek yang mirip dalam satu *cluster* (kelas). Dalam melakukan proses analisis, data yang dikelompokkan ke dalam satu cluster sehingga objek di dalam satu cluster memiliki kemiripan yang tinggi satu sama lain dengan memiliki perbedaan yang tinggi dibandingkan objek pada cluster lain (Muttaqin & Defriani, 2020).

*Kmean-s clustering* adalah salah satu metode dalam clustering non hierarki yang mempartisi data ke dalam satu atau lebih cluster atau kelompok, berdasarkan atribut menjadi  $k$  partisi, dimana  $k < n$ , dengan parameter input sebanyak  $k$  dan membagi kumpulan  $n$  objek kedalam  $k$  *cluster* sehingga tingkat kemiripan antara anggota dalam satu *cluster* tinggi sedangkan untuk kemiripan pada anggota *cluster* lain sangat rendah. Kemiripan pada anggota *cluster* diukur dengan kedekatan objek terhadap nilai mean pada cluster yang disebut dengan centeroid cluster (Suhanda et al., 2020).

*K-medoids* merupakan versi umum dari dari algoritma *k-means* yang dimana algoritma *k-medoids* dapat mengukur jarak dan memiliki komputasi yang lebih intensif. Algoritma *K-medoids* juga biasa disebut dengan *Partition around medoids* yang juga merupakan versi dari algoritma *k-means*. Pada algoritma *k-means* memiliki kelemahan pada *outlier* yang *sensitive* terhadap objek dengan nilai besar yang dimana memungkinkan mengalami penyimpangan dari distribusi data untuk itu metode *k-medoids* di rancang agar dapat mengatasi kelemahan yang dimiliki metode *k-means* (Argasah & Gustian, 2022).

Berdasarkan pemaparan diatas menjadi pertimbangan bagi penulis untuk dapat melakukan analisa evaluasi kinerja karyawan dengan memanfaatkan teknik data mining menggunakan algoritma *K-medoids*. Berdasarkan dari penelitian sebelumnya yang dilakukan oleh Akhmad Upi Fitriyadi dan Ana Kurniawati pada tahun 2021 dengan judul “Analisis algoritma *K-Means* dan *K-Medoids* untuk

*Clustering* data Kinerja Karyawan Pada Perusahaan Perumahan Nasional” dengan hasil perhitungan yang telah dilakukan dengan menggunakan metode *confusion matrix* dengan menggunakan 3 parameter yaitu *accuracy*, *recall* dan *precision* menyimpulkan bahwa algoritma *k-means* lebih baik jika dibandingkan dengan algoritma *k-medoids* karena memiliki nilai akurasi dan *recall* yang lebih tinggi (Fitriyadi & Kurniawati, 2021). Penelitian yang dilakukan oleh Anisa Fira dkk dengan judul “Komparasi algoritma *K-Means* dan *K-Medoids* Untuk Pengelompokan Penyebaran Covid-19 di Indonesia" diperoleh kesimpulan hasil *cluster* yang dievaluasi dengan menggunakan nilai *Silhouette Coefficient*. Algoritma *k-medoids* memiliki nilai *silhouette* sebesar 0,347 sedangkan algoritma *k-means* memiliki nilai *silhouette* sebesar 0,207. Sehingga algoritma *k-medoids* dapat memberikan hasil pengelompokan yang lebih baik daripada algoritma *k-means* (Fira & Rozikin, 2021). Dan penelitian selanjutnya yang dilakukan oleh Mira orisa dan Ahmad Faisol dengan judul penelitian “Analisis Algoritma *Partion Around Medoid* untuk Penentuan Klasterisasi” yang memperoleh hasil kesimpulan Algoritma *k-medoids* dapat melakukan klasterisasi terhadap dataset yang memiliki *outliers* dengan membaginya menjadi 3 cluster. Dengan hasil evaluasi kualitas *cluster* menggunakan metode *silhouette* yang menghasilkan struktur lemah karena objek antar cluster saling timpang tindih(Orisa & Faisol, 2021).

Berdasarkan penelitian yang telah dilakukan sebelumnya dalam pengelompokan (*clustering*) yang telah dipaparkan diatas terlihat bahwa algoritma *k-means* dan *k-medoids* digunakan pada berbagai dataset yang berbeda. Jumlah *k cluster* awal yang ditentukan secara acak dapat mempengaruhi hasil clustering baik penggunaanya pada algoritma *k-means* ataupun algoritma *k-medoids*.

Pada penelitian ini bermaksud untuk dapat melakukan pengelompokan kinerja karyawan berdasarkan karakteristik yang terbentuk dengan menggunakan algoritma *k-medoids* yang dimana algoritma *k-medoids* dapat mengatasi kelemahan algoritma *k-means* yang sensitive terhadap *outliers* pada dataset dan menggunakan metode *elbow* untuk dapat menentukan *cluster* optimal yang digunakan. Adapaun untuk menguji hasil kualitas dari klasterisasi algoritma *k-medoids* menggunakan

metode *silhouette coefficient* untuk mengetahui seberapa jauh jarak antar cluster yang terbentuk.

### **1.2 Rumusan Masalah**

Berdasarkan latar belakang diatas, permasalahan yang harus diselesaikan dalam penelitian ini adalah :

1. Bagaimana hasil pengelompokan (*clustering*) dengan menggunakan metode *K-Medoids* ?
2. Bagaimana hasil pengujian yang dilakukan dengan menggunakan metode *Silhouette Coefficient*?

### **1.3 Batasan Masalah**

1. Data yang digunakan merupakan data karyawan
2. Jumlah data 8336
3. Data diperoleh dari data *public* (kaggle)
4. Menggunakan algoritma *k-medoids*
5. Jumlah satuan data dalam kurun waktu satu tahun
6. Menggunakan metode *silhouette* untuk menentukan hasil *cluster* terbaik
7. Membentuk cluster sebanyak empat cluster

### **1.4 Tujuan Penelitian**

Tujuan dari penelitian ini adalah mendapatkan informasi atau *insight* dari hasil analisis dalam mengelompokkan karyawan yang dibagi menjadi beberapa tingkatan keaktifan kerja menggunakan metode pengelompokan dengan algoritma *k-medoids*

### **1.5 Manfaat Penelitian**

Dari pemaparan tujuan diatas adapun manfaat yang didapat dari penelitian ini adalah dari hasil analisis yang dilakukan dapat digunakan perusahaan dalam melakukan pengambilan keputusan baik bagi karyawan maupun bagi perusahaan, serta dapat meningkatkan pencapaian tujuan perusahaan.