

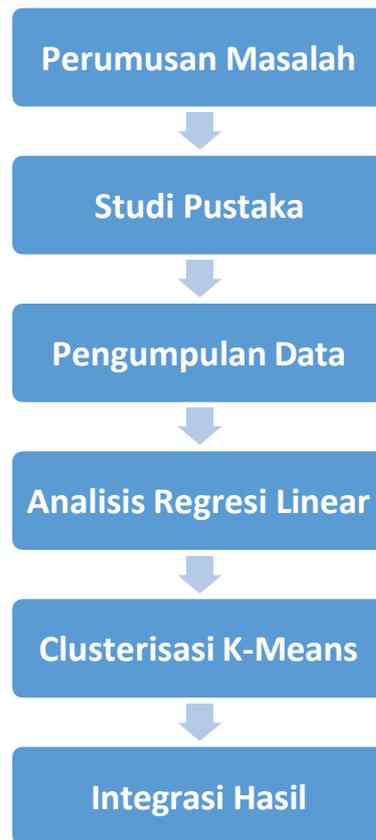
## **BAB III**

### **METODOLOGI PENELITIAN**

#### **1.1. Metodologi Penelitian**

Metodologi penelitian adalah suatu pendekatan sistematis yang digunakan untuk mengumpulkan, menganalisis, dan menginterpretasikan data dalam penelitian. Metodologi ini menjadi panduan bagi peneliti dalam merancang dan melaksanakan penelitian agar hasil yang diperoleh akurat, valid, dan dapat diandalkan. Pilihan metodologi yang digunakan bergantung pada tujuan penelitian, jenis data, dan pendekatan yang paling sesuai untuk menjawab pertanyaan penelitian. Berikut adalah langkah-langkah yang lebih rinci beserta uraian yang jelas untuk penelitian dengan judul "Analisa Kinerja Karyawan dengan Metode Regresi Linear Berganda dan Clusterisasi Menggunakan Metode K-Means" dengan fokus pada variabel motivasi kerja (insentif) dan kedisiplinan. Pembuatan alur penelitian merupakan sebagai tahapan dalam melakukan penelitian ini.

Berikut adalah alur penelitian yang akan dilakukan dalam penelitian ini.



Gambar 3. Alur Penelitian

### 1.1.1. Studi Pustaka

Tahap studi pustaka adalah kunci untuk memahami dan merumuskan landasan teoritis penelitian Anda. Dalam konteks penelitian "Analisa Kinerja Karyawan dengan Metode Regresi Linear dan K-Means dengan Variabel Kepemimpinan dan Motivasi Kerja

## 1.2. Alat dan Bahan

### 1.2.1. Alat

Pada penelitian ini alat yang digunakan sebagai berikut :

1. Laptop Acer
2. Prosesor : AMD E1-7010 APU with AMD Radeon R2 Graphics 1.5 GHz
3. Memory : 4,00 GB
4. System : 64-bit

### **1.2.2. Bahan**

Bahan yang digunakan dalam penelitian ini adalah perangkat lunak Microsoft Excel dan google form sebagai pengolah data.

### **1.3. Metode Pengumpulan data**

Data yang akan dianalisis merupakan data primer. Data primer berasal dari nara sumber kepala dan sekretaris LPPM. Metode pengumpulan data dengan menggunakan metode survei.[11]

#### **1.3.1. Pengolahan data**

##### **1. Data yang Digunakan**

Data yang terkumpul dikonversi ke format yang dapat diolah, seperti spreadsheet. Setiap observasi (pegawai) memiliki nilai-nilai untuk variabel kepemimpinan, motivasi kerja, dan kinerja. Sumber data dalam penelitian ini yaitu hasil kuisisioner tentang penilaian kinerja yang di isi oleh setiap karyawan pada PT. Mayora Kotabumi yang mencakup SOP kehadiran dan kinerja yang akan digunakan untuk menganalisa dan clustering kinerja setiap karyawan. Data tentang faktor-faktor yang mempengaruhi kinerja karyawan yaitu penerapan SOP, kehadiran, kedisiplinan, pelatihan dan insentif yang didapatkan. Data-data ini kemudian akan digunakan dalam analisis regresi linear berganda dan algoritma k-means untuk menganalisa dan mengclustering kinerja karyawan yang memungkinkan peneliti untuk mengetahui insentif yang didapat oleh setiap karyawan di masa mendatang berdasarkan faktor-faktor yang telah diidentifikasi. Penting untuk memilih data dengan hati-hati dan memastikan keandalan dan kualitasnya untuk mendukung keakuratan dan validitas hasil penelitian.

## 2. Pemeriksaan dan Pembersihan Data

Pemeriksaan dan pembersihan data merupakan tahapan yang penting dalam analisis data yang mencakup dua aktivitas utama:

1. **Pemeriksaan Data:** Pada langkah ini, data yang telah dikumpulkan diperiksa untuk menemukan ketidaksesuaian, ketidaklengkapan, atau inkonsistensi. Misalnya, data yang hilang (missing values), data duplikat, atau data yang berada di luar rentang yang diharapkan. Pemeriksaan ini bertujuan untuk memahami kualitas data, pola, dan potensi masalah yang ada sebelum data dianalisis lebih lanjut.
2. **Pembersihan Data:** Setelah masalah dalam data diidentifikasi, pembersihan dilakukan untuk memperbaiki atau menghapus data yang tidak sesuai atau cacat. Ini bisa mencakup penghapusan duplikasi, mengisi data yang hilang dengan nilai estimasi, mengoreksi kesalahan input, atau mengonversi format data agar konsisten. Pembersihan data bertujuan untuk memastikan bahwa data yang digunakan dalam analisis adalah akurat, lengkap, dan bebas dari kesalahan.

### 1.3.2. Pengolahan Dengan Metode

Data mining adalah proses mengekstrak informasi atau pola bermakna dari kumpulan data yang besar dan kompleks. Tujuannya adalah menemukan hubungan tersembunyi, pola, atau tren yang berguna untuk mendukung pengambilan keputusan. Data mining menggunakan teknik dari berbagai bidang, termasuk kecerdasan buatan, statistik, machine learning, dan database management. Tahapan dalam Data Mining:

1. **Pembersihan Data (Data Cleaning):** Menghilangkan data yang salah, hilang, atau tidak konsisten.
2. **Integrasi Data (Data Integration):** Menggabungkan data dari berbagai sumber untuk menciptakan satu dataset yang koheren
3. **Transformasi Data (Data Transformation):** Mengubah data menjadi format yang sesuai, seperti normalisasi atau agregasi.

4. Pemilihan Data (Data Selection): Memilih data relevan yang akan digunakan untuk analisis lebih lanjut.
5. Proses Data Mining: Tahapan utama di mana algoritma diterapkan untuk menemukan pola atau tren.
6. Evaluasi Pola (Pattern Evaluation): Mengidentifikasi pola yang paling relevan atau memiliki nilai informasi tinggi.
7. Presentasi dan Visualisasi: Menyajikan hasil dalam bentuk grafik atau visual yang mudah dipahami.

Identifikasi variabel independen (insentif) dan variabel dependen (kinerja karyawan).

### 1. Analisis Regresi Linear:

Regresi linier adalah metode statistika yang digunakan untuk membentuk model atau hubungan antara satu atau lebih variabel bebas X dengan sebuah variabel respon Y. Analisis regresi dengan satu variabel bebas X disebut sebagai regresi linier sederhana, sedangkan jika terdapat lebih dari satu variabel bebas X, disebut sebagai regresi linier berganda (Kurniawan, 2008). [12] Dalam konteks ini, peneliti akan memberikan langkah-langkah lengkap untuk melakukan regresi linear sederhana.

Langkah-langkah Regresi Linear Sederhana:

a) Perumusan Masalah:

Tentukan pertanyaan atau hipotesis penelitian yang ingin dijawab melalui analisis regresi linear sederhana.

b) Pengumpulan Data:

Kumpulkan data untuk variabel independen (X) dan variabel dependen (Y) dari sumber yang relevan. Pastikan data yang dikumpulkan bersifat kuantitatif.

## c) Pemeriksaan Data:

Periksa data untuk memastikan kebersihan dan kualitasnya. Identifikasi dan tangani nilai-nilai yang hilang atau outliers jika diperlukan.

## d) Visualisasi Data:

Visualisasikan data menggunakan grafik, seperti scatter plot, untuk melihat pola dan hubungan awal antara variabel X dan Y.

## e) Perhitungan Statistik Deskriptif:

Hitung nilai rata-rata, median, deviasi standar, dan lainnya untuk variabel X dan Y. Ini memberikan wawasan awal tentang distribusi dan karakteristik data.

## f) Pembagian Data:

Pisahkan data menjadi dua kelompok: satu kelompok untuk melatih model (training set) dan satu kelompok untuk menguji model (testing set). Biasanya, sekitar 70-80% data digunakan untuk melatih model dan sisanya untuk menguji model.

## g) Perhitungan Koefisien Regresi:

Hitung koefisien regresi (b) menggunakan rumus:

$$b = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{\sum(X_i - \bar{X})^2}$$

$\bar{X}$  dan  $\bar{Y}$  adalah rata-rata dari X dan Y, dan  $\sum$  adalah symbol sigma yang menunjukkan penjumlahan.

h) Perhitungan Intercept:

Hitung nilai intercept ( $a$ ) menggunakan rumus

$$a = \bar{Y} - b\bar{X}$$

i) Membentuk Persamaan Regresi:

Bentuk persamaan regresi linear sederhana:

$$Y = a + bX$$

j) Evaluasi model:

Evaluasi model menggunakan metrik seperti koefisien determinasi ( $R^2$ ), Mean Squared Error (MSE), atau Root Mean Squared Error (RMSE). Metrik ini memberikan informasi tentang sejauh mana model cocok dengan data.

k) Visualisasi garis regresi:

Visualisasikan garis regresi pada scatter plot untuk melihat sejauh mana model sesuai dengan pola data.

l) Uji statistik Signifikansi:

Lakukan uji statistik, seperti uji hipotesis koefisien regresi, untuk menentukan apakah hubungan antara X dan Y signifikan secara statistik.

## 2. Pengelohana menggunakan Algoritma K-Means:

K-Means merupakan salah satu algoritma dalam data mining yang bisa digunakan untuk melakukan pengelompokan/clustering suatu data. [13] Pengelompokan data dengan metode k-means dilakukan dengan :

8. Menentukan jumlah kelompok
9. Alokasi data ke dalam kelompok secara acak
10. Hitung pusat kelompok (*centroid*/rata-rata) dari data yang ada di masing-masing kelompok. Lokasi centroid setiap kelompok diambil dari rata-rata (mean) semua nilai data pada setiap fiturnya. Jika  $M$  menyatakan jumlah data dalam sebuah kelompok,  $i$  menyatakan fitur ke- $i$  dalam sebuah kelompok, dan  $p$  menyatakan dimensi data, maka persamaan untuk menghitung centroid fitur ke- $i$  digunakan persamaan 1. [13]

$$C_i = \frac{1}{M} \sum_{j=1}^M x_j$$

11. Alokasikan masing-masing data ke centroid/rata-rata terdekat. Ada beberapa cara yang dapat dilakukan untuk mengukur jarak data ke pusat kelompok, diantaranya adalah Euclidean. [13] Pengukuran jarak pada ruang jarak (distance space) Euclidean dapat dicari menggunakan persamaan 2.

$$a_{i1} = \begin{cases} 1 & d = \min \{D(x_i, c_1)\} \\ 0 & \text{lainnya} \end{cases}$$

$a_{i1}$  adalah nilai keanggotaan titik  $x_i$  ke pusat kelompok  $c_1$ ,  $d$  adalah jarak terpendek dari data  $x_i$  ke  $K$  kelompok setelah dibandingkan, dan  $c_1$  adalah centroid (pusat kelompok) ke-1. Fungsi objektif yang digunakan untuk metode K-Means ditentukan berdasarkan jarak dan nilai keanggotaan data dalam kelompok. Fungsi objektif menurut MacQueen (1967) dapat ditentukan menggunakan persamaan 4.

$$J = \sum_{i=1}^n \sum_{l=1}^k \alpha_{il} D(x_i, c_l)^2$$

$n$  adalah jumlah data,  $k$  adalah jumlah kelompok,  $\alpha_{il}$  adalah nilai keanggotaan titik data  $x_i$  ke kelompok  $c_l$  yang diikuti.  $\alpha$  mempunyai nilai 0 atau 1. Apabila data merupakan anggota suatu kelompok, nilai  $\alpha_{il} = 1$ . Jika tidak, nilai  $\alpha_{il} = 0$ .

12. Kembali ke langkah 3, apabila masih ada data yang berpindah kelompok atau apabila ada perubahan nilai centroid di atas nilai ambang yang ditentukan, atau apabila perubahan nilai pada fungsi objektif yang digunakan masih di atas nilai ambang yang ditentukan.

### 1. Analisis Kluster K-Means:

Terapkan algoritme K-Means untuk mengelompokkan data ke dalam kluster. Pilih jumlah kluster yang optimal dan identifikasi pola kluster yang muncul. Analisis Kluster K-Means adalah metode pembelajaran mesin yang digunakan untuk mengelompokkan data ke dalam beberapa kelompok (kluster) berdasarkan kesamaan karakteristiknya. Metode ini termasuk dalam kategori unsupervised learning karena tidak memerlukan data label atau target untuk melakukan pengelompokan. Berikut adalah langkah-langkah utama dalam analisis kluster K-Means:

1. Menentukan Jumlah Kluster (K): Tentukan terlebih dahulu berapa banyak kluster (K) yang diinginkan. Jumlah ini bisa ditentukan secara subjektif, atau dengan metode seperti Elbow Method atau Silhouette Score untuk menentukan jumlah kluster optimal berdasarkan distribusi data.
2. Inisialisasi Centroid: Algoritma akan memilih secara acak titik-titik awal (centroid) untuk setiap kluster.

3. Menghitung Jarak dan Mengelompokkan Data: Setiap titik data akan dihitung jaraknya dari setiap centroid, dan data tersebut akan dimasukkan ke kluster dengan centroid terdekat.
4. Memperbarui Centroid: Setelah semua data dikelompokkan, centroid setiap kluster akan diperbarui dengan menghitung rata-rata posisi titik-titik data yang berada dalam kluster tersebut.
5. Iterasi Sampai Konvergen: Langkah 3 dan 4 akan diulang sampai tidak ada perubahan signifikan pada posisi centroid, atau sampai data tidak berpindah kluster lagi. Pada tahap ini, proses dianggap sudah konvergen.
6. Evaluasi Kluster: Setelah kluster terbentuk, hasilnya dapat dievaluasi menggunakan metrik seperti inerti, silhouette score, atau visualisasi jika dimungkinkan.

### **1.3.3. Integrasi Hasil**

Integrasi hasil dalam konteks data mining adalah tahap di mana hasil analisis atau pola yang ditemukan dari proses data mining digabungkan dan diterjemahkan ke dalam format yang dapat dipahami, diinterpretasikan, dan digunakan oleh pihak yang berkepentingan. Integrasi hasil ini penting untuk memastikan bahwa informasi yang diperoleh dapat diterapkan secara efektif dalam pengambilan keputusan atau tindakan. Gabungkan hasil dari analisis regresi linear dan K-Means. Analisis ini dapat memberikan pemahaman yang lebih holistik tentang bagaimana disiplin dan motivasi kerja (insentif) mempengaruhi kinerja karyawan dan bagaimana karyawan dapat dikelompokkan

berdasarkan pola perilaku mereka.