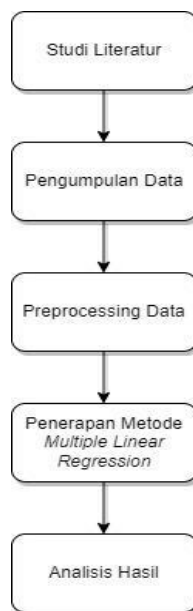


## BAB III

### METODOLOGI PENELITIAN

Metodologi penelitian tentunya berperan penting sebagai kerangka dan panduan proses penelitian. Metode analisis yang akan digunakan adalah metode *data mining* yaitu prediksi dengan menggunakan *tools software orange*. Tahapan-tahapan yang akan dilakukan sebagai berikut :



Gambar 3.1 Tahapan Penelitian

#### 3.1 Studi Literatur

Langkah pertama yang dilakukan adalah studi literatur dimana peneliti mencari referensi atau dokumen pendukung seperti buku, artikel, jurnal, dan sumber lainnya.

### **3.2 Pengumpulan Data**

Pengumpulan data yang dilakukan pada penulisan ini merupakan data sekunder yang berasal dari Dinas Pertanian Kabupaten Malaka, yang terdiri atas data padi, data jagung dan data kacang hijau dari masing-masing Kecamatan yang ada di Kabupaten Malaka. Data tersebut meliputi Kecamatan, luas tanam, luas panen, produksi dan produktivitas.

### **3.3 Pre-processing**

*Pre-processing* merupakan langkah awal yang dilakukan dalam pengolahan data untuk membantu metode yang digunakan agar dapat berjalan dengan baik dan nantinya pada tahap *pre-processing*, dilakukan proses *cleaning* data yang digunakan untuk menghilangkan informasi yang tidak diperlukan dalam proses *pre-processing*, yaitu menghilangkan atribut tertentu (Herwanto *et al.*, 2019).

### **3.4 Penerapan Algoritma Linear Regression**

Dalam penelitian ini, digunakan *multiple linear regression* karena terdapat lebih dari satu buah variabel independen. Data penelitian terdiri atas satu variabel dependen (Y) dan beberapa variabel independen (X). Variabel dependen tersebut adalah atribut produksi atau hasil panen, sedangkan variabel independen adalah atribut luas tanam, luas panen, produktivitas dan curah hujan (Herwanto *et al.*, 2019).

Regresi linier berganda adalah analisis regresi yang menjelaskan hubungan antara peubah respon (variabel dependen) dengan faktor-faktor yang

mempengaruhi lebih dari satu prediktor (variabel independen). Ketika suatu hasil/keluaran, atau kelas berupa numerik, dan semua atribut adalah numerik, regresi linear adalah teknik yang tepat untuk menyelesaikan (Triyanto *et al.*, 2019). Gunanya adalah untuk mengekspresikan kelas sebagai kombinasi linear dari atribut, dengan bobot yang telah ditentukan, dengan rumus sebagai berikut:

$$y = a + b_1X_1 + b_2X_2 + \dots + b_nX_n \dots\dots\dots(1)$$

Di mana

Y = Variabel tidak bebas (nilai yang diprediksikan)

X = Variabel bebas

a = Konstanta

b = Koefisien

Dimana nilai a, b<sub>1</sub>, b<sub>2</sub>... b<sub>n</sub> dapat dihitung dengan metode kuadrat terkecil yaitu:

$$b_1 = \frac{(\sum x_2^2)(\sum x_1 y) - (\sum x_2 y)(\sum x_1 x_2)}{(\sum x_1^2)(\sum x_2^2) - (\sum x_1 x_2)^2}$$

$$b_2 = \frac{(\sum x_1^2)(\sum x_2 y) - (\sum x_1 y)(\sum x_1 x_2)}{(\sum x_1^2)(\sum x_2^2) - (\sum x_1 x_2)^2} \dots\dots\dots(2)$$

$$a = \frac{\sum Y - (b_1 \sum x_1) - (b_2 \sum x_2)}{n}$$

Dimana:

$$\begin{aligned} \sum x_1^2 &= \sum X_1^2 - \frac{(\sum X_1)^2}{n} \\ \sum x_2^2 &= \sum X_2^2 - \frac{(\sum X_2)^2}{n} \\ \sum y^2 &= \sum Y^2 - \frac{(\sum Y)^2}{n} \quad \dots \dots \dots (3) \\ \sum x_1 y &= \sum X_1 Y - \frac{\sum X_1 \sum Y}{n} \\ \sum x_2 y &= \sum X_2 Y - \frac{\sum X_2 \sum Y}{n} \\ \sum x_1 x_2 &= \sum X_1 X_2 - \frac{\sum X_1 \sum X_2}{n} \end{aligned}$$

Ketika variabel bebas lebih dari 2, nilai konstanta dan koefisien regresi setiap variabel bebas dapat diperoleh dengan menggunakan matriks determinan. Contohnya adalah ketika terdapat 3 persamaan dengan 3 variabel yang tidak diketahui nilainya, yaitu a, b1, b2 dan b3, persamaan tersebut dapat dinyatakan dalam persamaan matriks sebagai berikut:

$$\begin{aligned} A &= N \quad \sum X_1 \quad \sum X_2 \quad \sum X_3 \\ &\quad \sum X_1 \quad \sum(X_1 X_1) \quad \sum(X_1 X_2) \quad \sum(X_1 X_3) \\ &\quad \sum X_2 \quad \sum(X_2 X_1) \quad \sum(X_2 X_2) \quad \sum(X_2 X_3) \\ &\quad \sum X_3 \quad \sum(X_3 X_1) \quad \sum(X_3 X_2) \quad \sum(X_3 X_3) \\ A_0 &= \sum Y \quad \sum X_1 \quad \sum X_2 \quad \sum X_3 \\ &\quad \sum(X_1 Y) \quad \sum(X_1 X_1) \quad \sum(X_1 X_2) \quad \sum(X_1 X_3) \\ &\quad \sum(X_2 Y) \quad \sum(X_2 X_1) \quad \sum(X_2 X_2) \quad \sum(X_2 X_3) \\ &\quad \sum(X_3 Y) \quad \sum(X_3 X_1) \quad \sum(X_3 X_2) \quad \sum(X_3 X_3) \end{aligned}$$

$$\begin{aligned}
A1 &= N \quad \Sigma Y \quad \Sigma X_2 \quad \Sigma X_3 \quad \dots \dots \dots (4) \\
&\quad \Sigma X_1 \quad \Sigma(X_1 Y) \quad \Sigma(X_1 X_2) \quad \Sigma(X_1 X_3) \\
&\quad \Sigma X_2 \quad \Sigma(X_2 Y) \quad \Sigma(X_2 X_2) \quad \Sigma(X_2 X_3) \\
&\quad \Sigma X_3 \quad \Sigma(X_3 Y) \quad \Sigma(X_3 X_2) \quad \Sigma(X_3 X_3) \\
A2 &= N \quad \Sigma X_1 \quad \Sigma Y \quad \Sigma X_3 \\
&\quad \Sigma X_1 \quad \Sigma(X_1 X_1) \quad \Sigma(X_1 Y) \quad \Sigma(X_1 X_3) \\
&\quad \Sigma X_2 \quad \Sigma(X_2 X_1) \quad \Sigma(X_2 Y) \quad \Sigma(X_2 X_3) \\
&\quad \Sigma X_3 \quad \Sigma(X_3 X_1) \quad \Sigma(X_3 Y) \quad \Sigma(X_3 X_3) \\
A3 &= N \quad \Sigma X_1 \quad \Sigma X_2 \quad \Sigma Y \\
&\quad \Sigma X_1 \quad \Sigma(X_1 X_1) \quad \Sigma(X_1 X_2) \quad \Sigma(X_1 Y) \\
&\quad \Sigma X_2 \quad \Sigma(X_2 X_1) \quad \Sigma(X_2 X_2) \quad \Sigma(X_2 Y) \\
&\quad \Sigma X_3 \quad \Sigma(X_3 X_1) \quad \Sigma(X_3 X_2) \quad \Sigma(X_3 Y)
\end{aligned}$$

Setelah dilakukan perhitungan menemukan matriks determinan dari masing-masing matriks A, A1, A2 dan A3, maka selanjutnya dilakukan perhitungan determinan menggunakan rumus sebagai berikut :

$$\begin{aligned}
a &= \frac{Det(A0)}{Det(A)} \\
b_1 &= \frac{Det(A1)}{Det(A)} \quad \dots \dots \dots (5) \\
b_2 &= \frac{Det(A2)}{Det(A)} \\
b_3 &= \frac{Det(A3)}{Det(A)}
\end{aligned}$$

### **3.5 Analisis Hasil**

Dalam penelitian ini, dilakukan prediksi terhadap hasil pertanian tanaman pangan di Kecamatan Malaka Tengah, Kabupaten Malaka. Mayoritas penduduk di kecamatan ini menggantungkan mata pencahariannya sebagai petani. Metode yang digunakan untuk prediksi adalah *multiple linear regression*, yang bertujuan menghasilkan estimasi jumlah hasil panen tanaman pangan. Hasil prediksi ini memiliki peran sebagai evaluasi bagi pemerintah daerah. Dengan memiliki perkiraan jumlah hasil panen tanaman pangan untuk periode tertentu di masa depan, pemerintah daerah dapat memahami situasi yang mungkin terjadi. Informasi ini sangat berharga untuk pengambilan keputusan yang berorientasi pada meminimalkan risiko dan mengoptimalkan potensi keuntungan, yang pada gilirannya dapat mendukung pertumbuhan ekonomi daerah. Selain itu, pemerintah daerah dapat menentukan prioritas dalam peningkatan produksi pangan dengan fokus pada jenis tanaman tertentu.