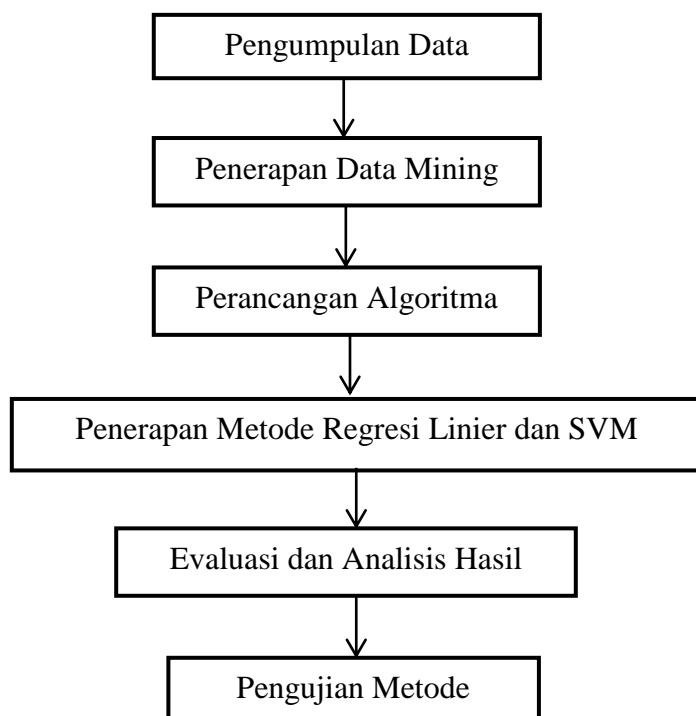


BAB III

ANALISA

3.1 Desain Penelitian

Desain Penelitian dalam penelitian ini dirancang untuk memastikan setiap tahapan analisis dan perancangan berjalan secara terstruktur dan sistematis.



Gambar 3.1 Desain Penelitian

Tahap-tahap sebagai berikut:

1. Pengumpulan Data

Tahap pertama dalam arsitektur sistem adalah pengumpulan data. Pada tahap ini, data yang relevan dan sesuai dengan tujuan penelitian, yaitu untuk meningkatkan keuntungan Amanda Brownies, dipilih dari berbagai sumber yang tersedia. Data yang dikumpulkan mencakup variabel-variabel

seperti harga jual, jumlah produk terjual, promosi, dan faktor lainnya yang dapat memengaruhi keuntungan. Data ini akan menjadi bahan dasar untuk proses analisis lebih lanjut.

2. Penerapan Data Mining

Setelah data terkumpul, tahap selanjutnya adalah penerapan teknik data mining untuk menggali pola atau informasi yang berguna dari data. Pada tahap ini, berbagai algoritma digunakan untuk menganalisis data, seperti klasifikasi, regresi, clustering, atau asosiasi. Proses ini bertujuan untuk menemukan hubungan tersembunyi dalam data yang mungkin tidak teridentifikasi dengan analisis tradisional.

3. Perancangan Model dan Klasifikasi.

Setelah data diproses, tahap selanjutnya adalah perancangan model dan klasifikasi. Pada tahap ini, dua metode utama yang digunakan untuk membangun model klasifikasi adalah:

- a. *Regresi Linier* Digunakan untuk menganalisis hubungan linier antara variabel-variabel independen (seperti jumlah produk terjual dan total pendapatan) dan variabel dependen (keuntungan).
- b. *Support Vector Machine (SVM)* Digunakan untuk melakukan klasifikasi yang lebih kompleks dan prediksi berdasarkan pola-pola dalam data.

Model-model ini kemudian diuji menggunakan dataset yang dibagi menjadi Data *Training* dan Data *Testing*.

- a. Data *Training* digunakan untuk melatih model agar bisa memprediksi keuntungan atau klasifikasi berdasarkan variabel yang ada.
- b. Data *Testing* digunakan untuk menguji kemampuan model dalam memprediksi atau mengklasifikasikan data yang belum pernah dilatih sebelumnya.

4. Evaluasi Model

Setelah kedua model dilatih menggunakan data training, tahap berikutnya adalah evaluasi model. Evaluasi dilakukan untuk mengukur efektivitas dan akurasi dari kedua metode yang digunakan. Beberapa metrik evaluasi yang digunakan antara lain:

- a. *Mean Squared Error* (MSE) Mengukur seberapa jauh prediksi model dari nilai yang sebenarnya. Semakin rendah MSE, semakin baik performa model.
- b. *Root Mean Squared Error* (RMSE) Mengukur akar kuadrat dari MSE, memberikan gambaran yang lebih intuitif mengenai seberapa besar kesalahan prediksi model.

Hasil evaluasi ini memberikan gambaran yang jelas mengenai keunggulan dan kelemahan dari setiap metode yang digunakan, serta membantu dalam memilih model yang paling efektif untuk meningkatkan keuntungan Amanda Brownies.

5. Pengambilan Keputusan Strategis

Dengan menggunakan arsitektur sistem yang sistematis ini, proses analisis data dapat dilakukan secara optimal. Hasil evaluasi dan perbandingan antara regresi linier dan SVM akan memberikan dasar yang kuat bagi pengambilan keputusan strategis. Keputusan ini bisa mencakup rekomendasi terkait penentuan harga jual, atau strategi promosi yang lebih efektif untuk meningkatkan keuntungan Amanda Brownies.

3.1.1 Lokasi dan Waktu Penelitian

Penelitian ini dilaksanakan di Amanda Brownies Cabang Rantauprapat yang berlokasi di Jl. Sirandorung, Rantau Prapat, Kabupaten Labuhanbatu, Sumatera Utara, dengan kode pos 21414. Penelitian dijadwalkan untuk dilaksanakan pada bulan Desember 2024, dengan fokus utama pada analisis data penjualan dan faktor-faktor yang memengaruhi keuntungan.

3.1.2 Populasi dan Sampel

Populasi dalam penelitian ini mencakup seluruh data penjualan Amanda Brownies selama periode tiga bulan terakhir, sedangkan sampel, yang dipilih menggunakan metode *purposive sampling*, terdiri dari data penjualan harian selama September hingga November 2024 untuk mewakili keseluruhan populasi dan dianalisis lebih lanjut.

3.2 Tahapan Proses KDD

Pada penelitian ini, pendekatan *Knowledge Discovery in Databases* (KDD) digunakan untuk mendukung tujuan meningkatkan keuntungan Amanda

Brownies melalui perbandingan metode Regresi Linier dan *Support Vector Machine* (SVM). KDD merupakan proses sistematis untuk mengekstraksi informasi yang berguna dari data besar, yang terdiri dari beberapa tahapan.

3.2.1 Selection (Pemilihan Data)

Tahap ini berfokus pada pemilihan data yang relevan dan sesuai dengan kebutuhan penelitian. Data yang dipilih diambil dari sumber yang terpercaya untuk memastikan keakuratan serta relevansi informasi yang akan digunakan. Dalam penelitian ini, data yang dianalisis berasal dari laporan penjualan Amanda Brownies Rantauprapat selama periode September hingga November 2024. Data penjualan dari bulan September hingga November memberikan gambaran mengenai pola penjualan yang bervariasi, yang nantinya akan dianalisis untuk memahami faktor-faktor yang mempengaruhi keuntungan dan mencari solusi untuk meningkatkan profitabilitas Amanda Brownies. Adapun data yang dipilih meliputi beberapa aspek utama, yaitu:

1. Tanggal Penjualan

Informasi mengenai tanggal transaksi penjualan yang dicatat secara harian. Data ini penting untuk analisis tren penjualan dalam periode waktu tertentu tertentu yang menunjukkan peningkatan aktivitas penjualan.

2. Jumlah Produk Terjual

Data mengenai jumlah produk yang terjual setiap harinya, yang dipecah berdasarkan variasi produk seperti ORI (Original), SP (Srikaya pandan) dan varian lainnya yang ditawarkan oleh Amanda Brownies. Informasi ini akan digunakan untuk mengevaluasi performa masing-masing varian.

3. Harga per Unit

Informasi harga jual setiap unit produk berdasarkan varian.data ini digunakan untuk menghitung total pendapatan penjualan harian.

4. Keuntungan Harian

Data ini dihitung berdasarkan total pendapatan penjualan harian dikurangi dengan biaya operasional harian, seperti biaya tenaga kerja, fasilitas, dan lain-lain. Informasi ini akan memberikan gambaran mengenai profitabilitas harian cabang Amanda Brownies selama periode penelitian.

Berikut adalah data dari bulan September hingga November 2024:

Tabel 3.1 Penjualan dan Pendapatan Bulan September 2024

| TANGGAL | NAMA PRODUK | | | | | | | | JUMLAH | PENDAPATAN |
|--------------|-------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|--------|---------------|
| | ORI | SP | TIRA | GANCHE | BLU | STAW | GRENN | T.M | | |
| 01-Sep-24 | 22 | 11 | 0 | 1 | 0 | 0 | 1 | 1 | 36 | Rp 2.579.000 |
| 02-Sep-24 | 30 | 10 | 0 | 1 | 0 | 0 | 1 | 3 | 45 | Rp 3.108.000 |
| 03-Sep-24 | 32 | 11 | 0 | 1 | 1 | 1 | 0 | 3 | 49 | Rp 3.287.000 |
| 04-Sep-24 | 20 | 11 | 1 | 0 | 0 | 0 | 1 | 2 | 35 | Rp 2.389.000 |
| 05-Sep-24 | 35 | 10 | 0 | 0 | 0 | 1 | 0 | 4 | 50 | Rp 3.371.000 |
| | | | ... | | | | | | | |
| | | | ... | | | | | | | |
| | | | ... | | | | | | | |
| 26-Sep-24 | 35 | 15 | 0 | 1 | 1 | 0 | 0 | 4 | 56 | Rp 3.781.000 |
| 27-Sep-24 | 29 | 12 | 1 | 0 | 0 | 1 | 1 | 4 | 48 | Rp 3.263.000 |
| 28-Sep-24 | 26 | 15 | 0 | 1 | 1 | 0 | 1 | 5 | 49 | Rp 3.413.000 |
| 29-Sep-24 | 24 | 12 | 0 | 1 | 1 | 1 | 1 | 5 | 45 | Rp 3.123.000 |
| 30-Sep-24 | 30 | 13 | 1 | 0 | 0 | 1 | 1 | 4 | 50 | Rp 3.493.000 |
| Jumlah | 818 | 345 | 9 | 13 | 14 | 17 | 17 | 113 | 1471 | Rp 93.907.000 |
| Harga Produk | Rp 55.000 | Rp 61.000 | Rp 72.000 | Rp 80.000 | Rp 66.000 | Rp 61.000 | Rp 61.000 | Rp 63.000 | - | - |

Sumber : Data penjualan Amanda Brownies

Data diambil dari data penjualan Amanda Brownies Rantauprapat bulan September 2024 yang totalnya berjumlah 1.612 penjualan dan total pendapatan berjumlah Rp. 93.907.000 per-bulannya.

Tabel 3.2 Penjualan dan Pendapatan Bulan Oktober 2024

| TANGGAL | NAMA PRODUK | | | | | | | | | |
|--------------|-------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|--------|---------------|
| | ORI | SP | TIRA | GANCHE | BLU | STAW | GRE | T.M | JUMLAH | PENDAPATAN |
| 01-Okt-24 | 20 | 8 | 0 | 1 | 1 | 0 | 1 | 3 | 34 | Rp 2.631.000 |
| 02-Okt-24 | 30 | 11 | 0 | 1 | 0 | 1 | 1 | 1 | 45 | Rp 3.132.000 |
| 03-Okt-24 | 35 | 14 | 0 | 1 | 0 | 2 | 1 | 4 | 57 | Rp 4.064.000 |
| 04-Okt-24 | 27 | 12 | 1 | 0 | 0 | 0 | 1 | 3 | 44 | Rp 3.134.000 |
| 05-Okt-24 | 25 | 12 | 0 | 1 | 1 | 0 | 2 | 3 | 44 | Rp 3.021.000 |
| | | | ... | | | | | | | |
| | | | ... | | | | | | | |
| | | | ... | | | | | | | |
| 26-Okt-24 | 29 | 11 | 0 | 1 | 1 | 0 | 1 | 3 | 46 | Rp 3.145.000 |
| 27-Okt-24 | 18 | 7 | 1 | 0 | 0 | 1 | 1 | 3 | 31 | Rp 2.274.000 |
| 28-Okt-24 | 15 | 4 | 0 | 1 | 1 | 0 | 0 | 3 | 24 | Rp 1.887.000 |
| 29-Okt-24 | 25 | 8 | 1 | 1 | 0 | 1 | 0 | 3 | 39 | Rp 2.799.000 |
| 30-Okt-24 | 22 | 7 | 1 | 0 | 1 | 0 | 0 | 3 | 34 | Rp 2.452.000 |
| 31-Okt-24 | 10 | 5 | 2 | 0 | 1 | 0 | 4 | 0 | 22 | Rp 1.431.000 |
| Jumlah | 729 | 324 | 14 | 18 | 16 | 14 | 20 | 87 | 1471 | Rp 85.801.000 |
| Harga Produk | Rp 55.000 | Rp 61.000 | Rp 72.000 | Rp 80.000 | Rp 66.000 | Rp 61.000 | Rp 61.000 | Rp 63.000 | - | - |

Sumber : Data penjualan Amanda Brownies

Data diambil dari data penjualan Amanda Brownies Rantauprapat bulan Oktober 2024 yang totalnya berjumlah 1.471 penjualan dan total pendapatan berjumlah Rp. 85.801.000 per-bulannya.

Tabel 3.3 Penjualan dan Pendapatan Bulan November 2024

| TANGGAL | NAMA PRODUK | | | | | | | | | |
|--------------|-------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|--------|---------------|
| | ORI | SP | TIRA | GANCHE | BLU | STAW | GRE | T.M | JUMLAH | PENDAPATAN |
| 01-Nop-24 | 29 | 10 | 0 | 2 | 2 | 0 | 2 | 10 | 55 | Rp 4.309.000 |
| 02-Nop-24 | 20 | 11 | 0 | 1 | 0 | 1 | 0 | 7 | 40 | Rp 2.981.000 |
| 03-Nop-24 | 27 | 12 | 1 | 0 | 0 | 1 | 2 | 6 | 49 | Rp 3.384.000 |
| 04-Nop-24 | 16 | 14 | 0 | 1 | 0 | 1 | 1 | 3 | 36 | Rp 2.475.000 |
| 05-Nop-24 | 34 | 14 | 0 | 1 | 0 | 1 | 2 | 5 | 57 | Rp 3.990.000 |
| | | | ... | | | | | | | |
| | | | ... | | | | | | | |
| | | | ... | | | | | | | |
| 26-Nop-24 | 18 | 7 | 0 | 0 | 0 | 0 | 0 | 5 | 30 | Rp 2.283.000 |
| 27-Nop-24 | 20 | 9 | 0 | 0 | 0 | 1 | 0 | 4 | 34 | Rp 2.513.000 |
| 28-Nop-24 | 16 | 6 | 0 | 0 | 0 | 0 | 0 | 3 | 25 | Rp 1.986.000 |
| 29-Nop-24 | 24 | 7 | 0 | 0 | 1 | 1 | 0 | 4 | 37 | Rp 2.677.000 |
| 30-Nop-24 | 18 | 5 | 0 | 0 | 0 | 0 | 0 | 3 | 26 | Rp 2.035.000 |
| JUMLAH | 695 | 304 | 4 | 13 | 8 | 16 | 13 | 138 | 1.459 | Rp 86.882.000 |
| Harga Produk | Rp 55.000 | Rp 61.000 | Rp 72.000 | Rp 80.000 | Rp 66.000 | Rp 61.000 | Rp 61.000 | Rp 63.000 | - | - |

Sumber : Data penjualan Amanda Brownies

Data diambil dari data penjualan Amanda Brownies Rantauprapat bulan November 2024 yang totalnya berjumlah 1.459 penjualan dan total pendapatan

berjumlah Rp. 86.882.000 per-bulannya. Jadi total seluruh data yang dipakai dalam penelitian ini adalah 4.542 penjualan.

3.2.2 Data Preprocessing

Data di bawah ini telah melalui tahap *preprocessing*, yaitu pembersihan data dengan menghapus duplikasi, menangani *missing values*, dan mengidentifikasi outlier menggunakan metode IQR, normalisasi data dengan menyesuaikan skala ke rentang 0-1 menggunakan metode *Min-Max Normalization*, transformasi data dengan mengubah kolom tanggal menjadi nilai numerik berupa jumlah hari dari awal periode data, serta *feature selection* dengan memilih fitur yang paling relevan, yaitu Tanggal_Num, Total Penjualan, dan Keuntungan.

Tabel 3.4 Data Preprocessing

| Tanggal_Num | Jumlah | Pendapatan |
|-------------|--------|------------|
| 0 | 44 | 2579000 |
| 1 | 54 | 3108000 |
| 2 | 57 | 3287000 |
| 3 | 41 | 2389000 |
| 4 | 59 | 3371000 |
| 5 | 46 | 2659000 |
| 6 | 54 | 3093000 |
| 7 | 53 | 3138000 |
| 8 | 32 | 1859000 |
| 9 | 58 | 3331000 |
| 10 | 60 | 3515000 |
| 11 | 54 | 3118000 |
| 12 | 57 | 3300000 |

| | | |
|----|----|---------|
| 13 | 34 | 1972000 |
| 14 | 67 | 3926000 |
| 15 | 67 | 3875000 |
| 16 | 55 | 3222000 |
| 17 | 66 | 3822000 |
| 18 | 46 | 2696000 |
| 19 | 43 | 2541000 |
| 20 | 49 | 2874000 |
| 21 | 66 | 3870000 |
| 22 | 49 | 2895000 |
| 23 | 56 | 3294000 |
| 24 | 53 | 3100000 |
| 25 | 65 | 3781000 |
| 26 | 56 | 3263000 |
| 27 | 58 | 3413000 |
| 28 | 53 | 3123000 |
| 29 | 60 | 3493000 |

Tahap-tahapan di jelaskan sebagai berikut:

1. Menghapus Duplikasi (*Remove Duplicates*)

Tujuan Menghindari data berulang yang dapat menyebabkan bias atau perhitungan yang salah dalam analisis. Prosesnya Data diperiksa untuk baris-baris yang memiliki nilai identik di semua kolom.

2. Penanganan *Missing Values*

Memastikan data tidak memiliki nilai kosong (*missing*) yang dapat memengaruhi analisis atau model prediktif. Prosesnya data diperiksa untuk kolom atau baris dengan nilai kosong (NaN). *Missing values* diisi menggunakan metode median (nilai tengah dari data kolom yang

bersangkutan) karena median lebih tahan terhadap outlier dibandingkan rata-rata. Hasilnya dataset kini lengkap tanpa nilai kosong.

3. Normalisasi Data (*Min-Max Normalization*)

Tujuannya Mengubah skala data menjadi rentang 0-1 untuk memastikan semua fitur memiliki bobot yang setara dalam analisis. Prosesnya :

Formula normalisasi Min-Max :

$$X_{min} = \frac{X - X_{min}}{X_{max} - X_{min}}$$

Proses ini diterapkan pada kolom Jumlah dan Pendapatan. Contoh : Jika nilai minimum dan maksimum pada kolom JUMLAH adalah 32 dan 67, maka nilai 44 dinormalisasi menjadi.

$$JUMLAH_norm = \frac{44-32}{67-32} = 0.34$$

Hasilnya Kolom baru Jumlah_Norm dan Pendapatan_Norm ditambahkan ke dataset, dengan nilai dalam rentang 0-1.

4. Transformasi Kolom Tanggal

Mengubah format tanggal menjadi angka numerik agar lebih mudah diolah dalam analisis. Prosesnya :

Kolom tanggal dikonversi menjadi tipe *datetime*. Jumlah hari dari tanggal awal (01-Sep-24) dihitung untuk setiap baris. Contohnya Jika tanggal adalah 03-Sep-24, maka nilai numeriknya adalah 2 (karena 3 hari dari awal periode data). Hasilnya Kolom baru Tanggal_Num berisi jumlah hari dari awal periode data.

3.2.3 Data Transformasi

Pada tahap Transformasi Data, data yang telah dikumpulkan diubah menjadi format yang siap untuk dimodelkan. Proses ini bertujuan untuk memastikan bahwa data dapat digunakan secara efektif dalam proses analisis, baik untuk regresi linier maupun metode klasifikasi seperti *Support Vector Machine* (SVM).

Tabel 3.5 Data Transformasi

| Tanggal | X1: Jumlah Terjual | X2: Total Pendapatan | Y: Keuntungan Harian | X1 (Norm) | X2 (Norm) |
|----------------|-----------------------------------|-------------------------------------|-------------------------------------|----------------------|------------------|
| 01-Sep-24 | 44 | 2.579.000 | Rp310.000 | 0,34 | 0,35 |
| 02-Sep-24 | 54 | 3.108.000 | Rp460.000 | 0,62 | 0,60 |
| 03-Sep-24 | 57 | 3.287.000 | Rp505.000 | 0,71 | 0,69 |
| 04-Sep-24 | 41 | 2.389.000 | Rp265.000 | 0,25 | 0,25 |
| 05-Sep-24 | 59 | 3.371.000 | Rp535.000 | 0,77 | 0,73 |
| 06-Sep-24 | 46 | 2.659.000 | Rp340.000 | 0,4 | 0,38 |
| 07-Sep-24 | 54 | 3.093.000 | Rp460.000 | 0,62 | 0,59 |
| 08-Sep-24 | 53 | 3.138.000 | Rp445.000 | 0,6 | 0,61 |
| 09-Sep-24 | 32 | 1.859.000 | Rp130.000 | 0 | 0 |
| 10-Sep-24 | 58 | 3.331.000 | Rp520.000 | 0,74 | 0,71 |
| 11-Sep-24 | 60 | 3.515.000 | Rp550.000 | 0,8 | 0,80 |
| 12-Sep-24 | 54 | 3.118.000 | Rp460.000 | 0,62 | 0,60 |
| 13-Sep-24 | 57 | 3.300.000 | Rp505.000 | 0,71 | 0,70 |

| | | | | | |
|-----------|----|-----------|-----------|-------|-------|
| 14-Sep-24 | 34 | 1.972.000 | Rp160.000 | 0,057 | 0,054 |
| 15-Sep-24 | 67 | 3.926.000 | Rp655.000 | 1 | 1 |
| 16-Sep-24 | 67 | 3.875.000 | Rp655.000 | 1 | 0,97 |
| 17-Sep-24 | 55 | 3.222.000 | Rp475.000 | 0,65 | 0,66 |
| 18-Sep-24 | 66 | 3.822.000 | Rp640.000 | 0,97 | 0,94 |
| 19-Sep-24 | 46 | 2.696.000 | Rp340.000 | 0,4 | 0,40 |
| 20-Sep-24 | 43 | 2.541.000 | Rp295.000 | 0,31 | 0,33 |
| 21-Sep-24 | 49 | 2.874.000 | Rp385.000 | 0,48 | 0,49 |
| 22-Sep-24 | 66 | 3.870.000 | Rp640.000 | 0,97 | 0,97 |
| 23-Sep-24 | 49 | 2.895.000 | Rp385.000 | 0,48 | 0,50 |
| 24-Sep-24 | 56 | 3.294.000 | Rp490.000 | 0,68 | 0,69 |
| 25-Sep-24 | 53 | 3.100.000 | Rp445.000 | 0,6 | 0,60 |
| 26-Sep-24 | 65 | 3.781.000 | Rp625.000 | 0,94 | 0,73 |
| 27-Sep-24 | 56 | 3.263.000 | Rp490.000 | 0,68 | 0,68 |
| 28-Sep-24 | 58 | 3.413.000 | Rp520.000 | 0,74 | 0,75 |
| 29-Sep-24 | 53 | 3.123.000 | Rp445.000 | 0,6 | 0,61 |
| 30-Sep-24 | 60 | 3.493.000 | Rp550.000 | 0,8 | 0,79 |

Transformasi kolom tanggal menjadi nilai numerik dilakukan untuk memudahkan model regresi dan SVM dalam menganalisis hubungan antar variabel, mengingat banyak model analitik yang lebih efektif saat bekerja dengan data numerik. Berikut adalah langkah-langkah transformasi yang dilakukan untuk variabel yang digunakan dalam penelitian ini.

1. Variabel

Variabel adalah faktor-faktor yang digunakan untuk memprediksi dalam penelitian ini, variabel yang digunakan sebagai berikut :

X₁ = Jumlah Produk Terjual

X₂ = Total Pendapatan.

(Y) = Variabel Dependen

Penerapan:

X₁ = Jumlah Terjual

Min(X₁) = 32 (terendah pada 09-Sep-24)

Max(X₁) = 67 (tertinggi pada 15-Sep-24 dan 16-Sep-24)

X₂ = Total Pendapatan

Min(X₂) = Rp1.859.000 (terendah pada 09-Sep-24)

Max(X₂) = Rp3.926.000 (tertinggi pada 15-Sep-24)

2. Normalisasi untuk Setiap Data

Hitung normalisasi untuk beberapa baris pertama dengan rumus yang sudah dijelaskan:

X₁ = (Jumlah Terjual) pada 01-Sep-24

$$X_{min} = \frac{X - X_{min}}{X_{max} - X_{min}}$$

$$X_1 \text{ norm} = \frac{44-32}{67-32} = \frac{12}{35} = 0.34$$

X_2 = (Total Pendapatan) pada 01-Sep-24:

$$X_1 \text{ norm} = \frac{2.579.000 - 1.859.000}{3.926.000 - 1.859.000} = \frac{720.000}{2.067.000} = 0.35$$

3. Transformasi Data

Pengkodean atau normalisasi data jika diperlukan, variabel kategorikal (misalnya, jenis produk) akan dikodekan menjadi numerik. Selain itu, normalisasi atau standarisasi dilakukan untuk memastikan bahwa skala variabel independen seperti jumlah produk terjual (X_1), dan total pendapatan (X_2) berada dalam rentang yang sama, sehingga tidak ada variabel yang mendominasi analisis. Penghitung Keuntungan Harian (Y):

Jika data mengenai keuntungan harian tidak tersedia langsung, maka dihitung terlebih dahulu dengan rumus:

$$Y = \text{Total Pendapatan} - \text{Biaya Operasional}$$

Setelah perhitungan ini dilakukan, data untuk variabel dependen (Y) menjadi siap digunakan dalam model regresi atau SVM.

3.2.4 Data Mining

Pada tahap Data Mining, proses analisis dilakukan menggunakan dua teknik utama yaitu Regresi Linier dan *Support Vector Machine* (SVM). Teknik-teknik ini digunakan untuk memprediksi keuntungan harian berdasarkan jumlah produk terjual dan total pendapatan. Regresi Linier dipilih karena kesederhanaannya dalam menggambarkan hubungan linier antara variabel independen dan dependen, sementara SVM dipilih untuk menangani hubungan yang lebih kompleks dan non-linier dalam data. Kedua metode ini memungkinkan

perbandingan efektivitas dalam meningkatkan prediksi keuntungan. Proses data mining bertujuan untuk menemukan pola dan hubungan yang signifikan di dalam data yang telah dipersiapkan pada tahap sebelumnya.

1. Regresi Linier

Regresi linier adalah metode statistik yang digunakan untuk memodelkan hubungan antara variabel dependen (keuntungan) dengan satu atau lebih variabel independen (jumlah produk terjual dan total pendapatan). Dalam penelitian ini, regresi linier digunakan untuk menghitung hubungan antara jumlah produk terjual (X_1) dan total pendapatan (X_2) terhadap keuntungan harian (Y).

Model Regresi Linier Sederhana :

$$y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \epsilon$$

Di mana:

y = Keuntungan (variabel dependen).

X_1 = Jumlah produk terjual.

X_2 = Total pendapatan.

β_0 = Intercept (nilai konstanta).

β_1, β_2 = Koefisien regresi.

ϵ (epsilo) = Error/residual.

2. *Support Vector Machine* (SVM)

SVM adalah metode klasifikasi dan regresi yang kuat. Dalam penelitian ini, SVM digunakan untuk melakukan regresi pada data dengan tujuan

untuk memprediksi keuntungan harian berdasarkan jumlah produk terjual dan total pendapatan. SVM mencoba menemukan *hyperplane* optimal yang memisahkan data ke dalam dua kelas atau mengurangi kesalahan prediksi.

Model *Support Vector Machine* :

$$y = \langle w, x \rangle + b$$

Di mana:

y = Keuntungan

$\langle w, x \rangle$ = Produk dot dari vektor bobot w dan vektor input x (yang terdiri dari jumlah produk terjual dan total pendapatan).

b = Bias.

w = vektor yang menggambarkan *hyperplane* yang memisahkan data.

3.2.5 Interpretation/Evaluation (Interpretasi/Evaluasi)

Setelah menjalankan model-model regresi linier dan SVM, tahap interpretasi akan dilakukan untuk mengevaluasi dan membandingkan hasil prediksi dari kedua metode tersebut.

1. Interpretasi Hasil dari Model Regresi Linier

Persamaan Regresi Linier berikut:

$$y = 350.000 + 0,8(X_1) + 0,1 (X_2)$$

Dengan penjelasan sebagai berikut:

y = Keuntungan (dalam Rp)

X_1 = Jumlah Produk Terjual (unit)

X_2 = Total Pendapatan (Rp)

Penjelasan koefisien :

- a. *Intercept* (350.000) adalah Keuntungan dasar yang diperoleh meskipun tidak ada penjualan. Bisa merepresentasikan keuntungan tetap dari langganan atau biaya tambahan lainnya.
- b. Koefisien X_1 (0,08) Setiap tambahan 1 unit produk yang terjual meningkatkan keuntungan.
- c. Koefisien X_2 (0,1) Setiap tambahan dalam total pendapatan meningkatkan keuntungan.

Contoh perhitungan penjualan 1 september sebagai berikut :

$$y = 350.000 + (0,8 \times 44) + (0,1 \times 2579)$$

$$y = 350.000 + 35,2 + 257,9$$

$$y = \text{Rp. } 643.100$$

Strategi Bisnis Berdasarkan Regresi Linier yaitu Meningkatkan jumlah produk terjual lebih efektif kontribusi X_1 lebih signifikan terhadap keuntungan, menargetkan keuntungan Rp. 643,100 per hari dapat dihitung dengan meningkatkan X_1 dan X_2 secara seimbang sesuai hubungan dalam model. Dan Strategi promosi harus diperhitungkan agar peningkatan penjualan tidak meningkatkan biaya operasional secara tidak efisien.

2. Interpretasi Hasil Model *Support Vector Machine* (SVM)

Misalkan model SVM setelah diuji menghasilkan *Root Mean Squared Error* (RMSE) = 50 yang berarti rata-rata kesalahan prediksi keuntungan adalah Rp 50.000 Contoh Prediksi dengan SVM sebagai berikut :

Jika model SVM memprediksi keuntungan untuk 44 unit produk terjual dan total pendapatan Rp 2.579.000, dan hasil prediksi keuntungan

menunjukan 650.000 maka kita bisa membandingkan dengan keuntungan sebenarnya yang dihitung dari regresi linier sebelumnya (Rp . 643.100). Jika nilai aktual keuntungan yang dicatat adalah Rp 630.000, maka kita bisa melihat Perbandingan prediksi keuntungan dengan regresi linier.

Tabel 3.6 Perbandingan

| Metode | Keuntungan | Selisih dengan Aktual |
|----------------|------------|-----------------------|
| Regresi Linier | Rp 643.100 | 13.100 |
| SVM | Rp 650.000 | 20.000 |

Regresi Linier memprediksi Rp1.360.784, sedangkan SVM memprediksi Rp 1.360.784 Selisih dengan nilai aktual :

$$\text{Regresi Linier} = 643.100 - 630.000 = 13.100$$

$$\text{SVM} = 650.000 - 630.000 = 20.000$$

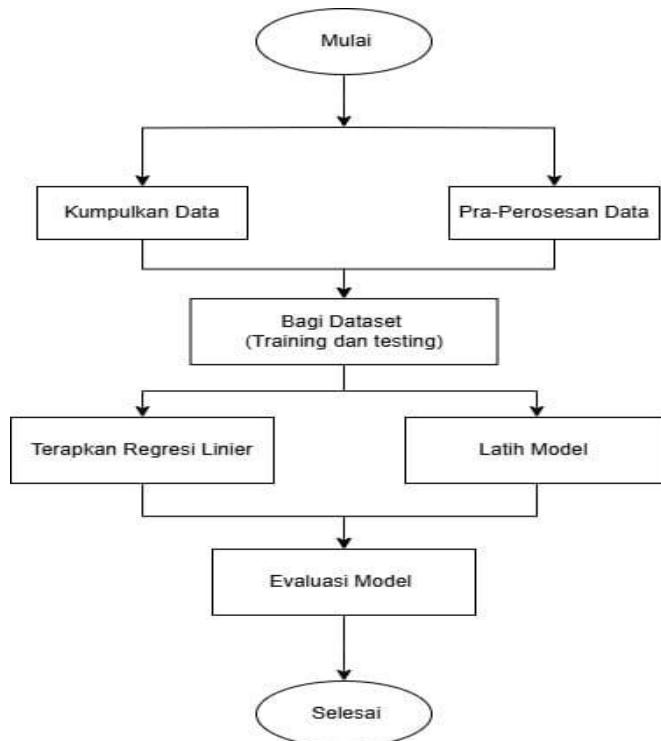
Dari sini, kita dapat menyimpulkan bahwa Regresi Linier memberikan prediksi yang sama dekat ke data aktual pada skenario ini. Namun, jika data lebih kompleks, SVM mungkin lebih akurat dalam menemukan pola tersembunyi.

Strategi Bisnis Berdasarkan Hasil Jika hasil SVM lebih akurat dalam beberapa kondisi, maka Amanda Brownies dapat menggunakan SVM untuk memprediksi keuntungan dalam situasi yang lebih kompleks, misalnya ketika ada pengaruh diskon atau promo, dan mengoptimalkan strategi harga: Jika pola harga dan jumlah produk terjual mempengaruhi keuntungan secara non-linier, maka SVM dapat memberikan gambaran yang lebih baik tentang titik optimal dalam penetapan harga..

3.3 Perancangan Algoritma

Pada penerapan data mining dengan menggunakan metode Regresi Linier dan Support Vector Machine (SVM) terdapat perencangan Algoritma yang dilakukan oleh penulis dalam bentuk flowchart yaitu sebagai berikut :

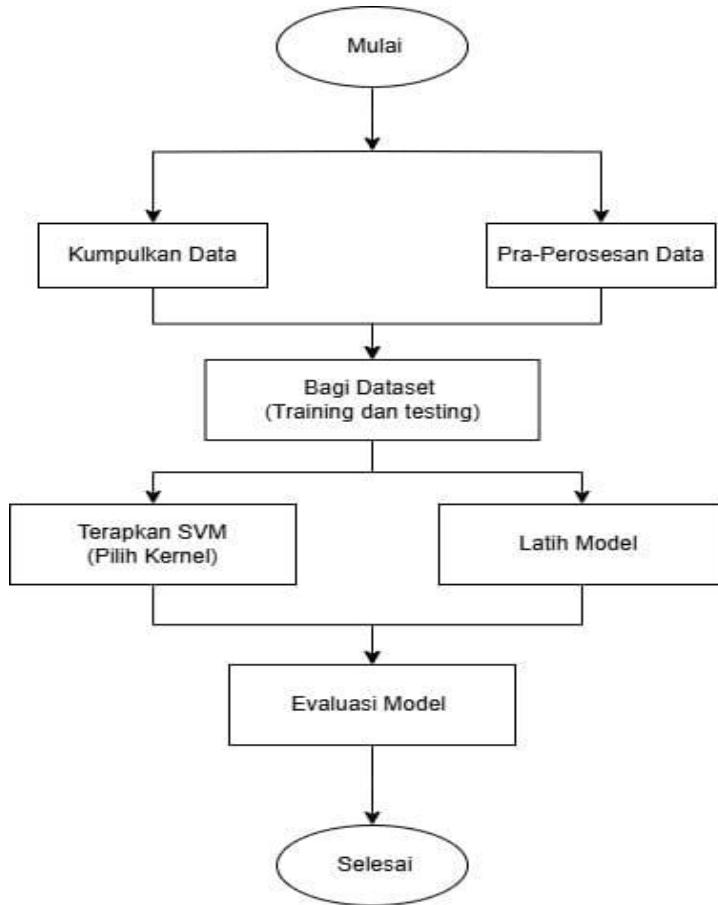
1. Regresi Linier



Gambar 3.2 Flowchart Metode Regresi Linier

Diagram ini menggambarkan langkah-langkah dalam penerapan Regresi Linier untuk membangun model prediksi. Sama seperti sebelumnya, proses dimulai dari pengumpulan dan pra-pemrosesan data, diikuti dengan pembagian dataset. Model kemudian dilatih menggunakan metode regresi linier sebelum dievaluasi untuk menilai performanya. Proses berakhir setelah evaluasi model selesai.

2. *Support Vector Machine (SVM)*



Gambar 3.3 Flowchart Metode (SVM)

Diagram ini menunjukkan alur proses dalam penerapan *Support Vector Machine* (SVM) untuk membangun model pembelajaran mesin. Dimulai dari pengumpulan dan pra-pemrosesan data, dataset dibagi menjadi training dan testing. Kemudian, model dilatih menggunakan SVM dengan pemilihan kernel tertentu. Setelah model selesai dilatih, dilakukan evaluasi untuk mengukur kinerjanya sebelum proses selesai.

3.4 Penerapan Metode Regresi dan SVM

Pada tahap ini, kedua algoritma diterapkan pada dataset yang telah melalui preprocessing. Prosesnya meliputi:

- 1. Pelatihan Model Regresi Linier**

Data dilatih menggunakan metode *Linear Regression* dan Hasil model divisualisasikan untuk mengidentifikasi tren linier.

- 2. Pelatihan Model SVM**

Data dilatih dengan kernel Polynomial untuk menangkap pola *non-linier* dan Validasi dilakukan untuk memastikan bahwa model tidak *overfitting*.

- 3. Prediksi**

Kedua model digunakan untuk memprediksi keuntungan berdasarkan variabel input tertentu.