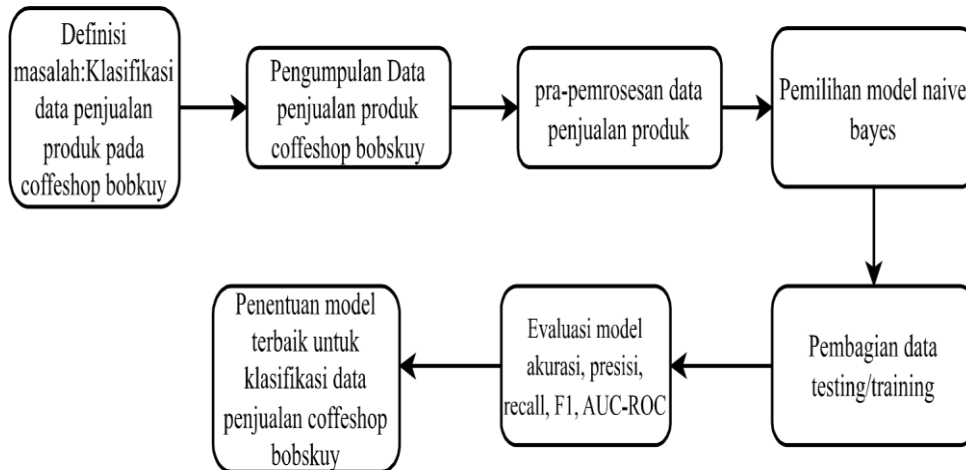


## BAB III

### METODE PENELITIAN



*Gambar 3. 1 Alur Penelitian*

#### 3.1 Metode Penelitian

Penelitian ini menggunakan metode penelitian kuantitatif, yang berfokus pada pengumpulan dan analisis data numerik untuk mendapatkan informasi yang mendalam dan relevan terkait dengan penjualan produk di Coffeeshop Bobskuy. Dalam penelitian ini, data yang digunakan berupa data historis transaksi penjualan yang mencakup informasi tentang produk yang terjual, harga, jumlah produk terjual, dan total pendapatan yang diperoleh dalam periode tertentu, yaitu dari bulan September hingga November 2025. Data transaksi ini akan dianalisis secara statistik dengan tujuan untuk mengidentifikasi pola-pola penjualan dan melakukan klasifikasi produk berdasarkan tingkat penjualannya, apakah termasuk produk yang "laris" atau "tidak laris".

Pendekatan penelitian kuantitatif dalam studi ini memanfaatkan algoritma *Naïve Bayes*, yang merupakan teknik pembelajaran mesin berbasis probabilitas untuk melakukan klasifikasi data. *Naïve Bayes* digunakan untuk memprediksi kemungkinan produk terjual berdasarkan atribut-atribut seperti harga produk, kategori produk, dan jumlah produk yang terjual. Dengan menggunakan data penjualan yang telah dikumpulkan, algoritma ini akan menghitung probabilitas masing-masing produk untuk termasuk dalam kategori produk yang laris atau tidak laris berdasarkan data historis yang ada. Salah satu keunggulan dari *Naïve Bayes* adalah kemampuannya untuk melakukan klasifikasi meskipun data yang digunakan tidak selalu lengkap atau sempurna, sehingga sangat cocok untuk analisis data di dunia bisnis yang sering kali memiliki ketidaksempurnaan dalam datanya.

### **3.2. Analisis Data**

Analisis data merupakan tahap yang melibatkan pengumpulan, pengolahan, evaluasi, dan interpretasi data untuk menghasilkan informasi yang dapat digunakan untuk pengambilan keputusan. Dalam penelitian ini, data yang digunakan mencakup periode penjualan dari September 2025 hingga November 2025 yang dikumpulkan melalui pengamatan langsung terhadap pencatatan dan pengolahan data penjualan di Coffeeshop Bobskuy.

Proses selanjutnya adalah mengubah data ke dalam format Excel untuk memudahkan pengolahan lebih lanjut. Data penjualan ini kemudian diproses berdasarkan per bulan, sehingga bisa dianalisis lebih mendalam. Pengolahan data

ini bertujuan untuk mengklasifikasikan dan memprediksi penjualan menggunakan metode algoritma *Naïve Bayes*.

### 3.3. Teknik Pengumpulan Data

Pada penelitian ini, wawancara akan dilakukan dengan pemilik atau pengelola Coffeeshop Bobskuy. Tujuan wawancara ini adalah untuk memperoleh informasi yang mendalam tentang proses penjualan, serta bagaimana pengelolaan data digunakan dalam proses pengambilan keputusan bisnis, seperti dalam perencanaan stok dan strategi promosi.

Dokumentasi juga akan dikumpulkan mencakup laporan penjualan dan catatan transaksi yang berkaitan. Dokumen-dokumen ini akan dianalisis menggunakan metode *Naïve Bayes* untuk menentukan pola-pola yang dapat digunakan untuk meningkatkan strategi pemasaran dan efisiensi operasional bisnis.

### 3.4. Praproses Data (Pre-processing)

Data yang akan digunakan pada penelitian ini didapat langsung dari data penjualan yang diberikan oleh pemilik coffeshop Bobskuy. Data yang diperoleh pada bulan September hingga November 2025.

*Tabel 3. 1 Data Penjualan september – November 2025*

BULAN SEPTEMBER					
NO	KATEGORI	NAMA BARANG	JUMLAH TERJUAL	HARGA BARANG	TOTAL
1	AWAL BULAN	Espresso	40	Rp10.000	Rp400.000
		Americano	23	Rp10.000	Rp230.000
		Sanger	60	Rp15.000	Rp900.000
		Coffee Latte	15	Rp15.000	Rp225.000
		Aren Latte	6	Rp15.000	Rp90.000
		Caramel Latte	19	Rp15.000	Rp285.000

		Choco Latte	53	Rp15.000	Rp795.000
		Boba Brown Sugar	31	Rp10.000	Rp310.000
		Kebab	15	Rp10.000	Rp150.000
		Burger	35	Rp10.000	Rp350.000
		Matcha	33	Rp10.000	Rp330.000
		Kentang Goreng	14	Rp10.000	Rp140.000
		Roti Bakar	12	Rp5.000	Rp60.000
		Dimsum	44	Rp10.000	Rp440.000
		Bakso Bakar	29	Rp10.000	Rp290.000
2	TENGAH BULAN	Espresso	56	Rp10.000	Rp560.000
		Americano	38	Rp10.000	Rp380.000
		Sanger	40	Rp15.000	Rp600.000
		Coffee Latte	21	Rp15.000	Rp315.000
		Aren Latte	0	Rp15.000	Rp0
		Caramel Latte	26	Rp15.000	Rp390.000
		Choco Latte	18	Rp15.000	Rp270.000
		Boba Brown Sugar	20	Rp10.000	Rp200.000
		Kebab	35	Rp10.000	Rp350.000
		Burger	40	Rp10.000	Rp400.000
		Matcha	40	Rp10.000	Rp400.000
		Kentang Goreng	25	Rp10.000	Rp250.000
		Roti Bakar	9	Rp5.000	Rp45.000
		Dimsum	34	Rp10.000	Rp340.000
Bakso Bakar	57	Rp10.000	Rp570.000		
3	AKHIR BULAN	Espresso	65	Rp10.000	Rp650.000
		Americano	17	Rp10.000	Rp170.000
		Sanger	59	Rp15.000	Rp885.000
		Coffee Latte	15	Rp15.000	Rp225.000
		Aren Latte	13	Rp15.000	Rp195.000
		Caramel Latte	21	Rp15.000	Rp315.000

		Choco Latte	28	Rp15.000	Rp420.000
		Boba Brown Sugar	10	Rp10.000	Rp100.000
		Kebab	55	Rp10.000	Rp550.000
		Burger	30	Rp10.000	Rp300.000
		Matcha	56	Rp10.000	Rp560.000
		Kentang Goreng	26	Rp10.000	Rp260.000
		Roti Bakar	0	Rp5.000	Rp0
		Dimsum	47	Rp10.000	Rp470.000
		Bakso Bakar	33	Rp10.000	Rp330.000

<b>BULAN OKTOBER</b>					
<b>NO</b>	<b>KATEGORI</b>	<b>NAMA BARANG</b>	<b>JUMLAH TERJUAL</b>	<b>HARGA BARANG</b>	<b>TOTAL</b>
1	AWAL BULAN	Espresso	45	Rp10.000	Rp450.000
		Americano	20	Rp10.000	Rp200.000
		Sanger	55	Rp15.000	Rp825.000
		Coffee Latte	18	Rp15.000	Rp270.000
		Aren Latte	8	Rp15.000	Rp120.000
		Caramel Latte	24	Rp15.000	Rp360.000
		Choco Latte	50	Rp15.000	Rp750.000
		Boba Brown Sugar	28	Rp10.000	Rp280.000
		Kebab	18	Rp10.000	Rp180.000
		Burger	30	Rp10.000	Rp300.000
		Matcha	30	Rp10.000	Rp300.000
		Kentang Goreng	16	Rp10.000	Rp160.000
		Roti Bakar	10	Rp5.000	Rp50.000
		Dimsum	39	Rp10.000	Rp390.000
		Bakso Bakar	31	Rp10.000	Rp310.000
		Espresso	60	Rp10.000	Rp600.000
		Americano	35	Rp10.000	Rp350.000
		Sanger	45	Rp15.000	Rp675.000
		Coffee Latte	20	Rp15.000	Rp300.000

2	TENGAH BULAN	Aren Latte	2	Rp15.000	Rp30.000
		Caramel Latte	28	Rp15.000	Rp420.000
		Choco Latte	19	Rp15.000	Rp285.000
		Boba Brown Sugar	22	Rp10.000	Rp220.000
		Kebab	30	Rp10.000	Rp300.000
		Burger	30	Rp10.000	Rp300.000
		Matcha	42	Rp10.000	Rp420.000
		Kentang Goreng	27	Rp10.000	Rp270.000
		Roti Bakar	8	Rp5.000	Rp40.000
		Dimsum	33	Rp10.000	Rp330.000
		Bakso Bakar	50	Rp10.000	Rp500.000
		3	AKHIR BULAN	Espresso	70
Americano	18			Rp10.000	Rp180.000
Sanger	50			Rp15.000	Rp750.000
Coffee Latte	13			Rp15.000	Rp195.000
Aren Latte	12			Rp15.000	Rp180.000
Caramel Latte	25			Rp15.000	Rp375.000
Choco Latte	30			Rp15.000	Rp450.000
Boba Brown Sugar	15			Rp10.000	Rp150.000
Kebab	40			Rp10.000	Rp400.000
Burger	35			Rp10.000	Rp350.000
Matcha	50			Rp10.000	Rp500.000
Kentang Goreng	28			Rp10.000	Rp280.000
Roti Bakar	6			Rp5.000	Rp30.000
Dimsum	50			Rp10.000	Rp500.000
Bakso Bakar	35			Rp10.000	Rp350.000

BULAN NOVEMBER					
NO	KATEGORI	NAMA BARANG	JUMLAH TERJUAL	HARGA BARANG	TOTAL
1	AWAL BULAN	Espresso	38	Rp10.000	Rp380.000
		Americano	15	Rp10.000	Rp150.000
		Sanger	47	Rp15.000	Rp705.000
		Coffee Latte	12	Rp15.000	Rp180.000

		Aren Latte	9	Rp15.000	Rp135.000
		Caramel Latte	20	Rp15.000	Rp300.000
		Choco Latte	40	Rp15.000	Rp600.000
		Boba Brown Sugar	28	Rp10.000	Rp280.000
		Kebab	22	Rp10.000	Rp220.000
		Burger	32	Rp10.000	Rp320.000
		Matcha	27	Rp10.000	Rp270.000
		Kentang Goreng	18	Rp10.000	Rp180.000
		Roti Bakar	14	Rp5.000	Rp70.000
		Dimsum	37	Rp10.000	Rp370.000
		Bakso Bakar	30	Rp10.000	Rp300.000
		2	TENGAH BULAN	Espresso	55
Americano	20			Rp10.000	Rp200.000
Sanger	50			Rp15.000	Rp750.000
Coffee Latte	17			Rp15.000	Rp255.000
Aren Latte	3			Rp15.000	Rp45.000
Caramel Latte	24			Rp15.000	Rp360.000
Choco Latte	19			Rp15.000	Rp285.000
Boba Brown Sugar	25			Rp10.000	Rp250.000
Kebab	40			Rp10.000	Rp400.000
Burger	12			Rp10.000	Rp120.000
Matcha	34			Rp10.000	Rp340.000
Kentang Goreng	29			Rp10.000	Rp290.000
Roti Bakar	7			Rp5.000	Rp35.000
Dimsum	39			Rp10.000	Rp390.000
Bakso Bakar	33			Rp10.000	Rp330.000
3	AKHIR BULAN	Espresso	64	Rp10.000	Rp640.000
		Americano	19	Rp10.000	Rp190.000
		Sanger	52	Rp15.000	Rp780.000
		Coffee Latte	14	Rp15.000	Rp210.000
		Aren Latte	16	Rp15.000	Rp240.000
		Caramel Latte	22	Rp15.000	Rp330.000
		Choco Latte	34	Rp15.000	Rp510.000

	Boba Brown Sugar	18	Rp10.000	Rp180.000
	Kebab	48	Rp10.000	Rp480.000
	Burger	26	Rp10.000	Rp260.000
	Matcha	40	Rp10.000	Rp400.000
	Kentang Goreng	24	Rp10.000	Rp240.000
	Roti Bakar	10	Rp5.000	Rp50.000
	Dimsum	42	Rp10.000	Rp420.000
	Bakso Bakar	38	Rp10.000	Rp380.000

### 3.5. Pembersihan Data

Langkah selanjutnya setelah proses data ialah menyatukan sumber data menjadi satu yang disebut juga dengan data *merging*.

*Tabel 3. 2 Data Merging*

No	Nama Barang	Jumlah Terjual	Harga Barang
1	Espresso	443	Rp10.000
2	Americano	167	Rp10.000
3	Sanger	458	Rp15.000
4	Coffee Latte	145	Rp15.000
5	Aren Latte	69	Rp15.000
6	Caramel Latte	189	Rp15.000
7	Choco Latte	273	Rp15.000
8	Boba Brown Sugar	177	Rp10.000
9	Kebab	263	Rp10.000
10	Burger	212	Rp10.000
11	Matcha	352	Rp10.000
12	Kentang Goreng	187	Rp10.000
13	Roti Bakar	76	Rp5.000
14	Dimsum	365	Rp10.000
15	Bakso Bakar	336	Rp10.000

No	Nama Barang	Jumlah terjual	Harga Barang	Total Pendapatan
1	Espresso	443	Rp10.000	Rp 4.430.000
2	Americano	167	Rp10.000	Rp 1.670.000
3	Sanger	458	Rp15.000	Rp 6.870.000

4	Coffee Latte	145	Rp15.000	Rp 2.175.000
5	Aren Latte	69	Rp15.000	Rp 1.035.000
6	Caramel Latte	189	Rp15.000	Rp 2.835.000
7	Choco Latte	273	Rp15.000	Rp 4.095.000
8	Boba Brown Sugar	177	Rp10.000	Rp 1.770.000
9	Kebab	263	Rp10.000	Rp 2.630.000
10	Burger	270	Rp10.000	Rp 2.700.000
11	Matcha	352	Rp10.000	Rp 3.520.000
12	Kentang Goreng	187	Rp10.000	Rp 1.870.000
13	Roti Bakar	76	Rp5.000	Rp 380.000
14	Dimsum	365	Rp10.000	Rp 3.650.000
15	Bakso Bakar	336	Rp10.000	Rp 3.360.000

Tahap berikutnya adalah mengidentifikasi nilai untuk masing-masing variabel dengan cara mengadakan wawancara bersama pemilik toko. Wawancara ini bertujuan untuk mengumpulkan informasi mengenai jumlah barang yang terjual, harga barang per unit, serta total harga yang tercatat untuk setiap variabel, sebagaimana yang ditampilkan dalam tabel berikut.

*Tabel 3. 3 Kategori Variabel*

Jumlah Barang Terjual	Nilai
Tinggi	201>500
Sedang	101-200
Rendah	0-100

Kategori	Nilai
Mahal	11.000 – 15.0000
Sedang	6.000 – 10.000
Murah	0 – 5.000

Jumlah	Nilai
Banyak	>Rp2.630.000,00
Sedikit	<Rp2.630.000,00

Perubahan atribut ini membantu memperjelas bagaimana data digunakan untuk menganalisis dan memprediksi kondisi di Coffeshop Bobskuy. Dengan data yang sudah tersusun rapi seperti ini, hubungan antar variabel menjadi lebih mudah dipahami.

Struktur data yang jelas dan konsisten sangat penting agar algoritma seperti *Naïve Bayes* dapat bekerja dengan baik. *Naïve Bayes* membutuhkan data yang teratur untuk mempelajari pola dan membuat prediksi yang akurat.

### 3.6. Pembagian Data (Split Data)

Pada penerapan algoritma *Naïve Bayes* untuk klasifikasi, dataset dibagi menjadi dua bagian utama, yaitu data pelatihan (*training*) dan data pengujian (*testing*). Pembagian ini bertujuan untuk memastikan bahwa model yang dibangun dapat mempelajari pola dari data historis dan diuji menggunakan data baru untuk mengevaluasi kinerjanya.

*Tabel 3. 4 Data Training*

Nama Barang	Jenis	Jumlah terjual	Harga Barang	Total Pendapatan	Promo	KLASI FIKAS I
Americano	Minuman	Sedang	Sedang	Sedikit	Tidak	Tidak Laris
Sanger	Minuman	Tinggi	Mahal	Banyak	Ya	Laris
Coffee Latte	Minuman	Sedang	Mahal	Sedikit	Tidak	Tidak Laris
Aren Latte	Minuman	Rendah	Mahal	Sedikit	Ya	Tidak Laris
Caramel Latte	Minuman	Sedang	Mahal	Banyak	Tidak	Tidak Laris
Choco Latte	Minuman	Tinggi	Mahal	Banyak	Ya	Laris
Boba Brown Sugar	Minuman	Sedang	Sedang	Sedikit	Ya	Tidak Laris

Kebab	Makanan	Tinggi	Sedang	Banyak	Tidak	Laris
Roti Bakar	Makanan	Rendah	Murah	Sedikit	Tidak	Tidak Laris
Dimsum	Makanan	Tinggi	Sedang	Banyak	Tidak	Laris
Bakso Bakar	Makanan	Tinggi	Sedang	Banyak	Ya	Laris

Data Training tersebut digunakan untuk data latih dalam klasifikasi naive bayes lalu dibagi per variabel untuk mencari probabilitas variabel yang akan digunakan.

*Tabel 3. 5 Data Testing*

<b>Nama Barang</b>	<b>Jenis</b>	<b>Jumlah Terjual</b>	<b>Harga Produk</b>	<b>Pendapatan</b>	<b>Promo</b>	<b>Klasifikasi</b>
<b>Espresso</b>	<b>Minuman</b>	<b>Tinggi</b>	<b>Sedang</b>	<b>Banyak</b>	<b>Ya</b>	<b>Laris</b>
<b>Kentang Goreng</b>	<b>Makanan</b>	<b>Sedang</b>	<b>Sedang</b>	<b>Sedikit</b>	<b>Ya</b>	<b>Tidak Laris</b>
<b>Burger</b>	<b>Makanan</b>	<b>Tinggi</b>	<b>Sedang</b>	<b>Sedikit</b>	<b>Tidak</b>	<b>Laris</b>
<b>Matcha</b>	<b>Minuman</b>	<b>Tinggi</b>	<b>Sedang</b>	<b>Banyak</b>	<b>Ya</b>	<b>Laris</b>

Data testing pada tabel 3.7 akan digunakan untuk menjadi data uji yang mana Model akan melakukan prediksi berdasarkan nilai-nilai yang telah dipelajari sebelumnya. Model kemudian akan memprediksi apakah produk tersebut termasuk dalam kategori “Laris” atau “Tidak Laris”, dan hasil prediksi tersebut akan dibandingkan dengan label asli untuk mengevaluasi seberapa baik performa model.

### 3.7. Perhitungan Probabilitas Likelihood

Probabilitas likelihood merupakan probabilitas kemunculan suatu kategori atribut terhadap kelas tertentu. Dalam algoritma Naïve Bayes, likelihood digunakan untuk mengukur seberapa besar kemungkinan suatu atribut muncul pada kelas Laris maupun Tidak Laris. Pada penelitian ini, atribut yang digunakan meliputi jumlah terjual, harga produk, pendapatan, promo, dan jenis produk. Setiap atribut dihitung berdasarkan frekuensi kemunculannya pada masing-masing kelas dalam data training.

Probabilitas Likelihood dihitung dengan rumus:

$$P(X_i | C) = \frac{\text{Jumlah Data Atribut } X_i \text{ Pada Kelas } C}{\text{Jumlah Seluruh Data Pada Kelas } C}$$

*Penjelasan:*

$P(X_i/C)$  = Probabilitas likelihood atribut  $X_i$  terhadap kelas  $C$

$X_i$  = Atribut atau kategori variabel

$C$  = Kelas Laris atau Tidak Laris

#### 3.7.1. Likelihood Variabel Jumlah Terjual

Berdasarkan data training, seluruh produk yang berlabel Laris berada pada kategori Tinggi, yaitu Sanger, Choco Latte, Kebab, Dimsum, dan Bakso Bakar. Sementara itu, pada kelas Tidak Laris tidak terdapat produk pada kategori Tinggi, terdapat 4 produk kategori Sedang yaitu Americano, Coffee Latte, Caramel Latte, dan Boba Brown Sugar, serta 2 produk kategori Rendah yaitu Aren Latte dan Roti

Bakar. Berdasarkan data tersebut, probabilitas likelihood untuk variabel jumlah terjual menunjukkan bahwa kategori Tinggi memiliki kecenderungan sangat kuat terhadap kelas Laris, sedangkan kategori Sedang dan Rendah lebih banyak muncul pada kelas Tidak Laris.

**Tabel 3. 6 Probabilitas Likelihood Barang Terjual**

Variabel	Kategori	Laris	Tidak Laris	P(Laris)	P(Tidak Laris)
Jumlah Barang Terjual	Tinggi	5	0	1	0
	Sedang	0	4	0	0,666666667
	Rendah	0	2	0	0,333333333
<b>Total</b>		<b>5</b>	<b>6</b>	<b>100%</b>	<b>100%</b>

### 3.7.2. Likelihood Variabel Harga Barang

Pada kelas Laris terdapat 2 produk dengan kategori harga Mahal, yaitu Sanger dan Choco Latte, serta 3 produk dengan kategori harga Sedang, yaitu Kebab, Dimsum, dan Bakso Bakar. Pada kelas Laris tidak terdapat produk dengan kategori harga Murah. Sementara itu, pada kelas Tidak Laris terdapat 3 produk dengan kategori harga Mahal, yaitu Coffee Latte, Aren Latte, dan Caramel Latte, 2 produk dengan kategori harga Sedang, yaitu Americano dan Boba Brown Sugar, serta 1 produk dengan kategori harga Murah, yaitu Roti Bakar.

**Tabel 3. 7 Probabilitas Likelihood Harga Barang**

Variabel	Kategori	Laris	Tidak Laris	P(Laris)	P(Tidak Laris)
Harga Barang	Mahal	2	3	0,4	0,5
	Sedang	3	2	0,6	0,333333333
	Murah	0	1	0	0,166666667
<b>Total</b>		<b>5</b>	<b>6</b>	<b>100%</b>	<b>100%</b>

### 3.7.3. Likelihood Variabel Pendapatan

Seluruh produk yang berlabel Laris berada pada kategori Banyak, yaitu Sanger, Choco Latte, Kebab, Dimsum, dan Bakso Bakar. Sementara itu, pada kelas Tidak Laris terdapat 1 produk kategori Banyak, yaitu Caramel Latte, dan 5 produk kategori Sedikit, yaituAmericano, Coffee Latte, Aren Latte, Boba Brown Sugar, dan Roti Bakar. Berdasarkan data tersebut, probabilitas likelihood untuk variabel total pendapatan menunjukkan bahwa kategori Banyak lebih dominan pada kelas Laris, sedangkan kategori Sedikit lebih banyak muncul pada kelas Tidak Laris.

**Tabel 3. 8 Probabilitas Pendapatan**

Variabel	Kategori	Laris	Tidak Laris	P(Laris)	P(Tidak Laris)
Pendapatan	Banyak	5	1	1	0,166666667
	Sedikit	0	5	0	0,833333333
<b>Total</b>		<b>5</b>	<b>6</b>	<b>100%</b>	<b>100%</b>

### 3.7.4. Likelihood Variabel Promo

Pada kelas Laris terdapat 3 produk dengan kategori promo Ya, yaitu Sanger, Choco Latte, dan Bakso Bakar, serta 2 produk dengan kategori promo Tidak, yaitu Kebab dan Dimsum. Sementara itu, pada kelas Tidak Laris terdapat 2 produk dengan kategori promo Ya, yaitu Aren Latte dan Boba Brown Sugar, serta 4 produk dengan kategori promo Tidak, yaituAmericano, Coffee Latte, Caramel Latte, dan Roti Bakar. Berdasarkan data tersebut, probabilitas likelihood untuk variabel promo menunjukkan bahwa pada kelas Laris promo Ya lebih banyak muncul dibandingkan promo Tidak, sedangkan pada kelas Tidak Laris promo Tidak lebih dominan.

*Tabel 3. 9 Probabilitas Likelihood Promo*

Variabel	Kategori	Laris	Tidak Laris	P(Laris)	P(Tidak Laris)
Promo	Ya	3	2	0,6	0,333333333
	Tidak	2	4	0,4	0,666666667
Total		5	6	100%	100%

### 3.7.5. Likelihood Variabel Jenis

Pada kelas Laris terdapat 2 produk kategori Minuman, yaitu Sanger dan Choco Latte, serta 3 produk kategori Makanan, yaitu Kebab, Dimsum, dan Bakso Bakar. Sementara itu, pada kelas Tidak Laris terdapat 5 produk kategori Minuman, yaituAmericano, Coffee Latte, Aren Latte, Caramel Latte, dan Boba Brown Sugar, serta 1 produk kategori Makanan, yaitu Roti Bakar. Berdasarkan data tersebut, probabilitas

likelihood untuk variabel jenis menunjukkan bahwa pada kelas Laris kategori Makanan lebih banyak muncul, sedangkan pada kelas Tidak Laris kategori Minuman lebih dominan.

**Tabel 3. 10 Probabilitas Likelihood Jenis**

Variabel	Kategori	Laris	Tidak Laris	P(Laris)	P(Tidak Laris)
Jenis	Makanan	2	5	0,42857143	0,375
	Minuman	3	1	0,57142857	0,625
Total		5	6	100%	100%

### 3.8. Perhitungan Probabilitas Prior

Probabilitas prior merupakan peluang awal dari masing-masing kelas sebelum mempertimbangkan atribut yang dimiliki oleh data. Dalam penelitian ini, kelas yang digunakan adalah Laris dan Tidak Laris. Perhitungan prior dilakukan berdasarkan jumlah data pada masing-masing kelas dalam data training. Berdasarkan data training yang digunakan pada penelitian ini, terdapat 11 data produk, dengan 5 produk termasuk ke dalam kelas Laris dan 6 produk termasuk ke dalam kelas Tidak Laris.

Probabilitas prior dihitung dengan rumus:

$$P(C) = \frac{\text{Jumlah data pada kelas } C}{\text{Jumlah Seluruh Data Training}}$$

Dengan:

$P(C)$  = probabilitas prior kelas

C = kelas, yaitu Laris atau Tidak Laris

**Tabel 3. 11 Probabilitas Prior Klasifikasi**

Variabel	Kategori	Jumlah	Laris/Tidak Laris
Klasifikasi	Laris	5	0,454545455
	Tidak Laris	6	0,545454545
Total		11	100%

Ketika data telah berhasil dikelompokkan maka langkah selanjutnya mengolah data menggunakan perhitungan manual algoritma naive bayes pada perangkat *software Microsoft Excel*.

**Tabel 3. 12 Hasil Prediksi**

Nama Barang	Jenis	Jumlah Terjual	Harga Produk	Pendapatan	Promo	Klasifikasi
Espresso	Minuman	Tinggi	Sedang	Banyak	Ya	Laris
Kentang Goreng	Makanan	Sedang	Sedang	Sedikit	Ya	Tidak Laris
Burger	Makanan	Tinggi	Sedang	Sedikit	Tidak	Laris
Matcha	Minuman	Tinggi	Sedang	Banyak	Ya	Laris
PREDICTION						
	Laris		Tidak Laris		Hasil	
Espresso	0,09818		0		Laris	
Kentang Goreng	0		0,028058361		Tidak Laris	
Burger	0,04364		0		Laris	
Matcha	0,09818		0		Laris	

**Tabel 3. 13 Confusion Matrix**

Laris	Tidak Laris
TP = 3	FP = 0
FN = 0	TN = 1

TP (*True Positive*): 3 (Produk yang benar-benar "Laris" dan diprediksi "Laris").

FP (*False Positive*): 0 (Tidak ada produk yang diprediksi "Laris" tetapi sebenarnya "Tidak Laris").

FN (*False Negative*): 0 (Tidak ada produk yang diprediksi "Tidak Laris" tetapi sebenarnya "Laris").

TN (*True Negative*): 1 (Produk yang benar-benar "Tidak Laris" dan diprediksi "Tidak Laris").

### 3.9. Confusion Matrix

#### 1. Accuracy

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$$

Dengan nilai  $\frac{3+1}{3+1+0+0} = \frac{4}{4} = 1$  atau 100%

#### 2. Precision

$$Precision = \frac{TP}{TP + FP}$$

Dengan nilai  $Precision = \frac{3}{3+0} = 1 = 100\%$

#### 3. Recall

$$Recall = \frac{TP}{TP + FN}$$

Dengan nilai  $Recall = \frac{3}{3+0} = 1 = 100\%$

#### 4. *F-1 score*

$$F-1 \text{ score} = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

Dengan nilai  $F-1 \text{ score} = 2 \times \frac{1 \times 1}{1+1} = 1 = 100\%$

**Tabel 3. 14 Hasil Evaluasi Model**

HASIL EVALUASI MODEL <i>NAIVE BAYES</i>	
Accuracy	1
Precision	1
Recall	1
F-1 score	1

Tabel di atas menunjukkan hasil evaluasi model *Naive Bayes* yang digunakan untuk klasifikasi produk yang diminati (Laris) dan tidak diminati (Tidak Laris). Berdasarkan hasil pengujian, model ini menunjukkan akurasi sebesar 100%, yang berarti bahwa seluruh data yang diuji dapat diprediksi dengan benar oleh model. Presisi yang tercatat sebesar 1 mengindikasikan bahwa setiap prediksi positif yang dilakukan oleh model benar-benar sesuai dengan kategori yang diinginkan, menunjukkan bahwa model sangat teliti dalam memprediksi kelas positif (produk "Laris").

Nilai *recall* yang tercatat sebesar 100% menunjukkan bahwa model ini berhasil menemukan seluruh produk yang seharusnya dikenali sebagai "Laris", tanpa ada yang terlewat (tanpa *False Negatives*).

Nilai *F1-Score* yang mencapai 100% menunjukkan keseimbangan yang sempurna antara presisi dan recall, dengan nilai yang mendekati angka ideal 1. Hal ini menunjukkan bahwa model *Naïve Bayes* tidak hanya akurat, tetapi juga sangat efisien dalam menangkap seluruh kelas positif dan menghasilkan prediksi yang tepat.

Secara keseluruhan, evaluasi ini mengindikasikan bahwa model *Naïve Bayes* yang digunakan memiliki performa yang sangat solid dalam mengklasifikasikan data, dengan tingkat ketepatan dan kemampuan yang sangat baik dalam mengenali kelas positif.