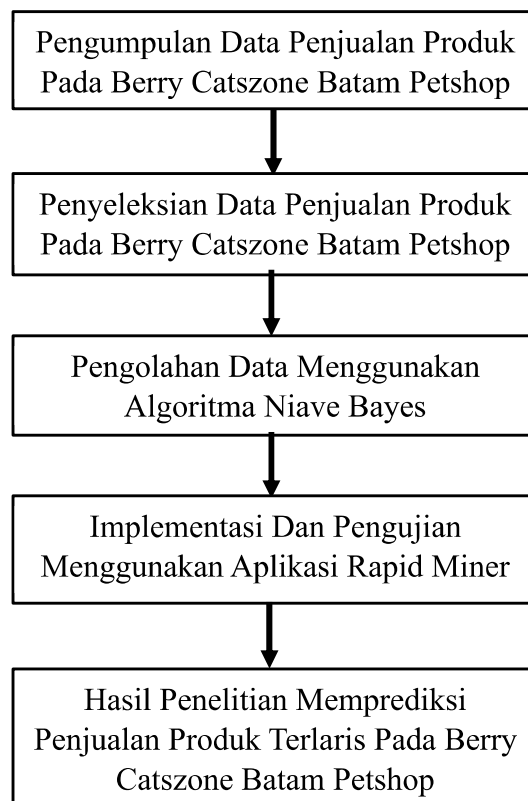


BAB III

METODE PENELITIAN

3.1 Desain Penelitian

Dalam melakukan penelitian, seorang peneliti telah merencanakan suatu desain penelitian yang menguraikan semua langkah yang akan dilakukan selama proses penelitian mengenai data mining mulai dari tahap awal hingga hasil akhir. Penjelasan mengenai desain penelitian data mining dalam penelitian ini adalah sebagai berikut:



Sumber: Data Penelitian (2024)

Berdasarkan desain penelitian yang telah dibuat, berikut penjelasannya:

1. Pengumpulan data

Data dikumpulkan untuk memperoleh informasi yang dibutuhkan dalam mencapai tujuan penelitian. Data diambil dari Berry Catszone Batam petshop dengan metode wawancara, observasi, dan studi literatur tentang data mining dan algoritma Naive Bayes, yang diperoleh melalui referensi dari buku dan jurnal, baik nasional maupun internasional.

2. Penyeleksian data

Peneliti melakukan seleksi data untuk memperoleh sampel yang sesuai dengan tujuan penelitian. Data yang dipilih mencakup penjualan produk di Berry Catszone Batam Petshop dari Januari hingga Desember 2023.

3. Pengolahan data

Data yang telah dipilih kemudian diproses menggunakan algoritma *naive bayes* dalam teknik data mining.

4. Implementasi

Penerapan pada desain penelitian ini dilakukan dengan memanfaatkan software RapidMiner untuk memprediksi penjualan produk terlaris pada Berry Catszone Batam petshop.

5. Hasil penelitian

Hasil penelitian yang telah dianalisis menggunakan teknik data mining dengan algoritma Naive Bayes dan diuji melalui *software* RapidMiner, menampilkan hasil prediksi penjualan produk terlaris di Berry Catszone Batam Petshop.

3.2 Teknik Pengumpulan Data

Dalam penelitian ini, peneliti menggunakan metode berikut untuk mengumpulkan data:

3.2.1 Observasi

Mengamati secara langsung ke toko Berry Catszone Batam petshop untuk memperoleh data yang dibutuhkan yaitu data penjualan produk yang sudah terjual satu tahun terakhir periode januari hingga desember Tahun 2023.

3.2.2 Wawancara

Dalam penelitian ini, wawancara dilakukan secara langsung kepada Manager toko Berry Catszone Batam Petshop. Pertanyaan yang diajukan terkait dengan jenis-jenis produk makanan kucing dan anjing. tujuan wawancara adalah untuk memperoleh data penjualan produk yang telah terjual. Jumlah raw data yang diberikan sebanyak 845 data dengan jumlah atribut sebanyak 17 tidak termasuk klasifikasi dengan rincian sebagai berikut :

Tabel 3. 1 Atribut Data Raw

No	Atribut	Keterangan
1	Nama Barang	Nama Barang ditoko Berry Catszone Batam Petshop
2	Kategori	Jenis-jenis Produk
3	Harga Satuan	Harga dari setiap produk makanan
4	Januari	Bulan Penjualan
5	Februari	Bulan Penjualan
6	Maret	Bulan Penjualan
7	April	Bulan Penjualan

No	Atribut	Keterangan
8	Mei	Bulan Penjualan
9	Juni	Bulan Penjualan
10	Juli	Bulan Penjualan
11	Agustus	Bulan Penjualan
12	September	Bulan Penjualan
13	Oktober	Bulan Penjualan
14	November	Bulan Penjualan
15	Desember	Bulan Penjualan
16	Total Qty	Jumlah produk yang terjual
17	Totoal Penjualan	Total penjualan satu tahun terakhir
18	Klasifikasi	Klasifikasi Penjualan Laris dan tidak laris

Sumber: Penelitian 2024

Tabel 3.1 berisi atribut data penjualan produk makanan kucing dan anjing di Berry Catszone Batam Petshop.

3.2.3 Studi Pustaka

Pada penelitian ini, peneliti mengumpulkan beragam referensi pendukung dari buku, jurnal, dan berbagai sumber lain yang relevan dengan objek penelitian.

3.3 Operasional Variabel

Operasional variabel yang diterapkan pada penelitian ini yaitu data penjualan produk terlaris pada Berry Catszone Batam Petshop untuk menentukan variabel dan indikator yang memiliki hubungan berbeda dalam sebuah penelitian. Untuk mendukung penelitian ini, peneliti mengambil data penjualan produk yang sudah terjual dalam satu tahun terakhir yaitu tahun 2023. Berikut ini variabel yang telah ditentukan dalam penelitian ini:

1. Nama Produk

Nama-nama produk yang sering dibeli oleh pelanggan digunakan sebagai dasar untuk pemilihan stok produk di Berry Catszone Batam Petshop.

2. Kategori

Kategori merupakan jenis-jenis produk yang ada pada Berry catszone batam petshop, dalam penelitian ini kategori produk yang diangkat sebagai variabel yaitu *Cat food*, *Cat food*, *Cat snack*, *Cat litter*, *Wet cat food*, *Accessories*, Shampo kucing, *Dog food*, dan *Cat dog food* yang sering diminati oleh pembeli produk petshop.

3. Harga

Harga dikelompokkan dalam dua bagian yakni Murah dan Mahal. Murah dikatakan apabila dibawah harga Rp 75.000 dan Mahal dikatakan apabila harga dibawah Rp 75.000.

4. Kuantitas

Kuantitas merupakan jumlah produk yang telah dinyatakan terjual pada Berry catszone batam petshop.

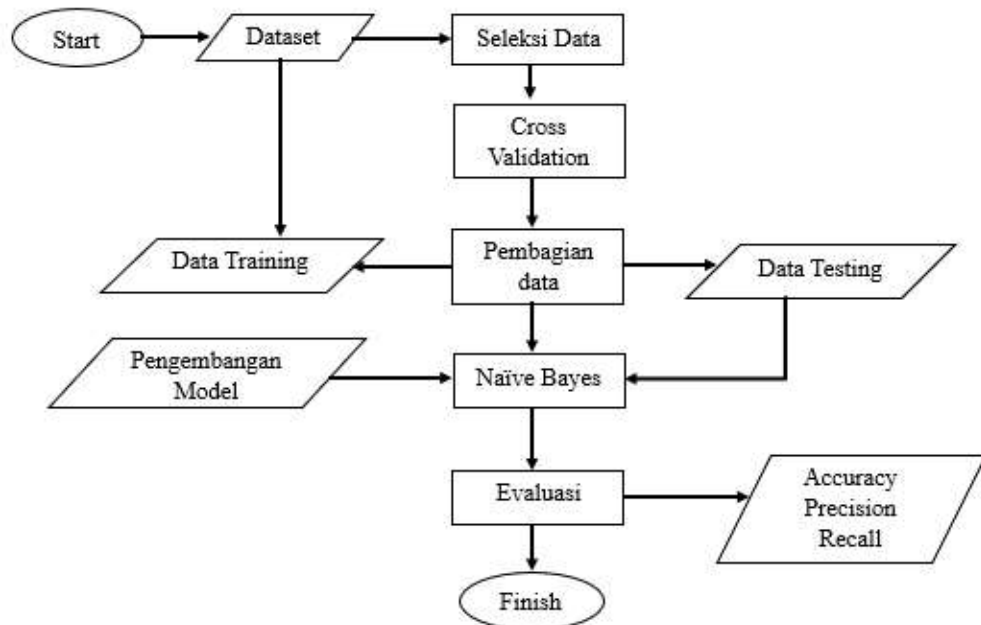
5. Klasifikasi

Klasifikasi merupakan suatu pengelompokkan dan pengkategorian yang didasarkan pengeluaran produk yang di nyatakan terjual dengan kategori laris dan tidak laris.

3.4 Teknik Analisis Data

Teknik dalam proses analisis data untuk memprediksi penjualan produk terlaris pada petshop dimulai dari pengumpulan data, preprocessing, proses dan

evaluasi terhadap hasil. Diagram alur pada penelitian ini dapat dilihat pada gambar 3.1 berikut.



Gambar 3. 1 Teknik Analisis Data

Sumber: Data Penelitian 2024

Dimulai dari Start, selanjutnya pengumpulan dataset dilakukan dari Berry Catszone Batam Petshop. Setelah dataset terkumpul, data di seleksi untuk menghilangkan duplikasi dan memastikan keakuratan data. Setelah proses di seleksi, data dibagi dalam dua bagian yaitu Data Training dan Data Testing. selanjutnya, dilakukan pengujian data menggunakan Algoritma Naïve Bayes. Setelah pengujian, hasilnya dievaluasi dengan menghitung nilai Akurasi, Presisi, dan Recall, dan kemudian ditarik kesimpulan. Akhirnya, proses pengujian selesai.

3.4.1 Seleksi Atribut

Pemilihan atribut pada dataset menggunakan teknik wight information gain pada rapid minner dengan tujuan mengurutkan nilai weight information gain pada dataset awal, atribut terbaik akan menjadi datset akhir untuk penelitian ini. Penliti memilih dengan jumlah atribut sebanyak 6 berdasarkan hasil yang paling tinggi pada Teknik information gain.

Rumus Gain:

Rumus 3. 1 Information Gain

$$IG(c, t) = S(c) + \sum_{j \in value(t)}^n \frac{|c_j|}{|c|} S(C_j)$$

$S(c)$ merupakan Entropy seluruh feature c (sebelum pemisahan),

$S(c_j)$ merupakan Entropy feature c untuk class $t = j$ (setelah pemisahan)

$value(t)$ merupakan himpunan nilai-nilai yang mungkin untuk class t,

n merupakan jumlah nilai-nilai yang mungkin untuk class t,

$|c_j|$ merupakan jumlah sample class dengan nilai = j,

$|c|$ merupakan jumlah sample untuk seluruh class.

Nilai Weight Information Gain yang tertinggi akan dijadikam sebagai fitur terbaik dalam memprediksi penjualan produk di Berry catszone Batam Petshop Batam.

3.4.2 Cross Validation

Validasi silang metode yang digunakan dalam pemodelan machine learning untuk mengukur seberapa baik model akan berfungsi pada data yang belum pernah dilihat sebelumnya. Teknik ini membagi data menjadi beberapa subset, melatih model pada sebagian data (data Training), dan mengujinya pada subset lainnya (data testing). Berikut ini merupakan contoh tabel dari cara kerja Cross Validation.

Tabel 3. 2 Cara Kerja Cross Validation

	Iterasi 1	Iterasi 2	Iterasi 3	Iterasi 4	Iterasi 5	Iterasi 6
Percobaan 1	Validation Subset	Training Subset	Training Subset	Training Subset	Training Subset	Training Subset
Percobaan 2	Training Subset	Validation Subset	Training Subset	Training Subset	Training Subset	Training Subset
Percobaan 3	Training Subset	Training Subset	Validation Subset	Training Subset	Training Subset	Training Subset
Percobaan 4	Training Subset	Training Subset	Training Subset	Validation Subset	Training Subset	Training Subset
Percobaan 5	Training Subset	Training Subset	Training Subset	Training Subset	Validation Subset	Training Subset
Percobaan 6	Training Subset	Training Subset	Training Subset	Training Subset	Training Subset	Validation Subset
Validation accuracy	82,19%	84,93	89,04%	87,67	90,41%	90,41%

Sumber: Penelitian 2024

Percobaan di atas merupakan ilustrasi dari 6-fold cross validation, yang berarti percobaan dilakukan dalam 6 tahap.

Data training pada penelitian ini merupakan peran penting dalam pengembangan dan pengujian model pembelajaran mesin di RapidMiner. Proses pengujian data training melibatkan beberapa langkah penting, mulai dari pemisahan

data hingga evaluasi model. Cara kerja data training dalam penelitian ini sebagai berikut:

1. Melakukan import data dari file CSV, kemudian membagi dataset dalam dua bagian yaitu data testing dan data training.
2. Menggunakan cross validation untuk membagi data menjadi beberapa subset.
3. Memilih operator naïve bayes sebagai algoritma pada pengujian dataset yang akan di uji.
4. Selanjutnya memilih operator *Apply* Model untuk mengaplikasikan model yang telah dilatih pada data uji.
5. Memilih Operator performance untuk mengevaluasi dan mengukur kinerja prediksi seperti akurasi, presisi, dan recall.

Tabel 3. 3 Data Training

Nama barang	Kategori	Harga satuan	Total qty	Total penjualan	Klasifikasi
Cat choize adult 800 Gram	Cat food	Murah	242	Rp 4.356.000,00	Laris
Cat choiz 1 kg	Cat food	Murah	290	Rp 7.250.000,00	Laris
Excel 500 gram	Cat food	Murah	129	Rp 1.806.000,00	Tidak Laris
Ori cat 1 kg	Cat food	Murah	135	Rp 2.970.000,00	Laris
Me-o	Cat snack	Murah	472	Rp 2.832.000,00	Laris
Liebao	Cat snack	Murah	127	Rp 635.000,00	Tidak Laris
Life cat	Cat snack	Murah	554	Rp 3.878.000,00	Laris
Furlove	Cat snack	Murah	279	Rp 1.674.000,00	Laris
One cat 5 liter	Cat litter	Murah	82	Rp 3.690.000,00	Tidak Laris
One cat 10 liter	Cat litter	Mahal	205	Rp 15.375.000,00	Laris

Nama barang	Kategori	Harga satuan	Total qty	Total penjualan	Klasifikasi
One cat 20 liter	Cat litter	Mahal	21	Rp 2.478.000,00	Tidak Laris
Goomi 20 liter	Cat litter	Mahal	49	Rp 5.880.000,00	Tidak Laris
Life cat 400 gram	Wet cat food	Murah	302	Rp 5.436.000,00	Laris
Fur love 400 gram	Wet cat food	Murah	126	Rp 2.016.000,00	Tidak Laris
Whiskas 400 gram	Wet cat food	Murah	279	Rp 6.696.000,00	Laris
Markotops 400 gram	Wet cat food	Murah	295	Rp 5.015.000,00	Laris
Kalung kucing	Accessories	Murah	282	Rp 1.410.000,00	Laris
Tali tuntun	Accessories	Murah	18	Rp 180.000,00	Tidak Laris
Kalung kucing pita mini	Accessories	Murah	21	Rp 147.000,00	Tidak Laris
Tenda kucing	Accessories	Mahal	5	Rp 775.000,00	Tidak Laris
Dog Choize	Dog Food	Murah	80	Rp 1.600.000,00	Tidak Laris
Bolt Dog	Dog Food	Mahal	136	Rp 31.280.000,00	Laris
Petopia Beef	Dog Food	Murah	134	Rp 2.546.000,00	Laris
Lightning Dog	Dog Food	Murah	91	Rp 1.729.000,00	Tidak Laris
Life Dog	Wet dog food	Murah	153	Rp 2.601.000,00	Laris
Favorite	Wet dog food	Murah	161	Rp 2.898.000,00	Laris
Petopia	Wet dog food	Murah	171	Rp 2.565.000,00	Laris
Ugo	Wet dog food	Murah	129	Rp 2.322.000,00	Tidak Laris
Cat'z shampo	Shampo kucing	Murah	50	Rp 700.000,00	Tidak Laris

Nama barang	Kategori	Harga satuan	Total qty	Total penjualan	Klasifikasi
Bestie Pet	Shampo kucing	Murah	65	Rp 845.000,00	Tidak Laris
Olive Essence	Shampo kucing	Murah	35	Rp 1.225.000,00	Tidak Laris
Pet Shampo	Shampo kucing	Murah	74	Rp 1.184.000,00	Tidak Laris
Bubu Shampo	Shampo kucing	Murah	44	Rp 1.100.000,00	Tidak Laris
Markotops 10 liter	Cat litter	Mahal	151	Rp 11.325.000,00	Laris
Markotops 5,5 liter	Cat litter	Murah	86	Rp 4.300.000,00	Tidak Laris
Maxi 30 liter	Cat litter	Mahal	66	Rp 10.560.000,00	Tidak Laris
Life cat all varian 7 liter	Cat litter	Murah	104	Rp 6.760.000,00	Tidak Laris
Grand panda 10 liter	Cat litter	Murah	146	Rp 8.614.000,00	Laris
Grand panda 25 liter	Cat litter	Mahal	61	Rp 6.405.000,00	Tidak Laris
Royal canin adult persian 500 Gram	Cat food	Mahal	132	Rp 11.088.000,00	Laris
I care fisrt 1,5 kg	Cat food	Mahal	60	Rp 7.920.000,00	Tidak Laris
Bolt 1 kg	Cat food	Murah	245	Rp 5.390.000,00	Laris
Ori cat 800 gram	Cat food	Murah	122	Rp 2.196.000,00	Tidak Laris
Ori cat	Cat snack	Murah	535	Rp 2.675.000,00	Laris
Coucou	Cat snack	Murah	295	Rp 1.475.000,00	Laris
Whiskas	Cat snack	Murah	544	Rp 3.808.000,00	Laris
Cat choize	Cat snack	Murah	555	Rp 3.607.500,00	Laris
BeaUty premium 1 kg	Cat food	Murah	253	Rp 7.337.000,00	Laris
Felibite 500 gram	Cat food	Murah	130	Rp 1.950.000,00	Tidak Laris
Markotop kitten/adult 1 kg	Cat food	Murah	247	Rp 6.422.000,00	Laris
Tas astronot kucing	Accessories	Mahal	9	Rp 990.000,00	Tidak Laris

Nama barang	Kategori	Harga satuan	Total qty	Total penjualan	Klasifikasi
Mangkok double kucing	Accessories	Murah	37	Rp 444.000,00	Tidak Laris
Sticky roll dan refill	Accessories	Murah	13	Rp 143.000,00	Tidak Laris
Bulu stick	Accessories	Murah	13	Rp 104.000,00	Tidak Laris
Sekop	Accessories	Murah	34	Rp 136.000,00	Tidak Laris
Litter box	Accessories	Murah	23	Rp 920.000,00	Tidak Laris
Gunting kuku kucing	Accessories	Murah	5	Rp 100.000,00	Tidak Laris
Sapu mini	Accessories	Murah	18	Rp 180.000,00	Tidak Laris
Me-o 400 gram	Wet cat food	Murah	285	Rp 5.130.000,00	Laris
Cici 400 gram	Wet cat food	Murah	184	Rp 2.944.000,00	Laris
Captain cat 400 gram	Wet cat food	Murah	298	Rp 4.768.000,00	Laris
Fur love 1 kg	Cat food	Murah	237	Rp 6.162.000,00	Laris
Whiskas 1,2 kg	Cat food	Murah	299	Rp 19.435.000,00	Laris
Senior 25 liter	Cat litter	Mahal	18	Rp 2.250.000,00	Tidak Laris
Senior 10 liter	Cat litter	Murah	38	Rp 2.660.000,00	Tidak Laris
Senior 5 liter	Cat litter	Murah	47	Rp 2.021.000,00	Tidak Laris
Markotops 25 liter	Cat litter	Mahal	40	Rp 5.080.000,00	Tidak Laris
Ori Dog	Dog Food	Murah	122	Rp 2.684.000,00	Tidak Laris
Im Organic 7,5 kg	Cat food	Mahal	35	Rp 11.550.000,00	Tidak Laris
Me-o 400 gram	Cat food	Murah	126	Rp 2.142.000,00	Tidak Laris
Happy cat 1 kg	Cat food	Murah	138	Rp 9.660.000,00	Laris
Cleo citten dan persian 400 gram	Cat food	Murah	123	Rp 2.214.000,00	Tidak Laris

Nama barang	Kategori	Harga satuan	Total qty	Total penjualan	Klasifikasi
Dog Menue	Wet dog food	Murah	206	Rp 3.296.000,00	Laris

Sumber: Penelitian 2024

3.4.3 Naive Bayes

Naïve Bayes dengan metode klasifikasi yang menggunakan teorema Bayes sebagai dasarnya. Metode ini sering digunakan untuk menghitung probabilitas suatu kelas. Secara sederhana, Naïve Bayes mengasumsikan bahwa keberadaan suatu fitur dalam kelas tertentu tidak bergantung pada fitur lainnya. Rumus teorema Bayes bisa dilihat berikut ini.

$$P(H|X) = \frac{P(X|H) \cdot P(H)}{P(X)}$$

Keterangan:

X = Data dengan kelas yang belum diketahui.

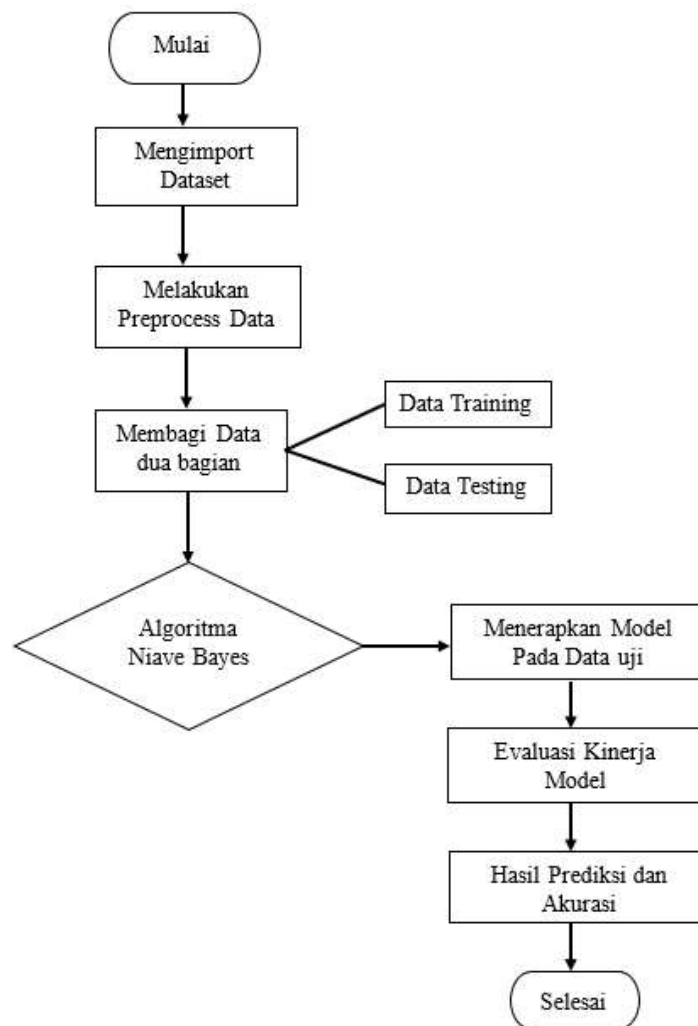
H = Hipotesis data X yang merupakan suatu kelas spesifik.

P(H) = Probabilitas hipotesis H (prior).

P(H|X) = Probabilitas hipotesis H berdasarkan kondisi X (posterior).

P(X) = Probabilitas X.

P(X|H) = Probabilitas X berdasarkan kondisi pada hipotesis H.



Gambar 3. 2 Alur Algoritma Naïve Bayes

Sumber: Data Penelitian 2024

Tahapan dalam melakukan klasifikasi dengan Naïve Bayes dapat dilihat pada gambar 3.2. Langkah pertama yaitu mengimport dataset, kemudian melakukan pra-pemrosesan data. Setelah itu, data dibagi menjadi dua bagian, yaitu data testing dan data training. Selanjutnya latih model naïve bayes dengan data pelatihan, setelah itu menerapkan model pada data uji. Setelah itu evaluasi kinerja model dan Kemudian hasil prediksi dan akurasi naïve bayes, selesai.

3.4.4 Evaluasi Performance

Pendekatan untuk mengukur peningkatan atau hasil yang dicapai dalam suatu periode tertentu

Rumus Accuracy:

Rumus 3. 2 Accuracy

$$Accuracy = \frac{(TN + TP)}{(TN + FN + TP + FP)} \times 100\%$$

$$Precision = \frac{TP}{TP + FP} \times 100\%$$

$$Recall = \frac{TP}{TP + FN} \times 100\%$$

Keterangan:

TP (True Positive) = Data positif yang diprediksi dengan benar

TN (True Negative) = Data negatif yang diprediksi dengan benar

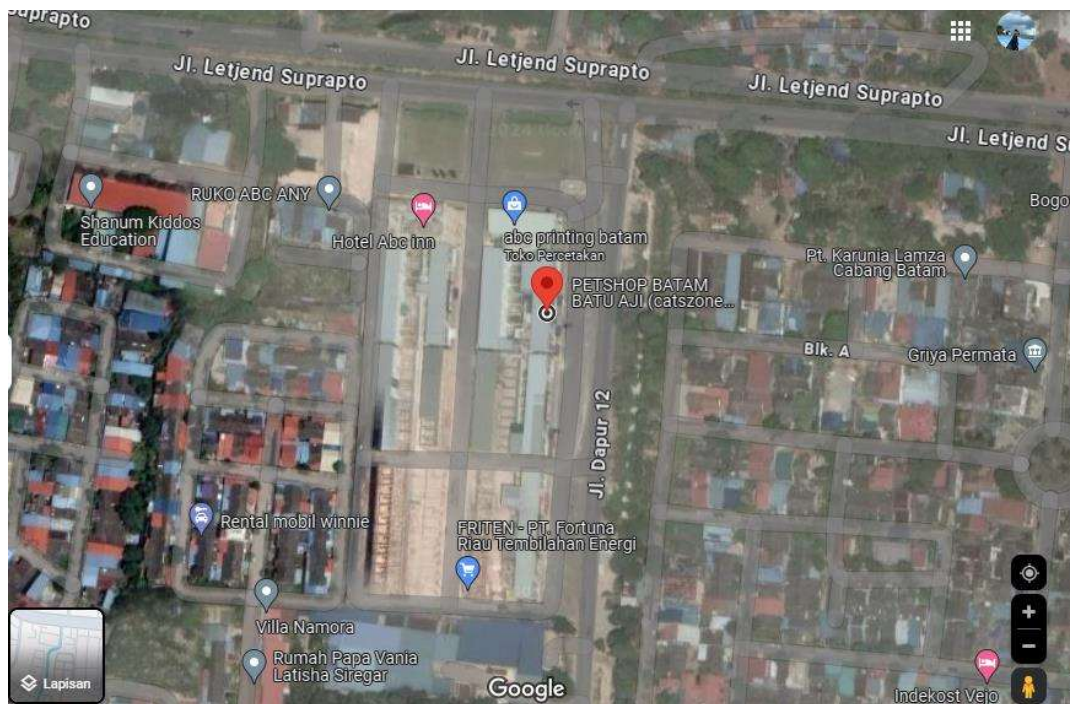
FP (False Positive) = Data negatif yang diprediksi sebagai positif

FN (False Negative) = Data positif yang diprediksi sebagai negatif

3.5 Lokasi dan jadwal Penelitian

3.5.1 Lokasi Penelitian

Penelitian ini dilakukan di Toko berry Catszone Batam Petshop di Ruko ABC Aji Bussiness center, Blok A No.12a, Sagulung kota Kec. Sagulung, Kota Batam, Kepulauan Riau 29438.



Gambar 3. 3 Tempat Penelitian

Sumber: Penelitian 2024

3.5.2 jadwal Penelitian

Penelitian ini diselenggarakan selama satu semester sesuai dengan agenda akademik yang telah ditetapkan oleh Universitas Putera Batam, setelah persetujuan judul penelitian diberikan. waktu penelitian ini mencakup periode mulai dari bulan Maret 2024 hingga bulan Juli 2024. Rencana jadwal yang telah disusun untuk penelitian ini dapat dilihat dalam tabel berikut.

Tabel 3. 4 Jadwal Penelitian

Nama kegiatan	Waktu Kegiatan																			
	Maret 2024				April 2024				Mei 2024				Juni 2024				Juli 2024			
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
Pengajuan judul	■	■	■																	
Penyusunan BAB 1				■	■	■														
Penyusunan BAB 2					■	■	■	■												
Penyusunan BAB 3									■	■	■	■	■	■	■	■	■	■	■	
Penyusunan BAB 4														■	■	■	■	■	■	
Penyusunan BAB 5																■	■	■	■	■
Penyelesaian Sikripsi																			■	■

Sumber: Penelitian 2024