

**PENDEKATAN DATA MINING UNTUK MEMILIH  
PRODUK TERLARIS MENGGUNAKAN  
ALGORITMA NAIVE BAYES**

**SKRIPSI**



**Oleh:  
Hidup Perjuangan Rajagukguk  
190210074**

**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS TEKNIK DAN KOMPUTER  
UNIVERSITAS PUTERA BATAM  
TAHUN 2023**

**PENDEKATAN DATA MINING UNTUK MEMILIH  
PRODUK TERLARIS MENGGUNAKAN  
ALGORITMA NAIVE BAYES**

**SKRIPSI**

**Untuk memenuhi salah satu syarat  
Guna memperoleh gelar Sarjana**



**Oleh:  
Hidup Perjuangan Rajagukguk  
190210074**

**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS TEKNIK DAN KOMPUTER  
UNIVERSITAS PUTERA BATAM  
TAHUN 2023**

## SURAT PERNYATAAN ORISINALITAS

Yang bertanda tangan di bawah ini saya:

Nama : Hidup Perjuangan Rajagukguk  
NPM : 190210074  
Fakultas : Teknik dan Komputer  
Program Studi : Teknik Informatika

Menyatakan bahwa "Skripsi" yang saya buat dengan judul:

### **PENDEKATAN DATA MINING UNTUK MEMILIH PRODUK TERLARIS MENGUNAKAN ALGORITMA NAIVE BAYES**

Adalah hasil karya sendiri dan bukan "duplikasi" dari karya orang lain. Sepengetahuan saya, di dalam naskah Skripsi ini tidak terdapat karya ilmiah atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis dikutip didalam naskah ini dan disebutkan dalam sumber kutipan dan daftar pustaka. Apabila ternyata di dalam naskah Skripsi ini dapat dibuktikan terdapat unsur-unsur PLAGIASI, saya bersedia naskah Skripsi ini digugurkan dan gelar akademik yang saya peroleh dibatalkan, serta diproses sesuai dengan peraturan perundang-undangan yang berlaku.

Demikian pernyataan ini saya buat dengan sebenarnya tanpa ada paksaan dari siapapun

Batam, 26 Juli 2023



Hidup Perjuangan Rajagukguk  
190210074

**PENDEKATAN DATA MINING UNTUK MEMILIH  
PRODUK TERLARIS MENGGUNAKAN  
ALGORITMA NAIVE BAYES**

**SKRIPSI**

**Oleh  
Hidup Perjuangan Rajagukguk  
190210074**

**Telah disetujui oleh Pembimbing pada tanggal  
seperti tertera di bawah ini**

**Batam, 26 Juli 2023**

  
**Rahmat Fauzi, S.Kom., M.Kom  
Pembimbing**

## ABSTRAK

Kemajuan teknologi pada saat ini dapat dimanfaatkan untuk mengolah data menjadi informasi yang lebih bermanfaat. Dalam pengumpulan data, pengumpulan informasi sangat berguna untuk memaksimalkan keuntungan dan mengembangkan strategi pemasaran. Salah satu cara untuk meningkatkan keuntungan adalah dengan menggunakan teknik *data mining* untuk membantu para pelaku bisnis dalam mengambil keputusan mengenai stok barang, peningkatan keuntungan dan lainnya. Matahari *Departement Store* adalah platform ritel terbesar di Indonesia, salah satu ritel yang berada di Batam adalah Matahari *Departement Store* Nagoya Hill Batam. Data transaksi pada Matahari *Departement Store* yang masih diolah tidak menggunakan sebuah metode menyebabkan pengolahan data penjualan produk menjadi kurang efektif dan kurang efisien. Dilihat dari banyaknya transaksi maka dibutuhkan sistem untuk memprediksi penjualan produk terlaris sehingga dapat menentukan stok yang tepat untuk produk yang dijual dan dapat meningkatkan keuntungan, penjualan dan pembelian produk. Untuk itu diadakannya penelitian ini dengan tujuan untuk menerapkan metode *data mining* menggunakan algoritma *Naive Bayes Classifier* guna memilih produk terlaris di outlet Matahari *Departement Store* Nagoya Hill Batam. Dengan menggunakan data penjualan yang terkumpul, sistem ini diharapkan dapat meningkatkan keuntungan secara stabil dan menghindari kekurangan stok produk. Melalui analisis menggunakan metode *Naive Bayes Classifier*, penelitian ini berhasil mencapai tingkat akurasi sebesar 67% dan mendapatkan hasil penjualan Tas menjadi penjualan terlaris selama Bulan Januari 2023 samapai Maret 2023 dengan mendapatkan persentase penjualan sebesar 20%.

Kata Kunci: *Data Mining*, *Naive Bayes*, Penjualan Terlaris.

## ABSTRACT

*Technological advances today can be exploited to process data into more useful information. In data collection, information collection is especially useful to maximize profits and develop marketing strategies. One way to increase profits is by using data mining techniques to help business actors in making decisions about stocks, increased profits and more. The Matahari Department Store is the largest retail platform in Indonesia, one of the retail stores located in Batam is the Matahari Department store Nagoya Hill Batam. The transaction data on the store that is still processed does not use a method that causes the processing of product sales data to be less effective and less efficient. Seeing from the number of transactions, a system is needed to predict the sale of the best-selling product as long as it can determine the correct stock for the products sold and can increase the profit, sale and purchase of the product. This research was conducted with the aim of applying data mining methods using the Naive Bayes Classifier algorithm to select the best-selling products in the outlet of the Matahari Nagoya Hill Batam Department Store. By using the collected sales data, the system is expected to increase profits steadily and avoid shortages of product stocks. Through analysis using the Naive Bayes Classifier method, the study achieved an accuracy of 67% and obtained a bag sales result to be the best-selling sale during January 2023 through March 2023 with a sales percentage of 20%.*

*Keywords: Data Mining, Naive Bayes, Bestselling Sales.*

## KATA PENGANTAR

Puji syukur kehadiran Tuhan Yang Maha Esa yang telah melimpahkan segala rahmat dan karuniaNya, sehingga penulis dapat menyelesaikan laporan tugas akhir yang merupakan salah satu persyaratan untuk menyelesaikan program studi strata satu (S1) pada Program Studi Teknik Informatika Universitas Putera Batam”

Penulis menyadari bahwa skripsi ini masih jauh dari sempurna. Karena itu, kritik dan saran akan senantiasa penulis terima dengan senang hati. Dengan segala keterbatasan, penulis menyadari pula bahwa skripsi ini takkan terwujud tanpa bantuan, bimbingan, dan dorongan dari berbagai pihak. Untuk itu, dengan segala kerendahan hati, penulis menyampaikan ucapan terima kasih kepada:

1. Rektor Universitas Putera Batam
2. Dekan Fakultas Teknik dan Komputer Universitas Putera Batam.
3. Kepala Program Studi Teknik Informatika Universitas Putera Batam.
4. Rahmat Fauzi, S.Kom., M.Kom selaku Dosen pembimbing Skripsi pada program studi Teknik Informatika yang telah memberikan bimbingan dan berbagai pengalaman kepada penulis.
5. Segenap Dosen Fakultas Teknik dan Komputer yang telah mendidik dan memberikan ilmu selama kuliah dan seluruh staf yang selalu sabar melayani segala administrasi selama proses penelitian ini.
6. Semua pihak yang telah membantu dan tidak dapat disebutkan satu persatu.

Semoga segala kebaikan dan pertolongan semuanya mendapat berkah dari Allah Swt. dan akhirnya saya menyadari bahwa skripsi ini masih jauh dari kata sempurna, karena keterbatasan ilmu yang saya miliki. Untuk itu saya dengan kerendahan hati mengharapkan saran dan kritik yang sifatnya membangun dari semua pihak demi membangun laporan penelitian ini.

Batam, 26 Juli 2023



(Hidup Perjuangan R)

## DAFTAR ISI

HALAMAN SAMPUL .....	i
HALAMAN JUDUL .....	ii
SURAT PERNYATAAN ORISINALITAS .....	iii
LEMBAR PENGESAHAN.....	iv
ABSTRAK .....	v
ABSTRACT .....	vi
KATA PENGANTAR .....	vii
DAFTAR ISI .....	viii
DAFTAR GAMBAR .....	x
DAFTAR TABEL .....	xi
DAFTAR RUMUS .....	xii
BAB I PENDAHULUAN .....	1
1.1 Latar Belakang .....	1
1.2 Identifikasi Masalah .....	3
1.3 Batasan Masalah.....	4
1.4 Rumusan Masalah .....	4
1.5 Tujuan Penelitian.....	4
1.6 Manfaat Penelitian.....	5
1.6.1 Manfaat Teroritis .....	5
1.6.2 Manfaat Praktis.....	5
BAB II TINJAUAN PUSTAKA .....	6
2.1 <i>Knowledge Discovery in Database</i> .....	6
2.2 <i>Data Mining</i> .....	8
2.2.1 Pengelompokan <i>Data Maining</i> .....	8
2.3 Klasifikasi .....	9
2.4 <i>Naive Bayes</i> .....	10
2.5 <i>Software</i> Pendukung.....	11
2.6 Penelitian Terdahulu.....	12
2.7 Kerangka Penelitian.....	18
BAB III METODE PENELITIAN .....	19
3.1 Desain Penelitian .....	19
3.2 Teknik Pengumpulan Data.....	20
3.2.1 Observasi.....	20
3.2.2 Wawancara .....	20
3.3 Operasional Variabel .....	23
3.4 Model Penelitian .....	24
3.4.1 <i>Data Selection</i> .....	24
3.4.2 <i>Preprocessing</i> .....	28
3.4.3 Transformasi Data .....	28
3.4.4 <i>Data Mining</i> .....	30
3.5 Lokasi dan Jadwal Penelitian .....	30
3.5.1 Lokasi Penelitian .....	30
3.5.2 Jadwal Penelitian .....	31
BAB IV HASIL DAN PEMBAHASAN .....	32
4.1 Hasil Penelitian .....	32



4.1.1 Dataset.....	32
4.1.2 Seleksi Data.....	39
4.2 Analisa Proses Algoritma .....	41
4.2.1 Probabilitas Kriteria Bulan.....	41
4.2.2 Probabilitas Kriteria Harga .....	42
4.2.3 Probabilitas Kriteria Kategori .....	42
4.3 Hasil Pengujian .....	43
4.4 Implementasi Tanagra .....	43
4.4.1 <i>Import data</i> .....	43
4.4.2 <i>Define Status</i> .....	44
4.4.3 <i>SPV Learning</i> .....	45
4.4.4 Perhitungan Hasil .....	46
BAB V KESIMPULAN DAN SARAN.....	48
5.1 Kesimpulan .....	48
5.2 Saran .....	49
DAFTAR PUSTAKA .....	50

## **LAMPIRAN**

Lampiran 1. Surat Balasan Izin Penelitian dari Matahari

Lampiran 2, Surat Penelitian dari UPB

Lampiran 3. Surat Keterangan Penerimaan Jurnal Comasie

Lampiran 4. Hasil Turnitin Skripsi

Lampiran 5. Hasil Turnitin Jurnal

## DAFTAR GAMBAR

<b>Gambar 2. 1</b> Logo <i>Software</i> Tanagra .....	11
<b>Gambar 2. 2</b> Tampilan Aplikasi Tanagra .....	12
<b>Gambar 2. 3</b> Kerangka Penelitian .....	18
<b>Gambar 3. 1</b> Desain Penelitian.....	19
<b>Gambar 3. 2</b> Lokasi Penelitian.....	30
<b>Gambar 4. 1</b> Probabilitas Kriteria Bulan .....	42
<b>Gambar 4. 2</b> Probabilitas Kriteria Harga.....	42
<b>Gambar 4. 3</b> Probabilitas Kriteria Kategori.....	43
<b>Gambar 4. 4</b> Hasil <i>Import</i> Data.....	44
<b>Gambar 4. 5</b> Parameter <i>Input</i> .....	45
<b>Gambar 4. 6</b> <i>Supervised Learning Naive Bayes</i> .....	45
<b>Gambar 4. 7</b> Parameter <i>Train-test</i> .....	46
<b>Gambar 4. 8</b> Hasil <i>Train-test</i> yang Sudah Dilakukan .....	47

## DAFTAR TABEL

<b>Tabel 3. 1</b>	Data Penelitian.....	20
<b>Tabel 3. 2</b>	Atribut pada Data.....	24
<b>Tabel 3. 3</b>	Data Penjualan.....	25
<b>Tabel 3. 4</b>	Hasil Data Transformasi .....	28
<b>Tabel 3. 5</b>	Jadwal Penelitian .....	31
<b>Tabel 4. 1</b>	<i>Dataset</i> .....	32
<b>Tabel 4. 2</b>	Seleksi Data.....	39
<b>Tabel 4. 3</b>	Probabilitas Kelas .....	41
<b>Tabel 4. 4</b>	Jumlah Produk Terjual .....	41
<b>Tabel 4. 5</b>	Tabel Class Prediksi.....	43

## DAFTAR RUMUS

<b>Rumus 2. 1</b> <i>Teorma Bayes</i> .....	10
---	----

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Kemajuan teknologi pada saat ini dapat dimanfaatkan untuk mengolah data menjadi informasi yang lebih bermanfaat. Dalam pengumpulan data, pengumpulan informasi sangat berguna untuk memaksimalkan keuntungan dan mengembangkan strategi pemasaran. Salah satu cara untuk meningkatkan keuntungan adalah dengan menyediakan produk sesuai dengan permintaan pasar supaya tidak terjadi adanya kekurangan stok dan terjadi kerugian.

Dan dengan kemajuan waktu, pemanfaatan teknik *data mining* telah tersebar di berbagai sektor, mulai dari industri bisnis dan perdagangan hingga sektor pendidikan dan telekomunikasi. Terdapat sebuah algoritma *data mining* bernama *Naive Bayes Classifier* yang menjadi sangat bermanfaat bagi para pelaku bisnis dalam mengambil keputusan mengenai stok barang, peningkatan keuntungan dan lainnya. Aplikasi dari algoritma tersebut dapat diterapkan di berbagai sektor bisnis.

Penggunaan *data mining* dapat digunakan untuk meningkatkan keuntungan, salah satu metode yang dapat digunakan adalah *Naive Bayes Classifier*. Metode ini ketika diterapkan pada basis data yang memiliki jumlah data yang besar terbukti memiliki tingkat keakuratan dan kecepatan yang sangat baik. Informasi tersebut dapat dimanfaatkan oleh toko untuk mengidentifikasi produk yang paling laris dibeli oleh konsumen, sehingga membantu toko untuk menemukan dan menargetkan pasar dengan lebih tepat dan akurat (Abdullah et al., 2022).

Matahari *Departement Store* adalah platform ritel terbesar di Indonesia, dengan 155 toko di 81 kota di seluruh Indonesia serta hadir secara online di Matahari.com. Selama lebih dari 60 tahun, Matahari telah menyediakan produk pakaian, kecantikan dan alas kaki yang berkualitas, modis dan terjangkau bagi kelas menengah Indonesia.

Matahari *Departement Store* Nagoya Hill Batam merupakan salah satu outlet / *showroom* Matahari *Departement Store* yang berada di Batam. Matahari *Departement Store* menjual berbagai macam kebutuhan konsumen, sehingga membuat konsumen tertarik melakukan pembelian. Matahari *Departement Store* juga telah menerapkan penjualan secara *online* dengan mempromosikan jualannya. Kategori produk yang di jual berupa sepatu, topi, tas, dan lainnya. Banyaknya transaksi menyebabkan penumpukan data. Data transaksi pada Matahari *Departement Store* yang masih diolah tidak menggunakan sebuah metode menyebabkan pengolahan data penjualan produk menjadi kurang efektif dan kurang efisien. Dilihat dari banyaknya transaksi maka dibutuhkan sistem untuk memprediksi penjualan produk terlaris sehingga dapat menentukan stok yang tepat untuk produk yang dijual dan dapat meningkatkan keuntungan, penjualan dan pembelian produk.

Berdasarkan permasalahan yang telah dipaparkan, maka direkomendasikan sebuah sistem *data mining* yang mampu membantu pihak outlet Matahari *Departement Store* Nagoya Hill Batam dalam penentuan produk terlaris, Untuk itu sistem ini diharapkan meningkatkan keuntungan setiap tahunnya, dapat menyediakan stok produk dengan tepat.

Dalam pembuatan sistem *data mining* menjadi lebih tepat guna, penelitian ini

menggunakan metode *Naive Bayes Classifier* karena dapat diterapkan pada pengolahan data dalam skala besar dan memiliki tingkat keakuratan yang sangat tinggi. Sistem ini dibuat dengan tujuan penjualan pada outlet Matahari *Departement Store* dapat meningkatkan keuntungan setiap tahunnya dengan stabil dan tidak mengalami kekurangan stok produk.

Oleh karena itu, dalam penelitian ini penulis tertarik untuk menjadikannya sebagai fokus penelitian dengan judul “Pendekatan *Data Mining* Untuk Memilih Produk Terlaris Menggunakan Algoritma *Naive Bayes*”.

## **1.2 Identifikasi Masalah**

Sesuai dengan latar belakang penelitian, bahwa dapat diidentifikasi penelitiannya adalah:

1. Permasalahan informasi penjualan di Matahari *Departement Store* Nagoya Hill Batam yang tidak memakai *data mining* sehingga tidak mudah untuk memilih produk terlaris.
2. Belum adanya metode yang digunakan Matahari *Departement Store* Nagoya Hill Batam untuk penentuan stok produk terlaris.
3. Kesulitan pengolahan informasi data transaksi di Matahari *Departement Store* Nagoya Hill Batam, membuat kurang efektif pengelolaan data transaksi penjualan produk.

### 1.3 Batasan Masalah

Dari identifikasi masalah yang sudah dijabarkan di atas, maka penulis membatasi masalah yang dapat diuraikan sebagai berikut:

1. Penelitian ini dilakukan di Matahari *Departement Store* Nagoya Hill Batam.
2. Penelitian ini menggunakan *data mining* dan menggunakan metode *Naive Bayes*.
3. Pembahasan sistem dibatasi pada penentuan penjualan produk terlaris.
4. Data masukan sistem diantaranya dari data penjualan produk pada Matahari *Departement Store* Nagoya Hill Batam.
5. Data keluaran sistem diantaranya laporan hasil pemilihan penjualan produk pada Matahari *Departement Store* Nagoya Hill Batam terlaris.
6. Sistem menggunakan Aplikasi Tanagra.

### 1.4 Rumusan Masalah

Sesuai dengan identifikasi masalah di atas, maka rumusan masalah pada penelitian ini adalah:

1. Bagaimana menerapkan *data mining* dalam memilih produk terlaris menggunakan metode *Naive Bayes*?
2. Bagaimana akurasi yang diperoleh dalam pengimplementasian *data mining* pada pemilihan produk terlaris menggunakan metode *Naive Bayes*?

### 1.5 Tujuan Penelitian

Tujuan dari penelitian ini adalah:

1. Untuk menerapkan *data mining* dalam memilih produk terlaris menggunakan metode *Naive Bayes* menggunakan aplikasi Tanagra.
2. Untuk mengetahui akurasi yang diperoleh dalam pengimplementasian *data*



*mining* pada pemilihan produk terlaris menggunakan metode *Naive Bayes* menggunakan aplikasi Tanagra.

## **1.6 Manfaat Penelitian**

Berdasarkan tujuan penelitian yang akan dicapai maka penelitian ini memiliki manfaat dalam pendidikan, baik secara formal maupun non-formal. Adapun manfaat penelitian ini adalah sebagai berikut:

### **1.6.1 Manfaat Teoritis**

1. Penelitian ini diharapkan mampu memberikan perkembangan ilmu pengetahuan data teknologi khususnya tentang pemilihan produk terlaris menggunakan metode *Naive Bayes*.
2. Penelitian ini dapat menginspirasi mahasiswa dan dapat menjadikannya sebagai referensi dalam melakukan pengolahan data.
3. Penelitian ini diharapkan mampu membantu pembaca yang tertarik melakukan penelitian lebih lanjut di bidang teknik informatika tentang *data mining*, terutama mengenai metode algoritma *Naive Bayes*.

### **1.6.2 Manfaat Praktis**

1. Hasil dari penelitian ini dapat digunakan untuk membantu mempermudah pemilihan produk terlaris dalam pengolahan data yang baik dan efektif.
2. Dengan penelitian ini, diharapkan dapat menjadi bahan rujukan untuk penelitian-penelitian di masa yang akan datang.
3. Dengan penelitian ini, diharapkan dapat meningkatkan pengetahuan bagi peneliti, membantu dalam memecahkan permasalahan secara sistematis.

## BAB II

### TINJAUAN PUSTAKA

#### *2.1 Knowledge Discovery in Database*

*Knowledge Discovery In Database* (KDD) merupakan metode untuk memperoleh pengetahuan dari *database* yang ada. Dalam *database* terdapat tabel - tabel yang saling berhubungan / berelasi. Hasil pengetahuan yang diperoleh dalam proses tersebut dapat digunakan sebagai basis pengetahuan (*knowledge base*) untuk keperluan pengambilan keputusan (Yuli Mardi, 2019).

Istilah *Knowledge Discovery in Database* (KDD) dan data mining seringkali digunakan secara bergantian untuk menjelaskan proses penggalian informasi tersembunyi dalam suatu basis data yang besar. Sebenarnya kedua istilah tersebut memiliki konsep yang berbeda, tetapi berkaitan satu sama lain, dan salah satu tahapan dalam keseluruhan proses KDD adalah data mining. Proses KDD secara garis besar dapat dijelaskan sebagai berikut:

##### *1. Data Selection*

Pemilihan (seleksi) data dari sekumpulan data operasional perlu dilakukan sebelum tahap penggalian informasi dalam *Knowledge Discovery in Database* (KDD) dimulai. Data hasil seleksi yang akan digunakan untuk proses data mining, disimpan dalam suatu berkas terpisah dari basis data operasional.

##### *2. Pre-processing / Cleaning*

Sebelum proses *data mining* dapat dilaksanakan, perlu dilakukan proses *cleaning* pada data yang menjadi fokus *Knowledge Discovery in Database*

(KDD). Proses *cleaning* mencakup antara lain membuang duplikasi data, memeriksa data yang inkonsisten, dan memperbaiki kesalahan pada data, seperti kesalahan cetak. Juga dilakukan proses *enrichment*, yaitu proses “memperkaya” data yang sudah ada dengan data atau informasi lain yang relevan dan diperlukan untuk *Knowledge Discovery in Database* (KDD), seperti data atau informasi eksternal lainnya yang diperlukan.

3. *Transformation*

*Coding* adalah proses transformasi pada data yang telah dipilih, sehingga data tersebut sesuai untuk proses data mining. Proses *coding* dalam *Knowledge Discovery in Database* (KDD) merupakan proses kreatif dan sangat tergantung pada jenis atau pola informasi yang akan dicari dalam basis data.

4. *Data Mining*

*Data mining* adalah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Teknik-teknik, metode-metode, atau algoritma dalam data mining sangat bervariasi. Pemilihan metode atau algoritma yang tepat sangat bergantung pada tujuan dan proses *Knowledge Discovery in Database* (KDD) secara keseluruhan.

5. *Interpretation / Evaluation*

Pola informasi yang dihasilkan dari proses data mining perlu ditampilkan dalam bentuk yang mudah dimengerti oleh pihak yang berkepentingan. Tahap ini merupakan bagian dari proses *Knowledge Discovery in Database* (KDD) yang disebut *interpretation*. Tahap ini mencakup pemeriksaan apakah pola atau informasi yang ditemukan bertentangan dengan fakta sebelumnya.

## 2.2 *Data Mining*

*Data mining* menurut David Hand, Heikki Mannila, dan Padhraic Smyth dari MIT adalah analisa terhadap data (biasanya data yang berukuran besar) untuk menemukan hubungan yang jelas serta menyimpulkannya yang belum diketahui sebelumnya dengan cara terkini dipahami dan berguna bagi pemilik data tersebut (Yuli Mardi, 2019).

*Data mining* bukanlah suatu bidang yang sama sekali baru. Salah satu kesulitan untuk mendefinisikan data mining adalah kenyataan bahwa data mining mewarisi banyak aspek dan teknik dari bidang-bidang ilmu yang dulu sudah mapan terlebih dulu, data mining memiliki akar yang panjang dari bidang ilmu yang berbeda seperti kecerdasan buatan (*artificial intelligent*), *machine learning*, statistik, *database*, dan juga *information retrieval* (Yuli Mardi, 2019).

### 2.2.1 Pengelompokan *Data Mining*

Berdasarkan kegunaannya, aktivitas-aktivitas *data mining* dapat dikelompokkan sebagai berikut (Wanto et al., 2020):

1. Klastering (*Clustering*)

Digunakan untuk mengemlompokkan atau mengidentifikasi data yang memiliki karakteristik tertentu. Contoh algoritma: *K-Means*, *K-Medoids*, dan lainnya.

2. Klasifikasi (*Classification*)

Digunakan untuk menemukan model atau fungsi yang menjelaskan atau membedakan konsep atau kelas data, dengan tujuan untuk dapat memperkirakan kelsa dari suatu object yang labelnya tidak diketahui. Contoh algoritma: *C4.5*, *K-Nearest Neighbor*, *Naive Bayes*, dan lainnya.

3. Asosiasi (*Association*)

Digunakan untuk mengatasi masalah bisnis yang khas, yakni dengan menganalisa tabel transaksi penjualan dan mengidentifikasi produk-produk yang sering kali dibeli bersamaan oleh *customer*. Contoh algoritma: *Apriori*, *Frequent Pattern Growth (FP-Growth)*, dan lainnya.

4. Estimasi (*Estimation*)

Digunakan untuk memperkirakan atau menilai sesuatu hal yang belum pernah ada sebelumnya yang disajikan dalam bentuk hasil kuantitatif (angka). Contoh algoritma: *Regresi Linier*, *Confidence Interval Estimations*, dan lainnya.

5. Prediksi (*Prediction*)

Digunakan untuk memperkirakan atau meramalkan suatu kejadian yang belum pernah terjadi. Contoh algoritma: *Decision Tree*, *KNN*, dan lainnya.

### 2.3 Klasifikasi

Klasifikasi adalah proses menemukan model atau fungsi yang menjelaskan atau membedakan konsep atau data kelas, dengan tujuan untuk dapat mengubah kelas dari suatu obyek yang labelnya tidak diketahui. Dalam mencapai tujuan tersebut, proses klasifikasi membentuk suatu model yang mampu membedakan data ke dalam kelas-kelas yang berbeda berdasarkan aturan atau fungsi tertentu (Siswandi & Fitriana, 2019).

Klasifikasi merupakan suatu pekerjaan menilai objek data untuk memasukkannya ke dalam kelas tertentu dari sejumlah kelas yang tersedia. Dalam klasifikasi ada dua pekerjaan utama yang dilakukan, yaitu : pertama, pembangunan model sebagai *prototype* untuk disimpan sebagai memori dan kedua, penggunaan model tersebut untuk melakukan pengenalan / klasifikasi / prediksi pada suatu

objek data lain agar diketahui di kelas mana objek data tersebut dalam model yang mudah disimpan (Ariyanti & Iswardani, 2020).

#### 2.4 *Naive Bayes*

*Naive Bayes* adalah algoritma klasifikasi untuk menghitung probabilitas dengan menghitung frekuensi dan kombinasi nilai dalam suatu data (Romadhon & Kurniawan, 2021). Metode ini dikemukakan oleh ilmuwan Inggris Thomas Bayes, yaitu memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya sehingga dikenal sebagai *Teorema Bayes* (Pratama et al., 2022).

*Naive Bayes* merupakan salah satu algoritma klasifikasi yang sederhana namun memiliki akurasi yang tinggi. *Naive Bayes* memiliki kelemahannya sangat sensitif dalam. Metode klasifikasi berbasis fitur yang dikembangkan dalam penelitian tersebut menghasilkan akurasi yang baik. *Naive Bayes* memiliki beberapa keunggulan yaitu sederhana, cepat dan memiliki akurasi yang tinggi (Putri et al., 2020). Adapun persamaan dari *Teorema Bayes* ditunjukkan pada rumus berikut ini (Saputra & Herdiansyah, 2022):

$$P(H|X) = \frac{P(X|H) \cdot P(H)}{P(X)} \quad \textbf{Rumus 2. 1 Teorma Bayes}$$

Keterangan:

X: data dengan class yang belum diketahui.

H: hipotesis data menggunakan suatu class spesifik.

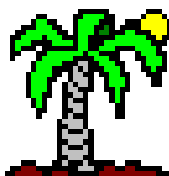
$P(H|X)$ : probabilitas hipotesis H berdasar kondisi X (parteriori probabilitas).

$P(X|H)$ : probabilitas X berdasarkan kondisi pada hipotesis H.

$P(H)$ : probabilitas hipotesis H (prior probabilitas).

$P(X)$ : probabilitas X.

## 2.5 Software Pendukung



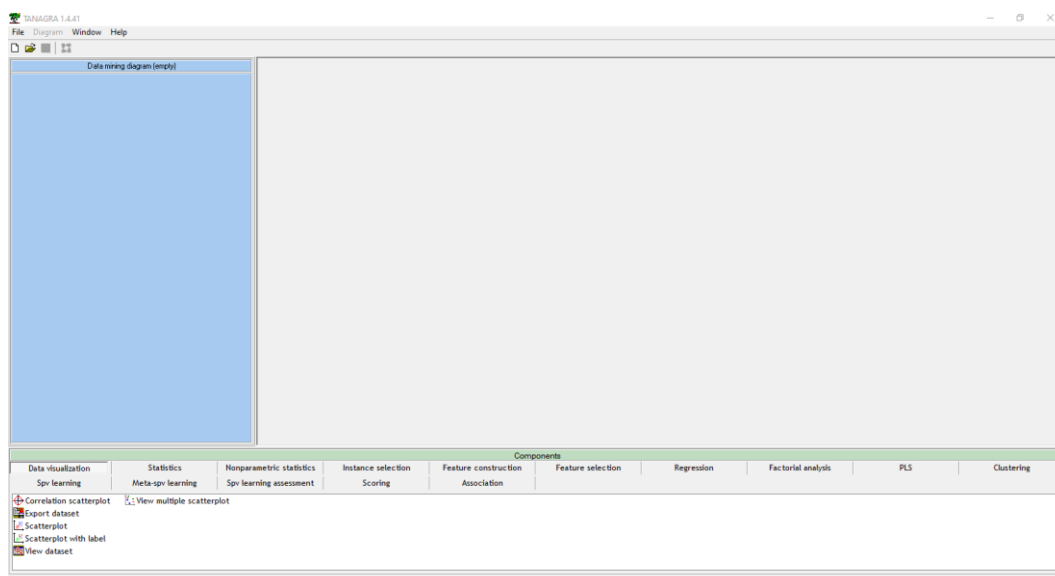
**Gambar 2. 1** Logo *Software Tanagra*

Tanagra adalah rangkaian perangkat lunak pembelajaran mesin gratis untuk tujuan penelitian dan akademik yang dikembangkan oleh Ricco Rakotomalala di Universitas Lumière Lyon 2, Prancis. Tanagra mendukung beberapa tugas penambangan data standar seperti itu sebagai: Visualisasi, Statistik deskriptif, pemilihan *Instance*, pemilihan fitur, konstruksi fitur, regresi, faktor analisis, pengelompokan, klasifikasi dan pembelajaran aturan asosiasi. Tanagra membuat kompromi yang baik antara pendekatan statistik (misalnya uji statistik parametrik dan nonparametrik), metode analisis multivariat (misalnya *factor analysis*, *correspondence analysis*, *cluster analysis*, *regression*) dan teknik pembelajaran mesin (misalnya. *neural network*, *support vector machine*, *decision trees*, *random forest*) (Naik & Samant, 2016).

Tanagra dengan logo seperti pada Gambar 2.1 merupakan perangkat lunak *open source*, sehingga setiap individu dapat dengan mudah mengakses kode sumbernya dan menambahkan algoritma mereka sendiri, asalkan mereka setuju dan mematuhi persyaratan lisensi distribusi perangkat lunak tersebut (Putra et al., 2019). Tampilan aplikasi Tanagra dapat dilihat pada Gambar 2.2.

Salah satu contoh penggunaan aplikasi Tanagra adalah pada penelitian milik (Simanjuntak et al., 2021) dengan judul “*Data Mining* Rekomendasi Pemakaian *Skincare*”. Aplikasi Tanagra pada penelitian ini digunakan untuk pengimplementasian *data mining* dengan metode *Naive Bayes*. Penelitian yang

dilakukan bertujuan untuk membuat rekomendasi penggunaan produk perawatan kulit di Kota Batam.



**Gambar 2. 2** Tampilan Aplikasi Tanagra

## 2.6 Penelitian Terdahulu

1. **Nama Pengarang:** (Rahmatullah et al., 2019), **judul:** **Data Mining Untuk Menentukan Produk Terlaris Menggunakan Metode *Naive Bayes*.**

**Tahun:** Vol.7 No.2 (2019) 57-64, ISSN: 2623 1247

PT. Cipta Niaga Semesta adalah salah satu perusahaan bagian dari Mayora group yang bergerak di bidang distribusi produk makanan dan minuman ringan yang ingin meningkatkan jumlah penjualan produk- produk yang mereka miliki. Untuk membantu perusahaan ini semakin maju diperlukan sebuah sistem yang akan membantu kemajuan perusahaan dalam memaksimalkan penjualan produk mereka, maka peneliti mencoba untuk melakukan suatu penelitian terhadap data produk PT. Cipta Niaga Semesta Sub Branch dengan menggunakan sebuah metode *Naive Bayes Classifier* Pada penelitian ini penulis menggunakan metode pengumpulan data metode



wawancara (*interview*), metode Observasi, metode Dokumentasi dan studi Literatur sedangkan metode pengembangan sistem menggunakan Metode *Waterfall*. Diimplementasi dengan menggunakan bahasa pemrograman menggunakan *Boland Delphi 7* dan *database Microsoft Acces 2010*. Aplikasi produk terlaris yang dibangun menggunakan *metode Naive Bayes Classifier* ini meliputi data produk, *cluster Naive Bayes* serta pelaporan. *Data Mining Untuk Menentukan Produk Terlaris Menggunakan Metode Naive Bayes Pada PT. Cipta Niaga Semesta Sub Branch Kotabumi* ini bertujuan untuk Sistem pengklasifikasian pada produk terlaris dan membantu Kepala *Area Operational Supervisor (AOS)* PT. Cipta Niaga Semesta Sub Branch Kotabumi dalam pemilihan produk terlaris. Sistem ini menghasilkan penentuan produk terlaris dengan 8 atribut yang diambil dari data penjualan selama 2 tahun terakhir.

2. **Nama Pengarang:** (Budiyanto & Dwiasnati, 2018), **judul:** *The Prediction of Best-Selling Product Using Naïve Bayes Algorithm (A Case Study at PT Putradabo Perkasa).*

**Tahun: Vol.5 No.6 (2018) 68-74, ISSN: 2394-2231**

*Data Mining is a technique for processing and extracting large data into information which can form new data. The technique is applied by using Knowledge Discovery Process in Database (KDD). The objective of the research is determining which product is the best-selling in 2018, so that the increase of customer's demand can be anticipated in the following year. In this research, the authors employ classification method in producing the best-selling product information by using algorithms naïve Bayes. There are some*

*variables involved, such as type of goods, brand of goods, quality of goods, price of goods, and Target. Rapid minner Studio 9.0 is a tool for assessing data which calculated to produce a model. The analysis results are expected to be used by the company for preparation supply of the best-selling products. The findings reveal that the level of data accuracy is is 78,33% and the best-selling product based on sales is the IP Camera product with type Infinity I-993V.*

3. **Nama Pengarang:** (Abdullah et al., 2022), **judul: Penerapan Data Mining untuk Memprediksi Jumlah Produk Terlaris Menggunakan Algoritma Naive Bayes Studi Kasus (Toko Prapti).**

**Tahun: Vol.13 No.1 (2022) 20-27, ISSN: 2477-3786**

Toko Prapti adalah perusahaan kecil milik pribadi yang menjual barang kebutuhan pokok,. Selama ini, toko prapti menghasilkan data penjualan setiap hari akan tetapi hasil yang diperoleh menunjukkan toko prapti belum memaksimalkan data tersebut sehingga menjadi penumpukan data. Maka dari itu, peneliti melakukan suatu penelitian terhadap data penjualan produk dengan memanfaatkan dan menerapkan data mining dengan menggunakan algoritma *naïve bayes classifier* untuk mengetahui minat pembelian barang di toko prapti. Penulis melakukan penelitian ini dengan menggunakan metode wawancara, *observasi*, dan studi pustaka tentang metode pengumpulan data. Dalam penelitian ini, penulis menggunakan metode pengembangan sistem *waterfall*.. Penulis mengimplementasikan penelitian ini menggunakan bahasa pemrograman *web* yaitu *PHP* dengan menggunakan *framework codeIgniter* dengan basis data *MySQL*. Sistem yang dibangun dengan algoritma *naïve*

*bayes* ini meliputi data penjualan produk, perhitungan *naïve* dari masing-masing atribut serta pelaporan. Sistem ini menghasilkan 4 atribut yang sangat mempengaruhi hasil klasifikasi. Atribut yang digunakan dalam penelitian ini yaitu atribut adalah triwulan 1, triwulan 2, triwulan 3 dan triwulan 4. Hasil prediksi yang diperoleh dengan menggunakan metode algoritma *naïve bayes* menghasilkan informasi yang dapat digunakan oleh toko untuk mengidentifikasi produk terlaris yang dibeli konsumen sehingga dapat membantu toko prapti untuk menemukan dan menentukan target pasar dengan lebih akurat. Sumber data yang diambil dari 1 tahun sebelumnya dengan keakuratan sistem menggunakan *confusion matrix* menghasilkan *accuracy* 83,3%, *precision* 84,2% dan *recall* 88,9%.

4. **Nama Pengarang:** (Pransiska et al., 2019), **judul: Penerapan Data Mining Prediksi Penjualan Barang Elektronik Terlaris Menggunakan Algoritma Naïve Bayes ( Study Kasus : Planet Cash And Credit Cabang Muara Enim ).Tahun: Vol.1 No.6 (2019) 2157-2169, ISSN: 2685-2683**

Dalam menghadapi persaingan pasar untuk menghasilkan peningkatan pendapatan toko, pihak terkait harus menentukan strategi pemasaran produk yang dijual. Para pelaku bisnis juga harus membutuhkan sebuah analisis untuk memprediksi barang yang paling banyak terjual. Analisis yang dibutuhkan harus bias membantu dan memudahkan para karyawan took untuk mengetahui barang yang mudah habis terjual dengan memprediksi data penjualan pada tahun-tahun sebelumnya dengan menerapkan *data mining* menggunakan algoritma *naive bayes* menggunakan data penjualan tahun 2014 sampai dengan tahun 2016 di PT Solusi Prima Artha planet *cash and*

*credit* cabang Muara Enim.

5. **Nama Pengarang:** (Pratama et al., 2022), **judul: Sistem Klasifikasi Penjualan Produk Alat Listrik Terlaris Untuk Optimasi Pengadaan Stok Menggunakan Naïve Bayes.**

**Tahun: Vol.6 No.4 (2022) 2135-2139, ISSN: 2614-5278**

Optimasi merupakan suatu proses penyelesaian suatu masalah sehingga dapat memberikan kondisi terbaik yang mampu memberikan nilai maksimum atau minimum. Dalam sebuah bisnis, optimasi pengadaan stok menjadi hal yang penting termasuk dalam hal penjualan produk. Jika stok suatu produk kosong maka potensi penjualan menurun. Oleh karena itu perlu suatu metode untuk mengoptimalkan stok sehingga dapat memenuhi permintaan konsumen dan akhirnya dapat meningkatkan penjualan. Data mining dapat diterapkan dalam sistem penjualan dengan membuat model klasifikasi penjualan produk terlaris. Dalam penelitian ini dilakukan klasifikasi penjualan produk terlaris pada toko elektronik menggunakan *Naïve Bayes*. Data yang digunakan dalam penelitian ini adalah data penjualan produk elektronik selama 3 bulan. Pada tahap awal dilakukan *preprocessing* yaitu dengan label *encoding*. Pengujian model dilakukan menggunakan *percentage split* dan *cross validation* dengan beberapa kali percobaan. Melalui penggunaan *percentage split* diperoleh akurasi terbaik sebesar 93,3% dengan perbandingan 30% data uji dan 70% data latih. Akurasi terbaik dengan menggunakan *cross validation* diperoleh sebesar 84% untuk *7-fold*. Sistem klasifikasi yang telah dibuat mampu melakukan klasifikasi produk terlaris setiap triwulan dalam setahun. Melalui penggunaan sistem klasifikasi produk terlaris tersebut maka pihak toko dapat

mengetahui stok produk yang terlaris sehingga stok tidak kosong. Dengan demikian pengadaan stok toko dapat lebih optimal dan penjualan menjadi lebih meningkat.

6. **Nama Pengarang:** (Wijaya & Dwiasnati, 2020), **judul:** **Data Mining dengan Algoritma Naïve Bayes pada Penjualan Obat.**

**Tahun:** Vol.7 No.1 (2020) 2135-2139, ISSN: 2355-6579

Jenis obat yang makin lama makin bervariasi, dari obat yang berharga murah sampai harga yang kalau dilihat sangat kurang masuk akal namun fungsinya sangat bagus. Meningkatnya peredaran jenis obat terutama vitamin, hal ini mendorong penulis untuk melakukan penelitian untuk menentukan produk vitamin mana yang LAKU atau TIDAK LAKU yang bisa di gunakan sebagai pedoman sebuah apotek menentukan jumlah stok barang yang harus ada pada gudang apotek tersebut. Penelitian ini bertujuan untuk mendapatkan nilai accuracy untuk data penjualan obat terutama jenis-jenis vitamin dengan menggunakan algoritma klasifikasi data mining yaitu algoritma Naïve Bayes. Penelitian ini menggunakan tools Rapidminer versi 8 sebagai media untuk menguji data yang akan diolah untuk mendapatkan hasil accuracy dan ROC.

7. **Nama Pengarang:** (Soepriyanto, 2021), **judul:** *Comparative Analysis of K-NN and Naïve Bayes Methods to Predict Stock Prices.*

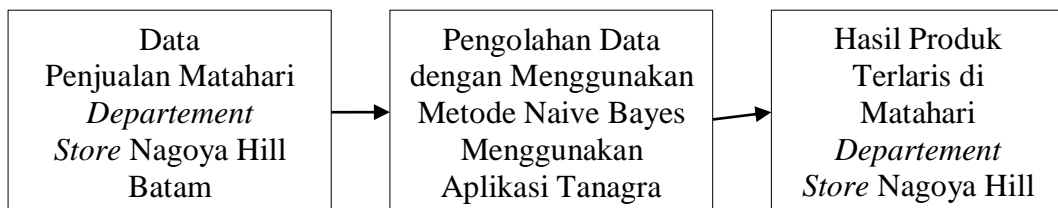
**Tahun:** Vol.2 No.2 (2021) 49-53, ISSN: 2745-9659

*Buying and selling shares is a transaction that is widely carried out at this time, especially buying and selling stocks online which are widely available in the market, to make buying and selling shares require ability or knowledge so that the buying and selling of shares are profitable, to be able to help*

*economic players predict prices. Profit shares or not purchased in the future, this research will conduct stock price predictions using classification methods, namely K-Nearest Neighbor and Naïve Bayes, to predict the stock price data used for one month in minute levels totalling 39065 data, based on prediction results. The highest results obtained were using Naïve Bayes with an accuracy value of 69.38 then the K-Nearest Neighbor method with a K = 5 value of 67.25%, based on these results it can be concluded that the use of the K-Nearest Neighbor and Naïve Bayes methods for prediction share price not yet owned I high accuracy, so it can be combined with other methods or by using other variable predictors.*

## 2.7 Kerangka Penelitian

Untuk mempermudah penelitian ini, peneliti telah membuat suatu kerangka penelitian seperti berikut ini:



**Gambar 2. 3** Kerangka Penelitian

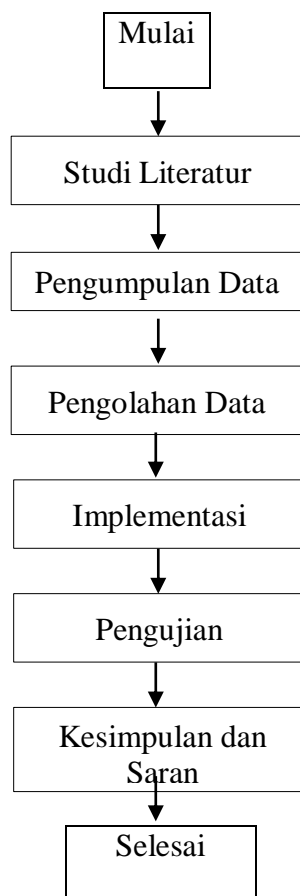
Pada Gambar 2.3 dapat dilihat bahwa data yang akan digunakan sebagai data *input* adalah data penjualan pada Matahari *Departement Store* Nagoya Hill Batam. Kemudian data akan diproses menggunakan metode *Naive Bayes* yang pengimplementasiannya menggunakan aplikasi Tanagra., maka akan mendapatkan hasil prediksi penjualan produk terlaris pada Matahari *Departement Store* Nagoya Hill Batam dan juga hasil akurasi untuk produk terlaris menggunakan *Naive Bayes*.

## BAB III

### METODE PENELITIAN

#### 3.1 Desain Penelitian

Desain penelitian dalam Pendekatan *Data Mining* Untuk Memilih Produk Terlaris Menggunakan Algoritma *Naive Bayes* dapat dilihat pada Gambar 3.1.



**Gambar 3. 1** Desain Penelitian

Pada Gambar 3.1 dapat dilihat urutan desain penelitian yang dimana dimulai dari studi literatur yang mendukung dari berbagai sumber. Kemudian dilanjutkan dengan pengumpulan data, pengolahan data dengan menggunakan metode *Naive Bayes*, dilanjutkan dengan implementasi dan pengujian, tahap terakhir dari penelitian ini yaitu penarikan kesimpulan dan saran.

## 3.2 Teknik Pengumpulan Data

Pengumpulan data dalam penelitian ini yaitu menggunakan teknik wawancara, observasi dan studi literatur. Dataset yang akan digunakan dalam penelitian ini sejumlah 540 data yang dimana sebanyak 90% akan dijadikan data *training* dan sebanyak 10% akan dijadikan data *testing*.

### 3.2.1 Observasi

Dalam penelitian ini, cara observasi yaitu langsung ke Matahari *Departement Store* Nagoya Hill Batam untuk melakukan pengamatan terhadap data transaksi pembelian dan persediaan produk.

### 3.2.2 Wawancara

Dalam penelitian ini, dilakukan wawancara secara langsung kepada karyawan Matahari *Departement Store* Nagoya Hill Batam dengan tujuan untuk memperoleh data penelitian.

Berikut ini data yang telah diperoleh dari Matahari *Departement Store* Nagoya Hill Batam pada Tahun 2023 yang dapat dilihat pada Tabel 3. 1.

**Tabel 3. 1** Data Penelitian

No	Bulan	Nama Produk	Harga	Kategori	Qty	Label
1	Januari	Connexion Slipper Sandals Double Strap	Rp131.940	Sepatu	49	Tidak Laris
2	Januari	Yongki Komaladi Nayla Heels	Rp149.750	Sepatu	47	Tidak Laris
3	Januari	Peter Keiza Heels	Rp299.500	Sepatu	83	Laris
4	Januari	Yongki Komaladi Olivia Heels	Rp183.770	Sepatu	13	Tidak Laris
5	Januari	Yongki Komaladi Elvina Heels	Rp149.750	Sepatu	10	Tidak Laris
6	Januari	Cole Classic Oxford Sepatu Formal Pria	Rp319.920	Sepatu	26	Tidak Laris



7	Januari	Cole Asn Sepatu Oxford Pria	Rp151.960	Sepatu	66	Tidak Laris
8	Januari	Watchout Sepatu Formal Detail Stitching	Rp325.454	Sepatu	54	Tidak Laris
9	Januari	Cole Slipper Sandals Men Peter2	Rp175.920	Sandal	20	Tidak Laris
10	Januari	Cole Flip Flops Men	Rp175.920	Sandal	77	Laris
11	Januari	Cole Slipper Sandals Men Peter3	Rp175.920	Sandal	67	Tidak Laris
12	Januari	Nevada Wide Strap Sandal Pria	Rp119.925	Sandal	69	Tidak Laris
13	Januari	Nevada Wide Strap Sandal Pria	Rp119.900	Sandal	80	Laris
14	Januari	Cole Flip Flops Sandal Game1	Rp134.925	Sandal	43	Tidak Laris
15	Januari	Nevada Chain Solid Backpack Wanita	Rp449.500	Tas	2	Tidak Laris
16	Januari	Connexion Backpack Casual Solid	Rp449.500	Tas	35	Tidak Laris
17	Januari	Bellezza Backpack Elegan	Rp386.910	Tas	88	Laris
18	Januari	Bellezza Backpack Elegan	Rp424.500	Tas	65	Tidak Laris
19	Januari	Amanda Jane's Sepatu LARRY-X20001	Rp199.000	Sepatu	40	Tidak Laris
20	Januari	Jackson Shoes Official Taro 4SG	Rp149.000	Sepatu	140	Laris
21	Februari	JACKSON SHOES OFFICI	Rp251.000	Sepatu	66	Tidak Laris
22	Februari	Cole Classic Oxford Sepatu Formal Pria	Rp319.920	Sepatu	41	Tidak Laris
23	Februari	Cole Asn Sepatu Oxford Pria	Rp151.960	Sepatu	31	Tidak Laris
24	Februari	Character Hello Kitty Slingback Sandals	Rp47.960	Sandal	26	Tidak Laris
25	Februari	Character Flip Flop Hello Kitty	Rp63.960	Sandal	23	Tidak Laris
26	Februari	Disney Minnie Slipper Sandals	Rp109.950	Sandal	31	Tidak Laris

27	Februari	Disney Frozen Flip Flops	Rp53.162	Sandal	91	Laris
28	Februari	Connexion Crossbody Bag Women Perry1	Rp189.950	Tas	58	Tidak Laris
29	Februari	Nevada Crossbody Bag Women Calla1	Rp174.950	Tas	83	Laris
30	Februari	Connexion Crossbody Bag Women Royall1	Rp249.000	Tas	37	Tidak Laris
31	Februari	Nevada Crossbody Bag Women	Rp244.930	Tas	22	Tidak Laris
32	Februari	Nevada Crossbody Bag Women Rush1	Rp249.000	Tas	36	Tidak Laris
33	Februari	Nevada Crossbody Bag Solid	Rp341.910	Tas	90	Laris
34	Februari	Arjuna Weda Busana Muslim Albany	Rp239.700	Pakaian	12	Tidak Laris
35	Februari	Arjuna Weda Busana Muslim Circle	Rp179.700	Pakaian	58	Tidak Laris
36	Februari	Nevada Socks Stripe	Rp29.900	Kaus Kaki	28	Tidak Laris
37	Februari	Nevada Socks Anchor	Rp29.900	Kaus Kaki	46	Tidak Laris
38	Februari	Nevada Caps Baseball Batman	Rp118.930	Topi	129	Laris
39	Februari	d&f Flat Cap Klasik	Rp164.925	Topi	31	Tidak Laris
40	Februari	D&F Topi Baseball Aksen Tekstur	Rp164.925	Topi	44	Tidak Laris
41	Maret	Cardinal Sepatu Juan 2	Rp622.930	Sepatu	17	Tidak Laris
42	Maret	Jim Joker Official Ares 1FA Sepatu Pria	Rp558.000	Sepatu	31	Tidak Laris
43	Maret	Jim Joker Official Ares 2FA Sepatu Pria	Rp558.000	Sepatu	23	Tidak Laris
44	Maret	Disney Mickey Flip Flops	Rp55.960	Sandal	93	Laris
45	Maret	Character Flip Flops Hello Kitty	Rp55.960	Sandal	35	Tidak Laris
46	Maret	Nevada Ankle Strap Sandals Ribbon Strap	Rp68.362	Sandal	141	Laris

47	Maret	Disney Slipper Mickey Sandal Perempuan	Rp109.950	Sandal	17	Tidak Laris
48	Maret	Disney Frozen Classic Flip Flops	Rp63.960	Sandal	45	Tidak Laris
49	Maret	WAIST BAG NYLON PARA	Rp95.920	Tas	130	Laris
50	Maret	President Chest Bag 995-26	Rp247.000	Tas	19	Tidak Laris
51	Maret	Cardinal Gorgeous Bags Hand	Rp319.120	Tas	47	Tidak Laris
52	Maret	Cardinal Shoulder Bag Casual Mars	Rp231.920	Tas	54	Tidak Laris
53	Maret	President Backpack 1588-1-26	Rp423.000	Tas	88	Laris
54	Maret	Nevada Basic Plain Sweater Wanita	Rp124.950	Pakaian	143	Laris
55	Maret	Details Ki Kemeja Rivet Feb Vapor	Rp168.935	Pakaian	61	Tidak Laris
56	Maret	Pipiniko Socks Dino Dasi	Rp14.320	Kaus Kaki	56	Tidak Laris
57	Maret	Pipiniko Socks Gajah Stripe	Rp14.320	Kaus Kaki	35	Tidak Laris
58	Maret	Creart Concept Casual Look Topi Pria	Rp69.000	Topi	17	Tidak Laris
59	Maret	Nevada Caps Bucket Hat Caps	Rp67.960	Topi	5	Tidak Laris
60	Maret	Cardinal Jeans Topi Pria Once Auth	Rp69.900	Topi	91	Laris

### 3.3 Operasional Variabel

Operasional variabel dalam penelitian ini yaitu data transaksi yang ada pada Matahari *Departement Store* Nagoya Hill Batam yang terjadi pada 3 bulan yakni pada tahun 2023..

**Tabel 3. 2** Atribut pada Data

Nama Atribut	Nilai Atribut
Bulan	Januari Februari Maret
Harga	Murah Dikatakan murah apabila harga di bawah Rp.300.000,00 Mahal Dikatakan mahal apabila harga di atas Rp.300.000,00
Kategori	Sepatu Sendal Tas Pakaian Kaos Kaki Topi
Label	Laris Dikatakan laris apabila <i>range</i> penjualan $> 75$ Tidak Laris Dikatakan tidak laris apabila <i>range</i> penjualan $< 75$

### 3.4 Model Penelitian

Model penelitian yang digunakan dalam penelitian ini yaitu, teknik *data mining* dengan menggunakan algoritma *Naive Bayes*. Penelitian ini mengikuti tahapan *Knowledge Discovery in Database* (KDD) untuk menghasilkan informasi sesuai dengan urutan yang sudah ditentukan, berikut adalah tahap-tahapannya:

#### 3.4.1 Data Selection

Data yang digunakan dalam penelitian ini berasal dari data penjualan pada Matahari *Departement Store* Nagoya Hill Batam.

**Tabel 3. 3** Data Penjualan

No	Bulan	Nama Produk	Harga	Kategori	Qty	Label
1	Januari	Connexion Slipper Sandals Double Strap	Rp131.94 0	Sepatu	49	Tidak Laris
2	Januari	Yongki Komaladi Nayla Heels	Rp149.75 0	Sepatu	47	Tidak Laris
3	Januari	Peter Keiza Heels	Rp299.50 0	Sepatu	83	Laris
4	Januari	Yongki Komaladi Olivia Heels	Rp183.77 0	Sepatu	13	Tidak Laris
5	Januari	Yongki Komaladi Elvina Heels	Rp149.75 0	Sepatu	10	Tidak Laris
6	Januari	Cole Classic Oxford Sepatu Formal Pria	Rp319.92 0	Sepatu	26	Tidak Laris
7	Januari	Cole Asn Sepatu Oxford Pria	Rp151.96 0	Sepatu	66	Tidak Laris
8	Januari	Watchout Sepatu Formal Detail Stitching	Rp325.45 4	Sepatu	54	Tidak Laris
9	Januari	Cole Slipper Sandals Men Peter2	Rp175.92 0	Sandal	20	Tidak Laris
10	Januari	Cole Flip Flops Men	Rp175.92 0	Sandal	77	Laris
11	Januari	Cole Slipper Sandals Men Peter3	Rp175.92 0	Sandal	67	Tidak Laris
12	Januari	Nevada Wide Strap Sandal Pria	Rp119.92 5	Sandal	69	Tidak Laris
13	Januari	Nevada Wide Strap Sandal Pria	Rp119.90 0	Sandal	80	Laris
14	Januari	Cole Flip Flops Sandal Game1	Rp134.92 5	Sandal	43	Tidak Laris
15	Januari	Nevada Chain Solid Backpack Wanita	Rp449.50 0	Tas	2	Tidak Laris
16	Januari	Connexion Backpack Casual Solid	Rp449.50 0	Tas	35	Tidak Laris
17	Januari	Bellezza Backpack Elegan	Rp386.91 0	Tas	88	Laris
18	Januari	Bellezza Backpack Elegan	Rp424.50 0	Tas	65	Tidak Laris

19	Januari	Amanda Jane's Sepatu LARRY-X20001	Rp199.000	Sepatu	40	Tidak Laris
20	Januari	Jackson Shoes Official Taro 4SG	Rp149.000	Sepatu	140	Laris
21	Februari	JACKSON SHOES OFFICI	Rp251.000	Sepatu	66	Tidak Laris
22	Februari	Cole Classic Oxford Sepatu Formal Pria	Rp319.920	Sepatu	41	Tidak Laris
23	Februari	Cole Asn Sepatu Oxford Pria	Rp151.960	Sepatu	31	Tidak Laris
24	Februari	Character Hello Kitty Slingback Sandals	Rp47.960	Sandal	26	Tidak Laris
25	Februari	Character Flip Flop Hello Kitty	Rp63.960	Sandal	23	Tidak Laris
26	Februari	Disney Minnie Slipper Sandals	Rp109.950	Sandal	31	Tidak Laris
27	Februari	Disney Frozen Flip Flops	Rp53.162	Sandal	91	Laris
28	Februari	Connexion Crossbody Bag Women Perry1	Rp189.950	Tas	58	Tidak Laris
29	Februari	Nevada Crossbody Bag Women Calla1	Rp174.950	Tas	83	Laris
30	Februari	Connexion Crossbody Bag Women Royall1	Rp249.000	Tas	37	Tidak Laris
31	Februari	Nevada Crossbody Bag Women	Rp244.930	Tas	22	Tidak Laris
32	Februari	Nevada Crossbody Bag Women Rush1	Rp249.000	Tas	36	Tidak Laris
33	Februari	Nevada Crossbody Bag Solid	Rp341.910	Tas	90	Laris
34	Februari	Arjuna Weda Busana Muslim Albany	Rp239.700	Pakaian	12	Tidak Laris
35	Februari	Arjuna Weda Busana Muslim Circle	Rp179.700	Pakaian	58	Tidak Laris
36	Februari	Nevada Socks Stripe	Rp29.900	Kaus Kaki	28	Tidak Laris
37	Februari	Nevada Socks Anchor	Rp29.900	Kaus Kaki	46	Tidak Laris
38	Februari	Nevada Caps Baseball Batman	Rp118.930	Topi	129	Laris

39	Februari	d&f Flat Cap Klasik	Rp164.925	Topi	31	Tidak Laris
40	Februari	D&F Topi Baseball Aksentuksur	Rp164.925	Topi	44	Tidak Laris
41	Maret	Cardinal Sepatu Juan 2	Rp622.930	Sepatu	17	Tidak Laris
42	Maret	Jim Joker Official Ares 1FA Sepatu Pria	Rp558.000	Sepatu	31	Tidak Laris
43	Maret	Jim Joker Official Ares 2FA Sepatu Pria	Rp558.000	Sepatu	23	Tidak Laris
44	Maret	Disney Mickey Flip Flops	Rp55.960	Sandal	93	Laris
45	Maret	Character Flip Flops Hello Kitty	Rp55.960	Sandal	35	Tidak Laris
46	Maret	Nevada Ankle Strap Sandals Ribbon Strap	Rp68.362	Sandal	141	Laris
47	Maret	Disney Slipper Mickey Sandal Perempuan	Rp109.950	Sandal	17	Tidak Laris
48	Maret	Disney Frozen Classic Flip Flops	Rp63.960	Sandal	45	Tidak Laris
49	Maret	WAIST BAG NYLON PARA	Rp95.920	Tas	130	Laris
50	Maret	President Chest Bag 995-26	Rp247.000	Tas	19	Tidak Laris
51	Maret	Cardinal Gorgeous Bags Hand	Rp319.120	Tas	47	Tidak Laris
52	Maret	Cardinal Shoulder Bag Casual Mars	Rp231.920	Tas	54	Tidak Laris
53	Maret	President Backpack 1588-1-26	Rp423.000	Tas	88	Laris
54	Maret	Nevada Basic Plain Sweater Wanita	Rp124.950	Pakaian	143	Laris
55	Maret	Details Ki Kemeja Rivet Feb Vapor	Rp168.935	Pakaian	61	Tidak Laris
56	Maret	Pipiniko Socks Dino Dasi	Rp14.320	Kaus Kaki	56	Tidak Laris
57	Maret	Pipiniko Socks Gajah Stripe	Rp14.320	Kaus Kaki	35	Tidak Laris
58	Maret	Creart Concept Casual Look Topi Pria	Rp69.000	Topi	17	Tidak Laris

59	Maret	Nevada Caps Bucket Hat Caps	Rp67.960	Topi	5	Tidak Laris
60	Maret	Cardinal Jeans Topi Pria Once Auth	Rp69.900	Topi	91	Laris

### 3.4.2 Preprocessing

Pada tahap ini, dilakukan penggabungan data untuk menghubungkan tabel penjualan. Setelah itu, dilakukan pembersihan data untuk menciptakan dataset yang bersih dan siap digunakan dalam tahap selanjutnya, yaitu *mining*. Tahap ini akan melakukan pembersihan data seperti data kosong dan data *missing*, jika tidak ada data yang kosong atau *missing*, data akan diproses ke tahap *transformasi*.

### 3.4.3 Transformasi Data

Pada tahap ini, data yang sudah dibersihkan akan ditransformasi kedalam bentuk-bentuk yang cocok dan sesuai dengan algoritma yang digunakan. Hasil dari data transformasi dapat dilihat pada di bawah ini.

**Tabel 3. 4** Hasil Data Transformasi

Bulan	Harga	Kategori	Label
Januari	Murah	Sepatu	Tidak Laris
Januari	Murah	Sepatu	Tidak Laris
Januari	Murah	Sepatu	Laris
Januari	Murah	Sepatu	Tidak Laris
Januari	Murah	Sepatu	Tidak Laris
Januari	Mahal	Sepatu	Tidak Laris
Januari	Murah	Sepatu	Tidak Laris
Januari	Mahal	Sepatu	Tidak Laris
Januari	Murah	Sandal	Tidak Laris
Januari	Murah	Sandal	Laris
Januari	Murah	Sandal	Tidak Laris
Januari	Murah	Sandal	Tidak Laris
Januari	Murah	Sandal	Laris
Januari	Murah	Sandal	Tidak Laris
Januari	Mahal	Tas	Tidak Laris
Januari	Mahal	Tas	Tidak Laris



Januari	Mahal	Tas	Laris
Januari	Mahal	Tas	Tidak Laris
Januari	Murah	Sepatu	Tidak Laris
Januari	Murah	Sepatu	Laris
Februari	Murah	Sepatu	Tidak Laris
Februari	Mahal	Sepatu	Tidak Laris
Februari	Murah	Sepatu	Tidak Laris
Februari	Murah	Sandal	Tidak Laris
Februari	Murah	Sandal	Tidak Laris
Februari	Murah	Sandal	Tidak Laris
Februari	Murah	Sandal	Laris
Februari	Murah	Tas	Tidak Laris
Februari	Murah	Tas	Laris
Februari	Murah	Tas	Tidak Laris
Februari	Murah	Tas	Tidak Laris
Februari	Murah	Tas	Tidak Laris
Februari	Mahal	Tas	Laris
Februari	Murah	Pakaian	Tidak Laris
Februari	Murah	Pakaian	Tidak Laris
Februari	Murah	Kaus Kaki	Tidak Laris
Februari	Murah	Kaus Kaki	Tidak Laris
Februari	Murah	Topi	Laris
Februari	Murah	Topi	Tidak Laris
Februari	Murah	Topi	Tidak Laris
Maret	Murah	Sepatu	Tidak Laris
Maret	Mahal	Sepatu	Tidak Laris
Maret	Mahal	Sepatu	Tidak Laris
Maret	Murah	Sandal	Laris
Maret	Murah	Sandal	Tidak Laris
Maret	Murah	Sandal	Laris
Maret	Murah	Sandal	Tidak Laris
Maret	Murah	Sandal	Tidak Laris
Maret	Murah	Tas	Laris
Maret	Murah	Tas	Tidak Laris
Maret	Murah	Tas	Tidak Laris
Maret	Mahal	Tas	Laris
Maret	Murah	Pakaian	Laris
Maret	Murah	Pakaian	Tidak Laris
Maret	Murah	Kaus Kaki	Tidak Laris
Maret	Murah	Topi	Laris

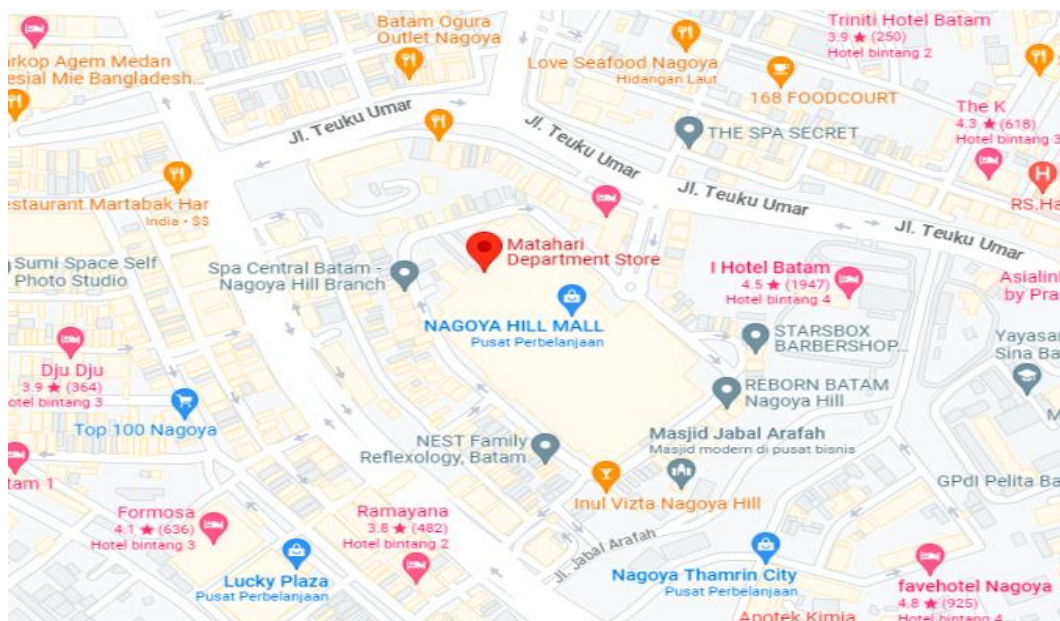
### 3.4.4 Data Mining

Setelah data melewati tahapan *preprocessing* dan transformasi, data akan dilakukan proses *data mining*. Pada tahap ini dilakukan pengujian yaitu dengan melakukan perhitungan dari Tanagra untuk mendapatkan hasil akurat.

## 3.5 Lokasi dan Jadwal Penelitian

### 3.5.1 Lokasi Penelitian

Penelitian ini dilakukan di salah satu *departement store* yang bernama Matahari *Departement Store* Nagoya Hill Batam yang beralamat Jl. Teuku Umar, Lubuk Baja Kota, Kec. Lubuk Baja, Kota Batam, Kepulauan Riau 29444. Lokasi penelitian dapat dilihat pada Gambar 3. 2.



Gambar 3. 2 Lokasi Penelitian

### 3.5.2 Jadwal Penelitian

Penelitian ini dilakukan dengan jangka waktu satu semester seperti yang dijadwalkan dalam kalender akademik Universitas Putera Batam, setelah pengajuan judul penelitian ini diterima. Penelitian ini dilakukan pada bulan April 2023 hingga Agustus 2023. Jadwal yang direncanakan pada penelitian ini dapat dilihat pada Tabel 3.5.2

**Tabel 3.5.2** Jadwal Penelitian

Nama Kegiatan	April				Mei				Juni				Juli				Agustus			
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
Pengajuan Judul	■																			
BAB 1	■	■																		
BAB 2			■																	
BAB 3				■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
BAB 4																	■	■	■	■
BAB 5																	■	■	■	■
Penyelesaian Skripsi																	■	■	■	■