

**IMPLEMENTASI DATA MINING UNTUK  
MEMPREDIKSI PENJUALAN PRODUK TERLARIS  
PADA PETSHOP MENGGUNAKAN ALGORITMA  
NAIVE BAYES**

**SKRIPSI**



**Oleh:**

**Ibrani Gaho**

**200210014**

**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS TEKNIK DAN KOMPUTER  
UNIVERSITAS PUTERA BATAM  
2024**

**IMPLEMENTASI DATA MINING UNTUK  
MEMPREDIKSI PENJUALAN PRODUK TERLARIS  
PADA PETSHOP MENGGUNAKAN ALGORITMA  
NAIVE BAYES**

**SKRIPSI**

**Untuk memenuhi salah satu syarat**

**Memperoleh gelar Sarjana**



**Oleh:**

**Ibrani Gaho**

**200210014**

**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS TEKNIK DAN KOMPUTER  
UNIVERSITAS PUTERA BATAM  
2024**

## SURAT PERNYATAAN ORISINALITAS

Yang bertanda tangan di bawah ini saya :

Nama : Ibrani Gaho  
NPM : 200210014  
Fakultas : Teknik dan Komputer  
Program Studi : Teknik Informatika

Menyatakan bahwa “Skripsi” yang saya buat dengan judul :

**“IMPLEMENTASI DATA MINING UNTUK MEMPREDIKSI PENJUALAN PRODUK TERLARIS PADA PESHOP MENGGUNAKAN ALGORITMA NAÏVE 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, 12 Juli 2024



**Ibrani Gaho**  
200210014

**IMPLEMENTASI DATA MINING UNTUK  
MEMPREDIKSI PENJUALAN PRODUK TERLARIS  
PADA PETSHOP MENGGUNAKAN ALGORITMA  
NAÏVE BAYES**

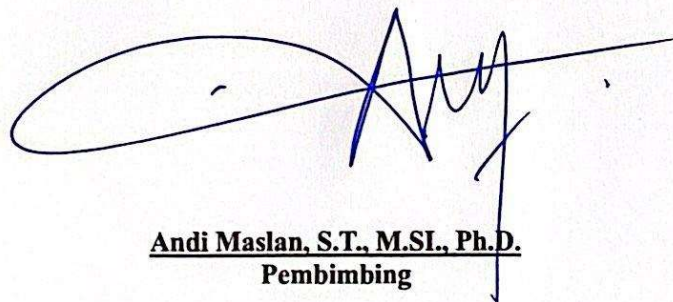
**SKRIPSI**

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

**Oleh :  
Ibrani Gaho  
200210014**

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

**Batam, 12 Juli 2024**



**Andi Maslan, S.T., M.SI., Ph.D.  
Pembimbing**



## ABSTRAK

Penelitian ini dilakukan untuk memprediksi penjualan produk terlaris di petshop Berry Catszone Batam menggunakan algoritma Naïve Bayes, sebuah metode data mining yang efektif untuk analisis pola dan tren data. Masalah utama yang dihadapi adalah sulitnya memprediksi penjualan produk terlaris karena beragamnya produk yang dijual dan analisis manual yang memakan waktu serta sumber daya besar. Dengan metode ini, penelitian bertujuan untuk mengelola stok produk lebih efisien, mengurangi biaya penyimpanan, dan meningkatkan kepuasan pelanggan. Data yang digunakan berasal dari penjualan satu tahun terakhir. Hasil penelitian menunjukkan bahwa implementasi algoritma Naïve Bayes dapat meningkatkan akurasi prediksi penjualan dan memberikan wawasan yang bermanfaat bagi manajemen petshop. Pengolahan data dengan RapidMiner menghasilkan akurasi prediksi sebesar 90,41%. Class precision untuk prediksi produk laris adalah 88,24%, sedangkan untuk produk yang tidak laris adalah 90,00%. Class recall untuk prediksi produk laris mencapai 90,91%, sementara untuk produk yang tidak laris mencapai 92,31%. Implementasi menggunakan RapidMiner efektif dalam memproses data penjualan dan menghasilkan model prediksi yang akurat, membantu pemilik petshop dalam mengelola stok produk, mengurangi biaya penyimpanan, dan meningkatkan rotasi stok serta kepuasan pelanggan.

kata kunci: Data Mining; Naive Bayes; Prediksi Penjualan; Petshop

## ABSTRACT

*This study was conducted to predict the best-selling products at Berry Catszone Batam petshop using the Naïve Bayes algorithm, an effective data mining method for analyzing patterns and trends in data. The main problem faced is the difficulty in predicting the best-selling products due to the variety of products sold and the manual analysis that consumes a lot of time and resources. Using this method, the research aims to manage product stock more efficiently, reduce storage costs, and improve customer satisfaction. The data used comes from sales over the past year. The research results show that the implementation of the Naïve Bayes algorithm can improve the accuracy of sales predictions and provide valuable insights for petshop management. Data processing with RapidMiner resulted in a prediction accuracy of 90.41%. The class precision for predicting popular products is 88.24%, while for non-popular products it is 90.00%. The class recall for predicting popular products reaches 90.91%, while for non-popular products it reaches 92.31%. The implementation using RapidMiner is effective in processing sales data and producing accurate predictive models, helping pet shop owners manage product stock, reduce storage costs, and improve stock rotation and customer satisfaction.*

*Keywords: Data Mining; Naive Bayes; Sales Prediction; Petshop*

## KATA PENGANTAR

Puji syukur dihadirkan kepada 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. Ibu Dr. Nur Elfi Husda, S.Kom., M.SI selaku Rektor Universitas Putera Batam.
2. Bapak Welly Sugianto, S.T., M.M selaku Dekan Fakultas Teknik dan Komputer.
3. Bapak Andi Maslan, S.T., M.SI, PH.D selaku Ketua Program Studi Teknik Informatika sekaligus menjadi pembimbing skripsi pada Program Studi Teknik Informatika Universitas Putera Batam.
4. Bapak Ellbert Hutabri, S.Kom., M.Kom selaku dosen pembimbing Akademik pada Program Studi Teknik Informatika Universitas Putera Batam.
5. Seluruh Dosen dan Staff Universitas Putera Batam.
6. Kedua orang tua tercinta Ayahanda Nasogoloda Gaho dan Ibunda Kasitia Gaho, yang tidak henti-hentinya memberikan kasih sayang dengan penuh cinta, motivasi, serta dukungan dan doa hingga penulis mampu menyelesaikan studinya sampai sarjana.



7. Abang Inspeksi Gaho, Kakak Mawartini P. Fau, Kakak Afri susanti Gaho, Adik Politenes Gaho dan Adik Epa proditus Gaho yang selalu memotivasi penulis.
8. Teman-teman mahasiswa teknik informatika Shintya Rahayu, Mayana Kris Monika, Isnaini Hutagalung, Yohana Ndoya, Elza Maudy Zahra yang telah memberikan masukan serta sarannya agar penelitian ini bisa sukses.
9. Isnaini Hutagalung selaku sahabat yang selalu menemani dan membantu penulis pada saat proses pembuatan sikripsi ini.
10. Teman baik kak Alissah, Zulfikar, Hendri, yang mau mendengarkan segala keluh kesah penulis dalam membuat sikripsi ini.
11. Diri sendiri, Ibrani Gaho karena telah mampu bertahan atas kerja keras dan berjuang sejauh ini dan semangat sehingga tidak pernah menyerah dalam mengerjakan tugas akhir ini dengan menyelesaikan sabaik dan semaksimal mungkin, ini merupakan pencapaian yang patut dibanggakan untuk diri sendiri.

Akhir kata, semoga Tuhan Yang Maha Esa. berkenan membalas segala kebaikan dan selalu mencurahkan hidayah serta taufik-Nya, Amin.

Batam, 12 Juli 2024  
Penulis



Ibrani Gaho

## DAFTAR ISI

SURAT PERNYATAAN ORISINALITAS.....	iii
HALAMAN PENGESAHAN.....	iv
ABSTRACK .....	vii
KATA PENGANTAR.....	viii
DAFTAR ISI.....	x
DAFTAR GAMBAR.....	xiii
DAFTAR TABEL .....	xiv
DAFTAR RUMUS.....	xiii
<b>BAB I PENDAHULUAN.....</b>	<b>1</b>
<b>1.1 Latar Belakang.....</b>	<b>1</b>
<b>1.2 Identifikasi Masalah .....</b>	<b>4</b>
<b>1.3 Batasan Masalah .....</b>	<b>4</b>
<b>1.4 Rumusan Masalah .....</b>	<b>5</b>
<b>1.5 Tujuan Penelitian .....</b>	<b>5</b>
<b>1.6 Manfaat Penelitian.....</b>	<b>5</b>
<b>1.6.1 Manfaat Teoritis.....</b>	<b>5</b>
<b>1.6.2 Manfaat Praktis .....</b>	<b>6</b>
<b>BAB II TINJAUAN PUSTAKA.....</b>	<b>7</b>
<b>2.1 Teori Dasar .....</b>	<b>7</b>
<b>2.2 Knowledge Discovery in Database (KDD).....</b>	<b>7</b>
<b>2.3 Data Mining.....</b>	<b>8</b>
<b>2.2.1 Tahap-tahap data mining.....</b>	<b>9</b>
<b>2.2.2 Pengelompokkan data mining .....</b>	<b>10</b>
<b>2.4 Metode Data Mining .....</b>	<b>10</b>
<b>2.5 Algoritma data mining.....</b>	<b>11</b>
<b>2.4.1 Algoritma Naive Bayes.....</b>	<b>11</b>
<b>2.4.2 Algoritma C4.5.....</b>	<b>13</b>
<b>2.4.3 Algoritma Apriori.....</b>	<b>14</b>
<b>2.4.4 Algoritma K-Nearest Neighbor (KNN).....</b>	<b>15</b>

2.4.5	Algoritma <i>Support Vector Machine</i> (SVM).....	15
2.4.6	Algoritma K-Means.....	16
2.6	<i>Software</i> Pendukung .....	18
2.7	Penelitian Terdahulu.....	19
2.8	Keraangka Pemikiran.....	30
<b>BAB III</b>	<b>METODE PENELITIAN</b> .....	<b>31</b>
3.1	Desaian Penelitian .....	31
3.2	Teknik Pengumpulan Data.....	33
3.2.1	Observasi .....	33
3.2.2	Wawancara.....	33
3.2.3	Studi Pustaka .....	34
3.3	Operasional Variabel .....	34
3.4	Teknik Analisis Data .....	35
3.4.1	Seleksi Atribut.....	37
3.4.2	Cross Validation.....	38
3.4.3	Naive Bayes .....	43
3.4.4	Evaluasi Performance .....	45
3.5	Lokasi dan jadwal Penelitian .....	45
3.5.1	Lokasi Penelitian .....	45
3.5.2	jadwal Penelitian .....	46
<b>BAB IV</b>	<b>HASIL DAN PEMBAHASAN</b> .....	<b>48</b>
4.1	Hasil Analisis data.....	48
4.1.1	Probabilitas Class .....	48
4.1.2	Menghitung Prediksi Probabilitas Class .....	53
4.1.3	Seleksi Data .....	60
4.1.4	Transormasi Data .....	62
4.1.5	Implementasi Data.....	65
4.2	Hasil Pengujian <i>Software RapidMinner</i> .....	70
4.2.1	<i>Import Data</i> .....	70
4.2.2	<i>Import Data Step 2</i> .....	71
4.2.3	<i>Import Data step 3</i> .....	72
4.2.4	<i>Step Akhir Import data</i> .....	73
4.2.5	Hasil Membangun Model Prediksi.....	74

4.2.6	<i>Icon Run</i> .....	76
4.2.7	Hasil Analisis Data.....	76
4.3	Pembahasan.....	79
<b>BAB V KESIMPULAN DAN SARAN</b> .....		80
5.1	Kesimpulan.....	80
5.2	Saran .....	81
<b>DAFTAR PUSTAKA</b> .....		82
<b>LAMPIRAN</b> .....		86

## DAFTAR GAMBAR

<b>Gambar 2.2</b> Rapidminer icon .....	18
<b>Gambar 3.1</b> Teknik Analisis Data.....	36
<b>Gambar 3.2</b> Alur Algoritma Naïve Bayes .....	44
<b>Gambar 3.3</b> Tempat Penelitian .....	46
<b>Gambar 4.1</b> Grafik Hasil Seleksi Data Informasi Gain .....	61
<b>Gambar 4.2</b> Import Data Training dan testing.....	71
<b>Gambar 4.3</b> Import Data Step 2.....	72
<b>Gambar 4.4</b> Import Data Step 3.....	73
<b>Gambar 4.5</b> Import Data Training dan testing Step Akhir.....	74
<b>Gambar 4.6</b> Proses Cross Validation .....	75
<b>Gambar 4.7</b> Subproses Operator Cross Validation .....	75
<b>Gambar 4.8</b> icon Run.....	76
<b>Gambar 4.9</b> Hasil Prediksi Cross validation.....	77
<b>Gambar 4.10</b> Hasil Accuracy.....	77

## **DAFTAR TABEL**

<b>Tabel 2. 1</b> Penelitian Terdahulu .....	21
<b>Tabel 3. 1</b> Atribut Data Raw .....	33
<b>Tabel 3. 2</b> Cara Kerja Cross Validation .....	38
<b>Tabel 3. 3</b> Data Training .....	39
<b>Tabel 3. 4</b> Jadwal Penelitian .....	47
<b>Tabel 4. 1</b> Probabilitas Klasifikasi .....	49
<b>Tabel 4. 2</b> Probabilitas Kategori .....	52
<b>Tabel 4.3</b> Probabilitas Harga.....	53
<b>Tabel 4.4</b> Hasil Prediksi Perhitungan Manual .....	56
<b>Tabel 4. 5</b> Hasil Seleksi data Information Gain .....	60

## DAFTAR RUMUS

<b>Rumus 2. 1</b> Naïve Bayes .....	12
<b>Rumus 2. 2</b> Algoritma C4.5 .....	13
<b>Rumus 2. 3</b> Algoritma Apriori.....	14
<b>Rumus 2. 4</b> Algoritma K-NN .....	15
<b>Rumus 2. 5</b> Algoritma SVM .....	16
<b>Rumus 2. 6</b> Algoritma K-Means .....	17
<b>Rumus 3. 1</b> Information Gain .....	37
<b>Rumus 3. 2</b> Accuracy .....	45