

**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 Mining</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

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

DAFTAR TABEL

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

DAFTAR RUMUS

Rumus 2.1 <i>Teorma Bayes</i>	10
--	----