

**PENERAPAN DATA *MINING* UNTUK KLASIFIKASI
PRODUK TERLARIS PADA BENGKEL ALDO MOTOR
MENGUNAKAN ALGORITMA *NAIVE BAYES***

RINGKASAN SKRIPSI



ALVIN JULIANTO

1914001

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS KATOLIK MUSI CHARITAS**

PALEMBANG

2024

PENERAPAN DATA *MINING* UNTUK KLASIFIKASI PRODUK TERLARIS PADA BENGKEL ALDO MOTOR MENGUNAKAN ALGORITMA *NAIVE BAYES*

Alvin Julianto

Program Studi Sistem Informasi Universitas Katolik Musi Charitas

Jl. Bangau No 60 - Palembang

Email: alvinjulianto080601@gmail.com

ABSTRACT

Aldo Motor workshop is one of the motorbike repair shops in the city of Palembang, South Sumatra. Aldo Motor Workshop provides motorbike repair services and sells motorbike products such as oil, lights, spark plugs, etc. This workshop has a problem, namely that products that are in demand are often out of stock, while for products that are not in demand, the product stock is piling up. To solve this problem, classification data mining analysis was carried out using the Naive Bayes algorithm using the Rapidminer application. The data used is sales data from May 2023 - October 2023. Based on the results of data processing using the rapidminer application, the accuracy value was 86.77%, the precision value was 85.39%, and the recall value was 76%. These results show that the Naive Bayes algorithm can be used to determine the best-selling products at the Aldo motor repair shop.

Keywords: Best selling products, RapidMiner, Naive Bayes, Classification

I. PENDAHULUAN

Perkembangan bisnis di era digital sekarang yang sangat cepat dapat memberikan pengaruh di berbagai bidang, sehingga persaingan antar bisnis juga semakin meningkat dan semakin ketat. Berbagai bidang bisnis yang mendapatkan pengaruh akibat perkembangan teknologi diantaranya yaitu pariwisata, keuangan, pertambangan, perdagangan, jasa, dan yang lain nya. Salah satu dari bidang bisnis yang mendapatkan pengaruh adalah bidang jasa, bidang jasa memiliki banyak bentuk bisnis diantaranya jasa potong rambut (*barbershop*), jasa perbaikan motor atau mobil (bengkel), jasa pengiriman barang, jasa fotografi, dan banyak lain nya. Dalam penelitian ini akan berfokus pada jasa

perbaikan motor atau mobil (bengkel) yang ada di Kota Palembang, Sumatera Selatan.

Salah satu bengkel yang ada di Kota Palembang, Sumatera Selatan adalah Bengkel Aldo Motor yang menyediakan jasa perbaikan untuk motor, bengkel Aldo Motor berlokasi di Jalan Amphibi no 1973D, Sekip Ujung, Palembang. Dalam kesehariannya bengkel Aldo Motor menyediakan stok barang yang digunakan untuk memperbaiki motor seperti oli, lampu, spion, rantai, ban luar, ban dalam, dan bagian motor lainnya. Permasalahan yang dialami oleh bengkel ini adalah pembelian produk yang laris sering kali mengalami kehabisan stok produk

sedangkan untuk barang yang kurang laris stok produknya mengalami penumpukan. Dalam hal ini bengkel memerlukan cara untuk mengatasi permasalahan tersebut agar stok barang dapat disediakan secara efektif dan efisien.

Pada penelitian ini akan menggunakan data *mining* seperti yang ada pada jurnal terdahulu yang ditulis oleh Rahmatullah *et al* (2019), Data *mining* dapat membantu bisnis dalam mengidentifikasi pola pembelian pelanggan mereka. Dengan menganalisis data pembelian pelanggan, bisnis dapat menentukan produk mana yang paling sering dibeli. Dalam Santosa (2007) disebutkan bahwa hasil analisis yang diperoleh dapat digunakan dalam pengambilan keputusan.

Menurut Han dan Kamber (2006) data *mining* adalah proses mengekstraksi atau menambang pengetahuan yang dibutuhkan dari sejumlah besar data. Dalam data *mining* terdapat berbagai fungsi, salah satu dari fungsi data *mining* itu adalah klasifikasi yang akan digunakan dalam penelitian ini. Dalam Fayyad *et al* (1996) klasifikasi (*classification*) adalah menggeneralisasi struktur yang diketahui untuk di aplikasikan pada data-data baru, klasifikasi juga memiliki berbagai contoh, diantaranya adalah klasifikasi penyakit jantung, klasifikasi produk terlaris dan tidak laris, klasifikasi jenis bunga mawar, dan klasifikasi penerima beasiswa pada sekolah menengah pertama.

Dalam data *mining* terdapat banyak algoritma yang bisa dipakai untuk menganalisis klasifikasi, seperti pohon keputusan, *K-nearest neighbor*, *neural networks*, *naïve bayes*, dan *support vector machines*. Pada

penelitian ini akan menggunakan algoritma *naive bayes*. Contoh penelitian yang menggunakan algoritma *naive bayes* untuk mengklasifikasi produk terlaris adalah penelitian H. Aini *et al* (2023), hasil penelitian menunjukkan bahwa perhitungan metode *naive bayes* menggunakan *rapidminer* menghasilkan nilai akurasi 97,50%, presisi 100,00%, dan *recall* 93,48%. Penelitian selanjutnya adalah penelitian Romli *et al* (2019), hasil perhitungan dengan *rapidminer* menghasilkan nilai akurasi 92,19%, presisi 98,39%, dan nilai *recall* 87,14%, ini menunjukkan bahwa pendekatan *naive bayes* cocok untuk penelitian ini.

Berdasarkan kedua penelitian di atas, dapat disimpulkan bahwa algoritma *naive bayes* dapat berfungsi dengan baik dalam proses klasifikasi. Dalam Han *et al* (2012) juga disebutkan kelebihan dari algoritma *naive bayes* adalah algoritma ini tidak membutuhkan data latih yang banyak untuk menentukan skala yang diperlukan pada saat klasifikasi.

II. TINJAUAN PUSTAKA

2.1. Studi Pustaka

2.1.1. Data

Menurut Hoffer *et al.* (2004) data adalah segala sesuatu yang mewakili hal-hal dan peristiwa-peristiwa yang dianggap penting dan bermakna oleh pengguna (*user*).

Menurut Elmasri dan Navathe (2004) data adalah fakta yang dapat dicatat dan memiliki arti.

Dapat disimpulkan dari kedua definisi diatas bahwa data adalah fakta yang dapat dicatat, mewakilkan suatu objek serta memiliki arti yang penting bagi pengguna (*user*).

2.1.2. Data Mining

Menurut Davies dan Paul Beyton (2004), data *mining* merupakan proses penemuan atau penambangan informasi baru dengan mencari pola dari kumpulan data yang besar. *Discovery of knowledge in database* (KDD) adalah istilah tambahan untuk data *mining*, yang melibatkan penggunaan data historis untuk mengidentifikasi pola dan keteraturan dalam data berukuran besar. Setiap pola yang ditemukan diharapkan memberikan manfaat atau keuntungan tertentu. Proses ini dapat berlangsung secara otomatis atau semi-otomatis.

2.1.3. Klasifikasi

Menurut Han dan Kamber (2006), klasifikasi adalah metode untuk mengembangkan model atau fungsi yang menerapkan model matematika pada data atau konsep untuk membuat inferensi statistik untuk data yang tidak diketahui.

2.1.4. Algoritma Naive Bayes

Menurut Bramer (2007) dalam buku yang berjudul “*Principles of Data Mining*” Algoritma *naive bayes* tidak memiliki aturan, metode ini menggunakan suatu cabang matematik yaitu konsep probabilitas. Konsep ini digunakan untuk menemukan probabilitas tertinggi dari proses klasifikasi yang dilakukan, dengan melakukan pengamatan terhadap kemunculan setiap klasifikasi dalam data pelatihan. Algoritma ini adalah salah satu yang populer yang dapat digunakan dalam pada penambangan data.

Dalam Bustami (2014) rumus teorema Bayes adalah :

$$P(H|X) = \frac{P(H|X)P(H)}{P(X)} \quad \text{Persamaan 2.1}$$

Penjelasan dari persamaan 1 sebagai berikut:

X : Data yang memiliki kelas yang tidak diketahui.

H : Hipotesis untuk data X yang termasuk dalam kelas tertentu.

P(H|X) : Probabilitas nilai hipotesis H berdasarkan kondisi X.

P(H) : Nilai probabilitas hipotesis H.

P(X|H) : Probabilitas nilai hipotesis X berdasarkan kondisi H.

P(X) : Nilai probabilitas hipotesis X

a. Perhitungan Prior

Perhitungan yang dilakukan dengan menghitung perbandingan jumlah anggota kelas dengan seluruh sampel data.

$$P = \frac{X}{A} \quad \text{Persamaan 2.2}$$

Keterangan :

P : Angka *prior*

X : Total data untuk setiap kelas

A : Kumpulan data untuk semua kelas

b. Perhitungan Likelihood

Proses menghitung nilai probabilitas/kemungkinan untuk setiap kelas probabilitas kemunculan karakteristik dari karakteristik yang dipilih.

$$L = \frac{F}{B} \quad \text{Persamaan 2.3}$$

Keterangan :

L : Nilai probabilitas

F : Total data atribut untuk setiap kelas

B : Total seluruh data tiap kelas

c. Perhitungan Posterior

Hasil dari perhitungan *likelihood* berupa probabilitas atribut untuk suatu

kelas, yang digunakan untuk menentukan probabilitas bahwa masing-masing atribut dimasukkan ke dalam kelas. Perhitungan kembali dihitung dengan mengalikan kemungkinan karakteristik input dengan kuadrat, dalam proses ini diperoleh probabilitas akhir dari hasil hasil tugas.

$$P(c) \pi P(a|c) \quad \text{Persamaan 2.4}$$

Keterangan :

P(C) : Angka *prior* tiap kelas

P(a|c) : Angka probabilitas

2.1.5. RapidMiner

Menurut C Aprilla *et al* (2013) dalam buku "Belajar Data Mining dengan RapidMiner", Analisis untuk data mining, text mining, dan prediksi disediakan oleh perangkat lunak *open source rapidminer*. Banyak metode deskriptif dan prediktif membantu orang membuat keputusan yang lebih baik. Kurang dari separuh operator penambahan data yang tersedia di *rapidminer* ditujukan untuk *input*, *output*, pra-pemrosesan, dan visualisasi data.

2.2. Studi Pustaka

2.2.1. Implementasi Data Mining dengan Algoritma Naïve Bayes pada Penjualan Obat

Herry dan Saruni (2020) menyatakan bahwa dalam penelitian ini membahas tentang implementasi data mining pada penjualan obat menggunakan algoritma *naive bayes*. Dalam pengumpulan informasi mengenai produk-produk yang tersedia di apotek, misalnya vitamin, seringkali sulit untuk menentukan produk mana yang tersedia di apotek dan mana yang tidak. Hal ini karena suplemen vitamin memiliki umur simpan yang sedikit

berbeda dengan vitamin yang tersedia di apotek. Untuk mengatasi permasalahan diatas maka obat-obatan yang dijual di apotek diklasifikasi berdasarkan variabel-variabel yang diperoleh. Dalam penelitian ini, alat *rapidminer* versi 8 digunakan untuk menguji data yang akan diolah untuk menghasilkan hasil akurasi dan nilai *ROC*. *Accuracy* hasilnya mencapai 88.00%.

2.2.2. Penentuan Tingkat Penjualan Mobil di Indonesia Dengan Menggunakan Algoritma Naive Bayes

Romli *et al* (2019) menyatakan bahwa dalam penelitian ini membahas tentang penerapan algoritma *naive bayes* untuk menghitung tingkat penjualan mobil di Indonesia. Ada banyak merek mobil yang masuk ke pasar otomotif, tetapi GAIKINDO sendiri tidak memberikan klasifikasi pembeda untuk merek mobil yang paling diminati dan paling laris di situs webnya. Oleh karena itu, klasifikasi yang berbeda diperlukan agar pelanggan dapat memilih merek mobil mana yang paling laris berdasarkan kategori yang paling diminati. Selain itu, mendorong para produsen mobil untuk membuat mobil sesuai dengan keinginan pelanggan. Hasil penelitian memberikan kategori dan klasifikasi pembeda merek mobil yang paling diminati oleh pelanggan. Dengan nilai akurasi 92,19%, presisi 98,39%, dan nilai *recall* 87,14%, Metode *naive bayes* dianggap bisa digunakan dalam penelitian ini.

2.2.3. Penerapan Data Mining Dengan Metode Naive Bayes Classifier Pada Penjualan

Barang Untuk Optimasi Strategi Pemasaran

Lubis dan Sitohang (2023) menyatakan bahwa dalam penelitian ini membahas tentang peningkatan strategi pemasaran menggunakan data *mining* dengan teknik klasifikasi *naive bayes*. Salah satu masalah yang dihadapi oleh toko ini adalah data transaksi penjualan hanya disimpan, yang berarti mereka mengumpulkan data tanpa mengetahui manfaatnya. Toko *Universal Homeware* saat ini tidak tahu bagaimana mengolah data untuk memahami pertumbuhan bisnis atau operasi yang sedang berlangsung. Selain itu, ketika data tidak tersedia, sering terjadi pembelian barang yang dibutuhkan pelanggan. Hasil analisis metode yang menggunakan algoritma *naive bayes* dengan 36 titik data menghasilkan akurasi sekitar 80%. Tercatat pada tahun 2020, 38% transaksi produk laris dan 30% produk tidak laris. Program *rapidminer* dapat membantu dalam memberikan bukti bahwa hasil pengujian dengan metode *naive bayes* menghasilkan nilai akurasi yang baik.

III. METODOLOGI PENELITIAN

3.1. Analisis Data

Pada penelitian ini akan digunakan data penjualan Mei 2023 - Oktober 2023 dari bengkel Aldo Motor. Berikut informasi kelas yang digunakan

Tabel 1 Informasi Data Penjualan

Data Penjualan	
Jumlah Data	1857
Atribut	6
Tahun Data	2023

Berdasarkan tabel diatas data penjualan memiliki jumlah 1857 *record* data dan jumlah atribut berjumlah enam atribut. Atribut-

atribut yang ada adalah nomor, tanggal penjualan, nama produk, jenis produk, harga, jumlah penjualan.

3.2. Data Selection

Data penjualan bengkel Aldo Motor dari Bulan Mei 2023 - Oktober 2023 adalah data yang akan dipakai dalam penelitian ini. Data ini didapatkan dari hasil merekap pembukuan beserta nota penjualan dari bengkel Aldo Motor. Data ini terdiri dari nomor, tanggal penjualan, nama produk, jenis produk, harga, dan total penjualan.

3.3. Preprocessing Data

3.3.1. Data Cleaning

Data yang berasal dari dunia nyata (*real world*) biasanya tidak bisa langsung digunakan. Data perlu dibersihkan terlebih dahulu karena adanya kesalahan manusia atau instrumen dalam menginput data.

Data yang sudah di *cleaning* bisa dilihat pada tabel 2 dibawah ini

Tabel 2 Data Yang Sudah Dibersihkan

No	Tanggal Penjualan	Nama Produk	Jenis Produk	Harga	Total Penjualan
1	01/05/2023	Enjin Matic	Oli	45000	1
2	01/05/2023	MPX Gear	Oli	20000	2
3	01/05/2023	Felera X	Oli	45000	4
4	01/05/2023	Filter Udara	Filter Udara	40000	1
5	01/05/2023	MPX 2	Oli	57000	6
6	01/05/2023	Yamalube Silver	Oli	47000	1
7	01/05/2023	Baut	Baut	8000	1
8	01/05/2023	Baterai	Baterai	25000	1
9	01/05/2023	Sheti Matic 2 (ii)	Oli	67000	1
10	01/05/2023	Bola Lampu Depan	Lampu	15000	2
11	01/05/2023	Yamalube Matic	Oli	47000	1
12	01/05/2023	Matic	Oli	47000	1
13	01/05/2023	Ban Dalam	Ban	45000	1
14	02/05/2023	Sheti X	Oli	45000	1
15	02/05/2023	Bola Lampu Depan	Lampu	15000	1
16	02/05/2023	Yamalube Silver	Oli	47000	1
17	02/05/2023	Sheti X	Oli	42000	1
18	02/05/2023	Kunci Kontak	Kunci Kontak	14000	1
19	02/05/2023	MPX 2	Oli	57000	9
20	02/05/2023	MPX Gear	Oli	20000	4
21	02/05/2023	Cincin Spion	Spion	20000	1
22	02/05/2023	Enduro 4T	Oli	45000	1
23	02/05/2023	Yamalube Sport	Oli	58000	1
24	02/05/2023	Yamalube Matic	Oli	47000	2
25	02/05/2023	Filter Udara	Filter Udara	40000	2
26	02/05/2023	Kihahar	Kihahar	40000	2
27	02/05/2023	Yamalube XP	Oli	36000	1
...
1857	31/10/2023	Sheti 17B (ii)	Oli	50000	1

Jumlah atribut pada dataset yang semula berjumlah enam atribut dan 1857 *record* data, tetap memiliki enam atribut dan 1857 *record* data. Hal ini terjadi dikarenakan tidak ada atribut ataupun data yang perlu dihilangkan dalam tahapan ini.

3.3.2. Data Integration

Dalam Han *et al.* (2012) Proses untuk menggabungkan data dari berbagai sumber menjadi satu

kumpulan data. Sebelum melakukan tahapan ini jumlah atribut data pada dataset berjumlah enam atribut dan 1857 record data setelah dilakukan tahapan ini jumlah atribut bertambah menjadi enam atribut dan satu kelas/label serta data berkurang menjadi 279 record data. Atribut yang bertambah yaitu kelas/label status penjualan yang berisi laris atau tidak laris. Kelas/label status penjualan didapatkan dari hasil perhitungan jumlah minimal penjualan produk dengan total penjualan produk. Hasil data *integration* bisa dilihat pada tabel 3 dibawah ini

Tabel 3 Data Yang Sudah Diintegrasikan

nama_produk	nama_produk	jenis_produk	harga	total_penjualan	status_penjualan
Mei	Yamalube Matic	Oli	48000	30	Laris
Mei	Yamalube Silver	Oli	48000	23	Laris
Mei	Yamalube Gold	Oli	51000	2	Tidak Laris
Mei	Yamalube Sport	Oli	60000	9	Tidak Laris
Mei	Yamalube Super Matic	Oli	37000	9	Tidak Laris
Mei	Yamalube XP	Oli	27000	9	Tidak Laris
Mei	Gear Yamalube 1	Oli	20000	8	Tidak Laris
Mei	Gear Yamalube 2	Oli	22000	2	Tidak Laris
Mei	Gear Yamalube 3	Oli	50000	11	Tidak Laris
Mei	MPX 1 (0.8)	Oli	84000	5	Tidak Laris
Mei	MPX 2	Oli	57000	77	Laris
Mei	MPX Gear	Oli	20000	58	Laris
Mei	Enduro 4T	Oli	60000	1	Tidak Laris
Mei	Enduro Racing	Oli	48000	13	Tidak Laris
Mei	Enduro Matic	Oli	47000	6	Tidak Laris
Mei	Mesran X	Oli	48000	35	Tidak Laris
Mei	Shell X	Oli	42000	20	Laris
Mei	Shell UTR (0.8)	Oli	35000	1	Tidak Laris
Mei	Shell UTR (1)	Oli	35000	6	Tidak Laris
Mei	Shell UTR (0.8)	Oli	58000	2	Tidak Laris
Mei	Shell UJ (0.8)	Oli	62000	4	Tidak Laris
Mei	Shell UJ (1)	Oli	67000	3	Tidak Laris
Mei	Shell Matic 2 (0.8)	Oli	80000	20	Laris
Mei	Shell Matic 2 (1)	Oli	52000	1	Tidak Laris
Mei	Federal X	Oli	50000	1	Tidak Laris
Mei	Federal Racing	Oli	50000	1	Tidak Laris
.....
Oktober	Air Radiator	Air Radiator	25000	3	Tidak Laris

3.3.3. Data Reduction

Dalam tahapan ini, data yang kurang relevan akan dikurangi atau dihapus sehingga memperlancar proses pengolahan data. Data yang sudah dikurangi dapat dilihat pada tabel 4 dibawah ini

Tabel 4 Data Yang Sudah Dikurangi

nama_produk	nama_produk	jenis_produk	harga	total_penjualan	status_penjualan
Mei	Yamalube Matic	Oli	48000	30	Laris
Mei	Yamalube Silver	Oli	48000	23	Laris
Mei	Yamalube Gold	Oli	51000	2	Tidak Laris
Mei	Yamalube Sport	Oli	60000	9	Tidak Laris
Mei	Yamalube Super Matic	Oli	37000	9	Tidak Laris
Mei	Yamalube XP	Oli	27000	9	Tidak Laris
Mei	Gear Yamalube 1	Oli	20000	8	Tidak Laris
Mei	Gear Yamalube 2	Oli	22000	2	Tidak Laris
Mei	Gear Yamalube 3	Oli	50000	11	Tidak Laris
Mei	MPX 1 (0.8)	Oli	84000	5	Tidak Laris
Mei	MPX 2	Oli	57000	77	Laris
Mei	MPX Gear	Oli	20000	58	Laris
Mei	Enduro Racing	Oli	48000	13	Tidak Laris
Mei	Enduro Matic	Oli	47000	6	Tidak Laris
Mei	Mesran X	Oli	48000	35	Tidak Laris
Mei	Shell X	Oli	42000	20	Laris
Mei	Shell UTR (0.8)	Oli	35000	1	Tidak Laris
Mei	Shell UTR (1)	Oli	35000	6	Tidak Laris
Mei	Shell UTR (0.8)	Oli	58000	2	Tidak Laris
Mei	Shell UJ (0.8)	Oli	62000	4	Tidak Laris
Mei	Shell UJ (1)	Oli	67000	3	Tidak Laris
Mei	Shell Matic 2 (0.8)	Oli	80000	20	Laris
Mei	Shell Matic 2 (1)	Oli	52000	1	Tidak Laris
Mei	Federal X	Oli	50000	1	Tidak Laris
Mei	Federal Racing	Oli	50000	1	Tidak Laris
.....
Oktober	Air Radiator	Air Radiator	25000	3	Tidak Laris

3.4. Data Transformation

Dalam tahapan ini data diklasifikasikan, data mentah yang ada diganti menjadi data dengan interval

yang lebih kecil. Konsep ini menyederhanakan data *real* dan membuat proses penambangan menjadi lebih efisien. Pola data yang dihasilkan pada proses data mining akan lebih mudah dimengerti.

- Proses transformasi pada atribut harga yang ditransformasikan menjadi atribut harga tetapi dengan kategori murah dan mahal. Hasil transformasi bisa dilihat pada tabel 5 dibawah ini.

Tabel 5 Transformasi Atribut Harga

Bulan	Nama Produk	Jenis Produk	Harga	Total Penjualan	Status Penjualan
Mei	Yamalube Matic	Oli	Murah	30	Laris
Mei	Yamalube Silver	Oli	Murah	23	Laris
Mei	Yamalube Gold	Oli	Murah	2	Tidak Laris
Mei	Yamalube Sport	Oli	Mahal	9	Tidak Laris
Mei	Yamalube Super Matic	Oli	Mahal	3	Tidak Laris
Mei	Yamalube XP	Oli	Murah	9	Tidak Laris
Mei	Gear Yamalube 1	Oli	Murah	9	Tidak Laris
Mei	Gear Yamalube 2	Oli	Murah	8	Tidak Laris
Mei	Gear Yamalube 3	Oli	Murah	2	Tidak Laris
Mei	MPX 1 (0.8)	Oli	Murah	11	Tidak Laris
Mei	MPX 1 (1)	Oli	Mahal	5	Tidak Laris
Mei	MPX 2	Oli	Murah	77	Laris
Mei	MPX Gear	Oli	Murah	58	Laris
Mei	Enduro 4T	Oli	Murah	9	Tidak Laris
Mei	Enduro Racing	Oli	Mahal	1	Tidak Laris
Mei	Enduro Matic	Oli	Murah	13	Tidak Laris
Mei	Mesran X	Oli	Murah	6	Tidak Laris
Mei	Shell X	Oli	Murah	35	Tidak Laris
Mei	Shell UTR (0.8)	Oli	Murah	20	Laris
Mei	Shell UTR (1)	Oli	Murah	1	Tidak Laris
Mei	Shell UTR (0.8)	Oli	Murah	6	Tidak Laris
Mei	Shell UJ (0.8)	Oli	Murah	9	Tidak Laris
Mei	Shell UJ (1)	Oli	Murah	2	Tidak Laris
Mei	Shell Matic 2 (0.8)	Oli	Mahal	1	Tidak Laris
Mei	Shell Matic 2 (1)	Oli	Mahal	3	Tidak Laris
Mei	Federal X	Oli	Murah	20	Laris
Mei	Federal Racing	Oli	Murah	1	Tidak Laris
.....
Oktober	Air Radiator	Air Radiator	Murah	3	Tidak Laris

Kategori harga dapat dilihat pada tabel 6 dibawah ini.

Tabel 6 Kategori Harga

Range harga	Kategori
Diatas Rp 60000	Mahal
Dibawah Rp 60000	Murah

- Proses transformasi pada atribut total penjualan yang di transformasikan menjadi atribut total penjualan tetapi dengan kategori banyak, sedang dan sedikit. Hasil transformasi dapat dilihat pada tabel 7 dibawah ini.

Tabel 7 Transformasi Atribut Total Penjualan

Bulan	Nama Produk	Jenis Produk	Harga	Total Penjualan	Status Penjualan
Mei	Yamalube Matic	Oli	Murah	30	Banyak
Mei	Yamalube Silver	Oli	Murah	23	Banyak
Mei	Yamalube Gold	Oli	Murah	2	Tidak Laris
Mei	Yamalube Sport	Oli	Mahal	9	Tidak Laris
Mei	Yamalube Super Matic	Oli	Mahal	3	Tidak Laris
Mei	Yamalube XP	Oli	Murah	9	Tidak Laris
Mei	Gear Yamalube 1	Oli	Murah	9	Tidak Laris
Mei	Gear Yamalube 2	Oli	Murah	8	Tidak Laris
Mei	Gear Yamalube 3	Oli	Murah	2	Tidak Laris
Mei	MPX 1 (0.8)	Oli	Murah	11	Tidak Laris
Mei	MPX 1 (1)	Oli	Mahal	5	Tidak Laris
Mei	MPX 2	Oli	Murah	77	Banyak
Mei	MPX Gear	Oli	Murah	58	Banyak
Mei	Enduro Racing	Oli	Mahal	1	Tidak Laris
Mei	Enduro Matic	Oli	Murah	13	Tidak Laris
Mei	Mesran X	Oli	Murah	6	Tidak Laris
Mei	Shell X	Oli	Murah	35	Tidak Laris
Mei	Shell UTR (0.8)	Oli	Murah	20	Laris
Mei	Shell UTR (1)	Oli	Murah	1	Tidak Laris
Mei	Shell UTR (0.8)	Oli	Murah	6	Tidak Laris
Mei	Shell UJ (0.8)	Oli	Murah	9	Tidak Laris
Mei	Shell UJ (1)	Oli	Murah	2	Tidak Laris
Mei	Shell Matic 2 (0.8)	Oli	Mahal	1	Tidak Laris
Mei	Shell Matic 2 (1)	Oli	Mahal	3	Tidak Laris
Mei	Federal X	Oli	Murah	20	Banyak
Mei	Federal Racing	Oli	Murah	1	Tidak Laris
.....
Oktober	Air Radiator	Air Radiator	Murah	3	Tidak Laris

Kategori total penjualan dapat dilihat pada tabel 8 dibawah ini.

Tabel 8 Kategori Total Penjualan

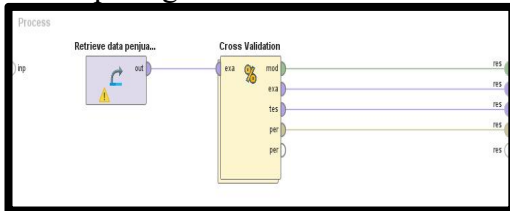
Range total penjualan	Kategori
> = 20	Banyak
10-19	Sedang
1-9	Sedikit

3.5. Data Mining

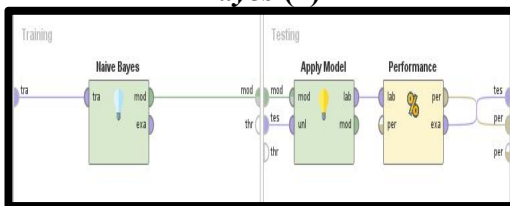
Penelitian ini menggunakan model klasifikasi dan algoritma *naive bayes*, Proses pemodelan dilakukan menggunakan aplikasi *RapidMiner* versi 10.3.

3.6. Pengujian Menggunakan Aplikasi *RapidMiner*

Pengujian selanjutnya akan menggunakan *tools* untuk mempermudah dalam proses pengujian algoritma yaitu aplikasi *rapidminer*. Pengujian algoritma *naive bayes* pada aplikasi *rapidminer* akan menggunakan seluruh data penjualan yang sudah didapatkan. Model yang dipakai dapat dilihat pada gambar 1 dan 2.



Gambar 1 Model Algoritma *Naive Bayes* (1)



Gambar 2 Model Algoritma *Naive Bayes* (2)

Gambar di atas menunjukkan model algoritma *naive bayes* yang digunakan untuk proses pengujian aplikasi *rapidminer*, hasil pengujian menggunakan *rapidminer* dapat dilihat pada gambar 3 di bawah ini.

accuracy: 86.77% +/- 5.04% (micro average: 86.74%)			
	true Laris	true Tidak Laris	class precision
pred. Laris	76	24	76.00%
pred. Tidak Laris	13	166	92.74%
class recall	85.39%	87.37%	

Gambar 3 Hasil Pengujian Algoritma *Naive bayes* dengan *Rapidminer*

Untuk mengukur performansi akan dilakukan dengan melakukan perhitungan menggunakan nilai yang dihasilkan dari tabel dibawah.

Tabel 9 *Confusion Matrix* Pengujian dengan *RapidMiner*

Prediction Sample			
Actual Sample		True	False
	True	76	24
	False	13	166

➤ Accuracy

$$Accuracy = \frac{76+166}{76+166+13+24} \times 100\% = 86,77\%$$

➤ Precision

$$Precision = \frac{76}{76+13} \times 100\% = 85,39\%$$

➤ Recall

$$Recall = \frac{76}{76+24} \times 100\% = 76\%$$

3.7. Interpretation/Evaluation

Berdasarkan hasil evaluasi performa algoritma *naive bayes* dengan *confusion matrix*, hasilnya adalah :

- accuracy 86,77% = klasifikasi yang baik (*good classification*).
- precision 85,39% = presisi yang tinggi
- recall 76% = recall yang tinggi

Dibawah ini akan diberikan hasil pengolahan data penjualan bengkel aldo motor dengan menggunakan algoritma *naive bayes* pada aplikasi *rapidminer*. Hasil bisa dilihat pada tabel 10 dibawah ini.

Tabel 10 Hasil Pengolahan dengan RapidMiner

...	Status Pe...	prediction(Statu...	confide...	confide...	Bulan	Nama Barang	Jenis Barang	Harga	Total Penjualan
1	Tidak Laris	Laris	0.695	0.305	Mei	Busi Censo	Busi	Murah	Sedang
2	Laris	Laris	1.000	0.000	Juni	MPX 2	Oli	Murah	Banyak
3	Tidak Laris	Tidak Laris	0.001	0.999	Juni	Enduro 4T	Oli	Murah	Sedikit
4	Tidak Laris	Tidak Laris	0.387	0.613	Juni	Busi Censo	Busi	Murah	Sedikit
5	Tidak Laris	Tidak Laris	0.000	1.000	Juli	Yamalube Gold	Oli	Murah	Sedikit
6	Tidak Laris	Tidak Laris	0.000	1.000	Juli	Enduro 4T	Oli	Murah	Sedikit
7	Laris	Laris	1.000	0.000	Juli	Shell Matic	Oli	Murah	Banyak
8	Tidak Laris	Tidak Laris	0.000	1.000	Juli	Federal Racing	Oli	Murah	Sedikit
9	Tidak Laris	Tidak Laris	0.000	1.000	Agustus	Yamalube Super Matic	Oli	Mahal	Sedikit
10	Laris	Laris	1.000	0.000	Agustus	MPX 2	Oli	Murah	Banyak
11	Tidak Laris	Tidak Laris	0.349	0.651	Agustus	Shell LTR (0.8)	Oli	Murah	Sedang
12	Tidak Laris	Laris	0.553	0.447	Agustus	Busi Censo	Busi	Murah	Sedang
13	Tidak Laris	Tidak Laris	0.000	1.000	Septem.	Yamalube Sport	Oli	Mahal	Sedikit
14	Tidak Laris	Tidak Laris	0.000	1.000	Septem.	Yamalube VP	Oli	Murah	Sedikit
15	Tidak Laris	Tidak Laris	0.000	1.000	Septem.	Shell Matic 2 (0.8)	Oli	Mahal	Sedikit

IV. PEMBAHASAN

Dari hasil olah data melalui aplikasi *rapidminer* dengan menggunakan algoritma *naive bayes* dapat disimpulkan bahwa atribut total penjualan memiliki pengaruh yang besar dalam menentukan laris dan tidak laris nya produk. Produk yang total penjualannya masuk kedalam kategori “banyak” memiliki kemungkinan “laris” lebih tinggi daripada kategori total penjualan lainnya, untuk produk yang masuk kedalam kategori total penjualan “sedang” maka akan dilihat pada atribut harga, jenis produk, dan bulan untuk menentukan laris dan tidak larisnya, begitu juga dengan kategori total penjualan “sedikit” maka akan dilihat pada atribut harga, jenis produk dan bulan untuk menentukan laris dan tidak laris.

Jumlah laris dan tidak laris pada hasil akhir dari pengolahan dengan menggunakan algoritma *naive bayes* pada *rapidminer* terbagi menjadi 100 produk laris dan 179 produk tidak laris, yang jika dijadikan persentase menjadi 35,85% produk laris, dan 64,15% produk tidak laris. Produk yang paling laris pada 5 teratas yaitu :

- MPX Gear (Oli)
- Shell Matic (Oli)
- MPX 2 (Oli)
- Yamalube Matic (Oli)
- Ban Dalam (Ban)

V. KESIMPULAN DAN SARAN

5.1. Kesimpulan

Penelitian untuk menentukan produk laris dan tidak laris dengan memakai metode klasifikasi dan algoritma *naive bayes* pada aplikasi *rapidminer* dapat disimpulkan bahwa :

1. Algoritma *naive bayes* dapat dipakai untuk melakukan penelitian terkait produk laris dan tidak laris pada bengkel Aldo Motor.
2. Oli motor pada kota Palembang memiliki tingkat penggunaan yang tinggi.

5.2. Saran

Berdasarkan hasil yang diperoleh dari hasil diatas, saran yang dapat dibuat adalah :

1. Untuk penelitian selanjutnya data yang digunakan bisa dilakukan penambahan pada berbagai data penjualan bengkel yang ada di Kota Palembang untuk melihat produk yang paling diminati oleh pelanggan bengkel di Kota Palembang sehingga hasilnya bisa menjadi lebih akurat dan valid.
2. Membandingkan hasil algoritma klasifikasi lainnya seperti ID3, C.45, *Neural Network*, *Support Vector Machine* (SVM), dan Apriori.

VI. DAFTAR PUSTAKA

- Aprilla, C Dennis *et al.*, (2013). Belajar *Data Mining* dengan *Rapidminer*, *e-book*, Jakarta.
- Bramer, Max. (2007). *Principles of Data Mining*, *Springer Science*.
- Bustami. (2014). Penerapan Algoritma *Naive Bayes* untuk Mengklasifikasi Data Nasabah Asuransi. *Jurnal Informatika*, Vol. 8, No. 1.

- Davies, and Paul Beynon. (2004). *Database System Third Edition, Palgrave Macmillan, New York.*
- Elmasri, R., & Navathe, S. B. (2004). *International Edition : Fundamentals of Database System Fourth Edition. Boston: Peason Education.*
- Fayyad, Usama. (1996). *From Data Mining to Knowledge Discovery in Database.*
- H Aini *et al.* (2023). Penerapan Algoritma *Naive Bayes* Untuk Menentukan Klasifikasi Produk Terlaris Pada Penjualan Voucher Kuota Di Edi Cell. *Journal of Social Science Research*, Vol. 3, No. 2.
- Han, J. & M. Kamber. (2006). *Data Mining Concepts and Techniques Second Edition.* San Francisco: Morgan Kaufmann.
- Han, J, Kamber, M, & Pei, J. (2012). *Data Mining: Concept and Techniques, Third Edition. Wattham: Morgan Kauffman Publishers.*
- Hoffer, Jeffrey A., Prescott, Marry B., & McFadden, Fred R. (2005). *Modern Database Management (7th ed., p5). Prentice Hall.*
- Lubis, Putra, Rama, Kahfi., & Sunarsan Sitohang. (2023). Penerapan Data Mining Dengan Metode *Naive Bayes Classifier* Pada Penjualan Barang Untuk Optimasi Strategi Pemasaran. *Jurnal COMASIE (Computer and Industrial Science Industrial Engineering)*, Vol. 9, No. 7.
- Rahmatullah, Sidik. Mukrim., Monica Nindy Pramitha. (2019). *Data Mining Untuk Menentukan Produk Terlaris Menggunakan Metode Naive Bayes*, Vol. 7, No. 2.
- Romli, Ikhsan. Esem Pusnawati., & Arif Siswandi. (2019). Penentuan Tingkat Penjualan Mobil Di Indonesia Dengan Menggunakan Algoritma *Naive Bayes*. *Prosiding Seminar Nasional Teknologi dan Sains (SNasTeks)*, Vol. 1, No. 1.
- Santosa, Budi. (2007). *Data Mining : Teknik Pemanfaatan Data untuk Keperluan Bisnis. Penerbit: Graha Ilmu, Yogyakarta.*
- Wijaya & Dwiasnati. (2020). Implementasi *Data Mining* dengan Algoritma *Naive Bayes* pada Penjualan Obat. *JURNAL INFORMATIKA*, Vol. 7, No. 1.