



Klasifikasi Penjualan Produk Terlaris Pada Kedai Ira Dengan Menggunakan Algoritma Naïve Bayes Dan Algoritma K-Nearest Neighbor

Intan Purwasih^{*1}, Kiki Setiawan², Frencis Matheos Sarimole³, Tundo⁴

^{*1, 2, 3, 4}Jurusan Sistem Informasi, Sekolah Tinggi Ilmu Komputer Cipta Karya Informatika, Jakarta, Indonesia

*Email Penulis Korespondensi: intanpurwasih780@gmail.com

Abstrak

Di dunia ritel dan teknologi saat ini, persaingan sangat kompetitif. Dengan pertumbuhan dan tersebarnya bisnis ritel di setiap wilayah, kebutuhan konsumen semakin meningkat, dan bisnis ritel berlomba-lomba untuk mengembangkan bisnis mereka dengan menggunakan teknologi saat ini. Data transaksi penjualan harian yang terus meningkat menyebabkan banyaknya penyimpanan. Toko Ira memiliki lebih dari 228 rekaman data transaksi penjualan dari tahun 2023 hingga 2024 yang belum digunakan. Data memerlukan banyak ruang penyimpanan. Selain itu, data tersebut belum digunakan dengan cara yang efektif. Berdasarkan masalah, tujuan penelitian ini adalah untuk menentukan barang mana yang paling laris dengan menggunakan data mining untuk mengklasifikasikan data transaksi penjualan. Studi kasus ini adalah studi kualitatif. Hasil penelitian ini menunjukkan bahwa algoritma K-Nearest Neighbors (KNN) dengan pembagian data 50:50 lebih efektif dalam memprediksi dan mengklasifikasikan penjualan produk laris dan tidak laris di toko kedai ira. Hasilnya menunjukkan bahwa algoritma Naive Bayes memiliki akurasi sebesar 89,91%, sedangkan algoritma K-Nearest Neighbors (KNN) dengan pembagian data 50:50 memiliki akurasi sebesar 89,91%.

Kata kunci— Toko Ira, Penjualan, Algoritma Naïve Bayes, Algoritma K-Nearest Neighbor

Abstract

In today's world of retail and technology, competition is fiercely competitive. With the growth and opening of retail businesses in every region, consumer needs are increasing, and retail businesses are competing to develop their businesses using current technology. Daily sales transaction data continues to increase, causing a lot of storage. Toko Ira has more than 228 sales transaction data records from 2023 to 2024 that have not been used. Data requires a lot of storage space. Additionally, the data has not been used in an effective way. Based on the problem, the aim of this research is to determine which goods are selling best by using data mining to classify sales transaction data. This case study is a qualitative study. The results of this research show that the K-Nearest Neighbors (KNN) algorithm with a 50:50 data division is more effective in predicting and classifying sales of best-selling and non-selling products in IRA stores. The results show that the Naive Bayes algorithm has an accuracy of 89.91%, while the K-Nearest Neighbors (KNN) algorithm with a 50:50 data division has an accuracy of 89.91%.

Keywords— *Ira Store, Sales, Naïve Bayes Algorithm, K-Nearest Neighbor Algorithm*

1. PENDAHULUAN

Dalam dunia bisnis yang semakin kompetitif, memahami produk yang paling diminati oleh konsumen menjadi sangat penting bagi perusahaan. Produk terlaris tidak hanya mencerminkan preferensi dan kebutuhan pasar saat ini, tetapi juga memberikan wawasan berharga tentang tren dan perilaku konsumen, khususnya di bidang kuliner. Suatu usaha yang sudah menjadi tren dan juga menjadi tantangan bagi para pembisnis pemula untuk membuat produk baru atau yang sudah lama berkecimpung dalam bisnis akan dituntut untuk mengembangkan produk baru agar dapat bertahan dalam persaingan bisnis. Produk terlaris adalah produk yang mencapai tingkat penjualan tertinggi dalam kategori tertentu selama periode waktu tertentu. Keberhasilan produk ini dipengaruhi oleh berbagai faktor seperti kualitas, rasa, harga, tren kesehatan, dan strategi pemasaran [1]. Di tengah persaingan yang ketat, mengetahui produk makanan mana yang paling diminati oleh konsumen dapat memberikan keuntungan kompetitif yang signifikan bagi produsen dan pengecer. Bisnis kuliner mie ayam saat ini tidak memiliki bisnis baru jika sudah lama ada. Karena itu, pemilik bisnis mie ayam mungkin menghadapi tantangan untuk sedikit berinovasi dalam membuat produk mereka. Produk terlaris didefinisikan sebagai barang atau jasa yang mencatatkan penjualan tertinggi dalam periode waktu tertentu. Produk ini dapat berasal dari berbagai kategori, mulai dari teknologi, fashion, makanan dan minuman, hingga layanan digital. Keberhasilan suatu produk menjadi terlaris biasanya dipengaruhi oleh kualitas, harga, strategi pemasaran yang efektif, serta kepuasan dan loyalitas pelanggan [2].

Produk tidak hanya mencerminkan preferensi dan kebutuhan pasar saat ini, tetapi juga memberikan wawasan berharga tentang tren dan perilaku konsumen. Dalam pendahuluan ini, kita akan membahas definisi produk terlaris, faktor-faktor yang mempengaruhi popularitas suatu produk, serta pentingnya produk terlaris bagi pertumbuhan dan keberlanjutan bisnis. Kelezatan dan variasi rasanya membuat mie ayam menjadi pilihan favorit banyak orang. Dalam industri kuliner, mengidentifikasi produk terlaris seperti mie ayam sangat penting untuk meningkatkan penjualan dan kepuasan pelanggan. Namun, di era digital ini, masih banyak usaha yang mengandalkan pengorderan secara manual, yang membutuhkan waktu dan rentan terhadap kesalahan manusia. Untuk menyelesaikan masalah ini, penggunaan teknologi dan algoritma analisis data menjadi solusi yang relevan dan efisien [3].

Dalam Penelitian Sebelumnya, Hasil dari penelitian tentang mengklasifikasikan data transaksi penjualan, produk diklasifikasikan menjadi kategori laris dan tidak laris. Dengan 6.043 data pelatihan, peneliti melakukan pengujian dan menemukan bahwa hasil validasi split terbaik adalah 0,7 dan keakuratan sebesar 93,82%. Untuk memenuhi keinginan konsumen, perusahaan harus dapat mengklasifikasikan jenis produk yang disukai dan paling sering dibeli konsumen. Hal ini sangat penting untuk dilakukan untuk menghindari tumpukan stok yang berisiko kerugian. Oleh karena itu, analisis data penjualan tahun sebelumnya dan analisis yang dapat digunakan diperlukan untuk mendapatkan informasi dari banyak data [4].

Dengan menggunakan algoritma *Naive Bayes*, perusahaan dapat mengidentifikasi pola konsumsi dan preferensi konsumen berdasarkan data historis. Hal ini memungkinkan pembisnis untuk merancang strategi pemasaran yang lebih efektif, mengoptimalkan pengelolaan stok, dan merespons tren pasar dengan lebih cepat. Selain itu, *Naive Bayes* dapat membantu dalam segmentasi pasar, memungkinkan perusahaan untuk menargetkan kelompok konsumen tertentu dengan produk yang paling sesuai dengan preferensi mereka.

2. METODE PENELITIAN

3.1 Data Penelitian

Data yang digunakan penelitian ini adalah kualitatif dan terdiri dari semua informasi yang diperoleh dari responden atau dari dokumen-dokumen lainnya untuk keperluan penelitian. Berapa tahapan penelitian dilakukan, yaitu:

- Observasi

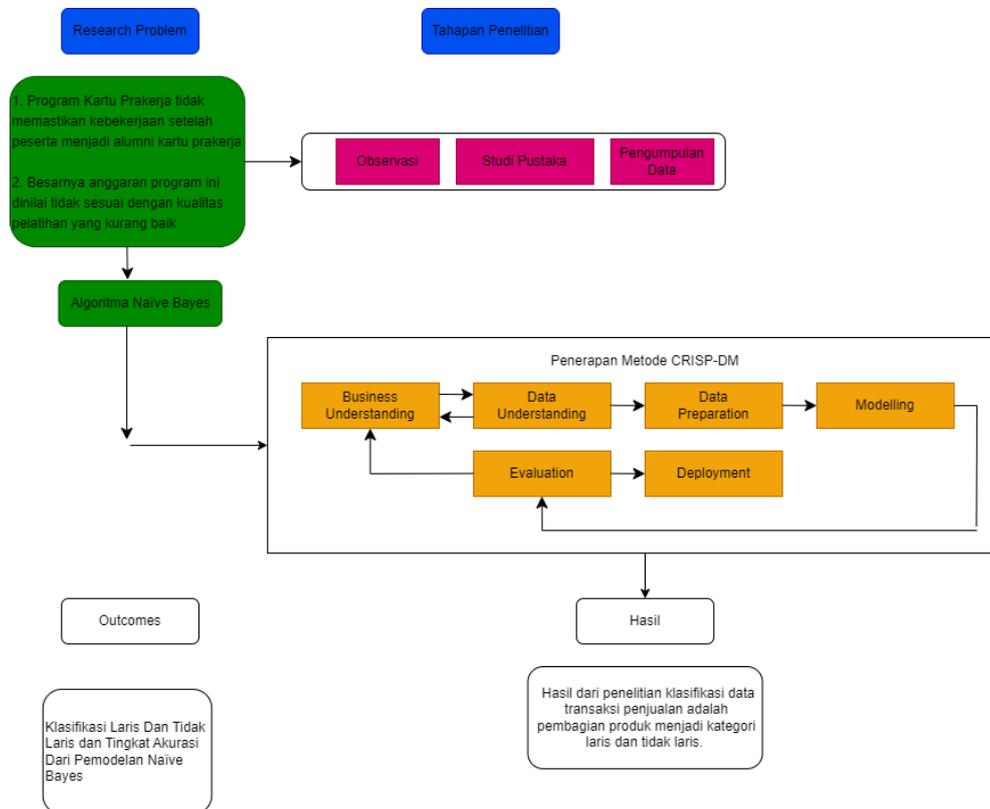
Dalam studi kasus ini, penulis melakukan observasi di Toko Ira.

- Waktu Pengambilan Data

Setelah observasi selesai, data diambil dari Toko Ira, dengan bantuan pihak penjual. Ini dilakukan pada bulan Mei 2024. Data yang dikumpulkan terdiri dari Tanggal, Jam, Menu, Qty, Harga, dan Kategori seperti dapat dilihat pada Tabel 1. Data ini disimpan dalam *notebook* dan terdiri dari 228 rekaman. Data ini dalam penelitian ini digunakan adalah *dataset* pribadi. Data kualitatif dalam penelitian ini adalah informasi teks tentang objek atau subjek yang tidak dapat diukur dengan angka tetapi dapat dilihat atau dirasakan [5].

Tabel 1 *Attribut Dataset*

<i>Attribut</i>	Keterangan
Tanggal	Pada Tanggal Berapa Pembelian.
Jam	Waktu Pembelian.
Menu	Pilihan Makanan.
<i>Qty</i>	Jumlah yang dibeli.
Harga	Harga yang dibayarkan.
Kategori	Laris dan Tidak Laris Penjualan.



Gambar 1 Tahapan Penerapan Metodologi

3.2 Penerapan Metodologi

Penelitian ini mengklasifikasikan status penjualan di Kedai Toko Ira dengan menggunakan Metode *Naive Bayes*. Pengumpulan data awal diikuti oleh pembersihan data dan kemudian transformasi data. Pembersihan data adalah proses membersihkan data, sedangkan transformasi data adalah mengubah skala data untuk mencapai distribusi yang diantisipasi. Selanjutnya, untuk memproses dan menguji data, digunakan model proses data mining standar Cross-Industry (CRISP-DM) [6]. *Understanding Business, Data Understanding, Data Preparation, Modeling, Evaluation, dan Deployment* adalah komponen CRISP-DM. Dengan menerapkan model ini, hasil dan akurasi *Klasifikasi* status penjualan di toko akan ditampilkan. Berikut adalah tahapan penelitian metodologi yang digunakan dapat dilihat pada Gambar 1.

3.3 Desain Penelitian

Metode Cross Industry Standard for Data Mining (CRISP-DM) adalah kerangka kerja pemrosesan data yang umum dan membantu pengguna memahami dan mengelola proyek data mining secara sistematis. Proses penelitian ini mengacu pada enam tahap dari kerangka kerja ini, yang terdiri dari:

3.3.1 Business Understanding

Berpusat pada tujuan dan kebutuhan. Selain itu, tahapan ini menunjukkan pemahaman tentang masalah yang akan diteliti adalah untuk menggunakan algoritma *Naive Bayes* mengklasifikasikan penjualan di toko ira. Selanjutnya kami mengukur tingkat akurasi algoritma untuk mengklasifikasikan penjualan di toko ira, dengan harapan bahwa ini akan memberikan informasi tentang tingkat akurasinya.

3.3.2 Pemahaman Data

Tahap ini memberikan dasar analitis untuk penelitian dengan menghasilkan kesimpulan (ringkasan) dan menemukan masalah potensial dalam data. Data sebanyak 228 diperoleh dari Toko Ira. Tanggal, Jam, Menu, Kuantitas, Harga, dan Kategori adalah enam atribut yang diperoleh dari data tersebut. Mereka akan diproses sebelum diproses untuk membersihkan dan mengubahnya.

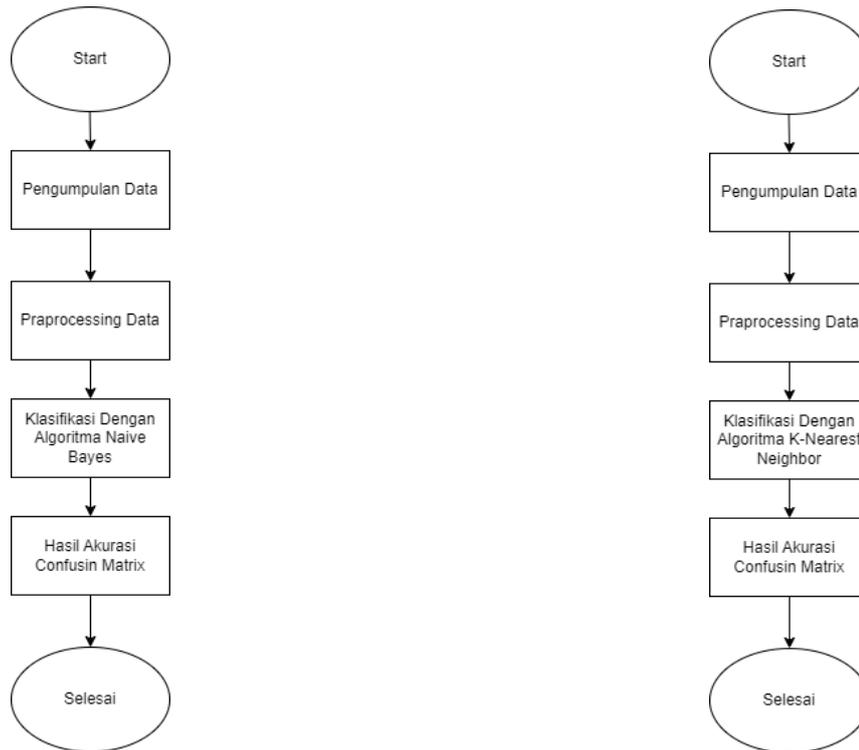
3.3.3 Persiapan Data

Selanjutnya adalah persiapan data, yang mencakup semua tindakan yang diperlukan untuk membuat data akhir (data yang akan digunakan dalam alat pemodelan) dari data mentah awal. Pada tahap ini, juga dilakukan pembersihan dan transformasi data. Tahapan persiapan data adalah sebagai berikut [8]:

1. Pengumpulan data: Pengumpulan data dimulai dengan rekaman penjualan dari toko ira.
2. Pembersihan data: Pembersihan data adalah proses pembersihan data. Data yang telah diperoleh harus dipilih kembali untuk menentukan data yang tidak akurat, tidak lengkap, atau tidak benar dan untuk meningkatkan kualitas data. Dalam proses *Klasifikasi*, nama, jenis kelamin, dan tanggal pengukuran akan dihapus dari data penelitian.

3.3.4 Pemodelan

Tahap Pemodelan: Metode *klasifikasi* akan digunakan dengan algoritma *Naive Bayes* dan *K-Nearest Neighbor* seperti diperlihatkan pada Gambar 2 dan Gambar 3. Setelah itu, akan dibuatkan model *klasifikasi* data yang terdiri dari pelatihan data dan pengujian data. Data akan diklasifikasi dengan algoritma *Naive Bayes*, dan tingkat akurasi akan diukur dengan *confusion matrix* [9].



Gambar 2 Tahapan Penelitian *Naïve Bayes* Gambar 3 Tahapan Penelitian *K-Nearest Neighbor*

3.3.5 Evaluasi

Evaluasi model *klasifikasi* dengan algoritma *Naïve Bayes* yang telah dibuat menggunakan *confusion matrix* dapat diketahui nilai *true positive*, *true negative*, *false positive* dan *false negative* pada suatu model. Hasil dari evaluasi setiap model klasifikasi berupa nilai *accuracy*, *precision*, dan *recall*. Kinerja algoritma dilakukan untuk menghitung tingkat performa untuk di dapatkan kesamaan atau kedekatan suatu hasil perhitungan dengan angka atau data yang sebenarnya sehingga dapat memprediksi penjualan di toko kedai ira dengan data sebenarnya [10]. Rumus yang digunakan untuk menghitung *accuracy* dapat dilihat pada formula (1), *precision* seperti pada formula (2), dan *recall* seperti pada formula (3).

Rumus *Accuracy*:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

$$Accuracy = \frac{\text{Jumlah Prediksi benar (positif dan negatif)}}{\text{Jumlah keseluruhan data}}$$

Rumus *Precision*;

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

$$Precision = \frac{\text{Jumlah Prediksi Benar (Positif)}}{\text{Prediksi benar(positif)+prediksi salah (salah)}}$$

Rumus *Recall*;

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

$$Recall = \frac{\text{Jumlah prediksi benar (positif)}}{\text{Prediksi benar(positif)+Prediksi salah(Negatif)}}$$

3.3.7 Deployment

Penyebaran hasil dari penelitian ini berupa pengetahuan dan data yang diperoleh disusun untuk penggunaan praktis. Proses penyebaran berupa penyusunan laporan dan jurnal pada proses datamining yang telah dimasukan oleh peneliti.

3.3.8 Metode Algoritma Naïve Bayes Dan K-Nearest Neighbor

Salah satu algoritma pembelajaran induktif yang paling efektif dan efisien untuk pengolahan data dan pembelajaran mesin adalah *Naive Bayes*. Meskipun menggunakan asumsi keindependenan atribut, algoritma ini tetap kompetitif dalam proses klasifikasi [7]. Algoritma *K-Nearest Neighbor* digunakan untuk memasukkan objek baru ke dalam kategori berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut seperti dapat dilihat pada formula (4). Untuk menemukan persamaan antara data baru dan lama, algoritma non-parametrik K-NN akan memasukkan data baru ke dalam kategori yang paling mirip dengan kategori sebelumnya [11].

Rumus *Naive Bayes*;

$$P(H|X) = \frac{p(p|h)p(h)}{p(x)} \quad (4)$$

X adalah data dari kelas yang belum diketahui; H adalah label kelas tertentu untuk hipotesis data X; P(H|X) adalah probabilitas hipotesis H berdasarkan kondisi X (probabilitas posterior); P(H) adalah probabilitas hipotesis H (probabilitas prior); P(X|H) adalah probabilitas X berdasarkan kondisi hipotesis H; P(X) adalah probabilitas X. sedangkan rumus yang digunakan dalam K-NN dapat dilihat pada formula (5).

$$d_i = \sqrt{\sum (x1 - y1)^2} \quad (5)$$

D adalah jarak kedekatan, x adalah data pelatihan, dan y adalah data pengujian. n adalah jumlah atribut individu antara 1 dan n. f adalah fungsi si militer atribut i antara kasus X dan Y.

3. HASIL DAN PEMBAHASAN

3.1 Perangkat Keras dan Perangkat Lunak

Penelitian ini menggunakan seperangkat *tools* untuk mendukung penelitian yang terdiri dari perangkat lunak (*software*) dan perangkat keras (*hardware*). Versi *software*, alat, dan fungsinya ditunjukkan dalam Tabel 2. Sedangkan alat perangkat keras (*hardware*), jenisnya, dan spesifikasinya ditunjukkan dalam Tabel 3.

Tabel 2 Perangkat Lunak (*Software*)

No	<i>Software</i>	<i>Versi</i>	Fungsi
1	<i>Microsoft Excel</i>	2013	Untuk mengidentifikasi perasaan yang dihasilkan dari hasil <i>crawling</i> data <i>Twitter</i> untuk dimasukkan ke dalam data latih dan disimpan.
2	<i>RapidMiner Studio</i>	10.0	untuk memulai penggunaan teknik <i>text mining</i> .
3	<i>Windows</i>	10 Pro	Untuk penggunaan bisnis, dengan fitur-fitur yang dapat meningkatkan keamanan dan produktivitas.

Tabel 3 Spesifikasi *Hardware*

Jenis <i>Hardware</i>	Spesifikasi
<i>Model</i>	HP Notebook
<i>Processor</i>	Intel Celeron N3060
<i>Memory</i>	4Gb DDR3
<i>GPU</i>	<i>Graphics</i> HD
<i>Storage</i>	500Gb HDD
<i>Display</i>	14 inch High Definition (HD)

3.2 Implementasi Dan Pengujian

Dibawah ini adalah serangkaian Implementasi dan pengujian dari penelitian ini yang telah di lakukan:

3.2.1 Implementasi Penelitian

Penelitian ini menggunakan data penjualan dari toko kedai Ira; mereka akan diproses dalam beberapa tahap, termasuk data perbaikan, integrasi, dan pemilihan. Kemudian, algoritma *K-Nearest Neighbor* digunakan untuk menghasilkan nilai yang paling akurat. Selain itu, pengujian dapat dilakukan melalui metode *Cross-Industry Standard Process for Data Mining (CRISP DM)*, yang terdiri dari enam tahap [12]:

a. Business Understanding

Dataset mentah yang dikumpulkan dari toko kedai Ira adalah kertas data yang diubah menjadi rekapan Excel, yang mengandung enam atribut: Tanggal, Jam, Menu, Qty, Harga, dan Kategori. Total data mentah adalah 228 data.

b. Data Understanding

Buku data toko kedai Ira masih digunakan untuk pengumpulan data. Setelah itu, data diidentifikasi, dipilih, dan dibersihkan untuk memilih atribut yang dibutuhkan dalam proses penelitian. Gambar 4 adalah gambar data mentah kertas dari toko kedai Ira dan Gambar 5 adalah file *Excel* yang menunjukkan hasil pengolahan data mentah.

Tanggal	Menu	Qty	Harga	Total	Tarjual
3 Maret 2023	Mie ayam Baku	14	35.000	Total: Rp. 89.000 Tarjual: 4. porsi	4. porsi
	Mie ayam Pangsit	1	0		
	Mie ayam Bakso	1	13.000		
	Mie ayam chili oil	14	52.000		
4 Maret 2023	Mie ayam Baku	1	12.000	Total: Rp. 90.000 Tarjual: 7. porsi	7. porsi
	Mie ayam Pangsit	11	26.000		
	Mie ayam Bakso	1	15.000		
	Mie ayam chili oil	11	39.000		
5 Maret 2023	Mie ayam Baku	11	44.000	Total: Rp. 195.000 Tarjual: 14. porsi	14. porsi
	Mie ayam Pangsit	11	26.000		
	Mie ayam Bakso	11	39.000		
	Mie ayam chili oil	11	65.000		
7 Maret 2023	Mie ayam Baku	11	36.000	Total: Rp. 218.000 Tarjual: 17. porsi	17. porsi
	Mie ayam Pangsit	11	59.000		
	Mie ayam Bakso	11	52.000		
	Mie ayam chili oil	11	91.000		
8 Maret 2023	Mie ayam Baku	11	60.000	Total: Rp. 216.000 Tarjual: 17. porsi	17. porsi
	Mie ayam Pangsit	11	26.000		
	Mie ayam Bakso	1	15.000		
	Mie ayam chili oil	11	115.000		

Gambar 4 Data Mentah Penjualan Toko Kedai Ira

1	Tanggal	Jam	Menu	Qty	Harga	Kategori
2	03/02/2023	11:00	Mie ayam pangsit	24	306.000	Laris
3	04/02/2023	11:00	Mie ayam bakso	39	497.000	Laris
4	05/02/2023	11:00	Mie ayam chili oil	67	859.000	Laris
5	14/02/2023	11:00	Mie ayam chili oil	38	491.000	Laris
6	15/02/2023	11:00	Mie ayam bakso	20	258.000	Laris
7	16/02/2023	11:00	Mie ayam pangsit	32	411.000	Laris
8	17/02/2023	11:00	mie ayam spesial	20	400.000	Laris
9	18/02/2023	11:30	Sate usus	25	125.000	Laris
10	01/03/2023	11:30	Chesee Roll	5	50.000	Tidak Laris
11	02/03/2023	11:30	Mie ayam chili oil	39	506.000	Laris
12	03/03/2023	11:30	Es tee	20	100.000	Laris
13	04/03/2023	11:30	Leci tea	7	84.000	Tidak Laris
14	05/03/2023	11:30	Mie ayam pangsit	14	178.000	Kadang Tidak Laris
15	07/03/2023	11:30	Mie ayam bakso	17	218.000	Tidak Laris
16	08/03/2023	11:30	mie ayam spesial	17	340.000	Tidak Laris
17	09/03/2023	11:00	Es jeruk	10	50.000	Tidak Laris
18	10/03/2023	11:00	Soda gembira	14	139.000	Kadang Tidak Laris
19	11/03/2023	11:00	Mie ayam chili oil	20	260.000	Laris
20	14/03/2023	11:00	Mie ayam bakso	9	117.000	Tidak Laris
21	15/03/2023	11:00	mie ayam polos	10	120.000	Kadang Tidak Laris
22	16/03/2023	11:00	Es campur	8	96.000	Tidak Laris

Gambar 5 Data Input Excel

Beberapa tahapan dapat data understanding, dilakukan dengan beberapa tahapan meliputi:

1. Pengumpulan Data Awal, data yang dikumpulkan terdiri dari data penjualan di toko kedai Ira yang dikumpulkan dari toko tersebut.
2. Mendeskripsikan data, sebanyak 228 data dikumpulkan untuk atribut penjualan di toko kedai Ira.
3. Pemilihan *atribut*, tanggal, jam, menu, qty, harga, dan kategori adalah enam atribut yang diperoleh berdasarkan data toko Ira.

c. Data Preparation

Selain itu, tahap ini mencakup proses pembersihan dan transformasi data agar dapat digunakan dalam tahap pemodelan, serta pemilihan atribut dan tabel data [13]. Proses persiapan data terdiri dari tiga halaman utama, yaitu pilihan data, data *preprocessing*, dan data *Transformasi*.

d. Modeling

Tahap penelitian ini melibatkan pemilihan dan penerapan model *klasifikasi*. Penelitian ini akan menggunakan model KNN (*K-Nearest Neighbors*) dan *Naive Bayes* untuk memprediksi penjualan produk [14]. Dengan *RapidMiner*, dapat melakukan hal-hal berikut:

1. Pilih Model: Dari katalog operator *RapidMiner*, pilih model KNN dan *Naive Bayes*.
2. Konfigurasi Model Tentukan parameter seperti jumlah tetangga (K), metrik jarak, dan Probabilitas *Naive Bayes*.
3. Train Model: Latih model menggunakan data pelatihan.

e. Evaluation

Tahap evaluasi Ini dilakukan untuk menjamin bahwa model yang dibuat dapat mengategorikan penjualan produk sebagai laris, tidak cukup laris, dan laris. Di akhir fase penelitian ini, metode *data mining* yang biasa digunakan, seperti *akurasi*, *presisi*, dan *recall*, harus ditentukan [15]. Proses yang dapat dilakukan dengan *RapidMiner*:

1. Evaluasi Model: melihat metrik evaluasi model dengan menggunakan operator seperti kinerja.
2. Validasi Silang: Melakukan validasi silang untuk memastikan keandalan model.
3. Penyetelan Model: Mengubah parameter model untuk meningkatkan kinerja jika perlu.

f. Deployment

Setelah model divalidasi dan dianggap mampu, selanjutnya dapat mengekspornya dan menggunakannya untuk pengambilan keputusan atau *Klasifikasi* produk di toko Kedai Ira. Proses yang dapat dilakukan dengan *RapidMiner*:

1. Export Model: ekspor model ke format yang sesuai untuk penggunaan jangka panjang.
2. Integration: integrasikan model ke dalam sistem yang ada untuk aplikasi praktis.

3.2.2 Pengolahan Data dengan Algoritma K-Nearest Neighbor Dan Algoritma Naive Bayes pada RapidMiner

Proses impor data ke aplikasi *RapidMiner* akan melibatkan import *dataset* atau data *Excel* yang telah disiapkan sebelumnya. Setelah itu, label dan tipe data akan ditentukan. Label dan tipe data setiap data harus diidentifikasi pada tahapan penelitian ini. Label: digunakan sebagai variable tujuan, atribut: digunakan sebagai variable prediksi seperti dapat dilihat pada Gambar 6.

Import Data - Format your columns.

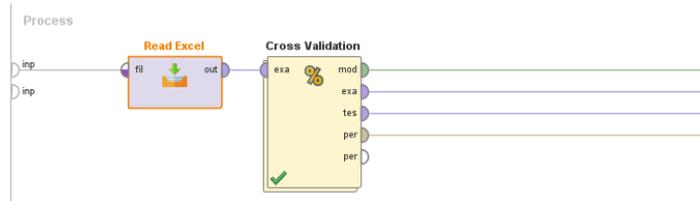
Format your columns.

Replace errors with missing values ⓘ

	Tanggal date	Jam time	Menu polynomial	Qty integer	Harga real	Kategori polynomial label
1	Feb 3, 2023	11:00:00 AM ICT	Mie ayam pangsit	24	306000.000	Laris
2	Feb 4, 2023	11:00:00 AM ICT	Mie ayam bakso	39	497000.000	Laris
3	Feb 5, 2023	11:00:00 AM ICT	Mie ayam chilli oil	67	859000.000	Laris
4	Feb 14, 2023	11:00:00 AM ICT	Mie ayam chilli oil	38	491000.000	Laris
5	Feb 15, 2023	11:00:00 AM ICT	Mie ayam bakso	20	258000.000	Laris
6	Feb 16, 2023	11:00:00 AM ICT	Mie ayam pangsit	32	411.000	Laris
7	Feb 17, 2023	11:00:00 AM ICT	mie ayam spesial	20	400000.000	Laris
8	Feb 18, 2023	11:30:00 AM ICT	Sate usus	25	125000.000	Laris
9	Mar 1, 2023	11:30:00 AM ICT	Chesee Roll	5	50000.000	Tidak Laris
10	Mar 2, 2023	11:30:00 AM ICT	Mie ayam chilli oil	39	506000.000	Laris
11	Mar 3, 2023	11:30:00 AM ICT	Es tee	20	100000.000	Laris
12	Mar 4, 2023	11:30:00 AM ICT	Lada	7	94000.000	Tidak Laris

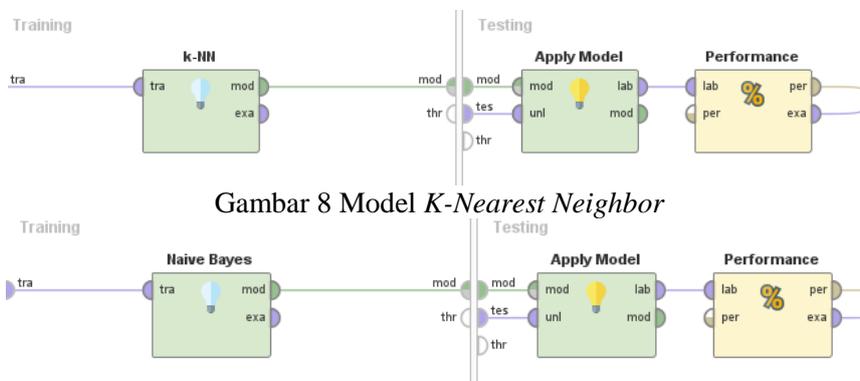
Gambar 6 Tampilan Data Yang Akan Menjadi Label

Setelah memilih atribut dan label *dataset*; simpan *dataset* pada menu *Repository*; kemudian, pilih *folder* data dan beri nama pada set; *drag set* data ke kolom proses *view*; kemudian, pergi ke menu *operator* dan pilih *Search Cross Validation*; dan *drag set* data ke bagian proses *View* seperti dapat dilihat pada Gambar 7.



Gambar 7 Tampilan Dataset Cross Validation

Setelah itu, klik dua kali pada *Cross Validation*, dan kolom pelatihan dan uji akan muncul dalam proses *view*. Cari Algoritma *K-Nearest Neighbors* dan *Naive Bayes* di *operator*, lalu *drag* ke kolom pelatihan. Selain itu, Setelah itu, klik dua kali pada *Cross Validation*, dan kolom pelatihan dan uji akan muncul dalam proses *view*. Cari Algoritma *K-Nearest Neighbors* dan *Naive Bayes* di *operator*, lalu *drag* ke kolom pelatihan seperti dapat dilihat pada Gambar 8 dan Gambar 9.



Gambar 8 Model K-Nearest Neighbor

Gambar 9 Model Naive Bayes

Gambar 10, Gambar 11, Gambar 12, dan Gambar 13 menunjukkan hasil pengukuran algoritma klasifikasi *K-Nearest Neighbors* dan *Naive Bayes* dari model yang telah dikembangkan.

accuracy: 60.18% +/- 10.36% (micro average: 60.09%)

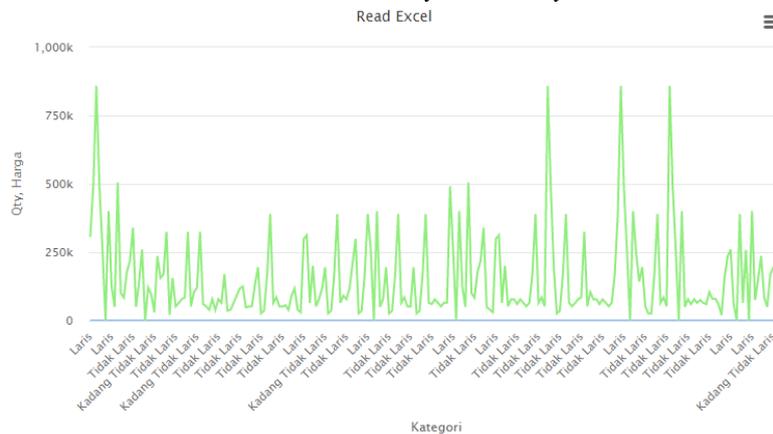
	true Laris	true Tidak Laris	true Kadang Tidak ...	true Laris	class precision
pred. Laris	33	13	3	1	66.00%
pred. Tidak Laris	28	90	29	0	61.22%
pred. Kadang Tida...	2	15	14	0	45.16%
pred. Laris	0	0	0	0	0.00%
class recall	52.38%	76.27%	30.43%	0.00%	

Gambar 10 Accuracy *K-Nearest Neighbor*

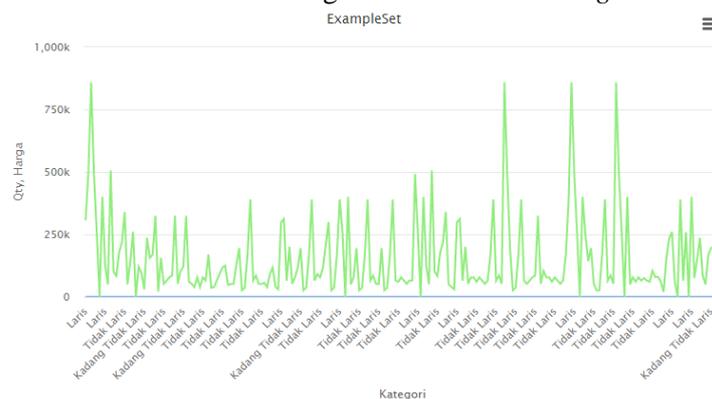
accuracy: 89.94% +/- 3.53% (micro average: 89.91%)

	true Laris	true Tidak Laris	true Kadang Tidak ...	true Laris	class precision
pred. Laris	62	3	6	1	86.11%
pred. Tidak Laris	0	109	6	0	94.78%
pred. Kadang Tida...	1	6	34	0	82.93%
pred. Laris	0	0	0	0	0.00%
class recall	98.41%	92.37%	73.91%	0.00%	

Gambar 11 Accuracy *Naive Bayes*



Gambar 12 Grafik Pengukuran *K-Nearest Neighbor*



Gambar 13 Grafik Pengukuran *Naive Bayes*

Hasil Akurasi algoritma KNN (*K-Nearest Neighbors*) dan *Naive Bayes* dari 228 data yang dikumpulkan ditunjukkan oleh pengujian *Rapid Miner*:

1. *K-Nearest Neighbor* sebesar 60,09% dan *Naive Bayes* sebesar 60,09%, perhitungan *true positif*, *true negatif*, *false positif*, dan *false negatif* memberikan akurasi total untuk data penjualan produk di toko kedai Ira.
2. Berdasarkan pengujian yang dilakukan oleh peneliti jika jumlah atribut ditambahkan, ditemukan bahwa algoritma *Naive bayes* lebih akurat dari algoritma *K-Nearest Neighbor*, tetapi dalam tidak keseluruhan data dalam *Klasifikasi* data.

4. KESIMPULAN

Hasil penelitian menunjukkan bahwa ada produk laris dan tidak laris di toko kedai Ira. Salah satu cara untuk menghindari keduanya adalah dengan memantau perkembangan data penjualan setiap bulan. Untuk membandingkan *Klasifikasi* penjualan toko di Kedai Ira, studi ini menggunakan algoritma KNN (*K-Nearest Neighbors*) dan metode *data mining Naive Bayes*. Hasil menunjukkan bahwa Algoritma KNN memiliki nilai akurasi 60,09% dan *Naive Bayes* memiliki nilai 89,91%. Penelitian ini diuji dengan *RapidMiner Studio* memiliki sampel data 228.

5. SARAN

Salah satu rekomendasi yang diajukan kepada peneliti adalah agar mereka menyadari bahwa penelitian ini masih memiliki kekurangan. Sehingga, peneliti menyarankan pembaca dan peneliti selanjutnya untuk melakukan hal-hal berikut:

1. Menambah data pengujian biasa untuk mendapatkan nilai akurasi yang lebih tinggi;
2. Menjadi sumber acuan yang lebih baik untuk penjualan; dan
3. Menjadi referensi untuk jurnal dan skripsi yang akan datang.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada tim redaktur Jurnal Teknik Politeknik Negeri Sriwijaya yang telah memberikan kesempatan untuk diterbitkan artikel ilmiah ini.

DAFTAR PUSTAKA

- [1] N. W. Wardani, P. G. S. C. Nugraha, dan G. S. Mahendra, "Implementasi Naïve Bayes Pada Data Mining Untuk MengKlasifikasikan Penjualan Barang Terlaris Pada Perusahaan Ritel," *JST (Jurnal Sains dan Teknologi)*, vol. 12, no. 3, Jan 2024, doi: 10.23887/jstundiksha.v12i3.38605.
- [2] N. Pransiska dan Ah. Mirza, "PENERAPAN DATA MINING PREDIKSI PENJUALAN BARANG ELEKTRONIK TERLARIS MENGGUNAKAN ALGORITMA NAÏVE BAYES (Study Kasus : Planet Cash And Credit Cabang Muara Enim)," *Bina Darma Conference on Computer Science*, 2023.
- [3] S. Rahmatullah *dkk.*, "DATA MINING UNTUK MENENTUKAN PRODUK TERLARIS MENGGUNAKAN METODE NAIVE BAYES," 2289.
- [4] Khaerul Mukhlisin, "Penerapan Algoritma Naïve Bayes Pada Produk Brownies Tape Dalam Menentukan Tingkat Penjualan Laris Dan Kurang Laris," *KOPERTIP : Jurnal Ilmiah Manajemen Informatika dan Komputer*, vol. 5, no. 2, hlm. 51–56, Okt 2021, doi: 10.32485/kopertip.v5i2.168.
- [5] U. Darwan *dkk.*, "Menentukan Menu Makanan Favorit di Outlet Barbar Sampit Menggunakan Algoritma Naive Bayes," 2022.
- [6] Ismasari Nawangsih dan Asti Setyaningsih, "PENERAPAN ALGORITMA NAIVE BAYES UNTUK MENENTUKAN KLASIFIKASI PRODUK TERLARIS PADA PENJUALAN PULSA," *SIGMA – Jurnal Teknologi Pelita Bangsa*, hlm. 195–207, 2020.
- [7] Nurhidayati dan Alimuddin, "Klasifikasi Penjualan Obat Pertanian Laris Dan Kurang Laris Pada UD Cahaya Tani Menggunakan Metode Decision Tree," hlm. 101–108, 2289.
- [8] F. Rizki, A. Faisol, dan F. Santi Wahyuni, "PENERAPAN METODE NAÏVE BAYES UNTUK MEMPREDIKSI PENJUALAN PADA UD. HIKMAH PASURUAN BERBASIS WEB," 2020.
- [9] I. R. Pratama, M. Maimunah, dan E. R. Arumi, "Sistem Klasifikasi Penjualan Produk Alat Listrik Terlaris Untuk Optimasi Pengadaan Stok Menggunakan Naïve Bayes," *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 6, no. 4, hlm. 2135, Okt 2022, doi: 10.30865/mib.v6i4.4418.

- [10] H. Paul, A. Sartika Wiguna, dan H. Santoso, “PENERAPAN ALGORITMA SUPPORT VECTOR MACHINE DAN *NAIVE BAYES* UNTUK *KLASIFIKASI* JENIS MOBIL TERLARIS BERDASARKAN PRODUKSI DI INDONESIA,” 2023.
- [11] M. N. Muhammad Arhami, *Data Mining Algoritma Dan Implementasi*. Yogyakarta, 2289.
- [12] Made Adrian Astalina Pramana, Wayan Sudiarsa, dan Putu Gede Surya Cipta Nugraha, “Penerapan Algoritma *Naive Bayes* Untuk Prediksi Penjualan Produk Terlaris Pada CV Akusara Jaya Abadi,” *Jurnal Teknik Informatika dan Sistem Informasi* , hlm. 518–534, 2023.
- [13] B. P. E. L. H. M. A. J. H. M. S. H. R. A. N. Much Aziz Muslim, *Data Mining Algoritma C4.5 Disertai contoh kasus dan penerapannya dengan program computer*. Semarang, 2289.
- [14] Novaldi, Deo Wirayuda, Andestan Dwintara, dan Ivan Mei, “Penerapan Algoritma *Naive Bayes* Untuk Menentukan *Klasifikasi* Produk Terlaris Pada Penjualan Voucher Kuota Di Edi Cell,” 2023.
- [15] ST. , MM. , I. A. N. H. STP. , M. S. E. F. STP. , M. S. Dr.Rina Fitriana, *Data Mining dan Aplikasinya Contoh Kasus di Industri Manufaktur dan Jasa*. Wawasan Ilmu Anggota IKAPI, 2022.