

PENERAPAN DATA MINING DENGAN METODE NAIVE BAYES CLASSIFIER PADA PENJUALAN BARANG UNTUK OPTIMASI STRATEGI PEMASARAN

Kahfi Rama Putra Lubis¹, Sunarsan Sitohang²

¹mahasiswa Program Studi Teknik Informatika, Universitas Putera Batam

²dosen Program Studi Teknik Informatika, Universitas Putera Batam

email:pb190210059@upbatam.ac.id

ABSTRACT

Along with technological developments, market developments are also increasing rapidly and competition is starting to occur between entrepreneurs and business people, which requires entrepreneurs to take advantage of existing technology. At this time the Universal Homeware Store does not realize that this data can be processed effectively to find out the development of a business or business that is being run, besides that the procurement of goods that customers need is often unavailable. One of the data processing technologies is by applying data mining. Data mining is a form of extracting information from data that is large enough to obtain data patterns so as to produce information data that is useful for making decisions and increasing profits in developing marketing strategies. One of the uses of data mining is using the Naive Bayes Classifier algorithm. Naive Bayes Classifier is one of the existing algorithms in data mining which is used to deal with uncertainty problems by emphasizing the concept of probability of use and measuring the level of accuracy so that it can predict purchasing strategies from past sales analysis. with research results Using 36 data and with rapid minner tools produces an accuracy value of 80%

Keywords: *Data Mining; Naive Bayes Classifier; Sales Strategy Optimization*

PENDAHULUAN

Seiring dengan perkembangan teknologi, perkembangan pasar juga semakin pesat dan mulai terjadinya persaingan antara pengusaha dan pelaku bisnis sehingga mengharuskan para pengusaha memanfaatkan teknologi yang ada. Hal tersebut berorientasi pada peningkatan penjualan dengan menarik konsumen dan

berupaya mempertahankan pelanggan sehingga lebih unggul dari pengusaha toko lainnya. Dalam setiap pengusaha harus memiliki strategi untuk mendapatkan rencana tindakan yang akan digunakan

dalam persaingan memperoleh posisi yang menyangkut konsep kecepatan dan kesiapan dalam mencapai kebutuhan pelanggan.

Toko *Universal Homeware* yang berada di Mega Legenda merupakan toko yang bergerak pada barang furnitur serta peralatan rumah tangga. Sebagaimana di ketahui salah satu tujuan dalam mendirikan sebuah usaha ialah untuk memperoleh laba atau keuntungan semaksimal mungkin dalam rangka meningkatkan pertumbuhan toko tersebut. Namun yang sering timbul permasalahan adalah data transaksi penjualan hanya di arsipkan sehingga mengalami penumpukan tanpa mengetahui manfaat dari data transaksi tersebut. Pada saat ini Toko *Universal Homeware* tidak menyadari data tersebut dapat di olah secara efektif untuk mengetahui

perkembangan usaha atau bisnis yang sedang di jalankan, selain itu pengadaan penyediaan barang yang di perlukan pelanggan sering sekali dalam keadaan tidak tersedia . Salah satu teknologi pengolahan data tersebut yaitu dengan cara penerapan *data mining*.

Data mining merupakan suatu bentuk penggalian informasi dari data yang cukup besar untuk memperoleh pola-pola data sehingga menghasilkan sebuah data informasi yang berguna untuk pengambilan keputusan serta meningkatkan keuntungan dalam menyusun strategi pemasaran (Cahyani and Rahayu 2020). Salah satu pemanfaatan *data mining* menggunakan algoritma *Naive Bayes Classifier*. *Naive Bayes Classifier* merupakan salah satu algoritma yang ada pada *data mining* yang dimanfaatkan untuk menangani masalah ketidakpastian dengan menekan pada konsep rprobabilitas kegunaanya dan mengukur tingkat keakurasian sehingga dapat memprediksi strategi pembelian dari analisis penjualan di masa lampau. Ciri utama *Naive Bayes Classifier* adalah asumsi yang sangat kuat (naif) akan independensi dari masing-masing kondisi/kejadian (Harahap et al. 2021).

Pada saat ini persaingan bisnis yang begitu ketat, Toko *Universal Homeware* memerlukan suatu strategi pemasaran yang lebih efektif dan efisien yang dapat menarik dan menjangkau lebih banyak pelanggan serta dapat memuaskan keinginan kebutuhan pelanggan. Oleh sebab itu maka di butuhkan sebuah optimasi. Optimasi merupakan suatu proses penyelesaian suatu masalah ataustrategi pemasaran sehingga dapat memberikan kondisi terbaik yang mampu memberikan nilai maksimum atau minimum (Pratama and Arumi 2022). Dalam sebuah bisnis, optimasi strategi pemasaran menjadi hal yang penting termasuk dalam hal penjualan produk sehingga mencapai penjualan produk yang diinginkan. Jika stok suatu

produk kosong maka potensi penjualan menurun.

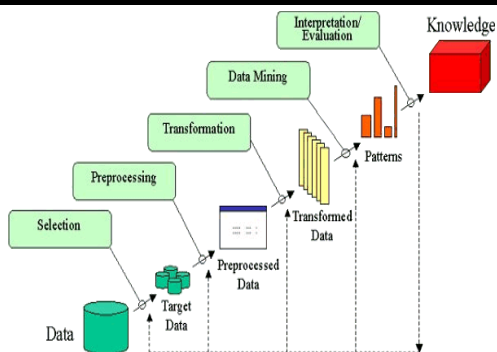
KAJIAN TEORI

2.1 KDD (*Knowledge Discovery in Database*)

Knowledge Discovery in Database (KDD) adalah metode dan cara mendapatkan sebuah informasi melalui basis data yang telah tersedia. Berikut tahap penyelesaiannya, yaitu (Putri, Ucha, Sanni Irawan, and Rizky 2021) :

1. *Data Selection* merupakan cara pengambilan data dari sebuah data set .
2. *Data preprocessing* merupakan tahap dalam melakukan proses pengolahan data dari awal untuk menghindari noise data..
3. *Transformation* merupakan tahap membuat pemodelan algoritma dalam pemrosesan.
4. *Data mining* merupakan pencarian pengetahuan hingga diperoleh suatu model informasi penting.

Evaluation merupakan hasil presentasi pengujian akurasi.



Gambar 1 KDD

Sumber: (Putri, Ucha, Sanni Irawan, and Rizky 2021)

2.2 Data Mining

Menurut (Harahap et al. 2021) data mining merupakan proses pencarian data dengan mencari pola atau informasi menggunakan teknik atau metode tertentu salah satu metode pada data mining adalah klasifikasi yaitu digunakan untuk memprediksi nilai dari atribut kategori target dengan menggunakan naive bayes classifier.

2.3 Klasifikasi

Menurut (Harahap et al. 2021) *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* bervariasi. Salah satu teknik *data mining* adalah klasifikasi. Teknik klasifikasi adalah teknik pembelajaran yang digunakan untuk memprediksi nilai dari atribut kategori target dengan menggunakan pendekatan machine learning. Salah satu metode klasifikasi yang dapat digunakan adalah *Naïve Bayes classifier*.

2.4 Naive Bayes

Naive Bayes Classifier merupakan salah satu algoritma yang ada pada *data mining* yang dimanfaatkan untuk menangani masalah ketidakpastian dengan menekan pada konsep probabilitas kegunaannya dan mengukur tingkat keakuratan sehingga dapat memprediksi strategi pembelian dari analisis penjualan di masa lampau. Ciri utama *Naive Bayes Classifier* adalah asumsi yang sangat kuat (naif) akan independensi dari masing-masing kondisi/kejadian (Harahap et al. 2021).

2.5 Optimasi

Optimasi merupakan suatu proses penyelesaian suatu masalah atau strategi pemasaran sehingga dapat memberikan kondisi terbaik yang mampu memberikan nilai maksimum atau minimum (Pratama and Arumi 2022). Dalam sebuah bisnis, optimasi strategi pemasaran menjadi hal yang penting termasuk dalam hal penjualan produk sehingga mencapai penjualan produk yang diinginkan. Jika stok suatu produk kosong maka potensi penjualan menurun

2.6 Software Pendukung

1. RapidMiner

Rapid Miner adalah sebuah lingkungan *machine learning data mining*, *text mining* dan *predictive analytics*. *Rapid Miner* merupakan perangkat lunak yang bersifat terbuka (*open source*). *Rapid Miner* adalah sebuah solusi untuk melakukan analisis terhadap *data mining*, *text mining* dan analisis prediksi. *Rapid Miner* menggunakan berbagai teknik deskriptif dan prediksi dalam memberikan wawasan kepada pengguna sehingga dapat membuat keputusan yang paling baik. *Rapid Miner* memiliki

kurang lebih 500 operator *data mining* termasuk operator input, output dan processing dan visualisasi (Indra Borman and Wati 2020).

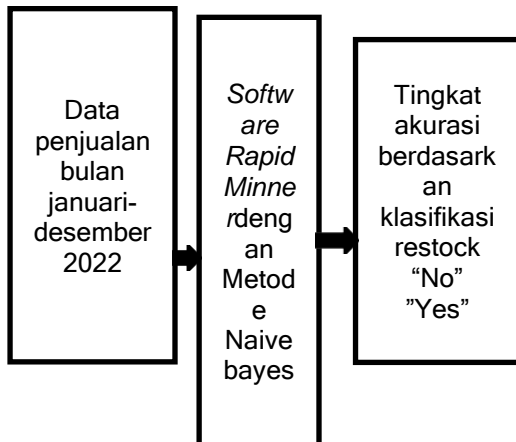


Gambar 2 RapidMiner

Sumber: (Indra Borman and Wati 2020).

2.7 Kerangka Pemikiran

Kerangka pemikiran merupakan ide dari peneliti dalam bentuk kerangka yang digunakan sebagai langkah awal penelitian yang berisi gambaran dari awal penelitian hingga akhir hasil penelitian



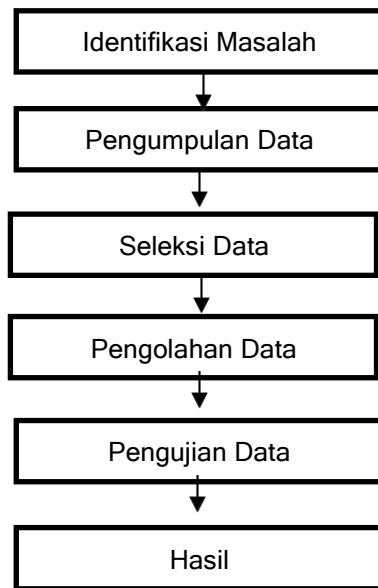
Gambar 3 Kerangka Pemikiran

Sumber Data penelitian 2023.

METODE PENELITIAN

3.1 Desain Penelitian

Desain penelitian adalah rangkaian prosedur dan metode yang digunakan untuk menganalisis dan menghimpun data untuk menentukan variabel yang akan menjadi topik penelitian. Desain penelitian juga didefinisikan sebagai strategi yang dilakukan peneliti untuk menghubungkan setiap elemen penelitian secara sistematis sehingga dapat menentukan fokus penelitian. Berikut di bawah ini merupakan desain penelitian pada penelitian ini:



Gambar 4 Desain Penelitian

Sumber: Data Penelitian 2023

3.2 Teknik Pengumpulan Data

Berikut di bawah ini merupakan beberapa teknik yang digunakan dalam teknik pengumpulan data:

1. Wawancara

Pada tahap ini, peneliti membuat daftar berisi pertanyaan-pertanyaan yang akan digunakan untuk mewawancarain narasumber yaitu selaku pemilik toko *Universal Homeware* sehingga peneliti dapat menerima hasil jawaban sesuai kebutuhan penelitian.

2. Observasi

Pada tahap ini, peneliti melakukan tinjauan kelapangan dengan arti langsung melihat teknik dan cara yang digunakan ditoko dalam pengolahan data. Dan kemudian meminta ijin untuk mendapatkan data tersebut untuk dijadikan data penelitian.

3. Studi Pustaka

Pada tahap ini, peneliti melakukan proses pengumpulan data dengan merangkum informasi-informasi mengenai variabel penelitian yang bersumber dari jurnal dan buku-buku untuk dijadikan sebagai referensi penyelesaian penelitian.

3.3 Operasional Variabel

Operasional variabel merupakan tahap yang diperlukan dalam penelitian sehingga dapat menemukan hubungan variabel dan indikator penelitian. Berikut ini variabel yang telah ditentukan dalam penelitian ini:

1. Bulan

Data yang digunakan sebagai data penelitian adalah data inventory yang ada pada toko *Universal Homeware* pada bulan Januari-Desember 2022. Dimana data tersebut terdiri atas atribut-atribut sehingga dapat ditentukan kombinasi antaritem set.

2. Nama Barang

Merupakan jenis atau nama barang yang ada pada toko *Universal Homeware* dan di pilih 10 item yang sama tiap bulanya untuk di jadikan sample penelitian.

3. Harga

Merupakan harga jual barang pada Toko *Universal Homeware* pada bulan tersebut.

4. Terjual

Merupakan transaksi jumlah penjualan barang yang telah terjual pada bulan tersebut.

5. Restock

Terdiri atas yes and no yang menandakan bahwa barang dengan nama yang telah ditentukan ada atau tidak pada Toko *Universal Homeware*.

HASIL DAN PEMBAHASAN

4.1 Hasil Penelitian

Merupakan proses pengaturan dan pengelompokan baik mengenai suatu informasi maupun suatu kegiatan berdasarkan fakta melalui usaha pemikiran peneliti dalam mengelolah dan menganalisa objek atau topik penelitian proses pengolahan data dan perhitungan dilakukan menggunakan microsoft excel dan menggunakan software pendukung yaitu *rapidminer*

1. Seleksi Data

Pada tahap ini adalah tahap dimana 36 data yang telah ditemukan akan dilakukan pengecekan apakah data tersebut layak dilakukan proses pengolahan data. Maka dari 36 data yang terdiri atas beberapa kolom data set pada tabel 1 menjadi 5 kolom

Tabel 1 Seleksi Data

Tahun	Kategori	Ukuran	Penjualan
2020	Murah	Kecil	Laku
	Murah	Besar	Tidak laku
	Murah	Besar	Tidak laku
	Sedang	Kecil	Laku
	Sedang	Besar	Tidak laku
	Sedang	Kecil	Tidak laku
	Mahal	Kecil	Tidak laku
	Mahal	Sedang	Laku
	Mahal	Besar	Laku
	Murah	Sedang	Laku
	Murah	Besar	Laku
2021	Murah	Kecil	Tidak laku
	Sedang	Besar	Laku
	Sedang	Besar	Tidak laku
	Sedang	Kecil	Laku
	Mahal	Besar	Tidak laku
	Mahal	Kecil	Laku
	Mahal	Kecil	Laku
	Murah	Sedang	Tidak laku
	Murah	Besar	Laku
	Murah	Sedang	Laku
	Sedang	Besar	Laku
2022	Sedang	Kecil	Laku
	Sedang	Besar	Laku
	Mahal	Besar	Tidak laku
	Mahal	Kecil	Laku
	Mahal	Besar	Tidak laku
	Murah	Kecil	Laku
	Murah	Kecil	Laku
	Murah	Sedang	Tidak laku
	Sedang	Besar	Laku
	Sedang	Sedang	Tidak laku
	Sedang	Besar	Tidak laku
Mahal	Kecil	Tidak laku	
Mahal	Besar	Laku	
Mahal	Kecil	Tidak laku	

Dari tabel diatas maka dengan melakukan perhitungan dengan algoritma

naive bayes berdasarkan probabilitas kelas hasil nya dalah sebagai berikut:

Tabel 2 Probabilitas kelas

Probabilitas kelas	
Status	Nilai
Laku	56%
Tidak Laku	44%
Jumlah	100%

1. Probabilitas Tahun

Berdasarkan data item set diatas dengan 36 data record maka di dapat diketahui bahwa terdapat 30 % dengan Penjualan Laku dan 38 % Tidak Laku pada

tahun 2020. terdapat 45 % dengan Penjualan Laku dan 19 % Tidak Laku pada tahun 2021. terdapat 25 % dengan Penjualan Laku dan 44 % Tidak Laku pada tahun 2022. Berikut dibawah ini probabilitas kriteria Tahun

Tabel 3 Probabilitas Tahun

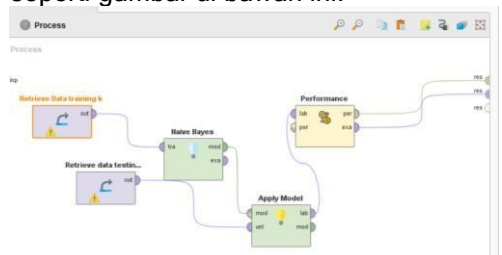
P (Tahun)	Laku	Tidak Laku
2020	30%	38%
2021	45%	19%
2022	25%	44%
Jumlah	100%	100%

4.2. Implementasi Rapidminer

1. Menghubungkan Operator


Setelah langkah-langkah *import data training* dan data *testing* dilakukan maka langkah selanjutnya yang akan di lakukan adalah dengan mencari operator naive bayes pada tempat pencarian dengan cara drag and drop operator naive bayes lalu tepatkan pada lembar kerja, setelah itu lakukan hal serupa pada apply model dan performance. Setelah seluruh operator yang akan di butuhkan ada berada lembar kerja maka selanjutnya langkah yang akan di lakukan adalah

dengan menghubungkan setiap operator seperti gambar di bawah ini:



Gambar 5 Menghubungkan operator
Sumber: Data Penelitian 2023

Perhitungan yang telah selesai di proses pada perangkat lunak rapid minner dengan menggunakan metode naive bayes untuk penjualan barang yang akan datang maka akan memberikan informasi mengenai data laku dan tidak laku Selanjutnya langkah yang akan di lakukan adalah menentukan tingkat keakuratan atau accuracy dengan cara klik tab *performance vector* yang ada di sebelah kanan maka akan muncul tabel view seperti gambar di bawah ini:



	true Tidak laku	true Laku	class precision
pred Tidak laku	1	0	100.00%
pred Laku	1	3	75.00%
class recall	50.00%	100.00%	

Gambar 6 Hasil pengujian
Sumber: Data Penelitian 2023

SIMPULAN

Bedasarkan penelitian yang telah dilakukan maka berikut ini adalah kesimpulan yang dapat diberikan oleh peneliti:

1. Penerapan data mining pada aplikasi rapidminner menggunakan algoritma naive bayes sangat sesuai untuk diterapkan pada pengklasifikasian data dengan penjumlahan data yang cukup banyak dan dengan cara menghitung probabilitas dari masing-masing kejadian.
2. Hasil analisa dari metode yang dilakukan dengan perhitungan algoritma naive bayes dengan pengujian menggunakan 36 data menghasilkan nilai akurasi 80% dengan diketahui bahwa terdapat 30 % dengan penjualan laku dan

38 % dengan penjualan tidak laku pada tahun 2020. Software rapidminner merupakan software yang dapat membantu untuk membuktikan bahwa hasil perhitungan dengan metode naive bayes menghasilkan nilai akurasi yang baik.

DAFTAR PUSTAKA

- Alghifari, Fauzan, and Didi Juardi. 2021. "Penerapan Data Mining Pada Penjualan Makanan Dan Minuman Menggunakan Metode Algoritma Naive Bayes."
- Anon. 2019. "SIGMA - Jurnal Teknologi Pelita Bangsa SIGMA - Jurnal Teknologi Pelita Bangsa." 10(September):40-47.
- Harahap, Fitriana, Nidia Enjelita Saragih, Elida Tuti Siregar, and Husin Sariangshah. 2021. "Penerapan Data Mining Dengan Algoritma Naive Bayes Classifier Dalam Memprediksi Pembelian Cat."
- Indra Borman, Rohmat, and Mina Wati. 2020. "Penerapan Data Maining Dalam Klasifikasi Data Anggota Kopdit Sejahtera Bandarlampung Dengan Algoritma Naive Bayes." *Jurnal Ilmiah Fakultas Ilmu Komputer* 09(01):25-34.
- Pratama, Irfan Reza, and Endah Ratna Arumi. 2022. "Sistem Klasifikasi Penjualan Produk Alat Listrik Terlaris Untuk Optimasi Pengadaan Stok Menggunakan Naive Bayes." 6:2135-39. doi: 10.30865/mib.v6i4.4418.
- Putri, Ayu Azlina. 2021. "Penerapan Data Mining Untuk Memprediksi Penjualan Buah Dan Sayur Menggunakan Metode K-Nearest Neighbor (Studi Kasus : PT . Central Brastagi Utama)." 1(6):354-61.
- Putri, Eka Ucha, Sanni Irawan, and Fitri Rizky. 2021. "Implementasi Data Mining Untuk Prediksi Penjualan Pesticida Pada CV MITRA ARTHA

SEJATI Menggunakan ALgoritma Naive Bayes.” *KESATRIA(Jurnal Penerapan Sistem Informasi Dan Manajemen* 2(1):39-46.

Putro, Hakam Febtadianrano, Retno Tri Vulandari, and Wawan Laksitoyuly Saptomo. 2020. “Penerapan Metode Naive Bayes Untuk Klasifikasi Pelanggan.” 8(2).

Sari, Rahayu Mayang, Fakultas Sains, Jurusan Sistem Komputer, Universitas Pembangunan, Panca Budi, and Kota Medan. 2022. “Implementasi Data Mining Untuk Memprediksi Penjualan Menggunakan Metode Penelitian Pendahuluan Pengumpulan Data Analisis Dan Perancangan Implementasi Pengujian.” 74-82.

Sosial, Bantuan, Studi Kasus, Klasifikasi Kemiskinan, Desa Lurah, and Odi Nurdiawan. 2023. “IMPLEMENTASI ALGORITMA C4 . 5 UNTUK MENENTUKAN PENERIMA.” 7(1):373-77.

	<p>Kahfi Rama Putra Lubis merupakan mahasiswa Prodi Teknik Informatika dari Universitas Putera Batam yang aktif dalam mendalami bidang teknologi dan informasi.</p>
	<p>Sunarsan Sitohang merupakan Dosen Prodi Teknik Informatika dari Universitas Putera Batam. Aktif sebagai tenaga kerja dan peneliti.</p>