

# Implementasi Data Mining dengan Algoritma Naïve Bayes pada Penjualan Obat

Herry Derajad Wijaya<sup>1</sup>, Saruni Dwiasnati<sup>2</sup>

<sup>1,2</sup> Universitas Mercu Buana Jakarta  
Fakultas Ilmu Komputer

Jl. Meruya Selatan 1 No. 1, Kembangan, 11650, Jakarta Barat  
e-mail: <sup>1</sup>herry.derajad@mercubuana.ac.id, <sup>2</sup>saruni.dwiasnati@mercubuana.ac.id

## Abstrak

Apotek merupakan tempat sarana pelayanan informasi obat dan perbekalan farmasi lainnya kepada masyarakat. Dewasa ini telah banyak apotek yang berdiri dan tersebar hingga ke pedesaan sehingga memudahkan masyarakat untuk mengakses peralatan farmasi dan obat-obatan. Tujuan penelitian ini adalah untuk menganalisis masalah-masalah pada penentuan sebuah produk mana yang dapat di kategorikan LAKU atau TIDAK LAKU. *Jenis obat yang makin lama makin bervariasi, dari obat yang berharga murah sampai harga yang kalau dilihat sangat kurang masuk akal namun fungsinya sangat bagus. Meningkatnya peredaran jenis obat terutama vitamin, hal ini mendorong penulis untuk melakukan penelitian untuk menentukan produk vitamin mana yang LAKU atau TIDAK LAKU yang bisa di gunakan sebagai pedoman sebuah apotek dalam menentukan jumlah stok barang yang harus ada pada gudang apotek tersebut. Informasi yang diinginkan dari penelitian ini adalah mendapatkan nilai accuracy untuk data penjualan obat-obatan terutama jenis-jenis vitamin yang sering menjadi pilihan dari nasabah yang membutuhkan obat-obatan tersebut dengan menggunakan algoritma klasifikasi data mining yaitu algoritma Naïve Bayes. Penelitian ini menggunakan tools Rapidminer versi 8 sebagai media untuk menguji data yang akan diolah untuk mendapatkan hasil accuracy dan nilai ROC. Nilai accuracy tersebut menunjukkan di nilai 88.00%.*

**Kata kunci : Datamining; Penjualan Obat; Naïve Bayes**

## Abstract

*A pharmacy is a place for drug information services and other pharmaceutical supplies to the community. Nowadays many pharmacies have been established and spread to the countryside, making it easier for the public to access pharmaceutical equipment and medicines. The purpose of this study is to analyze the problems in determining which product can be categorized as DO or NOT. The types of drugs which are increasingly varied, from low-priced drugs to prices that are seen as very unreasonable but whose function is very good. Increased circulation of drugs, especially vitamins, this prompts the authors to conduct research to determine which vitamin products are DOING or NOT DOING which can be used as a guideline for a pharmacy in determining the amount of stock that must exist in the pharmacy warehouse. The information desired from this research is to obtain accuracy values for sales data of drugs, especially the types of vitamins that are often the choice of customers who need these drugs by using a data mining classification algorithm, the Naïve Bayes algorithm. This research uses Rapidminer version 8 as a media tool to test the data to be processed to get accuracy and ROC values. The accuracy value shows that the value is 88.00%.*

**Keywords: Datamining; Drug Sales; Naïve Bayes**

## Pendahuluan

Dalam dunia kesehatan khususnya dunia obat-obatan di seluruh Indonesia sudah sampai pada titik yang mengesankan dan sangat menarik untuk terus ditelusuri dalam

kuantitas jenis dari berbagai merk yang ada. Pertumbuhan bisnis kesehatan khususnya apotek-apotek semakin berkembang dari waktu ke waktu, hal ini dapat dilihat dari banyaknya apotek yang tersebar baik di

perkotaan sampai ke pedalaman Indonesia. Dijaman sekarang, apotek tidak hanya menyediakan obat-obatan saja namun sudah sampai menyediakan alat-alat kosmetik, Karena hal tersebut, banyak apotek yang berlomba-lomba dalam menyediakan puluhan ragam varian obat-obatan untuk kesehatan serta vitamin agar dapat memudahkan pelanggannya dalam memenuhi kebutuhannya dengan variasi harga yang berbeda-beda. Apotek merupakan tempat sarana pelayanan informasi obat dan perbekalan farmasi lainnya kepada masyarakat. Dewasa ini telah banyak apotek yang berdiri dan tersebar hingga ke pedesaan sehingga memudahkan masyarakat untuk mengakses peralatan farmasi dan obat-obatan. Tujuan penelitian ini adalah untuk menganalisis masalah-masalah pada penentuan sebuah produk mana yang dapat di kategorikan LAKU atau TIDAK LAKU.

Dalam pemberian informasi mengenai data LAKU atau TIDAK LAKU sebuah apotek terhadap barang yang ada di apotek tersebut, seperti vitami sering mengalami kesulitan untuk menentukan barang mana yang LAKU atau TIDAK LAKU yang diminati oleh para konsumennya. Hal ini diakibatkan oleh menumpuknya stok vitamin yang kurang diminati di gudang, namun untuk vitamin yang diminati sama sekali tidak ada dalam persediaan pada apotek tersebut.

Untuk mengatasi masalah yang dijelaskan di atas, maka dilakukan pengklasifikasian obat yang LAKU dan TIDAK LAKU terjual di Apotek tersebut berdasarkan variabel-variabel yang di dapatkan. Algoritma yang digunakan dalam pengklasifian obat yang LAKU atau TIDAK LAKU menggunakan algoritma Naïve Bayes yang mampu memberikan informasi mengenai vitamn yang LAKU atau TIDAK LAKU pada apotek tersebut untuk meminimalisasikan stok yang menumpuk untuk vitamin yang di kategorikan TIDAK LAKU.

Ketersediaan data yang banyak dalam sebuah bidang dan kebutuhan untuk sebuah informasi atau pengetahuan dalam bidang tertentu dapat menjadi acuan sebagai pendukung pengambilan keputusan untuk membuat solusi bisnis dan dukungan infrastruktur di segala bidang merupakan cikal-bakal dari lahirnya teknologi *data mining*.

Penggunaan teknik *data mining* diharapkan dapat membantu percepatan proses pengambilan keputusan, memungkinkan perusahaan untuk mengelola informasi yang terkandung didalam data nasabah dan menjadi sebuah pengetahuan (*knowledge*) yang baru untuk menentukan keputusan yang dibutuhkan.

Oleh karena itu, untuk benar-benar mengetahui dan menetapkan jenis vitamin ini LAKU atau TIDAK LAKU perlu dilakukan forecasting (prediksi/estimasi) sehingga jenis vitamin tersebut akan sesuai dengan jenis vitamin yang diminati oleh pelanggan dari apotek tersebut. Setelah jenis obat ditetapkan, kemudian diestimasi berapa banyak permintaan jenis vitamin tersebut pada bulan berikutnya dengan menggunakan algoritma Naive Bayes. Data Mining adalah teknik untuk memproses dan mengekstraksi data besar menjadi informasi yang bisa membentuk data baru (Arif, B., Saruni, D).

Dalam melakukan estimasi (forecasting), banyak sekali metode-metode yang telah tersedia. Akan tetapi, dalam mengestimasi banyaknya penjualan pada bulan berikutnya terhadap jenis vitamin yang banyak di cari oleh pelanggan, metode forecasting yang cocok adalah algoritma Naive Bayes. Algoritma Naive Bayes cocok diterapkan pada data yang berskala ordinal. Jenis data Ordinal tersebut mempunyai variabel yang nilainya berupa symbol tetapi bisa diurutkan, tidak bisa diukur jaraknya dan tidak bisa dijumlahkan hasil dari semuanya. Algoritma Naive Bayes adalah suatu metode yang dapat menangani faktor prediksi pada bulan selanjutnya dan trend secara langsung. Keuntungan algoritma ini adalah memiliki kemampuan yang sangat baik dalam mengestimasi data yang memiliki pola trend dan prediksi untuk bulan selanjutnya. Oleh karena itu, algoritma Naive Bayes merupakan algoritma yang cocok dalam mengestimasi banyaknya penjualan pada bulan berikutnya terhadap jenis vitamin yang ada pada apotek yang diminati oleh pelanggan.

Berawal dari beberapa disiplin ilmu yang ada, data mining memiliki tujuan untuk memperbaiki teknik tradisional yang dapat menghasilkan sesuatu yang jauh lebih informatif sehingga bisa menangani: Jumlah data yang sangat besar, Dimensi data yang

tinggi dan Data yang heterogen dan berbeda sifat satu dengan yang lainnya (Ratniasih, 2019).

Setiap metode tersebut memiliki algoritma-algoritma yang terbaik yang dapat digunakan untuk memproses dataset yang ada. Namun dalam Penelitian ini diharapkan dapat membangun sebuah sistem dalam mengestimasi jenis vitamin untuk meningkatkan penjualan pada bulan berikutnya yang sesuai dengan data-data penjualan sebelumnya sehingga jenis vitamin yang diminati oleh pelanggan benar-benar tersedia pada apotek tersebut sehingga dapat membantu apotek dalam meningkatkan omset penjualannya menggunakan algoritma Naïve Bayes.

Data mining merupakan suatu langkah dalam melakukan Knowledge Discovery in Databases (KDD). Knowledge discovery sebagai suatu proses terdiri atas pembersihan data (data cleaning), integrasi data (data integration), pemilihan data (data selection), transformasi data (data transformation), data mining, evaluasi pola (pattern evaluation) dan penyajian pengetahuan (knowledge presentation). Data mining mengacu pada proses untuk menambang (mining) pengetahuan dari sekumpulan data yang sangat besar untuk menghasilkan sebuah pengetahuan baru dari bidang tertentu (Firdaus, 2017).

Pada Data Mining terdapat 3 bagian yaitu: Assosiasi, Klasifikasi dan Clustering. Assosiasi memiliki definisi sebuah proses yang digunakan untuk menemukan suatu hubungan yang terdapat pada nilai atribut dari sekumpulan data yang dimiliki, sedangkan Klasifikasi adalah teknik yang dilakukan untuk memprediksi class atau properti dari setiap instance data, dan Clustering sendiri memiliki makna pengelompokan data tanpa berdasarkan kelas data tertentu ke dalam kelas objek yang sama sesuai dengan topik yang diangkat.

Prediksi memiliki kemiripan definisi dengan klasifikasi, akan tetapi data yang digunakan pada klasifikasi berdasarkan perilaku atau nilai yang diperkirakan pada masa yang akan datang yang dilihat dari pola/nilai pada masa lalu. Contoh dari tugas prediksi misalnya untuk memprediksikan adanya pengurangan jumlah pelanggan dalam

beberapa waktu yang lalu dan prediksi harga saham dalam empat bulan ke depan.

Salah satu metode utama data mining adalah Supervised Learning (Pembelajaran dengan Guru), algoritma ini melakukan proses belajar berdasarkan nilai dari indikator target yang terhubung dengan indikator prediktor. Sedangkan metode analisis yang terdapat pada data mining salah satunya adalah Klasifikasi. Klasifikasi sendiri memiliki definisi sebuah proses yang digunakan untuk menemukan model atau fungsi yang mendeskripsikan dan membedakan data ke dalam kelas-kelas yang bertujuan untuk menemukan pola yang lebih memiliki nilai dari data yang berukuran relatif besar hingga sangat besar. Klasifikasi melibatkan proses pemeriksaan karakteristik dari objek dan memasukkan objek ke dalam salah satu kelas yang sudah didefinisikan sebelumnya. Dalam klasifikasi hanya terdapat satu atribut dari sekian banyaknya atribut yang bisa menjadi kemungkinan yang disebut atribut target, sedangkan atribut lainnya yang terdapat disebut atribut prediktor. Tiap kemungkinan nilai yang dimiliki oleh atribut target menunjukkan class yang diprediksi berdasarkan nilai-nilai dari atribut prediktor.

Naive Bayes merupakan teknik prediksi berbasis probabilistik sederhana yang berdasar pada penerapan teorema Bayes (aturan bayes) dengan asumsi independensi (ketidaktergantungan) yang kuat (naif). Dengan kata lain, dalam Naive Bayes model yang digunakan adalah model fitur independen (Prasetyo, 2012).

Sistem pengolahan data yang ada pada beberapa apotek yang ada di Indonesia masih belum dapat mengetahui dengan pasti produk apa saja yang LAKU atau TIDAK LAKU pada bulan sebelumnya agar mengetahui produk mana yang harus di stok lebih banyak lagi untuk bulan selanjutnya

Data yang akan digunakan pada penelitian ini adalah 150 dataset. Penelitian ini dilakukan karena atas masalah mengetahui produk pada bulan sebelumnya agar dapat menentukan stok produk mana yang mesti di perbanyak menggunakan algoritma Naïve Bayes agar mengetahui nilai accuracy nya.

Penelitian sebelumnya mengenai data mining dengan menggunakan algoritma Naïve Bayes adalah hasil penelitian yang dilakukan oleh (Dwiasnati, Devianto, 2018)

menghasilkan algoritma Naïve Bayes dapat dilakukan dalam penelitian di berbagai aspek bidang, seperti pada data set dalam memprediksi Nasabah Potensial pada sebuah perusahaan asuransi. Hasil penelitian yang dilakukan oleh ( Budiyanto, Dwiasnati, 2018) menyatakan bahwa algoritma Naïve Bayes tersebut mampu membantu penyelesaian permasalahan pemilihan produk terlaris di sebuah perusahaan yang bergerak dalam pengadaan dan layanan barang dari sistem CCTV bersama dengan sistem Kontrol Akses dengan tingkat akurasi memiliki kebenaran 81.36%. Selanjutnya penelitian yang dilakukan oleh (Prabowo, Kodar, 2019) dapat digunakan untuk memprediksi lama masa studi mahasiswa pada sebuah universitas baik fakultas eksak ataupun fakultas social. Penelitian yang dilakukan oleh (Bustami, 2014) dapat digunakan untuk memprediksi nasabah sebuah perusahaan asuransi untuk mengetahui lancar, kurang lancar atau tidak lancarnya nasabah tersebut dalam pembayaran pinjaman tersebut.

**Metode Penelitian**

Penelitian ini akan mengusulkan hasil accuracy baru untuk sebuah permasalahan memprediksi Produk LAKU atau TIDAK LAKU pada bulan sebelumnya untuk mnegetahui pola minat pelanggan terhadap pembelian vitamin pada sebuah apotek. Data yang digunakan dalam penelitian ini di ambil pada tahun 2018 bulan Januari-Mei, data yang di ambil kisaran 5.000 an namun untuk data testing untuk penelitian ini penulis hanya mengambil sampel data sebanyak 137. Model desain ini akan melakukan pemrosesan data training dan data testing untuk menguji metode algoritma yang digunakan. Tahapan yang akan dilalui dibagi menjadi 3 bagian, yaitu preprocessing, seleksi fitur (Feature Selection) dan validation yang didalamnya berisi sub proses training dan testing.

**a. Pengumpulan Data**

Data yang digunakan pada penelitian kali ini merupakan data sekunder, karena sumber data diperoleh melalui media perantara atau secara tidak langsung yang berupa buku, catatan, bukti yang telah tervalidasi, atau arsip baik yang dipublikasikan maupun yang tidak dipublikasikan secara umum yaitu 100 ribu data yang akan dipecah menjadi 80 persen data training dan 20 persen data testing. Data testing yang digunakan

sekitar 150 data yang telah diklasifikasikan berdasarkan variable yang ada.

Masalah yang harus dipecahkan pada penelitian kali ini adalah, bagaimana menghasilkan nilai accuracy yang jauh lebih optimal untuk mengklasifikasikan prediksi Calon Nasabah Potensial untuk ditawarkan produk lainnya. Dataset yang digunakan pada penelitian ini ada pada Gambar di bawah ini:

name	kustname	Sex	Age	Marital	tanggungan	Occupation	Wilayah	Keterangan Waktu	Target
2	MIK_SUPER_NYENVIHARDJANTO	P	46	kawin	1	Pegawai Swasta	Barat	Siang	Potensial
3	MIK_SUPER_NLUKMANHAKIM	L	47	kawin	3	Pegawai Swasta	Barat	Malam	Potensial
4	MIK_SUPER_NJAMBARIAL	L	47	kawin	1	Wiraswasta	Barat	Sore	Potensial
5	MIK_SUPER_NWAGINO	L	46	kawin	2	Polisi	Barat	Sore	Potensial
6	MIK_SUPER_NWINARTO	L	61	belum kawin	0	Wiraswasta	Barat	Siang	Potensial
7	MIK_SUPER_NSIHERI	L	59	kawin	2	Pegawai Swasta	Barat	Pagi	Potensial
8	MIK_SUPER_NADEINAMSUHARYANTO	L	57	kawin	1	Wiraswasta	Barat	Sore	Potensial
9	MIK_SUPER_NANDRI FERKA	L	61	kawin	1	Pegawai Swasta	Barat	Pagi	Potensial
10	MIK_SUPER_NEFFY ZULIFLIE	P	47	kawin	1	Wiraswasta	Barat	Sore	Potensial
11	MIK_SUPER_NNURJANNAH	P	57	kawin	3	Pegawai Swasta	Barat	Sore	Potensial
12	MIK_SUPER_NENDANG SRI SOESLOWAT P	P	44	kawin	3	Pegawai Swasta	Barat	Siang	Tidak Potensial
13	MIK_SUPER_NHARY AGUS WIBOWO	L	38	belum kawin	0	Pegawai Swasta	Barat	Siang	Tidak Potensial
14	MIK_SUPER_NSOEPRAHADI	L	40	belum kawin	0	Wiraswasta	Barat	Sore	Tidak Potensial
15	MIK_SUPER_NHAMONANGAN SIRAIT	L	45	kawin	2	Wiraswasta	Barat	Siang	Tidak Potensial
16	MIK_SUPER_NCHERYLASMARADEWI	L	38	kawin	2	Pegawai Swasta	Barat	Sore	Tidak Potensial
17	MIK_SUPER_NHARUN RASYID	L	40	belum kawin	0	Pendidik	Barat	Siang	Tidak Potensial
18	MIK_SUPER_NHADI SUWARNO	L	47	belum kawin	0	Pegawai Swasta	Barat	Sore	Potensial
19	MIK_SUPER_NZAINAL ARIFINS	L	47	belum kawin	0	Pegawai Swasta	Barat	Sore	Potensial

Gambar 1. Dataset

**b. Pengolahan Data**

Dataset ini dalam tahap preprocessing harus melalui 3 proses, yaitu:  
 Tahap ini akan melibatkan sub proses antara lain:

- a) Membuang duplikasi data
- b) Memeriksa data yang inkonsisten
- c) Memperbaiki kesalahan pada data.

Dan Hasil akhir dari data ini berupa kumpulan data yang sudah bersih atau tidak ada missing valuenya.

**c. Evaluasi dan Validasi Data**

Validasi dilakukan menggunakan 10 fold cross validation. Dimana dengan menggunakan teknik ini dengan membagi secara acak ke dalam tiap bagian dimana terdiri dari 10 bagian untuk setiap bagian akan dilakukan proses klasifikasi terlebih dahulu, Sedangkan pengukuran akurasi diukur dengan confusion matrix dan kurva ROC (Receiver Operating Characteristics) untuk mengukur nilai AUC.

AUC digunakan untuk mengukur kinerja diskriminatif dengan memperkirakan probabilitas output yang sudah di dapatkan hasilnya dari sampel yang dipilih secara acak dari populasi positif atau negatif, semakin besar nilai AUC, semakin kuat klasifikasi yang dihasilkan. Karena AUC merupakan bagian dari daerah unit persegi, nilainya yang dihasilkan akan selalu sama yang dihasilkannya, antara 0,0 dan 1,0.

Tabel 1. Confusion Matrix

Classification Observed Class	Predicted Class		
		Class= Yes	Class=No
	Class = Yes	a (True Positive - TP)	b (False Negative - FN)
Class = No	c (False Positive - FP)	d (True Negative -TFN)	

Gambar 2. Sumber : (Gorunescy, 2011)

### Hasil dan Pembahasan

Data yang digunakan dalam pengklasifikasi Calon Nasabah Potensial terdiri dari 5.000 dataset, 137 data yang digunakan untuk data testing berdasarkan variabel yang tersedia. Klasifikasi dilakukan dengan menggunakan Software Rapidminer dengan versi 8 untuk mengolah data yang sudah di tentukan

#### 1. Confusion Matrix algoritma Naïve Bayes

Gambar.3 merupakan perhitungan akurasi data training menggunakan algoritma Naïve Bayes yang menghasilkan accuracy 88.00%. Diketahui data training terdiri dari 150 record data, 10 data di klasifikasikan LAKU ternyata TIDAK LAKU, 2 data diprediksi TIDAK LAKU dan benar-benar TIDAK LAKU, 129 di prediksi LAKU ternyata benar-benar LAKU serta 8 data diprediksi TIDAK LAKU ternyata LAKU.

accuracy: 88.00% +/- 10.67% (mikro: 87.92%)

	true Tidak	true Ya	class precision
pred. Tidak	2	8	20.00%
pred. Ya	10	129	92.81%
class recall	16.67%	94.16%	

Gambar 3. Model Confusion Matrix algoritma Naïve Bayes

PerformanceVector:

accuracy: 88.00% +/- 10.67%  
 (mikro: 87.92%)

ConfusionMatrix:

True: Tidak Ya

Tidak: 2 8

Ya: 10 129

precision: 92.74% +/- 3.26%

(mikro: 92.81%) (positive class: Ya)

ConfusionMatrix:

True: Tidak Ya

Tidak: 2 8

Ya: 10 129

recall: 94.01% +/- 11.73% (mikro: 94.16%) (positive class: Ya)

ConfusionMatrix:

True: Tidak Ya

Tidak: 2 8

Ya: 10 129

AUC (optimistic): 0.514 +/- 0.330  
 (mikro: 0.514) (positive class: Ya)

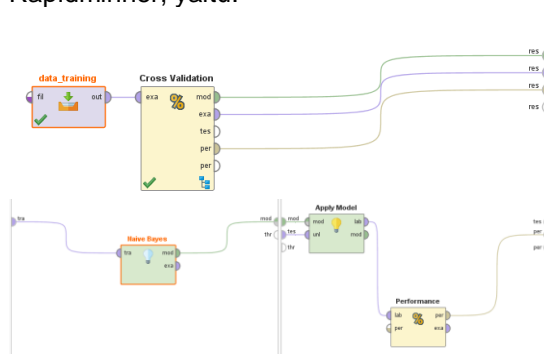
AUC: 0.494 +/- 0.329 (mikro: 0.494) (positive class: Ya)

AUC (pessimistic): 0.478 +/- 0.328 (mikro: 0.478) (positive class: Ya)

Gambar 4 .Text view model Confusion Matrix algoritma Naïve Bayes

Gambar 4. merupakan perhitungan accuracy data menggunakan algoritma Naïve Bayes. Diketahui data training terdiri dari 150 record data, 10 data di klasifikasikan LAKU ternyata TIDAK LAKU, 2 data diprediksi TIDAK LAKU dan benar-benar TIDAK LAKU, 129 di prediksi LAKU ternyata benar-benar LAKU serta 8 data diprediksi TIDAK LAKU ternyata LAKU.

Model yang dihasilkan pada algoritma Naïve Bayes yang dihasilkan menggunakan Rapidminer, yaitu:



Gambar 5. Model algoritma Naïve Bayes

## 2. Algoritma Naïve Bayes

Algoritma Naïve Bayes Bayesian classification adalah pengklasifikasian statistik yang dapat digunakan untuk memprediksi probabilitas keanggotaan suatu class. Bayesian classification didasarkan pada teorema Bayes yang memiliki kemampuan klasifikasi serupa dengan decision tree dan neural network. Bayesian classification terbukti memiliki akurasi dan kecepatan yang tinggi saat diaplikasikan ke dalam database dengan data yang besar [Supriyanto, 2013]. Metode Bayes merupakan pendekatan statistik untuk melakukan inferensi induksi pada persoalan klasifikasi. Pertama kali dibahas terlebih dahulu tentang konsep dasar dan definisi pada Teorema Bayes, kemudian menggunakan teorema ini untuk melakukan klasifikasi dalam Data Mining. Teorema Bayes memiliki bentuk umum sebagai berikut :

$$P(H|E) = \frac{P(E|H) \times P(H)}{P(E)}$$

Gambar 6. Rumus Dasar Algoritma Naïve Bayes

Keterangan :

X = Data dengan class yang belum diketahui  
 H = Hipotesis data X merupakan suatu class spesifik

P(H|X) = Probabilitas hipotesis H berdasarkan kondisi x (posteriori prob.)

P(H) = Probabilitas hipotesis H (prior prob.)

P(X|H) = Probabilitas X berdasarkan kondisi tersebut

P(X) = Probabilitas dari X

Tahapan proses Naive Bayes, yaitu:

1. Menghitung jumlah kelas / label
2. Menghitung Jumlah Kasus Per Kelas
3. Kalikan Semua Variable Kelas
4. Bandingkan Hasil Per Kelas

## 3. Penjualan

Penjualan merupakan sebuah proses dimana kebutuhan pembeli dan kebutuhan penjual dipenuhi, melalui antar pertukaran informasi dan kepentingan. Dalam proses penjualan, penjual atau penyedia barang dan jasa memberikan kepemilikan suatu komoditas kepada pembeli untuk suatu harga tertentu. Secara umum pengertian penjualan adalah sejumlah total yang dikenakan kepada pelanggan untuk barang atau jasa yang di jual.

## Kesimpulan

Sistem Klasifikasi Penjualan Obat ini digunakan untuk menentukan merk vitamin apa yang LAKU atau TIDAK LAKUnya pada sebuah apotek yang ada di Jakarta dengan menggunakan algoritma Naïve Bayes. Penelitian ini menggunakan 150 dataset. Penelitian ini dilakukan karena adanya masalah untuk mengetahui produk pada bulan sebelumnya agar dapat menentukan stok produk mana yang mesti di perbanyak lagi, dan stok barang mana yang harus lebih dipersedikitkan. Pengujian pada data rekapitulasi penjualan Obat dengan proses mining Algoritma Naïve Bayes menghasilkan tingkat accuracy dengan nilai 88.00%, dimana dalam pengujian model data, keseluruhan data set digunakan sebagai data testing. Penelitian ini dilakukan menggunakan tools Rapidminer pada dataset penjualan obat dengan menggunakan algoritma Naïve Bayes.

## Referensi

- Arif, B., Saruni, D. (2018). The Clasification Product Using Naïve Bayes Algorithm (A Case Study at PT Putradabo Perkasa). International Journal of Computer Techniques, 68-74.

- Bustami. (2014). Penerapan Algoritma Naive Bayes Untuk Mengklasifikasi Data Nasabah Asuransi. *Jurnal Informatika* Vol. 8, No. 1, Januari.
- Fajar, E.P., Achmad, K.(2019). Analisis Prediksi Masa Studi Mahasiswa Menggunakan Algoritma Naïve Bayes. *Jurnal Ilmu Teknik dan Komputer*, 147-151.
- Firdaus,Diky (2017). "Penggunaan Data Mining dalam Kegiatan Sistem Pembelajaran Berbantuan Komputer". *Jurnal FORMAT*, Vol 6 Nomor 2, ISSN : 2089-5615
- Prasetyo, E, 2012, *Data Mining : Konsep dan Aplikasi menggunakan Matlab*, Yogyakarta: Andi
- Ratniasih, Ni Luh, 2019. Optimasi Data Mining Menggunakan Algoritma Naïve Bayes dan C4.5 Untuk Klasifikasi Kelulusan Mahasiswa. *Jurnal Teknologi Informasi dan Komputer*, Volume 5, Nomor 1, Januari 2019.
- Saruni, D., Yudo, D. (2018). Naive Bayes Optimization Based On Particle Swarm Optimization To Predict The Decision Of Insurance Customer Candidate. *International Journal of Computer Techniques*, 1-7.
- Supriyanto, Catur. Purnama Parida. 2013. deteksi penyakit diabetes type ii dengan naive bayes berbasis particle swarm optimization. *Jurnal Teknologi Informasi*, Volume 9 Nomor 2, Oktober 2013.
- Angelina Permatasari. 2013. Perancangan Sistem Informasi Penjualan Pada Apotek. *ComTech*. Vol.4 No. 1 Juni 2013: 55-67
- Aditya Prihantara dan Berliana Kusuma Riasti. 2012. Design Dan Implementasi Sistem Informasi Apotek Pada Apotek Mitra Agung Pacitan. *Journal Speed – Sentra Penelitian Engineering dan Edukasi – Volume 4 No 3 - 2012 - ijns.org*