



PENERAPAN DATA MINING UNTUK MEMPREDIKSI PENJUALAN OBAT DI KLINIK HARAPAN KITA BATAM

Mangaratua Hutahaean¹ Koko Handoko²

¹ Mahasiswa Program Studi Teknik Informatika, Universitas Putera Batam

² Dosen Program Studi Teknik Informatika, Universitas Putera Batam
email : pb160210211@upbatam.ac.id

ABSTRACT

Clinics are places or facilities for providing information on drugs and other pharmaceutical supplies to the public. This resulted in uncertain results from sales at the clinic. Types of drugs that are increasingly varied, from drugs that are cheap to prices that if you look at it don't make sense but have very good functions, especially in the distribution of types of drugs. To overcome this problem, the classification of drugs that are sold or not sold at the clinic is based on the variables obtained using the Naive Bayes Algorithm which is able to provide information about the existing products in the clinic and minimize the stock that accumulates for the unsold category. The availability of a lot of data in an area of need for information becomes a reference in making decisions to make business solutions. The desired information in this study is to obtain an accuracy value for the sales data for these drugs, which are often an option. This research uses Rapidminer tools as a medium for testing data to be processed to obtain an ROC (Rank Order Centroid) value of 80% accuracy based on the importance or priority of the criteria.

Keywords: Data mining, Drug Sales Prediction, Naive Bayes Algorithm,

PENDAHULUAN

Data mining adalah rangkaian proses pendukung pengambilan keputusan dalam mengenali nilai tambah dalam suatu kumpulan data berupa pengetahuan yang selama ini tidak diketahui secara manual (Fajrin and Handoko 2018). Data mining dapat diterapkan pada berbagai bidang yang mempunyai sejumlah data. Menurut Daryl Pregibon data mining adalah campuran dari statistik, kecerdasan buatan dan riset basis data yang masih berkembang. Klinik merupakan tempat atau sarana pelayanan informasi obat

dan perbekalan farmasi lainnya kepada masyarakat. Saat sekarang ini banyak klinik yang berdiri dan tersebar hingga kepedesaan sehingga memudahkan masyarakat untuk mengakses peralatan farmasi dan obat-obatan (Wijaya and Dwiasnati 2020). Dalam dunia kesehatan khususnya obat-obatan yang berada di klinik sudah sangat mengesankan dan menarik untuk ditelusuri dalam kuantitas dan jenis merek yang ada. Pertumbuhan bisnis kesehatan di klinik semakin berkembang dari waktu ke waktu, hal ini dapat dilihat banyaknya klinik yang tersebar dimana aja. Dalam hal ini klinik

berlomba-lomba dalam menyediakan beragam-ragam jenis obat untuk kesehatan serta vitamin agar mempermudah pelanggan dalam memenuhi kebutuhannya. Terdapat permasalahan yang sering terjadi pada sebuah klinik yaitu dalam penentuan kategori laku atau tidak laku sebuah produk, misalnya dalam pemberian informasi mengenai minat para konsumen kurang untuk membeli produk tersebut sehingga mengalami penumpukan di gudang sedangkan minat membeli produk yang lain sering terjadi ketidaktersedian produk pada klinik. Ini mengakibatkan tidak menentunya hasil dari penjualan di klinik. Jenis obat yang makin lama makin bervariasi, dari obat yang berharga murah sampai harga yang kalau dilihat tidak masuk akal tapi memiliki fungsi yang sangat bagus terutama pada peredaran jenis obat.

Untuk mengatasi masalah tersebut maka dilakukan nya pengklasifikasian obat yang laku atau tidak laku terjual di klinik tersebut berdasarkan variabel-variabel yang di dapatkan dengan menggunakan Algoritma *Naive Bayes* yang mampu memberikan informasi mengenai produk yang ada di klinik dan meminimalisirkan stok yang menumpuk untuk kategori tidak laku. Ketersedian data yang banyak dalam sebuah bidang kebutuhan untuk informasi menjadi acuan dalam pengambilan keputusan untuk membuat solusi bisnis. Informasi yang diinginkan dalam penelitian ini adalah mendapatkan nilai akurasi untuk data penjualan obat-obatan tersebut yang sering menjadi pilihan. Penelitian ini menggunakan *tools Rapidminer* sebagai media pengujian data yang akan diolah untuk *mendapatkan hasil akurasi nilai ROC (Rank Order Centroid)* bernilai 80% berdasarkan tingkat kepentingan atau prioritas dari kriteria.

KAJIAN TEORI

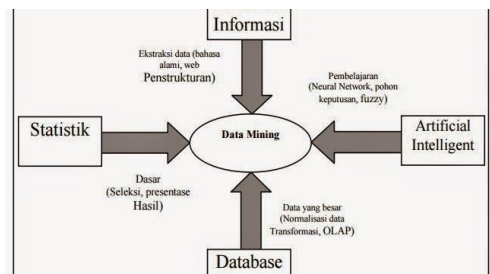
Knowledge Discovery in Database (KDD)

merupakan ilmu yang menemukan pengetahuan didalam database. Serta

didefinisikan sebagai proses ekstraksi atau identifikasi pola, pengetahuan dan informasi potensial dari sekumpulan data yang besar dan data yang dihasilkan bersifat sah, baru dan mudah dimengerti dan jika dilihat banyak konsep dan teknik yang digunakan dalam proses data mining.

Data Mining

Data mining adalah rangkaian proses pendukung pengambilan keputusan dalam mengenali nilai tambah dalam suatu kumpulan data berupa pengetahuan yang selama ini tidak diketahui secara manual (Elmayati 2017).



Gambar 1 Bidang ilmu data mining
Sumber : (Elmayanti, 2017)

Fungsi Data mining yang lainnya yaitu: *classification, association, clustering,*

Classification

Classification adalah tindakan untuk memberikan kelompok pada setiap keadaan. Setiap keadaan berisi sekelompok atribut, salah satunya adalah *class attribute*. Metode ini butuh untuk menemukan sebuah model yang dapat menjelaskan *class attribute* itu sebagai fungsi dari *input attribute*. Class adalah attribute *CollegePlans* yang berisi dua pernyataan, *Yes* dan *No* (Ulfha and Amin 2020).

Metode Algoritma Naive Bayes

Algoritma *Naive Bayes* merupakan teknik prediksi berbasis probalistik sederhana yang berdasarkan pada penerapan *teorema bayes* dengan asumsi independensi yang kuat dan suatu metode yang dapat menangani

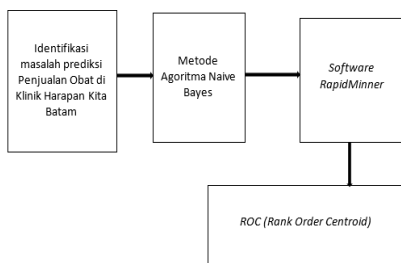
faktor prediksi pada bulan selanjutnya dan trend secara langsung. Keuntungan dari algoritma ini adalah memiliki kemampuan yang sangat baik dalam mengestimasi banyaknya penjualan pada bulan selanjutnya (Wijaya and Dwiasnati 2020).

Software Pendukung Rapid Miner

Rapidminer merupakan perangkat lunak yang bersifat terbuka (*open source*). *Rapidminer* adalah sebuah solusi untuk melakukan analisis terhadap data mining, *text mining* dan analisis prediksi. *Rapidminer* menggunakan berbagai teknik deskriptif dan prediksi dalam memberikan wawasan kepada pengguna sehingga dapat membuat keputusan yang paling baik (Efori Buulolo 2017).

Kerangka Pemikiran

Kerangka pemikiran merupakan sebuah gagasan atau ide yang di desain dalam bentuk alur diagram untuk menjelaskan urutan atau langkah dari awal proses hingga akhir proses penelitian. Kerangka penelitian dalam penelitian ini dibuat dari adanya masalah dalam prediksi penjualan obat. Berikut kerangka penelitian nya



Teknik Pengumpulan Data

Adapun teknik- teknik pengumpulan data yang peneliti lakukan dalam penelitian adalah sebagai berikut:

1. Teknik *Observasi*

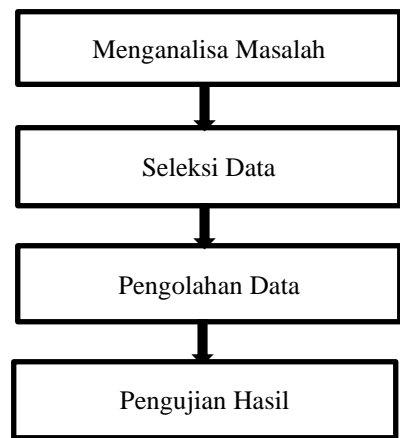
Pada tahap ini peneliti melakukan pengamatan secara langsung untuk mengetahui kejadian di Klinik Harapan Kita Batam, dalam hal ini peneliti mengambil data hasil transaksi

Gambar 2 Kerangka Pemikiran
Sumber : Data penelitian 2022

METODE PENELITIAN

Desain Penelitian

Dalam desain penelitian, peneliti memaparkan hasil desain penelitian yang telah dibuat dengan cara yang terstruktur agar dapat diambil sebagai acuan untuk melakukan peneliti dan membantu peneliti pada saat proses penelitian, desain penelitian juga sangat berperan penting sebagai pedoman bagi peneliti.



Gambar 3 Desain Penelitian
Sumber : Data penelitian 2022

penjualan yang ada pada database sesuai dengan data yang dibutuhkan dalam menyelesaikan masalah pada penelitian ini. Observasi juga memberikan data yang lebih real kepada peneliti karena peneliti bisa lebih leluasa dalam mengambil data di lapangan.

Operasional Variabel

Variabel harus didefinisikan secara operasional agar mudah dicari hubungannya antara satu variabel dengan lainnya dan pengukurannya. Tanpa operasional variabel, peneliti akan kesulitan dalam menentukan pengukuran hubungan antar variabel yang masih bersifat konseptual. Dalam penelitian ini operasional variabelnya sebagai berikut:

1. Tahun

Data dipengaruhi berdasarkan jumlah transaksi setiap produk yang akan menentukan kombinasi-kombinasi item set mana yang paling banyak terjual. Dari penelitian ini diambil dari tahun 2018-2020.

2. Jenis Obat

Jenis obat yaitu obat yang sering dibeli oleh konsumen berdasarkan penyakit yang dialami. Jenis obat yang paling sering diminati adalah obat nyeri, infeksi, vitamin dan demam.

3. Transaksi

Merupakan jumlah pembelian dari konsumen yang harus di sediakan pada klinik setiap bulan nya dengan kategori sedikit, hampir, banyak yang dilihat dari permintaan obat.

4. Penjualan Obat

Merupakan pengeluaran obat yang telah terjual dan di kategorikan dengan laku dan tidak laku.

HASIL DAN PEMBAHASAN

Hasil Penelitian

Hasil penelitian ini bertujuan untuk mengetahui nilai akurasi pada perhitungan algoritma naïve bayes dalam penjualan obat dengan cara mengolah data menggunakan *microsoft excel* , kemudian di urutkan dalam (sort & filter) lalu di bagi menjadi pembagian class, setelah itu data diuji dalam *Software Rapidminer*, hasil dari penelitian ini berupa proses perhitungan berdasarkan model naïve bayes.

Data set

Dataset penjualan terdiri dari beberapa atribut antara lain Tahun, Jenis obat, Transaksi, Persiapan data mencakup semua kegiatan untuk membangun dataset penjualan yang akan diterapkan ke dalam alat pemodelan, dari data mentah awal berupa dataset penjualan dan selanjutnya akan melakukan proses data mining. Berikut tabel 4.1 adalah data set yang akan di lakukan pengujian nya.

Tabel 1 Data set

Tahun	Tahun Data	Jenis Obat	Nama Obat	Transaksi	Permintaan	Penjualan Obat
2018	Jan-18	Nyeri	Acyclovir salp	Sedikit	270	Tidak
2018	Feb-18	Nyeri	Acyclovir tab 200 mg	Hampir	520	Laku
2018	Mar-18	Nyeri	Acyclovir tab 400 mg	Hampir	500	Laku
2018	Apr-18	Infeksi	Albendazol tab	Sedikit	100	Tidak
2018	Mei-18	Infeksi	Allopurinol	Hampir	740	Laku
2018	Jun-18	Infeksi	Alpara	Hampir	530	Laku
2018	Jul-18	Vitamin	Alpara syr	Sedikit	44	Tidak
2018	Agu-18	Vitamin	Alphamol drop	Sedikit	200	Tidak
2018	Sep-18	Vitamin	Ambeven tab	Sedikit	120	Tidak
2018	Okt-18	Demam	Ambroxol syr	Sedikit	78	Tidak
2018	Nov-18	Demam	Ambroxol tab	Hampir	1100	Tidak
2018	Des-18	Demam	Amlodipin 10 mg tab	Hampir	980	Laku

2019	Jan-19	Nyeri	Amlodipin 5 mg tab	Sedikit	130	Tidak
2019	Feb-19	Nyeri	Amoxan tab	Sedikit	50	Tidak
2019	Mar-19	Nyeri	Amoxicilin syr generik	Sedikit	45	Tidak
2019	Apr-19	Infeksi	Amoxicilin tab generik	Hampir	1300	Laku
2019	Mei-19	Infeksi	Ampicillin	Sedikit	450	Tidak
2019	Jun-19	Infeksi	Anacetin syr	Sedikit	180	Tidak
2019	Jul-19	Vitamin	Antasida tab	Sedikit	270	Tidak
2019	Agu-19	Vitamin	Antasida syr	Sedikit	77	Tidak
2019	Sep-19	Vitamin	Antihemoroid supp	Sedikit	43	Tidak
2019	Okt-19	Demam	Asam Folat	Sedikit	50	Tidak
2019	Nov-19	Demam	Aqua 25 ml	Sedikit	280	Tidak
2019	Des-19	Demam	Asam mefenamotab tab	Hampir	1250	Laku

Sumber: Data Penelitian 2022

Dari tabel data diatas maka dapat di tentukan formula perhitungan dengan menggunakan algoritma naive bayes

dan dapat ditentukan jumlah obat yang terjual berdasarkan class seperti tabel di bawah ini:

Tabel 2 Probalitas Kelas

Probalitas Kelas	
Prediksi	Nilai
Laku	31%
Tidak	69%
Jumlah	100%

Analisa Proses Agotitma

Pada tahap ini dilakukan pemodelan data, metode yang dipakai pada penelitian ini adalah probabilitas (prediksi) dengan menggunakan algoritma Naive Bayes. Data yang telah dikumpulkan, akan dikelola menggunakan probabilitas. Metode ini dapat digunakan dalam memprediksi peluang di masa depan berdasarkan

pengalaman di masa sebelumnya sebagai perbandingan. Data yang akan diujikan dibagi menjadi dua bagian yaitu training dan testing kemudian dianalisis menggunakan software Rapidminer. Untuk data training 36 record (tahun 2018, 2019, 2020 dan data testing 5 record.

Tabel 3 Data Training

Tahun	Jenis Obat	Transaksi	Penjualan Obat
2018	Nyeri	Sedikit	Tidak
2018	Nyeri	Hampir	Laku
2018	Nyeri	Hampir	Laku

2018	Infeksi	Sedikit	Tidak
2018	Infeksi	Hampir	Laku
2018	Infeksi	Hampir	Laku
2018	Vitamin	Sedikit	Tidak
2018	Vitamin	Sedikit	Tidak
2018	Vitamin	Sedikit	Tidak
2018	Demam	Sedikit	Tidak
2018	Demam	Hampir	Tidak

Sumber : Data Penelitian 2022

Tabel 4 Data Testing

Tahun	Jenis Obat	Transaksi	Penjualan Obat
2018	Nyeri	Hampir	Laku
2019	Nyeri	Sedikit	Tidak
2020	Infeksi	Sedikit	Tidak
2020	Demam	Hampir	Laku
2020	Demam	Banyak	Laku

Sumber:Data Penelitian 2022

Dari tabel data training pada tabel 4 diatas maka dilakukan pencarian probalitas dengan metode agoritma naive bayes seperi dibawah ini:

$$P(T|X) = \frac{P(X|T)P(T)}{P(X)}$$

Probabilitas Kriteria Tahun

Berdasarkan data penjualan obat pada data set diatas, diketahui bahwa

terdapat 36 data tiap perbulan nya. Dari 36 data tersebut terdapat 45% dengan penjualan obat laku dan 28% dengan penjualan obat tidak laku pada tahun 2018, terdapat 18% dengan penjualan obat laku dan 40% dengan penjualan obat tidak laku pada tahun 2019, terdapat 36% dengan penjualan obat laku dan 32% dengan penjualan obat tidak laku pada tahun 2020. Berikut tabel Probalitas kriteria Tahun:

Tabel 5 Probabilitas Kriteria Tahun

P(T)=	Laku	Tidak
2018	45%	28%
2019	18%	40%
2020	36%	32%
Jumlah	100%	100%

Sumber: Data penelitian 2022

Hasil Pengujian 1

Hasil pengujian 1 yang diberikan 5 data testing atau data uji fungsinya untuk membuktikan apakah model sudah sesuai dalam artian akurat

menggunakan microsoft excel sehingga mendapatkan class prediksi dalam proses pencarian nilai dan mendapatkan nilai *accuracy* 80%. Berikut tabel dibawah ini adalah tabel data testing:

Tabel 6 Hasil Data Testing

Tahun	Tahun Data	Jenis Obat	Nama Obat	Transaksi	Permintaan	Penjualan Obat	Class Prediction	Laku	Tidak
2018	Mar-18	Nyeri	Acyclovir tab 400 mg	Hampir	500	Laku	LAKU	3,4%	0,2%
2019	Apr-19	Infeksi	Amoxicilin tab generik	Hampir	1300	Laku	LAKU	1,8%	0,2%
2020	Nov-20	Demam	Buscopan inj	Hampir	700	Laku	LAKU	3,7%	0,2%
2019	Jun-19	Infeksi	Anacetin syr	Sedikit	180	Tidak	TIDAK	0,0%	5,3%
2020	Jan-20	Nyeri	Asam tranexamat tab	Sedikit	400	Tidak	TIDAK	0,0%	5,1%

Sumber: Data Penelitian 2022

Setelah data testing sudah dicari dan didapatkan nilainya maka proses selanjutnya

adalah mencari nilai akurasi. Berikut nilai akurasi yang di dapat dari tabel diatas.

Tabel 7 Class Prediksi

Class		
Prediksi	Laku	Tidak
Laku	3	1
Tidak	0	1

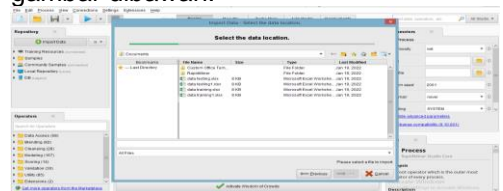
Sumber: Data Penelitian 2022

$$Accuracy = 80\%$$

Implementasi Rapidminer

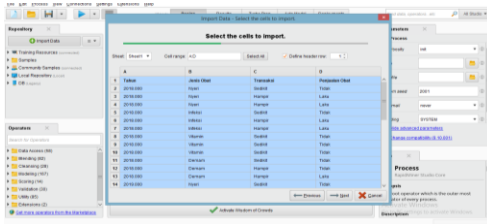
Pada tahap ini penulis mengambil data training tahun 2018, 2019, 2020 dan data testing sebanyak 5 record. Atribut yang digunakan sebagai label adalah penjualan obat. yang dipakai sebagai data testing dengan format excel untuk memprediksi penjualan obat yang mendapatkan laku dan tidak agar bisa

diakses menggunakan software RapidMiner. Selanjutnya lakukan importing data terlebih dahulu yaitu data training dan data testing yang ada pada komputer atau laptop seperti pada gambar dibawah:



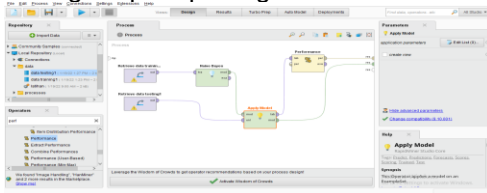
Gambar 4 Import Data Training
Sumber : Data Penelitian 2022

Setelah data training telah dipilih langkah selanjutnya yaitu klik Next dan akan muncul form data import wizard step 2. Pada step 2 ini tidak dilakukan perubahan apapun maka langsung ke data import wizard step 3 dengan cara klik Next maka akan muncul form data import wizard step 3, terlihat pada gambar 5.



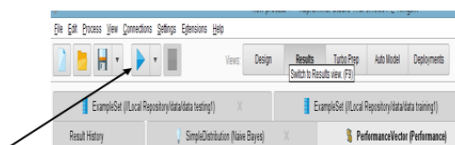
Gambar 5 Data import step 3
Sumber : Data Penelitian 2022

Kemudian setelah kedua data sudah di import maka step selanjutnya yaitu drag and drop operators Naive bayes, apply model, dan performance lalu hubungkan dengan data training dan testing. Seperti pada gambar 6.



Gambar 6 Menghubungkan operator
Sumber : Data Penelitian 2022

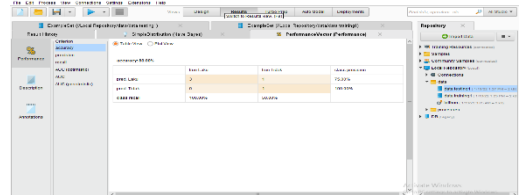
Setelah semuanya terhubung, klik icon run pada toolbar, seperti pada gambar 4.6 untuk menampilkan hasil. Tunggu beberapa saat, komputer memerlukan waktu untuk menyelesaikan perhitungan.



Gambar 7 Icon Run
Sumber : Data Penelitian 2022

Dari hasil proses perhitungan menggunakan rapid miner dengan metode prediksi menampilkan hasil dari

data testing dan training yang telah diuji. Kolom ini memberikan informasi tentang data penjualan obat yang mendapatkan laku dan tidak. Selanjutnya untuk mengetahui tingkat accuracy algoritma naive bayes, klik tab performanceVector yang terletak disebelah kanan, sehingga table view seperti gambar 8..



Gambar 4. 1 Hasil Perhitungan
Sumber: Data Penelitian 2022

SIMPULAN

Dari hasil analisa dan pembahasan maka didapat kesimpulan sebagai berikut:

1. Data penjualan obat diambil dari Klinik Harapan Kita Batam dari tahun 2018-2020, dengan mengambil 36 record data sebagai data set yang akan dilakukan pengujian nya, maka hasil yang didapatkan dari proses data mining sangat berguna untuk mengetahui hubungan pola frekuensi Laku atau Tidak penjualan jenis obat yang paling sering dibeli konsumen.
2. Dari 36 data set yang ada maka diambil 5 data untuk dijadikan data testing dengan dilakukan pengujian menggunakan Metode Naive Bayes dan perhitungan melalui microsoft excel dan rapidminer berhasil memprediksi penjualan obat dengan persentase keakuratan atau *accuracy* sebesar 80 %.
3. Dengan metode algoritma naive bayes dapat disimpulkan bahwa jenis obat yang paling banyak terjual adalah obat nyeri, infeksi, vitamin dan demam Dari 4 data tersebut terdapat kriteria nyeri 27% dengan penjualan obat laku dan 24% dengan penjualan obat tidak laku, terdapat kriteria infeksi 36% dengan penjualan obat laku dan 20% dengan penjualan obat

- tidak laku, terdapat kriteria vitamin 0% dengan penjualan obat laku dan 36% dengan penjualan obat tidak laku, terdapat kriteria Demam 36% dengan penjualan obat laku dan 20% dengan penjualan obat tidak laku.

DAFTAR PUSTAKA

Efori Buulolo, Stmik Budi. 2017. “Implementasi Algoritma Apriori Pada Sistem Persediaan Obat (Studi Kasus : Apotik Rumah Sakit Estomih Medan) Implementasi Algoritma Apriori Pada Sistem Persediaan Obat (Studi Kasus : Apotik Rumah Sakit Estomih Medan).” *Jurnal Teknik Informatika* 4(August 2013):71–83.

Elmayati, Elmayati. 2017. “Data Mining Dengan Metode Clustering Untuk Pengolahan Informasi Persediaan Obat Pada Klinik Srikandi Medika Berbasis Web.” *Pelita Informatika: Informasi Dan Informatika* 16(4):357–62.

Fajrin, Alfanisa Annurrullah, and Koko Handoko. 2018. “Penerapan Data Mining Untuk Mengolah Association Rule Tata Letak Buku Dengan Metode.” *Jurnal Ilmiah Informatika (JIF)* 2:60–65.

Ulfha, Nadya Febrianny, and Ruhul Amin. 2020. “Implementasi Data Mining Untuk Mengetahui Pola Pembelian Obat Menggunakan Algoritma Apriori.” *Komputasi: Jurnal Ilmiah Ilmu Komputer Dan Matematika* 17(2):396–402. doi: 10.33751/komputasi.v17i2.2150.

Wijaya, Herry Derajad, and Saruni Dwiasnati. 2020. “Implementasi Data Mining Dengan Algoritma Naïve Bayes Pada Penjualan Obat.” *Jurnal Informatika* 7(1):1–7. doi: 10.31311/ji.v7i1.6203.

	<p>Mangaratua Hutahaeen merupakan mahasiswa Prodi Teknik Informatika dari Universitas Putera Batam yang aktif dalam mendalami bidang teknologi dan informasi.</p>
	<p>Koko Handoko, merupakan Dosen Prodi Teknik Informatika dari Universitas Putera Batam. Aktif sebagai tenaga kerja dan peneliti.</p>