

Klasifikasi Data Penggunaan Obat Dengan Algoritma Naïve Bayes Di Rumah Sakit Umum Daerah (Rsud) Kota Pagar Alam

Nurmaleni^{1*}, Risnaini Masdalipa²

^{1,2}Teknik Informatika, Institut Teknologi Pagar Alam, Pagar Alam, Indonesia

Email: ¹lenipga@gmail.com, ²risnainipga@gmail.com

Email Penulis Korespondensi: ¹lenipga@email.com

Abstrak—Mengklasifikasikan data obat menggunakan algoritma Naive Bayes pada rumah sakit umum (RSUD) daerah Kota Pagaralam. Pengolahan data obat yang belum optimal dalam menentukan obat yang digunakan di Rumah Sakit Umum Daerah (RSUD) Kota Pagar Alam yang masih dilakukan secara manual dan belum tersruktur sehingga mengalami kesulitan dalam penyampaian informasi obat. Maka perlu dilakukan klasifikasi data obat menggunakan algoritma Naïve Bayes agar data obat terstruktur dan mudah dipahami dengan hasil klasifikasi berupa grafik. Dalam penelitian ini klasifikasi Data obat dengan bentuk obat yang paling sering digunakan berdasarkan data pengurangan yang menampilkan hasil berupa grafik dimana bentuk obat yang paling sering digunakan yang pertama yaitu Tablet, kolf, ampul, box, vial, botol, dan tube. Metode pengembangan pada klasifikasi ini adalah menggunakan metode Microsoft's Team Data Science Process dengan tahapan yaitu : pemahaman bisnis, memahami perolehan data, pemodelan, penyebaran dan hasil untuk menjadi panduan pengerjaan klasifikasi data obat. Untuk pengujian model klasifikasi dilakukan dengan menggunakan metode confusional matrix untuk mengukur akurasi dalam pengujian ini di dapat akurasi 98.99%. Hasil penelitian ini menunjukkan bahwa klasifikasi terdapat 7 warna yang membedakan classnya yang pertama yaitu warna biru muda Tablet berjumlah 55,200, hijau kolf berjumlah 20,254, oren ampul 19,400, hitam box 19,000, biru tua vial 11,750, pink botol 4,727, dan kuning tube 354.

Kata Kunci: Klasifikasi, Data obat RSUD, Naïve Bayes, Rumah Sakit,

Abstract—This study aims to classify drug data using the Naive Bayes algorithm at the Pagaralam City Regional General Hospital (RSUD). Drug data processing that is not optimal in determining the drugs used at the Pagaralam City Regional General Hospital (RSUD) is still done manually and is not structured so that it is difficult to convey drug information. So it is necessary to classify drug data using the Naive Bayes algorithm so that the drug data is structured and easy to understand with the classification results in the form of graphs. In this study, the classification of drug data with the most frequently used drug forms is based on reduced data that displays the results in the form of graphs where the most frequently used drug forms are tablets, kolf, ampules, boxes, vials, bottles, and tubes. The development method in this classification is to use the Microsoft's Team Data Science Process method with the following stages: understanding the business, understanding data acquisition, modeling, distribution and results to be a guide for working on drug data classification. For testing the classification model, the confusional matrix method was used to measure accuracy in this test, an accuracy of 98.99% was obtained. The results of this study indicate that the classification of drug forms in the form of tablets is 55,200, kolf is 20,254, ampoules are 19,400, boxes are 19,000, vials are 11,750, bottles are 4,727, and tubes are 354.

Keywords: Classification, Hospital medicine data, Naïve Bayes algorithm, Hospital

1. PENDAHULUAN

1.1 Latar Belakang

Dengan perkembangan teknologi yang semakin pesat disemua bidang baik dari sektor Pemerintahan, Pendidikan, Pertanian dan khususnya di lingkungan kesehatan, saat ini di era globalisasi teknologi dapat memberikan informasi yang cepat dan akurat baik untuk tim kesehatan, dokter, perawat bahkan pasien sendiri agar lebih mudah mengontrol kesehatan mereka sendiri [1]. Obat adalah zat yang digunakan untuk diagnosis, mengurangi rasa sakit, serta mengobati atau mencegah penyakit pada manusia atau hewan. Masyarakat saat mengonsumsi obat, harus memperhatikan beberapa hal yang tertera pada kemasan. produk, bahan yang terkandung di dalamnya, kategori obat, petunjuk penggunaan, dosis, efek samping dan tanggal kadaluarsa [2].

Obat merupakan salah satu unsur penting dalam pelayanan kesehatan yang harus selalu tersedia dan tidak tergantikan pada pelayanan kesehatan. Obat dapat merugikan kesehatan, bila digunakan secara tidak tepat atau bila disalah gunakan. Terkait dengan adanya penggolongan obat maka pemberian obat kepada konsumen ada aturan-aturan yang harus diterapkan [3]. Pada era globalisasi sekarang perkembangan Teknologi Informasi sudah sangat berkembang pesat, menuntut setiap kegiatan dalam kehidupan sehari-hari akan selalu menggunakan teknologi berbasis komputer. Perkembangan Teknologi Informasi semakin pesat dengan munculnya berbagai software dan hardware pendukung yang terbilang canggih. Dalam perkembangannya Teknologi Informasi sudah banyak diterapkan di berbagai bidang baik di bidang kesehatan, ekonomi, maupun pemerintahan. Pada bidang kesehatan penggunaan Teknologi Informasi sangat diperlukan contohnya pada apotek-apotek Teknologi Informasi diperlukan untuk melakukan pengelolaan data, seperti pengelolaan data obat, obat masuk, obat keluar, pembuatan laporan, maupun pembelian obat menggunakan data mining [4].



Data mining adalah proses menganalisa data dari sudut pandang yang berbeda dan mengubahnya menjadi informasi-informasi penting yang dapat digunakan untuk meningkatkan keuntungan, mengurangi biaya atau dari keduanya. Data mining merupakan aktivitas yang melibatkan pengumpulan data dan penggunaan data historis untuk menemukan keteraturan, pola atau hubungan dalam data besar kemudian mengekstrak data tersebut menjadi informasi-informasi yang nantinya dapat digunakan. Klasifikasi merupakan salah satu model data mining yang sering digunakan untuk menyelesaikan masalah penentuan dalam mengambil keputusan, klasifikasi juga disebut supervised learning yakni memasukkan klasifikasi nilai respon kategori untuk memisahkan data ke dalam kelas-kelas tertentu [5].

Metode Naïve Bayes dengan prinsip teorema memiliki kesamaan dalam perhitungan satu sama lain. Pendekatan yang digunakan Teorema Bayes adalah menghitung probabilitas sebuah kejadian dalam kondisi tertentu. Sedangkan analisis Regresi Logistik adalah analisis digunakan untuk memprediksi hasil variable yang terkait bersifat kategori yang nilainya berdasarkan satu atau dua variable yang bebas. Teorema Bayes menggunakan asumsi independensi (ketidaktergantungan) sehingga dalam algoritma Naïve Bayes menggunakan model fitur yang independent [6]. Teorema Bayes menerangkan hubungan antara probabilitas terjadinya peristiwa A dengan syarat peristiwa B telah terjadi dan probabilitas terjadinya peristiwa B dengan syarat peristiwa A telah terjadi. Teorema ini didasarkan pada prinsip bahwa tambahan informasi dapat memperbaiki probabilitas. Teorema Bayes ini bermanfaat untuk mengubah atau memutakhirkan (meng- update) probabilitas yang dihitung dengan tersedianya data dan informasi tambahan [7].

Berdasarkan hasil observasi yang dilakukan penulis di Rumah Sakit Umum Daerah (RSUD) Kota Pagar Alam peneliti menemukan masalah pengolahan data obat yang belum optimal dalam menentukan obat yang digunakan di Rumah Sakit Umu Daerah (RSUD) Kota Pagar Alam berdasarkan nama obat, bentuk dan pengurangan, yang masih dilakukan secara manual dan belum tersruktur sehingga mengalami kesulitan dalam penyampaian informasi obat. Maka perlu dilakukan klasifikasi data obat menggunakan algoritma Naïve bayes agar data obat tersruktur dan mudah dipahami dengan hasil klasifikasi berupa grafik. Dengan menggunakan algoritma Naïve Bayes dapat dilalukan pengklasifikasian data obat yang ada di apotik Rumah Sakit Umum Daerah (RSUD) Kota Pagar Alam, Untuk menghitung porbabilitas dan nilai kemungkinan untuk menentukan kelas-kelas tertentu dari sebuah data yang digunakan.

Berdasarkan penelitian yang berjudul “Klasifikasi Kategori Obat Menggunakan Algoritma Support Vector Machine” dapat di simpulkan bahwa Algoritma Support Vector Machine telah berhasil di implementasikan dan menunjukkan bahwa akurasi terbaik dari ketiga kernel SVM didapatkan ketika menggunakan kernel linear dan polinomial. berbeda dengan kernel RBF yang memiliki akurasi yang lebih rendah dari kedua kernel tersebut. kernel linear dan polinomial menghasilkan akurasi sebesar 95.0% sedangkan RBF menghasilkan akurasi sebesar 94.5% [8].

Berdasarkan penelitian yang berjudul “Penerapan Data Mining Dengan Algoritma C4.5 Dalam Pemesanan Obat Guna Meningkatkan Keuntungan Apotek” dapat di simpulkan bahwa Algoritma C4.5 telah berhasil di implementasikan dan mendapatkan Tingkat akurasi hasil yaitu sebesar 87.56%, Presisi 78.72%, Recal 95.03% dan Error rate sebesar 12.44% [9]. Berdasarkan penelitian yang berjudul “Rancang Bangun Aplikasi Data Mining untuk Klasifikasi Pemakaian Obat dengan Metode Naïve Bayes pada Puskesmas Bandar baru” dapat di simpulkan bahwa Algoritma Naïve Bayes telah berhasil di implementasikan dan Penelitian ini berhasil menerapkan metode Naïve Bayes ke dalam aplikasi. Hasil presentase akurasi pengujian pertama adalah 78,58%, pengujian kedua adalah 76,67%, pengujian ketiga adalah 84% (Fachrina, 2020). Hubungan antara ke tiga penelitian ini dengan penelitian yang penulis lakukan yaitu penelitian ini sama-sama menggunakan naive bayes untuk melakukan klasifikasi data obat.

Dari Latar belakang permasalahan dan tiga penelitian terdahulu di atas, dapat di tarik kesimpulan bahwa klasifikasi data penggunaan obat di Rumah sakit Umum daerah (RSUD) Kota Pagar Alam sangat penting bagi RSUD untuk menampilkan sebuah grafik informasi agar mempermudah pihak rumah sakit dalam mengelola data obat. Oleh karena itu Peneliti mengangkat judul penelitian yang diharapkan dapat membantu Rumah Sakit Umum Daerah (RSUD) Kota Pagar Alam dalam mengelola data obat untuk mengetahui informasi bentuk obat yang paling sering digunakan. Dengan judul Penelitian “KLASIFIKASI DATA PENGGUNAAN OBAT DENGAN ALGORITMA NAÏVE BAYES DI RUMAH SAKIT UMUM DAERAH (RSUD) KOTA PAGAR ALAM”.

2. METODOLOGI PENELITIAN

2.1 Metode Pengumpulan Data

2.2.1 Observasi



Observasi adalah mengemukakan observasi merupakan sebuah pengamatan secara langsung terhadap suatu objek yang ada di lingkungan baik itu yang sedang berlangsung atau masih dalam tahap yang meliputi berbagai aktivitas perhatian terhadap suatu kajian objek yang menggunakan pengindraan. Dan merupakan dari suatu tindakan yang dilakukan secara sengaja atau sadar dan juga sesuai urutan [10].

2.2.2 Wawancara

Wawancara adalah teknik pengumpulan data melalui proses tanya jawab lisan yang berlangsung satu arah, artinya pertanyaan datang dari pihak yang mewawancarai dan jawaban diberikan oleh yang diwawancarai. Menurut Hopkins, wawancara adalah suatu cara untuk mengetahui situasi tertentu di dalam kelas dilihat dari sudut pandang yang lain. Wawancara adalah bentuk komunikasi langsung antara peneliti dan responden [11].

Tabel 1. Wawancara

NO.	Pertanyaan	Jawaban Narasumber
1.	Bagaimana cara menentukan stok obat?	
2.	Bagaimana rumah sakit mengelolah data obat saat ini?	
3.	Bagaimana pengelompokan data obat yang ada di rumah sakit?	
4.	Ada berapa jenis golongan data obat?	

2.2.3 Teknik Dokumentasi

Dokumentasi merupakan suatu teknik mengumpulkan data yang berbentuk tulisan, gambar, atau karya-karya monumental dari seseorang. Dokumen yang berbentuk tulisan misalnya catatan harian, sejarah kehidupan, cerita, biografi, peraturan dan kebijakan [12].

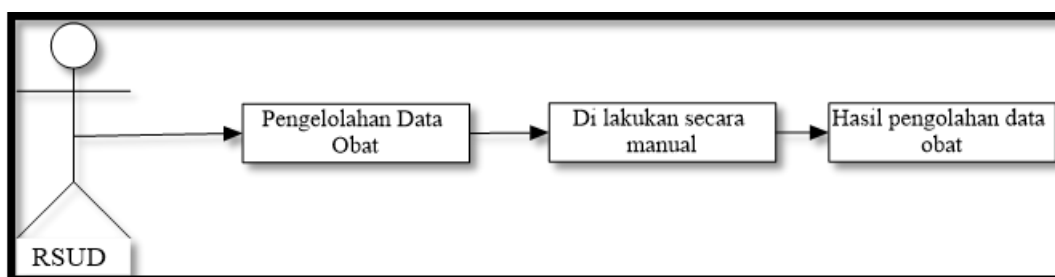
2.2.4 Teknik Studi Pustaka

Yaitu metode yang datanya diambil dari literatur dan jurnal-jurnal yang mendukung penelitian. Data-data di sini bersifat siap pakai yang artinya peneliti tidak terjun langsung kelapangan [13].

2.2.5 Observasi

Observasi adalah mengamati dan mencatat fakta yang dibutuhkan peneliti. Alat untuk mengobservasi dapat berupa lembar observasi atau checklist [14].

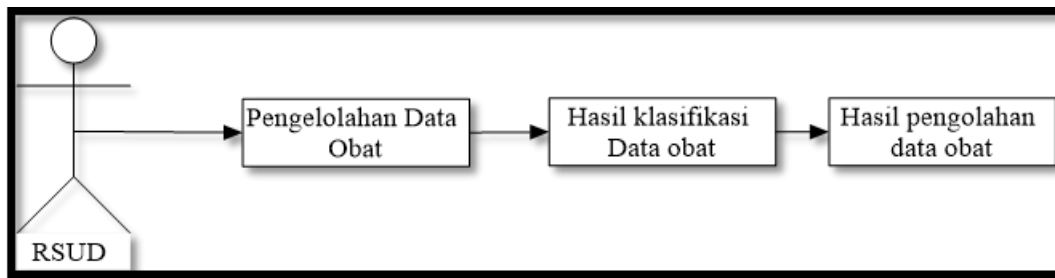
2.2 Sistem Yang Berjalan



Gambar 1. Sistem Yang berjalan

Metode yang berjalan saat ini Rumah Sakit Umum Daerah (RSUD) kota Pagaram melakukan pengolahan data secara manual.

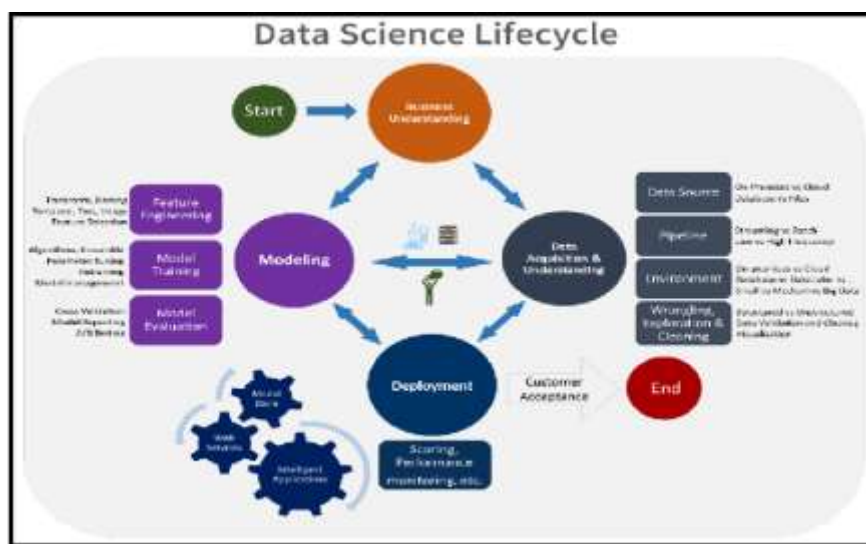
2.3 Sistem Yang Diusulkan



Gambar 2. Sistem Yang Diusulkan

Pengelolaan data obat berdasarkan hasil klasifikasi yang dilakukan data obat Rumah Sakit Umum Daerah (RSUD) kota Pagar Alam agar lebih mudah dikenali.

2.4 Metode Microsoft's Team Data Science Process



Gambar 3. Metode Data Science Process

Microsoft's Team Data Science Process adalah metodologi data science yang agile dan itertarif yang dibangun di atas best practice dari Microsoft untuk memfasilitasi keberhasilan implementasi proyek data science. Siklus Hidup TDSP terdiri dari 5 tahap, yaitu Pemahaman Bisnis (Business Understanding), Akuisisi & Pemahaman Data (Data Acquisition & Understanding), Pemodelan (Modelling), dan Penyebaran (Deployment) [15].

Siklus hidup Team Data Science Process (TDSP) di modelkan sebagai urutan langkah-langkah literasi yang membagikan panduan tentang tugas yang dibutuhkan untuk menggunakan model prediktif. Ada 5 tahapan siklus hidup Team Data Science Process TDSP yaitu :

a. Bussines understanding

Adalah tahap untuk menentukan tujuan dan mengidentifikasi masalah bisnis yang perlu di klasifikasi oleh analisis. Tujuan melakukan penelitian ini membagi data obat berdasarkan kelas-kelas tertentu berdasarkan nama obat, bentuk dan pengurangan. Untuk menampilkan sebuah grafik informasi agar mempermudah pihak rumah sakit dalam melihat bentuk obat yang paling sering digunakan dalam mengelolah data obat yang ada di Rumah Sakit Umum Daerah (RSUD) Kota Pagar Alam. Masalah yang akan diselesaikan dalam penelitian ini adalah mengklasifikasi data obat di Rumah sakit umum daerah (RSUD) Kota Pagar Alam menggunakan Algoritma Naïve Bayes agar data obat trsetruktur dan mudah di pahami.

b. Data aquision & understanding

Adalah tahapan untuk memasukan data, menjelajahi data dan menyiapkan alur data. Data yang akan di masukan pada penelitian ini berasal dari data obat di Rumah sakit umum daerah (RSUD) Kota Pagar Alam yaitu sebagai berikut :

1. Data Nama obat
2. Data bentuk obat



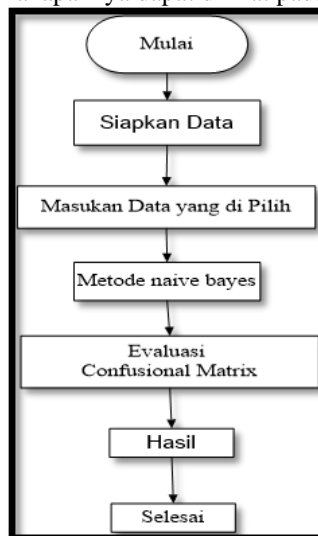
3. Data pengurangan

Setelah data dimasukan kemudian masuk ke langkah selanjutnya menjelajahi data, pada fase ini peneliti akan melakukan peroses pengkondisian data menjadi data set dan melakukan pembersihan data integrasi data, dan transformasi data. Setelah peroses menjelajahi data selesai masuk ke peroses selanjutnya yaitu peroses menyiapkan alur data, pada tahap ini data akan di ubah ke bentuk excel atau .csv.

c. Modelling

Adalah tahap membuat fitur data, membuat model training dan membuat model evaluasi. Data yang digunakan pada penelitian ini berjumlah 329 data sampel yang berasal dari dari data obat di Rumah sakit umum daerah (RSUD) Kota Pagar Alam. Pada tahap ini, peroses mengklasifikasi metode naive bayes akan di mulai. Tahap-tahap dalam perosesnya yaitu filtrur data pada penelitian ini adalah mempersiapkan data mentah untuk memfalidasi pelatihan model. Data yang di gunakan dalam penelitian ini adalah 90% untuk data training dan 10% untuk data testing dan menggunakan 3 variabel independen yang di gunakan untuk mengklasifikasi yaitu data obat yang ada di RSUD, nama obat, bentuk dan pengurangan. Data menta tersebut digabungkan dan dapat berupa file excel atau .csv.

Pelatihan model adalah peroses pemilihan dan penerapan teknik pemodelan dan beberapa parameternya akan disesuaikan untuk mendapat nilai yang optimal menggunakan aplikasi RaipidMiner di penelitian ini peneliti menggunakan Metode Naive Bayes. Tahapannya dapat dilihat pada gambar dibawah ini :



Gambar 4. Flowchart Naive Bayes

Keteran gan:

X : Data dengan class yang belum diketahui

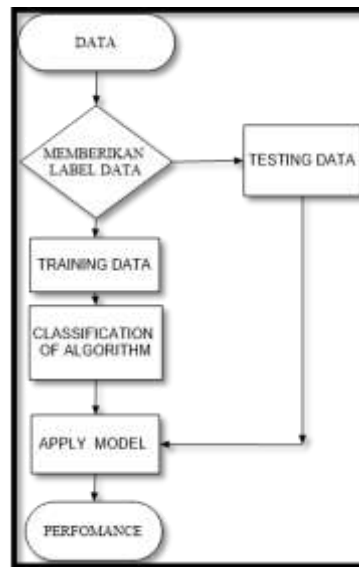
H : Hipotesis data X merupakan suatu class spesifik

(H|X): Probabilitas hipotesis H berdasar kondisi X (posterior probability)

(H): Probabilitas hipotesis H (prior probability)

(X|H): Probabilitas X berdasarkan kon- disi pada hipotesis H

(X): Probabilitas X



Gambar 5. Flowchart Rapidminer

Adapun langkah-langkah dalam perhitungan naïve bayes adalah sebagai berikut :

1. Menyiapkan data yang akan diklasifikasi yaitu data set data obat yang ada di Rumah Sakit Umum Daerah (RSUD) di Kota Pagar Alam jumlah data yang digunakan 329 sampel data dari bulan agustus 2023 sampai dengan januari 2024.

Tabel 2. Data Set Data Obat

No.	Kode	Nama Obat	Bentuk	Saldo Awal Bulan	Penambahan	Pengurangan	Saldo Akhir
1.	20278	Parasetamol 500 mg	TABLET	110400	24000	55200	79200
2.	20311	Spironolakton 25 mg	TABLET	63200	0	31600	31600
3.	20266	Omeprazole 20 mg Kapsul	TABLET	57640	30000	28820	58820
4.	20230	Metformin 500 mg	AMPUL	55400	0	27700	27700
5.	20224	Mecobalamin 500mg kapsul	TABLET	53800	25300	26900	52200
6.	20169	Furosemid 40 mg	TABLET	52400	20100	26200	46300
7.	20211	Lansoprazole 30 mg	TABLET	44640	20000	22320	42320
8.	20335	Vitamin B Kompleks tablet	TABLET	41000	40000	20500	60500
9.	20062	Candesartan Cilexotil 8mg tablet	TABLET	40920	43300	20460	63760
....
329.	20303	Ryzodeg	VIAL	0	100	0	100

d. Evaluasi Model

Evaluasi model di lakukan dengan menggunakan metode confusion matrix yaitu data di bagi secara acak dengan ukuran yang sama untuk menguji model metode yang di gunakan dan pada penelitian ini menggunakan aplikasi RapidMiner serta teknik confusion matrix untuk melihat akurasi dari model yang di gunakan.

Tabel 3. Contoh Confusion matrix multi-kelas

	Classes	a	b	c	d	Total
<i>Actual Classification</i>	a	6	0	1	2	9
	b	3	9	1	1	14
	c	1	0	10	2	13
	d	1	2	1	12	16





Total 11 11 13 17 52

Formulasi perhitungan adalah sebagai berikut:

$$accuracy = \frac{Tp + Tn}{Tp + Tn + Fp + Fn} \times 100\%$$

Tabel 4. *Confusional matrix multi class*

classes	Predicted Classification						
	Tablet	Ampul	Kolf	Box	Vial	Botol	Tube
Tablet	TP	FP	TN	TN	TN	TN	TN
Ampul	TN	TP	TN	TN	TN	FN	TN
Kolf	TN	TN	TP	TN	TN	TN	TN
Box	TN	TN	TN	TP	TN	TN	TN
Vial	TN	TN	TN	TN	TP	FN	TN
Botol	TN	TN	TN	TN	TN	TP	TN
Tube	TN	TN	TN	TN	TN	TN	TP

Tabel 5. *Confuional matrix multi class*

Classes	Predicted Classification							total
	Tablet	Ampul	Kolf	Box	Vial	Botol	Tube	
Tablet	136	1	0	0	0	0	0	137
Ampul	0	45	0	0	0	1	0	46
Kolf	0	0	17	0	0	0	0	17
Box	0	0	0	7	0	0	0	7
Vial	0	0	0	0	19	1	0	20
Botol	0	0	0	0	0	52	0	52
Tube	0	0	0	0	0	0	17	17
Total	136	46	17	7	19	54	17	296

TP : 136+45+17+7+19+52= 293

FP: 1

FN : 1+1= 2

TN: 0

Perhitungan akurasi klasifikasi

$$accuracy = \frac{Tp + Tn}{Tp + Tn + Fp + Fn} \times 100\%$$

$$accuracy = \frac{136 + 45 + 17 + 7 + 19 + 52 + 17}{293 + 1 + 2 + 0} \times 100\%$$

$$accuracy = \frac{293}{296} \times 100\%$$

accuracy = 0, 98 % x 100%

accuracy = 98 %

$$Precision = \frac{Tp}{Tp + Fp} \times 100\%$$

$$Precision = \frac{293}{293 + 1} \times 100\%$$

$$Precision = \frac{293}{294}$$

Precision = 0, 99 % x 100%

Precision = 99 %

$$Recall = \frac{Tp}{Tp + FN} \times 100\%$$

$$Recall = \frac{293}{293 + 2} \times 100\%$$





$$Recall = \frac{293}{295}$$

$$Recall = 0,99\% \times 100\%$$

$$Recall = 99\%$$

$$F1 = \frac{2 \times Recall \times precision}{Precision + Recall} \times 100\%$$

$$F1 = \frac{2 \times 0,99 \times 0,99}{0,99 + 0,99} \times 100\%$$

$$F1 = \frac{1,9602}{1,98}$$

$$F1 = 0,99\% \times 100\%$$

$$F1 = 99\%$$

e. Deployment

Adalah tahap pelaporan pemodelan akhir dengan detail penyebaran. Pada fase ini hasil dari model sudah keluar dan dapat digunakan untuk pengetahuan dan sudah di ketahui klasifikasi data obat di Ruma Sakit Umum Daerah (RSUD) Kota Pagar Alam menggunakan metode Naive Bayes.

f. Customer acceptance

Adalah tahap konfirmasi bahwa model yang di terapkan sudah memenuhi kebutuhan. Pada fase ini sudah dihasilkan kesimpulan hasil dari klasifikasi data obat menggunakan metode Naive Bayes yang akan di tunjukan dengan menggunakan grafik.

3. HASIL DAN PEMBAHASAN

3.1 Hasil Penelitian

Hasil dari penelitian ini adalah klasifikasi Data obat berdasarkan bentuk obat yang paling sering di gunakan dari bulan agustus 2023 – januari 2024. Penerapan data mining menggunakan metode algoritma naive bayes dapat mengklasifikasi data obat yang ada di Rumah sakit Umum Daerah (RSUD) Kota Pagar Alam yang dilakukan berdasarkan metode Team Data Science Process (TDSP) yang memiliki beberapa fase yaitu Business Understanding, Data Aquistion & Understanding, Modelling, Deployment Dan Customer Acceptance.

Penulis juga berhasil mengklasifikasi Data obat di Rumah sakit Umum Daerah (RSUD) Kota Pagar Alam dengan menggunakan data set berjumlah 329 data. Metode Naive Bayes menggunakan aplikasi RapidMiner dengan nilai akurasi 98.99%. Penjelasan hasil penelitian berdasarkan langkah-langkah Team Data Science Process.

3.2 Tahapan Team Data Science Process (TDSP)

3.2.1 Bussines Understanding (Pemahaman Bisnis)

Pada pemahaman bisnis dilakukan langkah menentukan tujuan dan mengidentifikasi masalah yang perlu di klasifikasi oleh analisis. Tujuan melakukan penelitian ini membagi data obat berdasarkan kelas- kelas tertentu. Untuk mengklasifikasi data obat yang di ambil dari file excel data tersebut hanya digunakan untuk melihat katagori bentuk obat yang paling sering di gunakan yang ada di Rumah Sakit Umum Daerah (RSUD) Kota Pagar Alam. Dari proses mengklasifikasi data obat di Rumah Sakit Umum Daerah (RSUD) Kota Pagar Alam menggunakan Algoritma Naive Bayes untuk mengetahui bentuk obat mana yang sering di gunakan.

3.2.2 Data Aquistion dan Understanding

Data yang di gunakan dalam penelitian ini yaitu data obat dari bulan agustus 2023 – januari 2024 di RSUD Kota Pagar Alam berupa satu dataset berbentuk excel. Setelah mendapatkan data dengan jumlah data 329 data yang di peroleh dari Rumah Sakit umum Daerah (RSUD). Data ini akan digunakan untuk menghasilkan hasil klasifikasi data obat berdasarkan bentuk obat mana yang paling sering di gunakan dengan menggunakan Algoritma Naive Bayes. Bentuk dataset tersebut dapat dilihat pada tabel berikut ini.



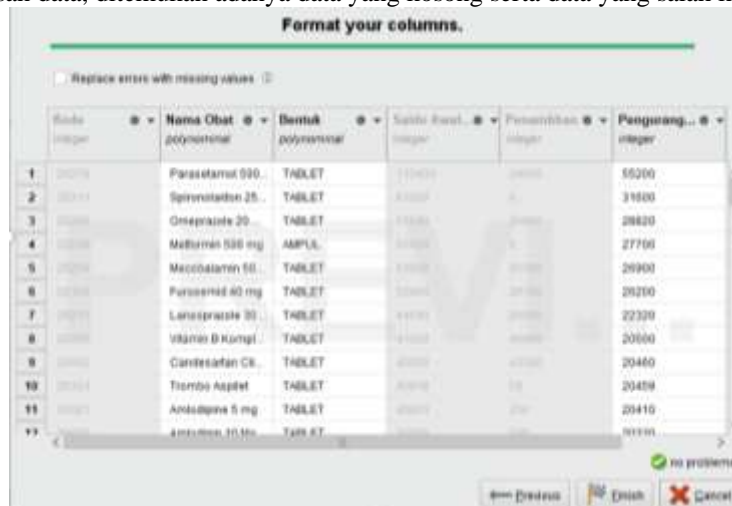
Tabel 6. Data Set Data Obat di Rumah Sakit

No.	Kode	Nama Obat	Bentuk	Saldo Awal Bulan	Penambahan	Pengurangan	Saldo Akhir
1.	20278	Parasetamol 500 mg	TABLET	110400	24000	55200	79200
2.	20311	Spirolakton 25 mg	TABLET	63200	0	31600	31600
3.	20266	Omeprazole 20 mg Kapsul	TABLET	57640	30000	28820	58820
4.	20230	Metformin 500 mg	AMPUL	55400	0	27700	27700
5.	20224	Mecobalamin 500mg kapsul	TABLET	53800	25300	26900	52200
6.	20169	Furosemid 40 mg	TABLET	52400	20100	26200	46300
7.	20211	Lansoprazole 30 mg	TABLET	44640	20000	22320	42320
8.	20335	Vitamin B Kompleks tablet	TABLET	41000	40000	20500	60500
9.	20062	Candesartan Cilexotil 8mg tablet	TABLET	40920	43300	20460	63760
.....
329.	20303	Ryzodeg	VIAL	0	100	0	100

Selanjutnya melakukan pengolahan data yaitu tahap klasifikasi data obat berdasarkan bentuk obat yang akan digunakan peneliti untuk klasifikasi ada 3 atribut diantaranya yaitu nama obat, bentuk, dan, pengurangan, data yang di terima berbentuk excel.

a. Data Cleaning

Bertujuan untuk mengidentifikasi dan memperbaiki kesalahan dalam data. Selama proses pemrosesan data, ditemukan adanya data yang kosong serta data yang salah input.



Gambar 6. Data Cleaning

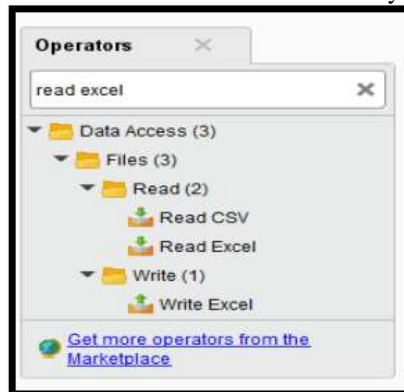
b. Data Integration

Bertujuan untuk menggabungkan data dari berbagai sumber menjadi satu kesatuan yang utuh dan konsisten. Selama proses penelitian, ditemukan ketidaksesuaian antara data yang berasal dari sumber yang berbeda. Oleh karena itu, dilakukan integrasi data agar data yang digabungkan menjadi serasi dan konsisten, sehingga dapat digunakan secara efektif untuk analisis selanjutnya.

3.2.3 Modelling

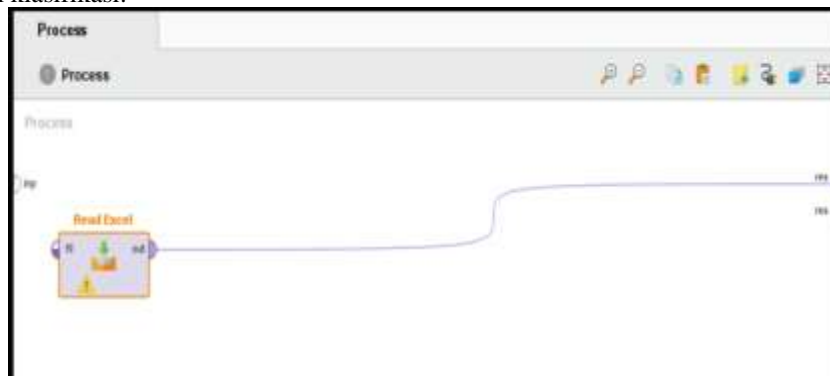


Pada tahap ini akan di lakukan proses permodelan klasifikasi metode Naive bayes pada data obat di Rumah sakit umum Daerah (RSUD) Kota Pagar Alam menggunakan Rapidminer. Tahap-tahap dalam prosesnya yaitu fitur data pada penelitian ini adalah mempersiapkan data mentah untuk memvalidasi pelatihan model.tahap pertama yaitu memilih operator read excel untuk memasukan file excel yang akan di klasifikasi.



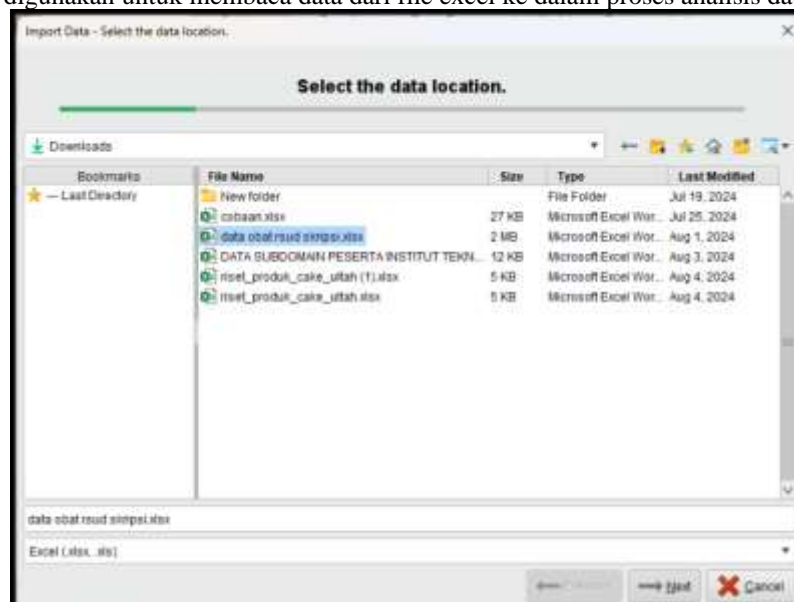
Gambar 7. Operator Rapidminer

Memilih operator read excel untuk di masukan ke proses dalam rapidminer untuk memilih data yang akan digunakan dalam klasifikasi.



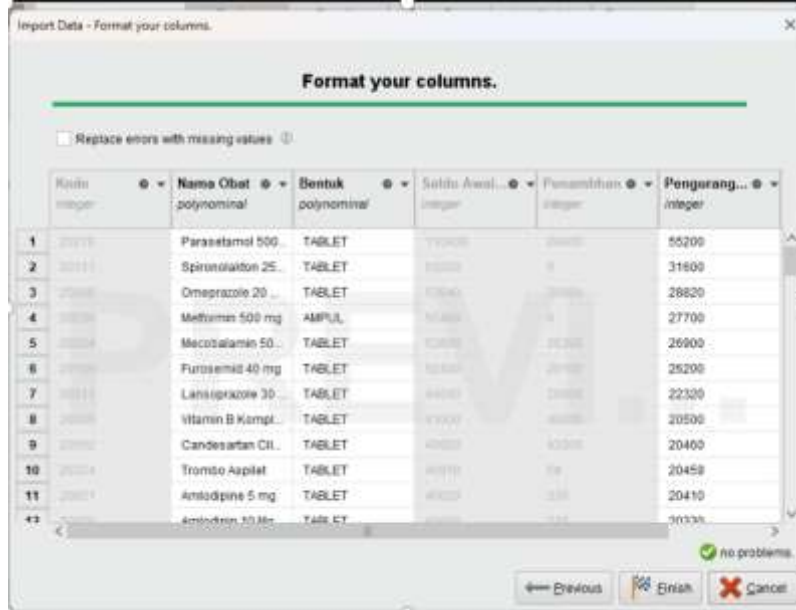
Gambar 8. Read Excel

Read excel digunakan untuk membaca data dari file excel ke dalam proses analisis data di rapidminer.



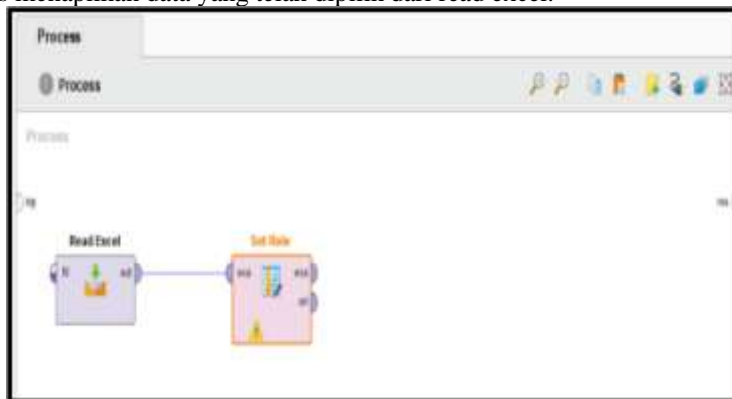
Gambar 9. Proses Pemilihan Data Excel

Gambar di atas menampilkan proses pemilihan data excel yang akan di gunakan dalam klasifikasi data obat-obatan Ruma Sakit Umum Daerah Pagar Alam (RSUD).



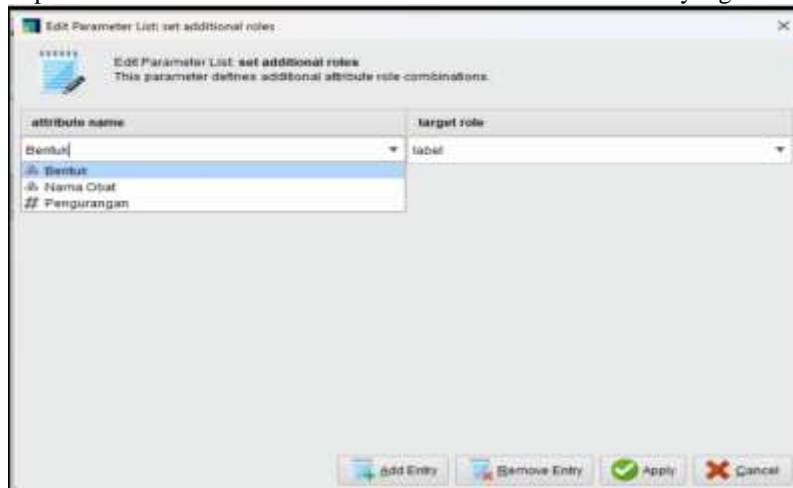
Gambar 10. Data dari Read Excel

Gambar di atas menampilkan data yang telah dipilih dari read excel.



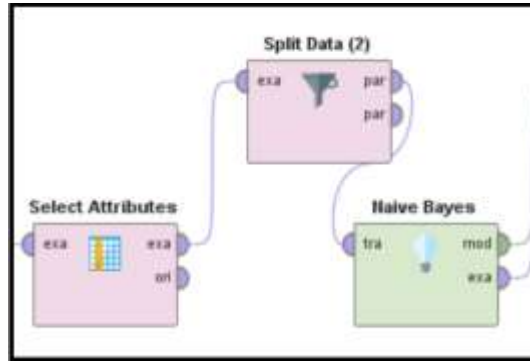
Gambar 11. Set Role

Memasukan operator set role untuk menentukan label dari sebuah data set yang akan di klasifikasi.



Gambar 12. Label dan Atribute

Gambar di atas menampilkan atribut yang di jadikan label dalam klasifikasi dimana pada model ini Bentuk menjadi labelnya.



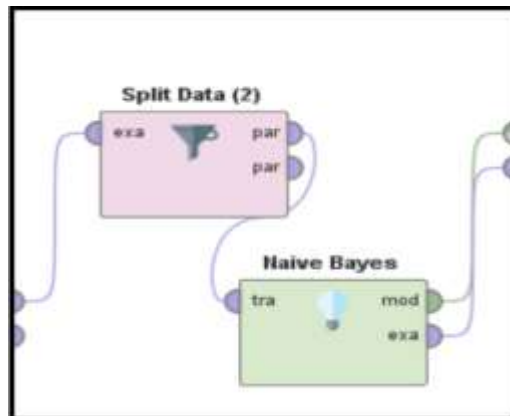
Gambar 13. Select Attributes

Memasukan operator select atribut untuk menentukan atribut yang akan di gunakan dalam klasifikasi.



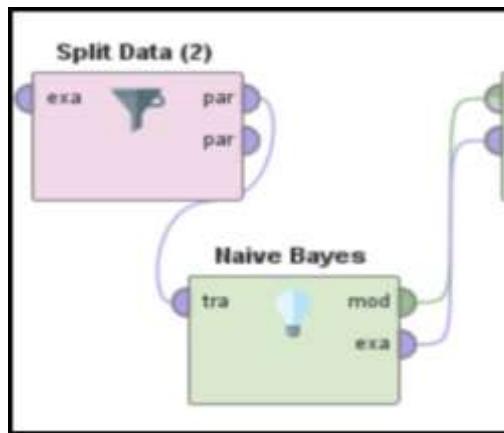
Gambar 14. Attributes

Gambar di atas menampilkan isi dari select atribut dima kita menentukan atribut yang akan di gunakan dalam klasifikasi.



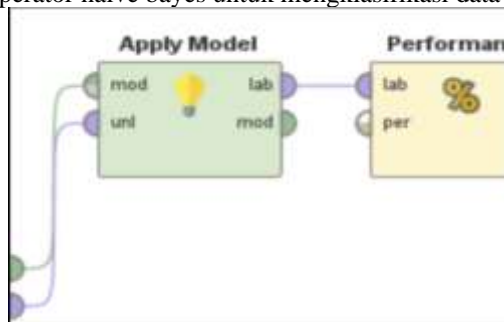
Gambar 15. Split Data

Memasukan operator split data dimana data di bagi menjadi data training 90% dan data testing 10%.



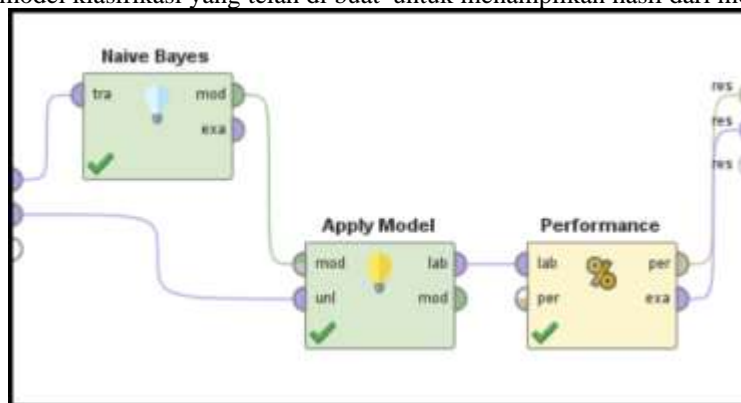
Gambar 16. Proses *Naive Bayes*

Kemudian memasukan operator naïve bayes untuk mengklasifikasi data set.



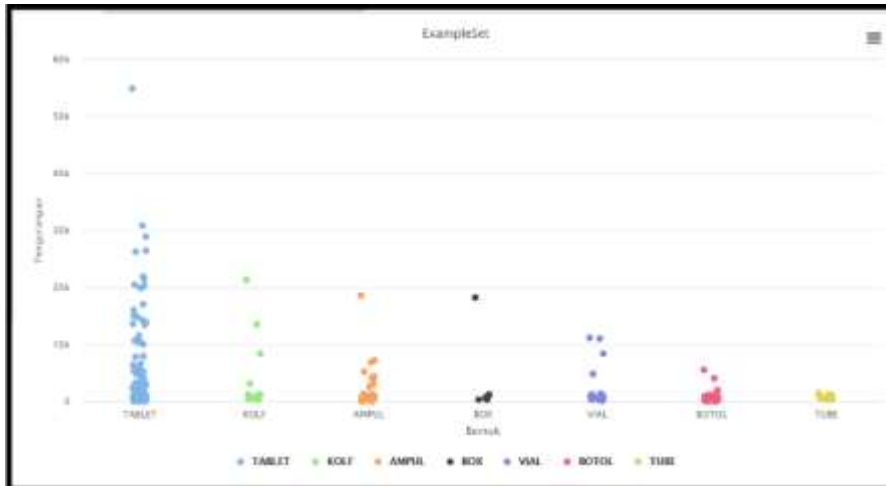
Gambar 17. *Apply Model*

Menjalankan model klasifikasi yang telah di buat untuk menampilkan hasil dari model klasifikasi.



Gambar 18. *Performance*

Berdasarkan grafik di bawah ini dapat di lihat bentuk obat yang paling sering digunakan yang di bagi menjadi 7 warna di bawah ini:



Gambar 19. Grafik bentuk Obat yang paling sering digunakan

Grafik di atas menampilkan hasil klasifikasi obat dalam bentuk yang paling sering di gunakan di rumah sakit umum daerah kota Pagar Alam dari bulan agustus 2023 – januari 2024, terdapat 7 warna yang membedakan classnya yang pertama yaitu warna biru muda Tablet berjumlah 55,200, hijau kolf berjumlah 20,254, oren ampul 19,400 , hitam box 19,000 , biru tua vial 11,750, pink botol 4,727, dan kuning tube 354.

3.2.4 Deployment

Tahap deployment merupakan proses melaporkan hasil penelitian yang telah melalui proses data mining. Hal-hal yang akan dipresentasikan mencakup pengetahuan yang diperoleh selama proses tersebut berjalan, sehingga dapat dipahami oleh pengguna. Tahapan pembuatan laporan dilakukan setelah melalui proses modelling dari penelitian terhadap algoritma Naïve Bayes untuk mengklasifikasi data obat di RSUD Kota Pagaralam. Setelah semua perintah operator berhasil di jalankan,akurasi dari Algoritma Naive Bayes di tampilkan dengan nilai akurasi yang dapat di tampilkan sebagai berikut:

accuracy: 98.99%								
	true TABLET	true AMPUL	true KOLF	true BOX	true VIAL	true BOTOL	true TUBE	class precision
pred. TABLET	136	1	0	0	0	0	0	99.27%
pred. AMPUL	0	45	0	0	0	1	0	97.83%
pred. KOLF	0	0	17	0	0	0	0	100.00%
pred. BOX	0	0	0	7	0	0	0	100.00%
pred. VIAL	0	0	0	0	19	1	0	95.00%
pred. BOTOL	0	0	0	0	0	52	0	100.00%
pred. TUBE	0	0	0	0	0	0	17	100.00%
class recall	100.00%	97.83%	100.00%	100.00%	100.00%	96.30%	100.00%	

Gambar 20. Hasil Kurasi

Hasil yang diperoleh dari tabel di atas menunjukkan bahwa performa metode yang digunakan oleh peneliti dalam klasifikasi data obat di Rumah Sakit Umum Daerah (RSUD) Kota Pagaralam mencapai akurasi 98.99%. Dengan demikian, metode ini diandalkan sebagai alat klasifikasi untuk menentukan Bentuk obat mana yang paling sering digunakan.

```

PerformanceVector
PerformanceVector:
accuracy: 98.99%
ConfusionMatrix:
True:  TABLET  AMPUL  KOLF  BOX  VIAL  BOTOL  TUBE
TABLET: 136    1    0    0    0    0    0
AMPUL:  0    45    0    0    0    1    0
KOLF:   0    0    17    0    0    0    0
BOX:    0    0    0    7    0    0    0
VIAL:   0    0    0    0    19    1    0
BOTOL:  0    0    0    0    0    52    0
TUBE:   0    0    0    0    0    0    17
    
```

Gambar 21. Performance Vector

Merupakan penghitungan accuracy menggunakan algoritma naive bayes. Diketahui data training terdiri dari 296 record data, jumlah obat dalam bentuk tablet 136 data yang dikira tablet ternyata bukan tablet tetapi Ampul adalah 1, dalam bentuk ampul berjumlah 45 data yang dikira ampul ternyata botol ada 1 data, obat dalam bentuk kolf berjumlah 17 data, obat dalam bentuk box berjumlah 7 data, obat dalam bentuk vial berjumlah 19 data yang dikira vial ternyata botol ada 1 data, obat dalam bentuk botol berjumlah 52 data, obat dalam bentuk tube 17 data.

3.2.5 Customer Acceptance

Tahap Customer acceptance Adalah tahap konfirmasi bahwa model yang diterapkan sudah memenuhi kebutuhan. Pada fase ini sudah dihasilkan kesimpulan hasil dari klasifikasi menggunakan metode Naïve Bayes berupa grafik yang menunjukkan obat mana yang sering di gunakan berdasarkan bentuk.

4. KESIMPULAN

Algoritma Naïve Bayes berhasil diimplementasikan untuk mengklasifikasikan data obat. Model ini mampu mengelompokan obat berdasarkan bentuk obat yang paling sering digunakan, dilihat dari data pengurangan dengan menggunakan metode TDSP (Team Data Science Proses). Hasil penelitian menunjukkan dari hasil klasifikasi bahwa obat mana yang paling sering digunakan berdasarkan bentuknya terdapat 7 warna yang membedakan classnya yang pertama yaitu warna biru muda Tablet berjumlah 55,200, hijau kolf berjumlah 20,254, oren ampul 19,400, hitam box 19,000, biru tua vial 11,750, pink botol 4,727, dan kuning tube 354.

Perhitungan menggunakan Rapidminer dengan algoritma Naïve Bayes didapatkan yaitu didapat hasil akurasi sebesar 98,99% Dengan demikian dari hasil pengujian model di atas dapat disimpulkan bahwa klasifikasi peminatan dengan metode naive bayes menghasilkan akurasi yang tinggi yaitu 98,99% menggunakan aplikasi RapidMiner. Kedepannya di sarankan memasukkan atau menggunakan data yang lebih besar lagi saat menguji model algoritma. selanjutnya dapat mencoba algoritma lain yang menghasilkan tingkat akurasi yang tinggi.

UCAPAN TERIMAKASIH

Terima kasih disampaikan kepada pihak-pihak yang telah mendukung terlaksananya penelitian ini.

- Tempat Penelitian Rumah Sakit Daerah Besemah Pagar Alam beserta tim.
- Kampus yang sudah mendukung dan terlaksana penelitian.

DAFTAR PUSTAKA

- Abdullah, L., Tamin, R., & Qashlim, A. A. (2021). Klasifikasi Penerimaan Beasiswa Menggunakan Algoritma Naive Bayes Di Universitas Al Asyariah Mandar Kabupaten Polewali Mandar. *Journal Peqguruang: Conference Series*, 3(1), 183. <https://doi.org/10.35329/jp.v3i1.1399>
- Ademariana, K., Lumbanraja, F. R., & Andrian, R. (2021). *Jurnal Pepadun Clustering K-Means Jenis Kata Pada Laporan Kegiatan Kuliah Kerja Nyata (KKN) Universitas Lampung Menggunakan Word2vec* © 2021 Ilmu Komputer Unila Publishing Network All Rights Reserved *Jurnal Pepadun*. 3(2), 221–228.
- Andreansyah, A. (2020). Clustering, Klasifikasi Obat Medis Berdasarkan Ekstraksi Ciri Menggunakan K-Means. *Agus Andreansyah*, 9(1), 33–41.
- Arfanda, I., Ramdhan, W., & Yusda, R. A. (2021). *Digital Transformation Technology (Digitech) | E-ISSN : 9999-9999 Naive Bayes Dalam Menentukan Penerima Bantuan Langsung Tunai Digital Transformation Technology (Digitech) | E-ISSN : 9999-9999*. 1(1), 9–16.
- Arif, A., & Arif, A. (2023). *Komparasi Regresi Linear Berganda Dan Support Vector Machine Untuk Prediksi Kopi Robusta Hasil Produksi Kopi Robusta Pagar Alam*. 03, 713–726.
- Ariyanti, D., & Iswardani, K. (2020). Teks Mining Untuk Klasifikasi Keluhan Masyarakat Pada Pemkot Probolinggo Menggunakan Algoritma Naïve Bayes. *Jurnal IKRA-ITH Informatika*, 4(3), 125–132.
- Binang, A. A. (2021). Pagar Alam Kota Perjuangan Dalam Perspektif Sejarah Lokal (1854-1867). *Tesis*, 1–108. <http://repository.radenfatah.ac.id/16666/>
- Fachrina, I. (2020). *Rancang Bangun Aplikasi Data Mining Untuk Klasifikasi Pemakaian Obat Dengan Metode Naive Bayes Pada Puskesmas Bandar Baru*. 1–9.
- Fadli, M. N., Damani, I. S., & Irawan, E. (2021). *Penerapan Metode Naive Bayes Dalam Menentukan Tingkat Kenyamanan Pada Rumah Sakit Terhadap Pasien*. 2(3), 117–122.
- Fajar Fauzan, M., Irma Purnamasari, A., & Dwilestari, G. (2023). Penerapan Data Mining Untuk Menganalisis Penjualan Air Minum Dalam Kemasan Selama Masa Pandemi Covid-19. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 7(1), 700–706. <https://doi.org/10.36040/jati.v7i1.6290>
- Fitria, R. I., & Ujianto, N. T. (2022). *Komparasi Algoritma Logistic Regression Dan Naive Bayes Untuk Penerimaan Siswa*





- Baru Pada Smk Di Brebes. ... , *Artificial Intelegence And Internet Of Thing*, 1, 43–54. <https://Jiaii.Upstegal.Ac.Id/Index.Php/Jiaii/Article/View/15%0Ahttps://Jiaii.Upstegal.Ac.Id/Index.Php/Jiaii/Article/Download/15/9>
- Ikhsan Romli, R. F. P. D. (2021). *Indonesian Journal Of Business Intelligence*. 4(1), 10–15.
- Imam, M., Udariansyah, D., Komputer, F. I., & Darma, U. B. (N.D.). *Penerapan Algoritma Naive Bayes Untuk Mengklasifikasi*. 919–926.
- Informatika, J. T., Dan, P. A. K., Persero, P. T. P., Mega, J., Barat, K., Kawasan, L., & Kuningan, M. (2021). *Berdasarkan Tingkat Kepentingan Pada Prodi / Jurusan D4 Teknik Informatika Politeknik Pos Indonesia Pt Pertamina (Persero) Abstrak Pt . Pertamina (Persero) Memiliki Beberapa Fungsi Salah Satunya Terdapat Fungsi Health And Medical Management . Fungsi H*. 13(2), 1–8.
- Leny, L. (2022). *PROSIDING Vol.1 No.1 2022*. 1(1), 38–49.
- Mardewi, Nuru Yarkuran, Sofyan, And F. A. (2023). *Journal Pharmacy And Application Support Vector Machine Drug Category Classification Using Support Vector Machine*. 1(1), 27–32.
- Miharso, S. (2021). *Pertimbangan Hukum Hakim Terhadap Pemilikan Dan Pengedaran Obat Keras Tanpa Resep Dokter Stevanus*. 6(10).
- Nabila, Z., Rahman Isnain, A., & Abidin, Z. (2021). Analisis Data Mining Untuk Clustering Kasus Covid-19 Di Provinsi Lampung Dengan Algoritma K-Means. *Jurnal Teknologi Dan Sistem Informasi (JTISI)*, 2(2), 100. <http://Jim.Teknokrat.Ac.Id/Index.Php/JTISI>
- Nofitri, R., & Irawati, N. (2019). Integrasi Metode Neive Bayes Dan Software Rapidminer Dalam Analisis Hasil Usaha Perusahaan Dagang. *JURTEKSI (Jurnal Teknologi Dan Sistem Informasi)*, 6(1), 35–42. <https://Doi.Org/10.33330/Jurteksi.V6i1.393>
- Nursyahfitri, R., Maharadja, A. N., Farissa, R. A., & Umaidah, Y. (2019). *Klasifikasi Penentuan Jenis Obat Menggunakan Algoritma Decision Tree*. 53–60.
- Permana, A. R. R. (2023). Penerapan Data Mining Untuk Menentukan Keputusan Produksi Berdasarkan Data Penjualan Menggunakan Algoritma Apriori. *Jurnal Teknorama (Informatika Dan ...)*, 1(1), 7–13. <https://Jurnal.Stikomelrahma.Ac.Id/Index.Php/Teknorama/Article/View/2%0Ahttps://Jurnal.Stikomelrahma.Ac.Id/Index.Php/Teknorama/Article/Download/2/2>
- Prisusanti, R. D. (2022). *Jurnal Peduli Masyarakat*. 4, 351–354.
- Purnamasari, I. (2020). Analisa Klasifikasi Loyalitas Siswa Lembaga Pendidikan Tari Dengan Metode Naïve Bayes. *Jurnal Ilmiah Komputasi*, 19(1), 59–68. <https://Doi.Org/10.32409/Jikstik.19.1.157>
- Safira, M. S., Rahaningsih, N., Dana, R. D., Informatika, T., Akuntansi, K., Informatika, M., Cirebon, K., Penjualan, J., Pranoto, G. T., Algoritma, D., & Untuk, K. N. (2024). *Penerapan Data Mining Untuk Klasifikasi Penjualan Obat Menggunakan Metode K-Nearest Neighbor*. 8(1), 380–385.
- Setiyani, L., Wahidin, M., Awaludin, D., & Purwani, S. (2020). Analisis Prediksi Kelulusan Mahasiswa Tepat Waktu Menggunakan Metode Data Mining Naïve Bayes : Systematic Review. *Faktor Exacta*, 13(1), 35. <https://Doi.Org/10.30998/Faktorexacta.V13i1.5548>
- Siswa, P., Pгри, S. M. P., Insani, A., Sari, R., Fitriyani, A., & Prabandari, R. D. (2020). *Optimalisasi Penggunaan MS . Word Dan MS . Excel*. 3(2), 95–104.
- Susanto, F., & Mahendra, G. S. (2022). *Penerapan Data Mining Dengan Algoritma C4 . 5 Dalam*. 5, 237–245.
- Syahrani. (2023). *Microsoft Excel-Based Bookeeping Training At The “Kusuma” Savings And Credit Cooperative Gemarang Cash Office Syahrani*. 1(6), 1428–1437.
- Syahputri, A. Z., Fallenia, F. Della, & Syafitri, R. (2023). *Kerangka Berfikir Penelitian Kuantitatif*.
- Wahyudi, H. (2021). *EXPLORE – Volume 11 No 2 Tahun 2021 Terakreditasi Sinta 5 SK No : 23 / E / KPT / 2019 Rancang Bangun Sistem Informasi Pengelolaan Stok Obat Pada Apotek Karya Husada EXPLORE – Volume 11 No 2 Tahun 2021 Terakreditasi Sinta 5 SK No : 23 / E / KPT / 2019*. 11(2), 146–152.
- Watratan, A. F., B, A. P., Moeis, D., Informasi, S., & Makassar, S. P. (2020). *Journal Of Applied Computer Science And Technology (Jacost) Implementasi Algoritma Naive Bayes Untuk Memprediksi Tingkat Penyebaran Covid-19 Di Indonesia*. 1(1), 7–14.
- Wijaya, H. D., & Dwiasnati, S. (2020). *Implementasi Data Mining Dengan Algoritma Naïve Bayes Pada Penjualan Obat*. 7(1), 1–7.
- Yanti, S., Vera, Y., Farmasi, P., Aufa, S., Di, R., & Padangsidempuan, K. (2020). Penyuluhan Tentang Cara Penggunaan Obat Yang Baik Dan Benar Di Desa Mangunggang Jae. *Journal Education And Development*, 8(1), 26–28. [Http://Bidhuan.Id/Apoteker](http://Bidhuan.Id/Apoteker)
- Zainuri, M., Fahmi, M. H., & Hamdhana, R. A. (2022). Komparasi Metode Klasifikasi Algoritma C5.0 Dan Naïve Bayes Untuk Menentukan Jurusan Siswa. *Jurnal Sistem Informasi Dan Informatika (JUSIFOR)*, 1(1), 1–8. <https://Doi.Org/10.33379/Jusifor.V1i1.1277>

