

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

SKRIPSI



**Oleh:
Mangaratua Hutahaean
160210211**

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK DAN KOMPUTER
UNIVERSITAS PUTERA BATAM
2022**

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

SKRIPSI

**Untuk memenuhi salah satu syarat
Guna memperoleh gelar Sarjana**



**Oleh:
Mangaratua Hutahaean
160210211**

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK DAN KOMPUTER
UNIVERSITAS PUTERA BATAM
2022**

SURAT PERNYATAAN ORISINALITAS

Yang bertanda tangan di bawah ini saya:

Nama : Mangaratua Hutahaeen

NPM : 160210211

Fakultas : Teknik dan Ilmu Komputer

Program Studi: Teknik Informatika

Menyatakan bahwa “Skripsi” yang Penulis buat dengan judul:

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

Ini adalah karya sendiri dan bukan “duplikasi” dari karya orang lain. Sepengetahuan saya, di dalam naskah skripsi ini tidak terdapat karya ilmiah atau pendapat yang pernah di tulis atau di terbitkan oleh orang lain, kecuali yang secara tertulis di kutip di dalam naskah ini dan di sebutkan dalam sumber kutipan dan daftar pustaka.

Apabila ternyata di dalam naskah Skripsi ini dapat di buktikan terdapat unsur-unsur PLAGIASI, saya bersedia naskah Skripsi ini di gugurkan dan gelar akademik yang saya peroleh di batalkan, serta di proses sesuai dengan peraturan perundang-undangan yang berlaku.

Demikian pernyataan ini saya buat dengan sebenarnya tanpa ada paksaan dari siapapun.

Batam, 01 Agustus 2022



Mangaratua Hutahaeen
160210211

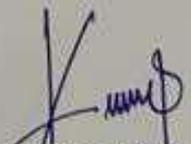
**PENERAPAN *DATA MINING* UNTUK
MEMPREDIKSI PENJUALAN OBAT DI KLINIK HARAPAN
KITA BATAM**

SKRIPSI

**Untuk memenuhi salah satu syarat
Guna memperoleh gelar Sarjana**

**Oleh:
Mangaratua Hutahaean
160210211**

**Telah disetujui oleh Pembimbing pada tanggal
Seperti tertera di bawah ini
Batam, 01 Agustus 2022**


**Koko Handoko, S.Kom., M.Kom
Pembimbing**

ABSTRAK

Bergeraknya dengan sangat pesat teknologi berpengaruh bagi setiap aspek kehidupan masyarakat luas. Perkembangan teknologi tidak hanya digunakan sebagai sumber berbagi informasi tetapi juga sudah dimanfaatkan dalam skala besar. Terkadang penyampaian berita tidak sebanding sama kepentingan berita yang tinggi sehingga memerlukan pendalaman informasi lebih dalam dari data yang jumlahnya besar. Dalam proses pengambilan data ini menggunakan disebut komputasi, komputasi, teknologi metode pembelajaran mesin dan analisis data untuk memperoleh informasi berharga dan informasi terintegrasi dari berbagai kelompok besar. Klinik adalah tempat atau organisasi yang memberikan informasi kepada masyarakat tentang obat-obatan dan produk kesehatan lainnya. Saat ini sudah banyak klinik di pedesaan yang memudahkan masyarakat untuk berobat. Terdapat permasalahan yang dialami oleh klinik yaitu dalam menentukan apakah produk obat tersebut laku atau tidak dalam waktu ke depan. Akibat kurang pemahaman mengenai minat pembeli terhadap merek produk obat dapat menimbulkan penumpukan obat digudang ini mengakibatkan kerugian yang signifikan dan dapat menyebabkan barang obat-obatan kadawarsa. Pemahaman yang baik mengenai barang yang sering terjual dapat meningkatkan jumlah pasokan produk obat yang paling banyak diminati atau di cari. Untuk mengatasi permasalahan demikian maka akan dilakukan pengelompokan obat yang laku atau tidak laku pada klinik tersebut berlandaskan variabel yang telah ditemukan menggunakan Algoritma *Naive bayes* yang dapat memberikan informasi mengenai produk yang ada di klinik serta meminimalisis penumpukan stok yang tidak laku dan menjadikan acuan bagi pemilik untuk mengambil keputusan dalam menyuplai produk obat-obatan yang baru. Hasil yang di peroleh sebagai informasi di dapatkan dari pemerolehan nilai kemungkinan 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 pada obat-obatan tersebut melalui pilihan. Data yang telah di dapat akan di kelola secara manual dan di input kedalam software *rapidminer* sehingga data yang di peroleh mendapatkan hasil nilai akurat dari *ROC(Rank Order Centroid)* dengan nilai 80% berdasarkan tingkat pengelompokan dan prioritas.

Kata kunci : Algoritma *Naive bayes*, *Data mining*, Prediksi Penjualan Obat

ABSTRACT

Technological developments are not only used as a source of information sharing but have also been utilized on a large scale. Sometimes the delivery of news is not proportional to the high importance of the news so that it requires deeper information deepening from large amounts of data. In this data retrieval process, the so-called computing, computing, machine learning method technology and data analysis are used to obtain valuable information and integrated information from various large groups. A clinic is a place or organization that provides information to the public about medicines and other health products. Currently, there are many clinics in rural areas that make it easier for people to seek treatment. There are problems experienced by the clinic, namely in determining whether the drug product sells or not in the future. Due to a lack of understanding of the buyer's interest in the brand of medicinal products, this can lead to accumulation of drugs in the warehouse resulting in significant losses and can lead to expired drug goods. A good understanding of frequently sold items can increase the supply of the most sought-after or sought-after medicinal products. To overcome these problems, a grouping of drugs that are sold or not sold will be carried out at the clinic based on the variables that have been found using the Naive bayes Agorithm which can provide information about the products in the clinic and minimize the accumulation of stock that does not sell and make reference for owners to take decisions in supplying new medicinal products. The results obtained as information are obtained from obtaining accuracy values for these drugs through choice. The data that has been obtained will be managed manually and inputted into the Rapidminner software so that the data obtained will get an accurate value of ROC (Rank Order Centroid) with a value of 80% based on the level of grouping and priority.

Keyword: Data mining, Drug Sales Prediction, Naive bayes Agorithm,

DAFTAR ISI

	Halaman
HALAMAN SAMPUL	i
HALAMAN JUDUL	ii
SURAT PERNYATAAN ORISINALITAS	iii
ABSTRAK	v
ABSTRACT	vi
DAFTAR ISI	vii
DAFTAR GAMBAR	viii
DAFTAR TABEL	ix
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Identifikasi Masalah	3
1.3 Batasan Masalah	3
1.4 Rumusan Masalah	4
1.5 Tujuan Penelitian	4
1.6 Manfaat Penelitian	5
BAB II TINJAUAN PUSTAKA	6
2.1 <i>Knowladge Discovery in Database (KDD)</i>	6
2.2 <i>Data mining</i>	6
2.3 <i>Metode Naive bayes</i>	11
2.4 <i>Software</i> Pendukung	12
2.5 Objek Penelitian.....	14
2.6 Penelitian Terdahulu	14
2.7 Kerangka Pemikiran.....	17
BAB III METODE PENELITIAN	18
3.1. Desain Penelitian.....	18
3.2. Teknik pengumpulan data	19
3.3. <i>Operasional Variabel</i>	20
3.4 Metode Perancangan sistem	23
3.5 Lokasi Dan Jadwal Penelitian.....	23
BAB IV HASIL DAN PEMBAHASAN	25
4.1 Hasil Penelitian	25
4.2 Analisa Proses Agotitma	29
4.3 Hasil Pengujian.....	34
4.4. <i>Implementasi Rapidminner</i>	35
BAB V KESIMPULAN DAN SARAN	41
5.1 Kesimpulan.....	41
5.2 Saran.....	42
DAFTAR PUSTAKA	43

DAFTAR GAMBAR

Gambar 2. 1 Ilmu <i>data mining</i>	7
Gambar 2. 2 Rapidminer	13
Gambar 2. 3 Kerangka Pemikiran	17
Gambar 2. 4 Tempat Penelitian	23
Gambar 3. 1 Desain Penelitian	18
Gambar 4. 1 Import Data Training	35
Gambar 4. 2 Data import step 3.....	36
Gambar 4. 3 Data import step 4.....	37
Gambar 4. 4 Data training step akhir	37
Gambar 4. 5 Menghubungkan operator	38
Gambar 4. 6 Icon Run	38
Gambar 4. 7 Hasil Perhitungan.....	39

DAFTAR TABEL

Tabel 3. 1 Data Penelitian	21
Tabel 3. 2 Jadwal Penelitian.....	24
Tabel 4. 1 Data set.	25
Tabel 4. 2 Seleksi Data	27
Tabel 4. 3 Probalitas Kelas.....	28
Tabel 4. 4 Jumlah Obat Terjual	29
Tabel 4. 5 Data Training	29
Tabel 4. 6 Data Testing	31
Tabel 4. 7 Probabilitas Kriteria Tahun.....	32
Tabel 4. 8 Probabilitas Kriteria Jenis Obat	33
Tabel 4. 9 Probabilitas Kriteria Transaksi.....	33
Tabel 4. 10 Hasil Data Testing	34
Tabel 4. 11 Class Prediksi	34

KATA PENGANTAR

Puji syukur kehadiran Tuhan Yang Maha Esa yang telah melimpahkan segala ramhat dan karunia-Nya, sehingga penulis dapat menyelesaikan laporan tugas akhir yang merupakan salah satu persyaratan untuk menyelesaikan program studi strata satu (S1) pada program studi Teknik Informatika Putera Batam. Penulis menyadari bahwa skripsi ini masih jauh dari kata sempurna. Karena itu, kritik dan saran akan senantiasa penulis terima dengan senang hati.

Dengan segala keterbatasan, penulis menyadari pula bahwa skripsi ini takkan terwujud tanpa bantuan, bimbingan, dan dorongan dari berbagai pihak. Untuk itu dengan segala kerendahan hati, penulis menyampaikan ucapan terima kasih kepada:

1. Ibu Dr.Nur Elfi Husda,S.Kom.,M.SI selaku Rektor Universitas Putera Batam
2. Bapak Welly Sugiyanto,S.T.,M.M selaku Dekan Fakultas Teknik dan Komputer Universitas Putera Batam
3. Bapak Andi Maslan,S.T.,M.SI selaku Ketua Program Studi Teknik Informatika.
4. Bapak Koko Handoko, S.Kom., M.Kom. selaku pembimbing Skripsi.
5. Bapak Nopriadi, S.Kom., M.Kom. selaku pembimbing akademik.
6. Dosen dan Staff Universitas Putera Batam
7. Kedua orang tua penulis yang selalu mendoakan dan menyemangati penulis hingga penulisan skripsi ini selesai.
8. Keluarga penulis yang selalu mendoakan dan memberikan motivasi kepada penulis.
9. Teman-teman seperjuangan yang bersedia membagi ilmunya dan sharing pendapat.
10. Serta pihak-pihak lain yang tidak dapat penulis sebutkan satu per satu.

Semoga Tuhan Yang Maha Esa membalas kebaikan dan selalu mencurahkan berkatNya atas kebaikan yang penulis terima, Amin.

Batam, 01 Agustus 2022



Mangaratua Hutahaean

BAB I

PENDAHULUAN

1.1 Latar Belakang

Bergeraknya dengan sangat pesat teknologi berpengaruh bagi setiap aspek kehidupan masyarakat luas. Perkembangan teknologi tidak hanya digunakan sebagai sumber berbagi informasi tetapi juga sudah dimanfaatkan dalam skala besar. Terkadang penyampaian berita tidak sebanding sama kepentingan berita yang tinggi sehingga memerlukan pendalaman informasi lebih dalam dari data yang jumlahnya besar. Pendalaman atau penggalian informasi memerlukan data dalam jumlah besar dan memanfaatkan gudang data sebagai sumber pengambilan keputusan. Dalam proses pengambilan data ini menggunakan disebut komputasi, komputasi, teknologi metode pembelajaran mesin dan analisis data untuk memperoleh informasi berharga dan informasi terintegrasi dari berbagai kelompok besar.

Kumpulan sistem suatu informasi penunjang dalam pengambilan keputusan pada data besar disebut juga dengan *data mining* (Fajrin & Handoko, 2018). *Data mining* dapat diukur di tempat yang berbeda dengan data yang berbeda. Menurut Daryl Pregibon, pengembangan data adalah kombinasi dari banyak sistem penyimpanan dan pengambilan data yang cerdas. Salah satu metode yang digunakan dalam penambangan data adalah menyempurnakan *riset* pasar untuk menciptakan strategi pemasaran atau pasar yang efektif menggunakan data pasar yang tersedia untuk klinik atau perusahaan.

Klinik adalah tempat atau organisasi yang memberikan informasi kepada masyarakat tentang obat-obatan dan produk kesehatan lainnya. Saat ini sudah banyak klinik di pedesaan yang memudahkan masyarakat untuk berobat. (Wijaya & Dwiasnati, 2020). Khususnya obat-obatan pada bidang kesehatan yang terdapat pada klinik. Meningkatnya kebutuhan obat dapat dilihat dari jumlah *quantity* yang diminati pembeli dapat dijadikan suatu bisnis baru sehingga banyak pebisnis bergerak cepat dalam menyuplai jenis-jenis obat.

Terdapat permasalahan yang dialami oleh klinik yaitu dalam menentukan apakah produk obat tersebut laku atau tidak dalam waktu ke depan. Akibat kurang pemahaman mengenai minat pembeli terhadap merek produk obat dapat menimbulkan penumpukan obat di gudang ini mengakibatkan kerugian yang signifikan dan dapat menyebabkan barang obat-obatan kadaluarsa. Pemahaman yang baik mengenai barang yang sering terjual dapat meningkatkan jumlah pasokan produk obat yang paling banyak diminati atau dicari.

Untuk mengatasi permasalahan demikian maka akan dilakukan pengelompokan obat yang laku atau tidak laku pada klinik tersebut berlandaskan variabel yang telah ditemukan menggunakan Algoritma *Naive bayes* yang dapat memberikan informasi mengenai produk yang ada di klinik serta meminimalisir penumpukan *stock* yang tidak laku dan menjadikan acuan bagi pemilik untuk mengambil keputusan dalam menyuplai produk obat-obatan yang baru. Hasil yang di peroleh sebagai informasi di dapatkan dari pemerolehan nilai akurasi pada obat-obatan tersebut melalui pilihan. Data yang telah di dapat akan di kelola secara manual dan di input kedalam *software rapidminer* sehingga data yang di peroleh

mendapatkan hasil nilai akurat dari *ROC(Rank Order Centroid)* dengan nilai 80% berdasarkan tingkat pengelompokan dan prioritas. Berlandaskan uraian latar belakang diatas maka penulis berminat mengambil judul penelitian yaitu “**Penerapan *Data mining* untuk Memprediksi Penjualan Obat di Klinik Harapan Kita Batam**”.

1.2 Identifikasi Masalah

Berdasarkan latar belakang yang ada diatas maka dapat ditarik kesimpulan tentang identifikasi masalah sebagai berikut:

1. Dalam membagikan informasi mengenai laku atau tidak laku nya suatu produk mempengaruhi *stock* penjualan.
2. Kesulitan dalam menangani prediksi minat pembelian suatu obat untuk bulan selanjutnya.
3. Obat yang dijual berbeda, dari penjualan harga rendah hingga harga tinggi terutama pada peredaran jenis obat vitamin,nyeri,demam.

1.3 Batasan Masalah

Berdasarkan identifikasi masalah yang di dapat dan di jabarkan di latar belakang maka berikut batasan masalah yang diambil:

1. Data yang diperoleh untuk penelitian di dapat dari Klinik Harapan Kita Batam.
2. Penelitian ini menggunakan *Data mining* dengan metode algoritma *Naive bayes* dalam prediksi pemasaran obat.

3. Menggunakan *software Rapidminer* sebagai aplikasi pendukung uji data yang diharapkan memperoleh nilai akurasi *ROC(Rank Order Centroid)*.

1.4 Rumusan Masalah

Berdasarkan uraian dari latar belakang penelitian serta identifikasi masalah maka dirumuskan masalah penelitian sebagai berikut:

1. Bagaimana implementasi dalam prediksi pemasaran obat yang terdapat pada klinik Harapan Kita Batam?
2. Bagaimana data yang diambil bisa menghasilkan suatu pengetahuan baru dalam memprediksi penjualan obat?
3. Bagaimana hasil dari perhitungan metode *Agoritma Naive bayes* sesuai dengan perhitungan *software Rapidminer*?

1.5 Tujuan Penelitian

Berlandaskan uraian latar belakang diatas, dengan ini penelitian memiliki tujuan sebagai berikut:

1. Untuk prediksi pemasaran obat menggunakan *data mining* sebagai pendukung keputusan sehingga dapat mengatasi penumpukan di gudang dan ketersediaan produk penjualan meningkat.
2. Untuk hasil dari pengolahan data dan jenis obat yang di jual pada Klinik Harapan Kita Batam menghasilkan suatu prediksi yang tepat.

3. Untuk penerapan perhitungan Algoritma *Naive bayes* yaitu dengan menggunakan *software Rapidminer* sehingga diharapkan dapat memperoleh hasil nilai akurasi *ROC(Rank Order Centroid)* bernilai 80-100%.

1.6 Manfaat Penelitian

Adapun manfaat penelitian ini adalah sebagai berikut:

1.6.1 Manfaat Teoritis

1. Bagi Peneliti

Menambah pengetahuan dan menguatkan teori tentang *data mining* dan penggunaan algoritma *naive bayes* dalam memprediksi penjualan obat.

2. Bagi Pembaca

Untuk membagikan informasi sebagai ilmu kepada pembaca mengenai *data mining*.

3. Bagi Akademisi

Tujuan dari penelitian ini adalah untuk memberikan gambaran atau informasi bagi peneliti yang ingin mendalami *data mining* khususnya bagi peneliti yang sedang mengembangkan *data mining* dalam memprediksi penjualan farmasi.

1.6.2. Manfaat Praktis

1. Dengan di buat nya karya ilmiah ini di harapkan dapat membatu para penjualan obat klinik dalam memprediksi obat yang akan di seddiakan di bulan berikutnya.

2. Karya tulis ilmiah ini di harapkan menabahnya pengetahuan bagi penulis tentang penerapan *data mining* untuk memprediksi penjualan obat

BAB II

TINJAUAN PUSTAKA

2.1 *Knowledge Discovery in Database (KDD)*

KDD adalah bagian dari ilmu pengetahuan yang mencakup *database* serta dapat diuraikan sebagai proses pengenalan pola kepada pengetahuan yang memiliki sekelompok besar data dan menghasilkan informasi yang mudah untuk dipahami. KDD memiliki beberapa proses untuk mendapatkan data yang ingin di capai seperti pemurnian data, integrasi data, dan pemilihan data dan presentasi pengetahuan (Elmayati, 2017)

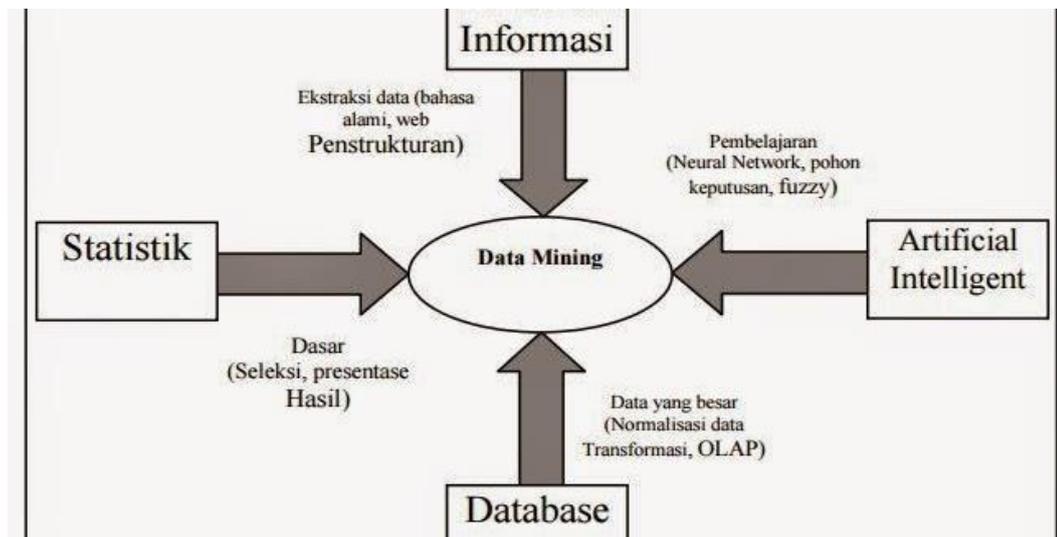
Data mining adalah akuisisi pengetahuan dalam praktiknya meliputi pembersihan data, integrasi data, pemilihan data, konversi data, analisis data. Pengumpulan data mengacu pada proses penggalian informasi dari sejumlah besar data untuk mendapatkan informasi baru tentang situs tertentu. (Fajrin & Handoko, 2018).

2.2 *Data mining*

Data mining merupakan kumpulan dari proses Kumpulan sistem suatu informasi penunjang dalam pengambilan keputusan pada data besar disebut juga dengan *data mining* (Fajrin & Handoko, 2018). *Data mining* dapat diukur di tempat yang berbeda dengan data yang berbeda. Menurut Daryl Pregibon, penambangan data adalah kombinasi dari banyak sistem penyimpanan dan pengambilan data yang cerdas. Salah satu metode yang digunakan dalam penambangan data adalah

menyempurnakan *riset* pasar untuk menciptakan strategi pemasaran atau pasar yang efektif menggunakan data pasar yang tersedia untuk klinik atau perusahaan.

Pengambilan informasi dapat digunakan di berbagai situs, termasuk informasi. Menurut Daryl Prigiban, pengolahan data merupakan kombinasi dari statistik, keahlian dan pengumpulan data. Salah satu metode yang digunakan dalam pencarian informasi adalah aturan perusahaan yang menganalisis produk untuk menetapkan strategi praktis untuk penjualan atau periklanan dengan menggunakan informasi penjualan yang tersedia di klinik atau di perusahaan.



Gambar 2. Ilmu *data mining*
Sumber : (Elmayanti, 2017)

Pada *data mining* terdapat bagian yang menjadi fungsi utamanya yaitu fungsi *descriptive* dan fungsi *predictive*.

1. *Descriptive*

Descriptive berfungsi sebagai informasi dari analisis data untuk mendapatkan pemahaman yang lebih baik dari informasi yang dipantau atau yang telah diamati. Setelah proses, berharap untuk mengetahui sifat data. Dari data yang telah di proses

tersebut maka dapat di lihat jenis informasi yang terkait. Dengan menggunakan fungsi penggalian data, dapat menemukan pola lain yang tersembunyi dalam rekaman. Dengan kata lain, ketika model diulang dan sesuai, elemen data dapat diidentifikasi.

2. *Estimasion*

Model ini disusun menggunakan *database* yang lengkap dan didasarkan pada nilai tukar yang diharapkan. Selain itu, dalam analisis lain, *varians* yang diestimasi didasarkan pada *varians* yang diharapkan. Misalnya, kami mengevaluasi tekanan darah sistolik pada pasien rumah sakit berdasarkan usia pasien, jenis kelamin, berat badan, dan kadar natrium darah. Korelasi antara variabel prognostik dalam tekanan darah sistolik dan proses pembelajaran memungkinkan terciptanya model hasil. Model penilaian yang dihasilkan dapat digunakan untuk masalah baru lainnya.(Harahap et al., 2022).

3. *Predictive*

Prediktif adalah bagian dari bagaimana satu pola menangkap data untuk pola lain. Pola-pola ini dapat diidentifikasi dengan perubahan yang berbeda dalam data. Setelah pola menemukan modelnya, maka pola dapat menggunakan model hasil untuk memprediksi beberapa variasi dalam nilai atau jenis yang tidak diketahui. Oleh karena itu, karya ini dikatakan sebagai karya prediksi dan peramalan. Fungsi ini juga dapat digunakan untuk mengidentifikasi variabel tertentu yang tidak ada dalam data. Tugas ini mudah dan nyaman bagi siapa saja yang membutuhkan informasi akurat untuk memperbaiki masalah penting ini.

Pada *data mining* terdapat fungsi yang perlu diketahui, yaitu klasifikasi, asosiasi dan *cluster*.

1. Klasifikasi

Klasifikasi adalah bagian dari aktivitas yang menyusun suatu sistem atau data dalam bentuk berkelompok atau golongan dengan standar yang telah ditetapkan. Metode ini adalah metode yang paling umum digunakan metode ini bertujuan untuk memperkirakan kelas dari suatu objek yang labelnya belum diketahui. Kelas adalah jenis tertinggi di mana dua kalimat adalah "ya" dan "tidak."(Ulfha & Amin, 2020).

2. Asosiasi

Asosiasi merupakan teknik yang digunakan untuk menemukan suatu hubungan yang menunjukkan suatu kondisi di dalam satu set data yang beberapa nilai atribut akan muncul secara bersamaan. Analisis hubungan mencakup studi tentang "apa dan apa". Contoh: 90% konsumen membeli vitamin di apotek dan 60% konsumen membeli keduanya.

Analisis asosiasi ini akan menghasilkan aturan asosiasi (*association rules*). Contoh: 90% orang yang berbelanja di suatu klinik yang membeli obat juga membeli vitamin, dan 60% dari semua orang yang berbelanja membeli keduanya. Sekumpulan *record*, masing-masing berisi kumpulan elemen tertentu, menciptakan aturan untuk mengamati salah satu dari empat elemen dari sesuatu yang terjadi. Aturan umumnya sering disebut sebagai riset pasar, karena menganalisis data transaksi pelanggan untuk menentukan produk mana yang dijual.(Despitaria et al., 2016).

3. *Clauster*

Clauster adalah topik sederhana untuk mengklasifikasikan data ke dalam satu kategori objek yang tidak didasarkan pada kumpulan data tertentu. *Cluster* memiliki tujuan untuk mengelompokkan suatu kelas ke dalam beberapa segmen berdasarkan atribut yang telah ditentukan. Mengklasifikasikan, memperkirakan, atau memprediksi nilai dari suatu variabel yang diestimasi. Namun, *algoritma clustering* mencoba untuk membagi semua data menjadi segmen-segmen yang identik, di mana data di satu bagian lebih berharga, tetapi di bagian lain datanya kurang penting.(Elmayati, 2017).

2.2.1 Tingkatan-tingkatan *Data mining*

Menurut (Efori Buulolo, 2017) ada beberapa tingkatan pada *data mining* yaitu:

1. Pembersihan data adalah metode yang digunakan untuk menghapus data yang tidak akurat dan tidak sesuai.
2. *Integrasi* data dapat digunakan dalam satu *database*.
3. Pilihan data yang akan diperoleh harus konsisten dengan hasil analisis
4. Memperbarui data Pemrosesan data memerlukan format dan konten *Internet* yang tepat.
5. Proses pengiriman.
6. Selama proses peninjauan, coba tunjukkan pola yang menarik kontras dengan informasi tersembunyi.

7. Manifestasi kebijaksanaan. Ini adalah proses yang memberikan informasi dan pengalaman yang membantu pengguna tetap mendapat informasi tentang hasil data mereka

2.2.2 Manfaat *Data mining*

Menurut (Efori Buulolo, 2017) dilihat dari sudut pandang terdapat dua manfaat yaitu:

1. Sudut Pandang Komersial.

Dari sudut pandang komersial, *data mining* dapat digunakan untuk menentukan jumlah data. Cara lain untuk bersaing adalah dengan menggunakan metode *data mining* untuk menangkap data dan hasil.

2. Perspektif ilmiah

Data mining merupakan bagian yang berfungsi sebagai analisis data yang dapat menyimpan data besar.

2.3 Metode *Naive bayes*

Naive bayes merupakan salah satu metode yang terdapat pada *data mining*. Algoritma *naive bayes* adalah teknik klasifikasi berdasarkan penerapan logika bayes dengan asumsi yang kuat bahwa semua prediktor independen satu sama lain dengan kata lain bahwa adanya kehadiran fitur kelas tidak tergantung pada kehadiran fitur kelas lain dikelas yang sama. Algoritma *naive bayes* banyak digunakan dalam proses prediksi karena memiliki kemampuan prediksi di bulan yang akan datang (Wijaya & Dwiasnati, 2020)

2.4 *Software* Pendukung

Software pendukung merupakan suatu perangkat lunak yang digunakan oleh peneliti untuk mendukung penelitiannya, berikut dibawah ini merupakan *software* pendukung yang digunakan oleh peneliti:

1. *Rapid Miner*

Rapid miner merupakan salah satu cara yang sering digunakan dalam menemukan solusi pada *data mining* karena mampu di jalankan pada berbagai sistem operasi. *Rapid miner* adalah sebuah perangkat lunak yang memiliki sifat *open source*. Sebelum dikenal dengan nama *rapid miner*, perangkat ini memiliki nama lain yaitu *YALE* (Yet Another Learning Environment) yang dikembangkan pada tahun 2001 oleh Ralf Klinkenberg.

Menurut Kdnuggets *RapidMiner* adalah bagian yang teratas dalam sebagai *software data mining*. *Rapidminer* menyediakan *GUI* (*Graphic Interface User*) untuk desain detail *pipeline*. *GUI* ini membuat file *XML* (bahasa penandaan tambahan) yang menentukan metode penyortiran yang harus digunakan pengguna untuk data. File ini dibaca oleh *Rapidminer* dan diaktifkan langsung dari penganalisis.

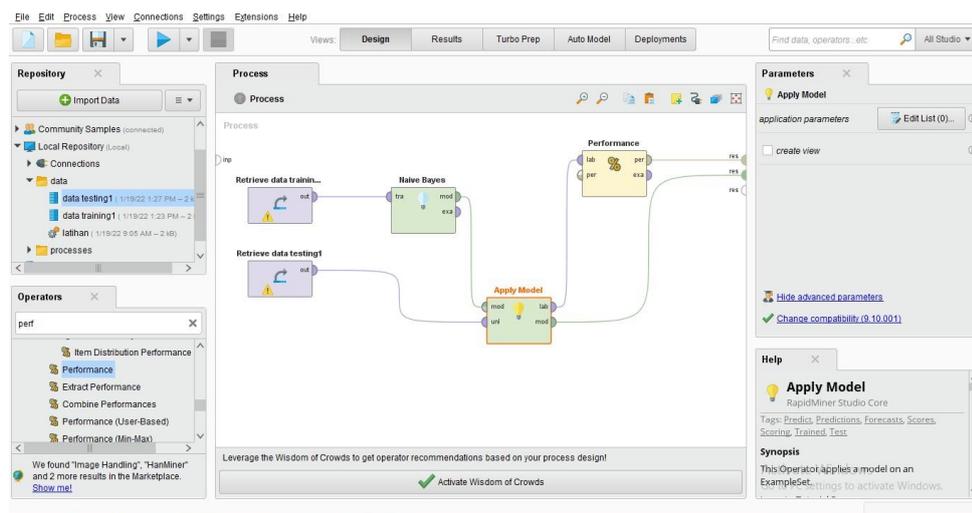
Rapidminer memiliki beberapa sifat sebagai berikut:

1. Java ditulis dalam bahasa pemrograman sehingga dapat berjalan di sistem operasi yang berbeda
2. Cara memperoleh informasi ada dalam model pohon
3. *XML Internal*
4. Bahasa *scripting* memungkinkan untuk eksperimen skala besar

5. Rancangan *multi-layer* untuk menjaga bentuk data yang efisien dan menjaga penanganan data.
6. Mempunyai *GUI*, *command line* mode dan *Java API* yang dapat dipanggil dari program lain.

Beberapa fitur dari *Rapidminer*, antara lain:

1. Seperti *decision tree* dan *selforganization map* adalah algoritma *data mining*.
2. *Tree chart* dan *3D scatter plots* bentuk grafis yang terbaru.
3. Seperti *text plugin* untuk melakukan analisis teks.
4. Menyuplai prosedur *data mining* dan *machine learning* termasuk: *ETL* (*extraction, transformation, loading*) *data preprocessing, visualisasi, modeling dan evaluasi*.
5. *Data mining* memiliki proses yang teratur dideskripsikan dengan *XML* dan dibuat dengan *GUI*.
6. Menggabungkan proyek *data mining* weka dan statistik R.



Gambar 2. 2 *Rapidminer*
Sumber: Data Penelitian 2022

2.5 Objek Penelitian

Pada Penelitian ini objek yang akan di ambil sebagai tempat penelitian adalah Klinik Harapan Kita Batam di Puri Legenda Blok D3 No 17 Batam Center, Kota Batam, Kepulauan Riau, 29464. Klinik Harapan Kita Batam Kota diresmikan sebagai Klinik Umum pada tanggal 17 November 2015 oleh PLT Gubernur Kepulauan Riau yang dalam pengelolaannya dibawah PT. Harapan Abadi Kencana (sebagai Representasi Pemilik) dengan persetujuan operasional yang diberikan oleh pemerintah kota Batam pada tanggal 06 November 2015 dengan nomor KPTS.330/HK/XI/2015.

Klinik Harapan Kita Batam Kota hadir untuk melengkapi impian masyarakat di Pulau Batam dan sekitarnya, pada tujuan penting pada jasa kesehatan pasien. Klinik ini dilengkapi dengan Fasilitas Kesehatan dan tenaga professional medis untuk siap melayani pasien.

2.6 Penelitian Terdahulu

Penelitian yang akan di lakukan berdasarkan penelitian terdahulu yang di kembangkan dengan upaya peneliti untuk menemukan kesamaan dan menemukan ide-ide baru tentang penelitian lain. Selain itu, dengan adanya penelitian terdahulu dapat di jadikan referensi tambahan bagi peneliti. Berikut merupakan penelitian terdahulu yang masih terkait dengan tema yang peneliti kaji:

1. Pada penelitian ini dilakukan oleh (Ulfha & Amin, 2020) dengan P-ISSN: 1693-7554, E-ISSN: 2654-3990 dengan judul “ **Implementasi *data mining* untuk mengetahui pola pembelian obat menggunakan algoritma apriori**”

Dalam penelitian ini, dimana menyatakan bahwa *Data mining*, sering juga disebut sebagai pengambilan pengetahuan dari *database (KDD)*, yang memiliki arti sebagai aktivitas yang melibatkan pengumpulan dan penggunaan data histori untuk menemukan *skrip*, model, atau hubungan dalam kumpulan data besar. Seringkali model atau aturan yang dihasilkan dapat digunakan sebagai panduan untuk memutuskan komposisi elemen berdasarkan kombinasi elemen yang diekstraksi dan elemen terkait.

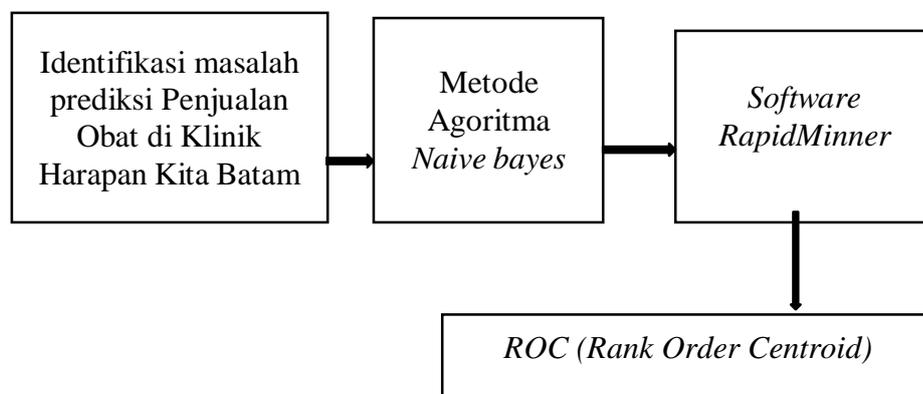
2. Pada penelitian ini dilakukan oleh (Wijaya & Dwiasnati, 2020) dengan ISSN: 2355-6579 | E-ISSN: 2528-2247 dengan judul “ **Implementasi *Data mining* dengan Algoritma *Naive bayes* pada Penjualan Obat**”. Dalam penelitian ini, dimana menyatakan Algoritma *Naive bayes* adalah sebuah algoritma yang mampu memperkirakan probabilitas yang paling sederhana menurut teorema bayes dengan hipotesis independen terkuat. Selain itu algoritma *naive bayes* mudah dan dapat menangani peristiwa prediksi di bulan-bulan mendatang. Keuntungan dari algoritma ini adalah ide yang bagus banyaknya penjualan pada bulan selanjutnya.
3. Pada penelitian ini dilakukan oleh (Despitaria et al., 2016) dengan judul “**Analisis asosiasi pada transaksi obat menggunakan *data mining* dengan algoritma *apriori***” Dalam penelitian ini menyatakan bahwa *Association* merupakan salah satu metode yang terdapat pada *data mining* yang dapat mengidentifikasi sebuah kaitan antara item. Analisis gabungan ini menentukan kondisi gabungan. Misalnya, 90% pengguna rumah sakit yang

membeli apotek membeli vitamin, dan 60% dari semua konsumen membeli semuanya bersama-sama..

4. Pada penelitian ini dilakukan oleh (Nainel et al., 2020) e-ISSN 2715-7393 (Media Online) dengan judul “ **Penerapan *data mining* untuk estimasi penjualan obat berdasarkan pengaruh *brand image* dengan algoritma *expectation maximation*”** menarik kesimpulan bahwa dengan *data mining* perusahaan dapat mengambil langkah-langkah yang tepat untuk meningkatkan target penjualan di masa depan.
5. Pada penelitian ini dilakukan oleh (Andriani et al., 2019) dengan e-ISSN: 2655-1438 p-ISSN: 2655-1632 dengan judul **Penerapan *Data mining* Menggunakan Algoritma Apriori Pada Website Indie Clothing**”. Dalam penelitian ini, *Data mining* adalah sistem yang menggunakan angka, matematika, pemodelan dan teknik pembelajaran mesin untuk mendeteksi dan mengamati data dan informasi yang berguna dikumpulkan dari database besar.
6. Berdasarkan pada penelitian ini yang dilakukan oleh (Fajrin & Handoko, 2018) berjudul “**Penerapan *data mining* untuk mengolah tata letak buku dengan metode *association rule*”** menarik kesimpulan bahwa *data mining* merupakan teknik yang sering digunakan dalam mendukung pengambilan proses keputusan dengan cara mengumpulkan serta mengenali semua nilai yang ada dalam kumpulan data besar.

2.7 Kerangka Pemikiran

Kerangka pemikiran adalah ide dari suatu pokok permasalahan yang akan dicari langkah dalam proses awal hingga sampai akhir yang digambarkan dalam bentuk desain alur urutan penyelesaian penelitian. Berikut kerangka pemikiran yang telah dibuat oleh peneliti:



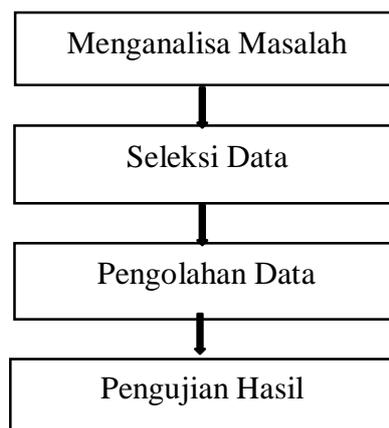
Gambar 2. 3 Kerangka Pemikiran
Sumber : Data penelitian 2022

BAB III

METODE PENELITIAN

3.1. Desain Penelitian

Ketika merancang sebuah penelitian, peneliti menggambarkan hasil model penelitian yang terstruktur secara sistematis sehingga dapat digunakan sebagai referensi untuk menemukan dan membantu peneliti dalam proses penelitian, serta sebagai panduan untuk merancang penelitian.



Gambar 3. 1 Desain Penelitian
Sumber : Data penelitian 2022

Adapun desain penelitian pada penelitian ini dapat dijelaskan sebagai berikut:

1. Menganalisa Masalah

Masalah yang telah ditemukan akan dianalisis maksud menginterpretasikan persoalan yang telah ditentukan sebelumnya yaitu masalah prediksi penjualan obat di klinik Harapan Kita Batam.

2. Seleksi Data

Merupakan proses meminimalkan jumlah data yang digunakan untuk proses *data mining* dengan tetap mempresentasikan data aslinya. Data yang diseleksi pada penelitian ini adalah data transaksi penjualan, obat di tahun 2018-2020.

3. Pengolahan Data

Data yang telah dianalisis dari data transaksi penjualan obat tahun 2018 - 2020 yang terdiri dari beberapa atribut tahun, jenis obat, transaksi, penjualan obat dan di proses dengan menggunakan metode *Naive bayes*.

4. Pengujian Data

Pada langkah terakhir peneliti akan membuat pengujian data terhadap klinik Harapan Kita Batam menggunakan perangkat lunak *Rapidminer* sehingga dapat terlihat apakah pengujian yang dilakukan sudah sesuai dengan yang diharapkan dalam memperoleh data dari suatu hubungan.

3.2. Teknik pengumpulan data

Berikut ini beberapa teknik pengumpulan data yang di gunakan peneliti dalam proses pengumpulan data yang di lakukan pada penelitian ini:

1. Teknik *Observasi*

Teknik ini merupakan teknik yang dilakukan peneliti dengan cara langsung terjun kelapangan atau langsung mengunjungi klinik Harapan kita batam dan melihat kondisi yang sedang terjadi pada klinik beserta mengumpulkan data dengan cara meminta data transaksi penjualan obat pada klinik tersebut sesuai data yang akan di butuhkan.

2. Teknik Wawancara

Teknik ini merupakan proses tanya jawab antara wawancara dan narasumber. Hasil wawancara di peroleh dari pertanyaan-pertanyaan yang telah di susun oleh peneliti atau pewawancara yang di tanyakan langsung ke pada pemilik atau staff penanggung jawab yang bekerja di klinik tersebut.

3. *Study Liteature*

Studi literatur adalah proses pengumpulan data yang merangkum informasi yang relevan dengan topik pembahasan. Dalam hal ini, peneliti belajar dengan mencari informasi dari sumber tertulis, seperti jurnal dan buku-buku yang relevan dengan penelitian, untuk mengumpulkan semua data yang diperlukan untuk penelitian.

3.3. *Operasional Variabel*

Operasional variabel merupakan hal yang sangat di perlukan dalam sebuah penelitian agar dapat menentukan indikator dan variabel yang di butuhkan dalam penelitian. *Operasional* variabel berisi teori secara ilmiah yang memiliki dasar utama dalam menentukan hubungan antara variabel dengan penelitian. Berikut ini variabel yang telah ditentukan dalam penelitian ini:

1. Tahun

Data yang di buat sebagai penelitian merupakan data yang diambil dari tahun 2018-2020, di mana data tersebut berdasarkan jumlah transaksi dari setiap produk yang paling banyak terjual sehingga dapat di tentukan kombinasi antar *item set*.

2. Jenis Obat

Jenis obat merupakan macam-macam obat yang ada pada klinik harapan kita batam, pada penelitian ini jenis obat yang diangkat sebagai variabel adalah jenis obat nyeri, infeksi, vitamin dan demam yang paling sering di minati oleh pembeli obat.

3. Transaksi

Transaksi merupakan jumlah pembelian dan penjualan obat yang terdapat pada klinik harapan kita batam, dengan kategori pengelompokan sedikit, hampir dan banyak permintaan obat.

4. Penjualan Obat

Penjualan obat merupakan pengeluaran obat yang dinyatakan terjual dengan kategori laku dan tidak laku.

Berikut tabel di bawah ini data yang telah diperoleh dari Klinik Harapan Kita Batam Tahun 2018-2020.

Tabel 3. 1 Data Penelitian

Tahun	Data Tahun	Jenis Obat	Nama Obat	Trans	Pr	PO
18	Jan-18	Nyeri	Diclofenac Potasium	Sedikit	270	Tidak
18	Feb-18	Nyeri	Natrium Diclofenac	Hampir	520	Laku
18	Mar-18	Nyeri	Asam Mefenamat Tab	Hampir	500	Laku
18	Apr-18	Infeksi	Clindamicyn	Sedikit	100	Tidak
18	Mei-18	Infeksi	Amoxicillin Tab	Hampir	740	Laku
18	Jun-18	Infeksi	Cefadroxil	Hampir	530	Laku
18	Jul-18	Vitamin	Thialicyn Syr	Sedikit	44	Tidak
18	Agu-18	Vitamin	Pritavid	Sedikit	200	Tidak
18	Sep-18	Vitamin	Caviplex Syr	Sedikit	120	Tidak
18	Okt-18	Demam	Sanmol Tab	Sedikit	78	Tidak

Tabel Lanjutan

18	Nov-18	Demam	Paracetamol Tab	Hampir	1100	Tidak
18	Des-18	Demam	Novagesic Tab	Hampir	980	Laku
18	Jan-18	Nyeri	Natrium Diclofenac	Sedikit	130	Tidak
19	Feb-19	Nyeri	Neuralgin	Sedikit	50	Tidak
19	Mar-19	Nyeri	Diclofenac Potasium	Sedikit	45	Tidak
19	Apr-19	Infeksi	Amoxicilin Tab	Hampir	1300	Laku
19	Mei-19	Infeksi	Cefadroxil	Sedikit	450	Tidak
19	Jun-19	Infeksi	Clindamicyn	Sedikit	180	Tidak
19	Jul-19	Vitamin	B Com C	Sedikit	270	Tidak
19	Agu-19	Vitamin	Pritavit Syr	Sedikit	77	Tidak
19	Sep-19	Vitamin	Thyalicin Syr	Sedikit	43	Tidak
19	Okt-19	Demam	Pamol Syr	Sedikit	50	Tidak
19	Nov-19	Demam	Sanmol Syr	Sedikit	280	Tidak
19	Des-19	Demam	Paracetamol Tab	Hampir	1250	Laku
20	Jan-20	Nyeri	Neuralgin Tab	Sedikit	400	Tidak
20	Feb-20	Nyeri	Asam Mefenamat Tab	Hampir	500	Laku
20	Mar-20	Nyeri	Novagesic Tab	Sedikit	380	Tidak
20	Apr-20	Infeksi	Eritromicin Tab	Sedikit	288	Tidak
20	Mei-20	Infeksi	Amoxicilin Syr	Sedikit	50	Tidak
20	Jun-20	Infeksi	Amoxicilin Tab	Hampir	1250	Laku
20	Jul-20	Vitamin	Pritavit Syr	Sedikit	100	Tidak
20	Agu-20	Vitamin	Caviplek Syr	Sedikit	60	Tidak
20	Sep-20	Vitamin	Thialisyn Syr	Sedikit	65	Tidak
20	Okt-20	Demam	Sanmol Syr	Sedikit	100	Tidak
20	Nov-20	Demam	Novagesic Tab	Hampir	700	Laku
20	Des-20	Demam	Paracetamol Tab	Banyak	1370	Laku

Sumber : Data Penelitian 2022

3.4 Metode Perancangan sistem

Pada penelitian ini metode perancangan sistem yang digunakan adalah dengan menggunakan metode algoritma *naive bayes* untuk memprediksi penjualan obat.

3.5 Lokasi Dan Jadwal Penelitian

3.5.1 Lokasi Penelitian

Penelitian ini dilakukan di Klinik Harapan Kita Batam beralamat di Puri Legenda Blok D3 No 17 Batam Center, Kota Batam, Kepulauan Riau, 29464.



Gambar 2. 4 Tempat Penelitian
Sumber : Data penelitian 2022

3.5.2. Jadwal Penelitian

Setiap penelitian pasti memiliki jadwal kegiatan yang akan dilaksanakan, jadwal kegiatan berisi tentang apa saja yang akan dilakukan selama penelitian. Berikut ini tabel jadwal kegiatan yang dilakukan selama penelitian berlangsung.

Tabel 3. 2 Jadwal Penelitian

No	Kegiatan	Tahun 2022																		
		Mar-22				Apr-22				Mei-22				Jun-22				Jul-22		
		1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3
1	Pengajuan Surat Penelitian																			
2	Nyusun Bab I																			
3	Nyusun Bab II																			
4	Nyusun Bab III																			
5	Nyusun Bab IV																			
6	Nyusun Bab V, Daftar Pustaka, Lampiran																			

Sumber : Data Penelitian 2022

BAB IV

HASIL DAN PEMBAHASAN

4.1 Hasil Penelitian

Penelitian ini menghasilkan nilai akurasi menggunakan metode algoritma *naive bayes* dalam memprediksi penjualan obat dengan melakukan perhitungan dan pengolahan data secara manual menggunakan *microsoft excel* dan menggunakan *software rapidminer*.

4.1.1 Data set

Dataset penjualan terdiri dari 8 atribut antara lain Tahun, Tahun Data, Jenis obat, Nama obat, Transaksi, Permintaan, dan Penjualan Obat, 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 data *set* yang akan di lakukan pengujian nya.

Tabel 4. 1 Data set

Tahun	Data Tahun	Jenis Obat	Nama Obat	Tr	Pr	PO
18	Jan-18	Nyeri	Diclofenac Potasium	Sedikit	270	Tidak
18	Feb-18	Nyeri	Natrium Diclofenac	Hampir	520	Laku
18	Mar-18	Nyeri	Asam Mefenamat Tab	Hampir	500	Laku
18	Apr-18	Infeksi	Clindamicyn	Sedikit	100	Tidak
18	Mei-18	Infeksi	Amoxicillin Tab	Hampir	740	Laku
18	Jun-18	Infeksi	Cefadroxil	Hampir	530	Laku
18	Jul-18	Vitamin	Thialicyn Syr	Sedikit	44	Tidak

18	Agu-18	Vitamin	Pritavid	Sedikit	200	Tidak
18	Sep-18	Vitamin	Caviplex Syr	Sedikit	120	Tidak
18	Okt-18	Demam	Sanmol Tab	Sedikit	78	Tidak
18	Nov-18	Demam	Paracetamol Tab	Hampir	1100	Tidak
18	Des-18	Demam	Novagesic Tab	Hampir	980	Laku
18	Jan-18	Nyeri	Natrium Diclofenac	Sedikit	130	Tidak
19	Feb-19	Nyeri	Neuralgin	Sedikit	50	Tidak
19	Mar-19	Nyeri	Diclofenac Potasium	Sedikit	45	Tidak
19	Apr-19	Infeksi	Amoxicilin Tab	Hampir	1300	Laku
19	Mei-19	Infeksi	Cefadroxil	Sedikit	450	Tidak
19	Jun-19	Infeksi	Clindamicyn	Sedikit	180	Tidak
19	Jul-19	Vitamin	B Com C	Sedikit	270	Tidak
19	Agu-19	Vitamin	Pritavit Syr	Sedikit	77	Tidak
19	Sep-19	Vitamin	Thyalicin Syr	Sedikit	43	Tidak
19	Okt-19	Demam	Pamol Syr	Sedikit	50	Tidak
19	Nov-19	Demam	Sanmol Syr	Sedikit	280	Tidak
19	Des-19	Demam	Paracetamol Tab	Hampir	1250	Laku
20	Jan-20	Nyeri	Neuralgin Tab	Sedikit	400	Tidak
20	Feb-20	Nyeri	Asam Mefenamat Tab	Hampir	500	Laku
20	Mar-20	Nyeri	Novagesic Tab	Sedikit	380	Tidak
20	Apr-20	Infeksi	Eritromicin Tab	Sedikit	288	Tidak
20	Mei-20	Infeksi	Amoxicilin Syr	Sedikit	50	Tidak
20	Jun-20	Infeksi	Amoxicilin Tab	Hampir	1250	Laku
20	Jul-20	Vitamin	Pritavit Syr	Sedikit	100	Tidak
20	Agu-20	Vitamin	Caviplek Syr	Sedikit	60	Tidak
20	Sep-20	Vitamin	Thialisyn Syr	Sedikit	65	Tidak
20	Okt-20	Demam	Sanmol Syr	Sedikit	100	Tidak
20	Nov-20	Demam	Novagesic Tab	Hampir	700	Laku
20	Des-20	Demam	Paracetamol Tab	Banyak	1370	Laku

Sumber: Data Penelitian 2022

4.1.2 Seleksi Data

Seleksi data merupakan penganalisan data-data yang relevan dari *database*. Pada database tidak semua data akan di jadikan sebagai data dalam pengujian maka perlunya melakukan seleksi data. Data tersebut di pilih dan di seleksi untuk dianalisis, dari semua data yang terdapat pada database maka terdapat empat data dari hasil seleksi untuk di jadikan indikator penelitian yaitu tahun, jenis obat, transaksi dan penjualan obat.

Tabel 4. 2 Seleksi Data

Tahun	Jenis Obat	Transaksi	Penjualan Obat
18	Nyeri	Sedikit	Tidak
18	Nyeri	Hampir	Laku
18	Nyeri	Hampir	Laku
18	Infeksi	Sedikit	Tidak
18	Infeksi	Hampir	Laku
18	Infeksi	Hampir	Laku
18	Vitamin	Sedikit	Tidak
18	Vitamin	Sedikit	Tidak
18	Vitamin	Sedikit	Tidak
18	Demam	Sedikit	Tidak
18	Demam	Hampir	Tidak
18	Demam	Hampir	Laku
19	Nyeri	Sedikit	Tidak
19	Nyeri	Sedikit	Tidak
19	Nyeri	Sedikit	Tidak
19	Infeksi	Hampir	Laku
19	Infeksi	Sedikit	Laku
19	Infeksi	Sedikit	Laku
19	Vitamin	Sedikit	Tidak
19	Vitamin	Sedikit	Tidak
19	Vitamin	Sedikit	Laku
19	Demam	Sedikit	Laku
19	Demam	Sedikit	Tidak

Tabel Lanjutan

19	Demam	Hampir	Laku
20	Nyeri	Sedikit	Tidak
20	Nyeri	Hampir	Laku
20	Nyeri	Sedikit	Tidak
20	Infeksi	Sedikit	Tidak
20	Infeksi	Sedikit	Tidak
20	Infeksi	Hampir	Laku
20	Vitamin	Sedikit	Tidak
20	Vitamin	Sedikit	Laku
20	Vitamin	Sedikit	Laku
20	Demam	Sedikit	Laku
20	Demam	Hampir	Laku
20	Demam	Banyak	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 4. 3 Probalitas Kelas

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

Sumber: Data Penelitian 2022

Tabel 4. 4 Jumlah Obat Terjual

Jumlah data	Kelas Laku	Kelas Tidak laku
36	11	25

Sumber: Data Penelitian 2022

4.2 Analisa Proses Algoritma

Pada tahap ini dilakukan pemodelan data, metode yang digunakan dalam penelitian ini adalah *probability* (perkiraan) dengan menggunakan algoritma *Naive bayes*. Data yang terkumpul akan diolah untuk memanfaatkan peluang tersebut. Metode ini dapat digunakan untuk melihat peluang masa depan berdasarkan pengalaman masa lalu 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 4. 5 Data Training

Tahun	Jenis Obat	Transaksi	Penjualan Obat
18	Nyeri	Sedikit	Tidak
18	Nyeri	Hampir	Laku
18	Nyeri	Hampir	Laku
18	Infeksi	Sedikit	Tidak
18	Infeksi	Hampir	Laku
18	Infeksi	Hampir	Laku
18	Vitamin	Sedikit	Tidak
18	Vitamin	Sedikit	Tidak
18	Vitamin	Sedikit	Tidak
18	Demam	Sedikit	Tidak

18	Demam	Hampir	Tidak
18	Demam	Hampir	Laku
19	Nyeri	Sedikit	Tidak
19	Nyeri	Sedikit	Tidak
19	Nyeri	Sedikit	Tidak
19	Infeksi	Hampir	Laku
19	Infeksi	Sedikit	Laku
19	Infeksi	Sedikit	Laku
19	Vitamin	Sedikit	Tidak
19	Vitamin	Sedikit	Tidak
19	Vitamin	Sedikit	Laku
19	Demam	Sedikit	Laku
19	Demam	Sedikit	Tidak
19	Demam	Hampir	Laku
20	Nyeri	Sedikit	Tidak
20	Nyeri	Hampir	Laku
20	Nyeri	Sedikit	Tidak
20	Infeksi	Sedikit	Tidak
20	Infeksi	Sedikit	Tidak
20	Infeksi	Hampir	Laku
20	Vitamin	Sedikit	Tidak
20	Vitamin	Sedikit	Laku
20	Vitamin	Sedikit	Laku
20	Demam	Sedikit	Laku
20	Demam	Hampir	Laku
20	Demam	Banyak	Laku

Sumber: Data Penelitian 2022

Tabel 4. 6 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 4.5 diatas maka dilakukan pencarian probalitas dengan metode agoritma *naive bayes* seperi dibawah ini:

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

Rumus 4. 1 Persamaan Agoritma *Naive bayes*

Keterangan:

X : Data dengan class yang belum diketahui

T : Hipotensis data X merupakan suatu class spesifik

P(T|X): Probalitas hipotensis T berdassarkan kondisi (*posteriori probability*)

P(T): Probalitas hipotensis T (*Perior probability*)

P(X|T): Probalitas X berdasarkan kondisi hipotesis T

P(X): Probalitas dari T

4.2.1 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 Probabilitas kriteria Tahun:

Tabel 4. 7 Probabilitas Kriteria Tahun

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

Sumber: Data penelitian 2022

4.2.2 Probabilitas Kriteria Jenis Obat

Berdasarkan data penjualan obat pada data set diatas, diketahui bahwa terdapat 4 jenis obat yaitu Nyeri, Infeksi, Vitamin, Demam data tiap perbulan nya. 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, Berikut tabel Probabilitas kriteria jenis obat:

Tabel 4. 8 Probabilitas Kriteria Jenis Obat

P(JO)=	Laku	Tidak
Nyeri	27%	24%
Infeksi	36%	20%
Vitamin	0%	36%
Demam	36%	20%
Jumlah	100%	100%

Sumber: Data penelitian 2022

4.2.3 Probabilitas Kriteria Transaksi

Berdasarkan data penjualan obat pada data set diatas, diketahui bahwa terdapat 4 transaksi yaitu Sedikit, Hampir, data tiap perbulan nya. Dari 3 data tersebut terdapat kriteria sedikit 0% dengan penjualan obat laku dan 96% dengan penjualan obat tidak laku, terdapat kriteria hampir 91% dengan penjualan obat laku dan 4% dengan penjualan obat tidak laku, terdapat kriteria banyak 9% dengan penjualan obat laku dan 0% dengan penjualan obat tidak laku, Berikut tabel Probabilitas kriteria transaksi:

Tabel 4. 9 Probabilitas Kriteria Transaksi

P(Tr)=	Laku	Tidak
Sedikit	0%	96%
Hampir	91%	4%
Banyak	9%	0%
Jumlah	100%	100%

Sumber: Data penelitian 2022

4.3 Hasil Pengujian

Dengan pemberian 5 data yang telah di jadikan data testing maka hasil yang di dapatkan dengan pengujian data menggunakan perhitungan *microsoft excel* menghasilkan *class* prediksi pencarian dengan nilai *accuracy* 80%. Berikut di bawah ini merupakan tabel data yang akan di jadikan sebagai data *testing*;

Tabel 4. 10 Hasil Data Testing

Thn	TD	Jenis Obat	Nama Obat	Tr	Pr	PO	Class Prediction	Lk	Tdk
2018	Mar-18	Nyeri	Asam Mefenamat Tab	Hampir	500	Laku	LAKU	3,40%	0,20%
2019	Apr-19	Infeksi	Amoxicilin Tab	Hampir	1300	Laku	LAKU	1,80%	0,20%
2020	Nov-20	Demam	Novagesic Tab	Hampir	700	Laku	LAKU	3,70%	0,20%
2019	Jun-19	Infeksi	Clindamicyn	Sedikit	180	Tidak	TIDAK	0,00%	5,30%
2020	Jan-20	Nyeri	Neuralgin Tab	Sedikit	400	Tidak	TIDAK	0,00%	5,10%

Sumber : data penelitian 2022

Setelah data testing sudah dicari dan didapatkan nilainya maka proses selanjutnya adalah mencari nilai acurasi. Berikut nilai *accuracy* yang di dapat dari tabel diatas.

Tabel 4. 11 Class Prediksi

Class		
Prediksi	Laku	Tidak
Laku	3	1
Tidak	0	1

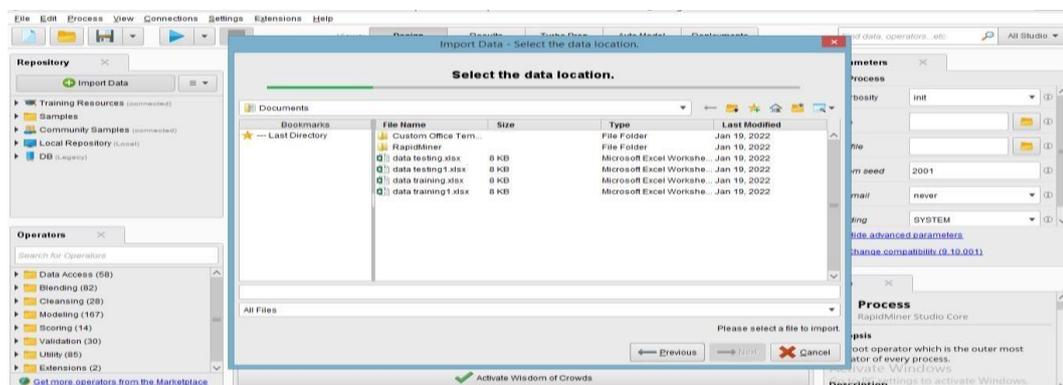
Sumber: Data Penelitian 2022

Accuracy= 80%

4.4. Implementasi Rapidminer

1. Import data

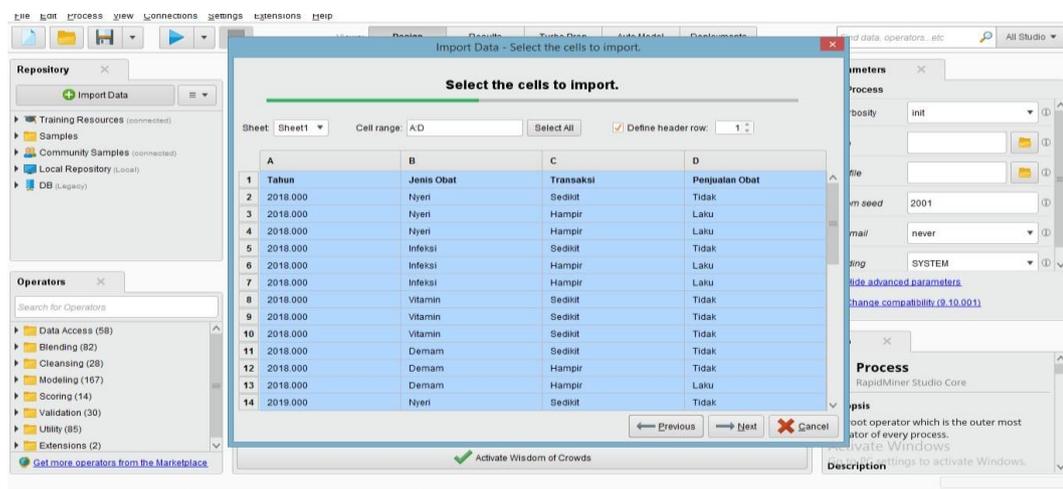
Data yang akan di ambil sebagai data training pada kasus ini yaitu prediksi penjualan obat yang dimulai pada tahun 2018,2019 sampai tahun 2020 dan dengan menjadikan 5 data *sample* sebagai data *testing*. Tidak hanya menentukan data *training* dan data *testing* perlu di ketahui bahwa penting nya atribut yang akan berperan sebagai label, oleh karena itu maka penjualan obat merupakan label pada penelitian ini. Proses implementasi dalam menentukan laku atau tidak lakunya penjualan obat akan dilakukan perhitungan menggunakan *microsoft excel* dan menggunakan perngkat lunak *rapid minner* dalam pengujian selanjutnya agar hasil yang di peroleh nanti nya lebih akurat dan sesuai dengan keinginan. Data *training* yang telah di dapatkan dari klinik Harapan Kita Batam akan di import terlebih dahulu kedalam *software rapid minner* dengan format *excel* agar data yang ada dapat di olah. Begitu juga demikian dengan data *testing* yang akan di jadikan sebagai *sample* penelitian pada kasus ini. Berikut gambar data training dan data testing yang telah di *import* ke *software rapid minner*.



Gambar 4. 1 *Import Data Training*
Sumber : Data Penelitian 2022

2. Data import step 3

Setelah data *training* dan data *testing* di *import* pada *software* maka langkah selanjutnya yang akan di lakukan adalah dengan mengolah data training terlebih dahulu dengan cara pilih data tarining lalu klik next dan akan muncul tampilan *wizard step 2*. Pada *step* ke 2 ini tidak ada perintah untuk mengubah apa pun maka klik *next* untuk melanjutkan ke *step* berikut nya maka akan muncul *wizard* baru pada *step 3* dan akan terlihat seperti gambar di bawah ini:

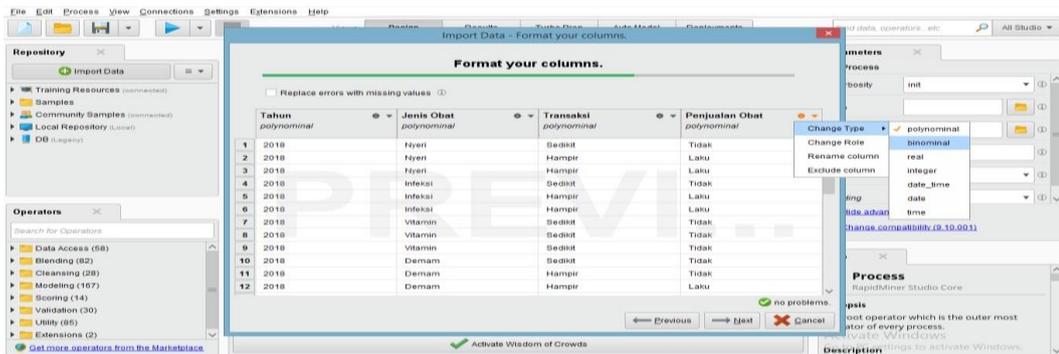


Gambar 4. 2 Data import step 3

Sumber : Data Penelitian 2022

3. Data import step 4

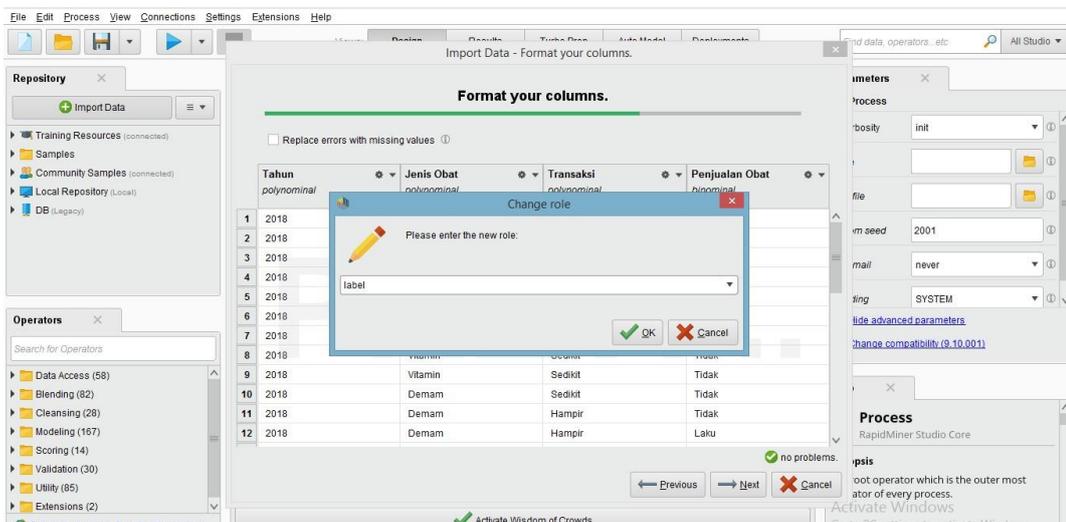
Pada tahap selanjutnya adalah dengan mengklik *next* untuk melanjutkan *step* ke 4. Pada *wizard step 4* ini akan melakukan perubahan tipe data yang sesuai dengan data training yang telah di inport. Pada column penjualan obat yang akan di jadikan sebagai atribut akan di ubah menjadi label dengan klik *change rule*. Perhatikan gambar berikut di bawah ini:



Gambar 4. 3 Data import step 4
Sumber : Data Penelitian 2022

4. Data training step akhir

Setelah ditentukan label yang dipilih maka klik ok dan finish seperti gambar 4.4. Setelah *import data training* selesai maka langkah selanjutnya yaitu melakukan *import wizard data testing* dengan menggunakan langkah-langkah yang sama.

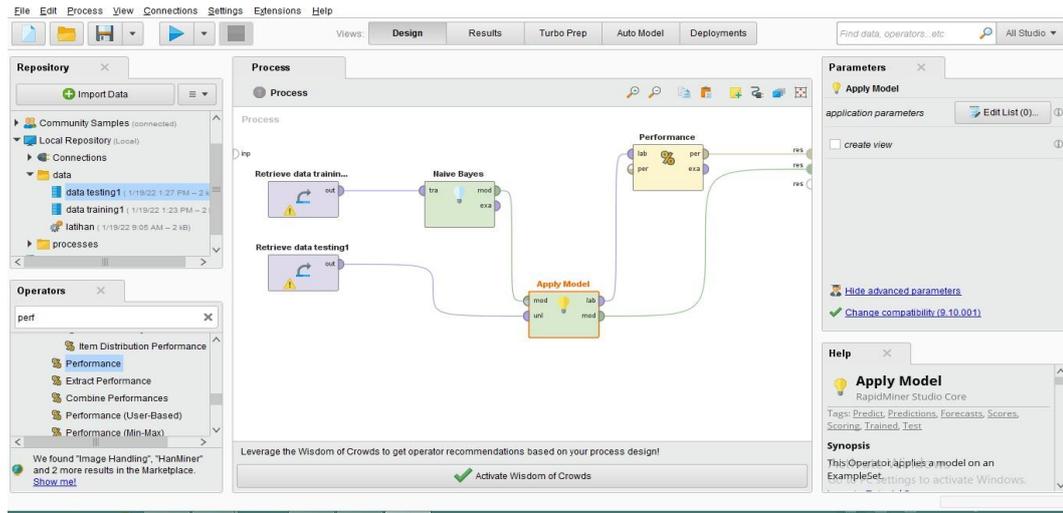


Gambar 4. 4 Data training step akhir
Sumber: Data Penelitian 2022

5. 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*

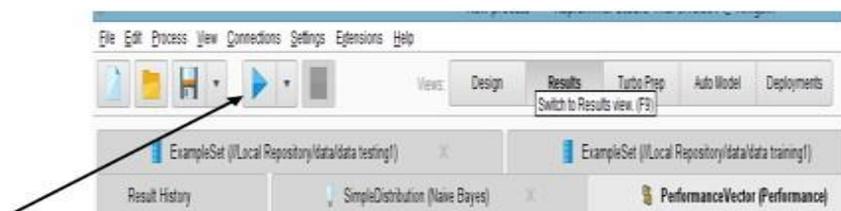
lalu tepatkan pada lembar kerja, setelah itu lakukan hal serupa pada *apply model* dan *performance*. Setelah seluruh operator yang akan di butuhkan ada berada lembar kerja maka selanjutnya langkah yang akan di lakukan adalah dengan menghubungkan setiap operator seperti gambar di bawah ini:



Gambar 4. 5 Menghubungkan *operator*
Sumber : Data Penelitian 2022

6. *Icon Run*

Pastikan semua data terhubung dnegan benar setelah itu klik *icon run* pada toolbar yang telah tersedia, kemudia tunggu beberapa saat sehingga proses perhitungan selesai dilakukan. Perhatikan gambar di bawah ini:



Gambar 4. 6 *Icon Run*
Sumber : Data Penelitian 2022

7. Perhitungan Hasil

Perhitungan yang telah selesai di proses pada perangkat lunak rapid minner dengan menggunakan metode *naive bayes* untuk memprediksi penjualan obat yang akan datang maka akan memberikan informasi mengenai data penjualan obat laku atau 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 Laku	true Tidak	class precision
pred Laku	3	1	75.00%
pred Tidak	0	1	100.00%
class recall	100.00%	50.00%	

Gambar 4. 7 Hasil Perhitungan
Sumber: Data Penelitian 2022

Dapat dilihat pada gambar 4.7 tingkat accuracy dari performanceVector yaitu 80.00%, class precision yaitu laku 75.00%, Tidak 100.00%, dan untuk class recall yaitu Laku 100.00 % sedangkan tidak 50.00%.

Secara Umum precision, recall, dan accuracy dapat dirumuskan sebagai berikut:

➤ Untuk kelas Laku

$$Precision = \left\{ \frac{3}{3 + 1} \right\};$$

$$Precision = 0.75 = 75\%$$

$$Recall = \left\{ \frac{3}{3 + 0} \right\};$$

$$Recall = 1 = 100\%$$

➤ Untuk kelas Tidak

$$Precision = \left\{ \frac{1}{1 + 0} \right\};$$

$$Precision = 1 = 100\%$$

$$Recall = \left\{ \frac{1}{1 + 1} \right\};$$

$$Recall = 0,5 = 50\%$$

➤ Untuk Tingkat Accuracy

$$accuracy = \left\{ \frac{3 + 1}{3 + 0 + 1 + 1} \right\};$$

$$accuracy = \left\{ \frac{4}{5} \right\};$$

$$accuracy = 0,8 = 80\%$$

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Dari hasil penelitian yang telah dilakukan pada penelitian ini, maka peneliti memberi kesimpulan dari hasil penelitian nya sebagai berikut:

1. Pada penelitian ini data yang di peroleh berasal dari Klinik Harapan Kita Batam yang berada di kota Batam di mulai tahun 2018-2020. Dengan menetapkan 36 *record* data sebagai data *set* yang akan di lakukan pengujian, sehingga hasil pengujian yang di lakukan menghasilkan data penjualan yang paling laku dan tidak laku di beli oleh konsumen diketahui bahwa 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.
2. 36 *set* data yang di gunakan sebagai data penelitian maka dapat di tentukan lima data yang akan menjadi data *testing* atau data *sample* untuk di jadikan sebagai data uji. Metode yang di gunakan untuk menguji adalah dengan metode agoritma *naive bayes* dan melakukan proses perhitungan secara manual menggunakan *microsoft excel* dan *software rapid minner* yang menghasilkan nilai peresentasi penjualan obat dengan akurasi 80%.

3. Penggunaan *naive bayes* dalam penyelesaian studi kasus tentang prediksi penjualan obat menarik kesimpulan bahwa obat yang di temui paling banyak laku atau terjual ialah obat nyeri, infeksi, vitamin dan demam. Dari ke empat jenis obat tersebut terdapat kriteria nyeri dengan 27 % penjualan laku dan 24% tidak laku terjual, untuk kriteria infeksi terdapat 36 % laku terjual dan 20 % tidak laku, untuk kriteria vitamin terdapat 0 % laku dan 36 % tidak laku, pada kriteria demam terdapat 36 % penjualan obat terjual dan terdapat 20% tidak laku terjual.

5.2 Saran

Penelitian ini tidak lepas dari batasan yang ada, berikut adalah beberapa saran yang di berikan untuk pengembangan penelitian selanjutnya yang diharapkan mampu mengembangkan penelitian ini yaitu:

1. Pengujian lebih baik dilakukan menggunakan metode lain untuk menentukan metode yang lebih akurat untuk memprediksi penjualan obat.
2. Jenis obat yang diambil peneliti pada klinik harapan kita batam belum terbilang lengkap karena hanya mengambil 4 jenis obat yaitu nyeri, infeksi, demam dan vitamin oleh sebab itu untuk penelitian selanjutnya diharapkan dapat menambah jenis obat, sehingga hasil yang didapat lebih baik.
3. Untuk penelitian selanjutnya berdasarkan analisis pengembangan perangkat atau program aplikasi sangat diperlukan agar dapat digunakan lebih mudah seperti dapat dikembangkan menjadi sebuah *web mining*.