

PENERAPAN DATA MINING MEMPREDIKSI PENJUALAN OBAT MENGUNAKAN METODE K-NEAREST NEIGHBOR (STUDI KASUS : APOTEK DIFANA)

RaECKy Meyzal Febrialdo¹⁾, Yumai Wendra²⁾

^{1,2}Program Studi Teknik Informatika, Fakultas Teknik, Universitas Nahdlatul Sumatera Barat

Corresponding Author: ¹raeckymeyzalf@gmail.com

Article Info

Article history:

Received: Oct 13, 2025

Revised: Oct 15, 2025

Accepted: Oct 17, 2025

Published: Oct 30, 2025

Keywords:

K-Nearest Neighbor (KNN);
Inventory Management;
Data Mining Pharmacy;
Inventory Management;
Web-Based System.

ABSTRACT

The development of information technology in the digital era has provided opportunities for various sectors, including the pharmaceutical industry, to improve operational efficiency and service quality. One of the main challenges faced by pharmacies is inventory management, where both stockouts and overstocks often occur due to the limitations of conventional prediction methods. This study aims to apply data mining techniques using the K-Nearest Neighbor (KNN) algorithm to predict drug sales at Apotik Difana. KNN is chosen because it is simple yet effective in recognizing sales patterns from historical data. A web-based prediction system was developed to facilitate accessibility and usability for the pharmacy. The scope of this study focuses on historical sales data without considering external factors, and only KNN is used without comparison to other algorithms. The results are expected to assist the pharmacy in determining the right type and quantity of drugs, optimizing inventory management, reducing losses from expired drugs, and improving customer service quality. Furthermore, this research provides a theoretical contribution to the development of data mining in sales prediction and offers a practical, technology-based solution for pharmaceutical inventory management.



This is an open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International (CC BY SA 4.0)

1. PENDAHULUAN

Pada era digital, perkembangan teknologi informasi membuka peluang bagi berbagai sektor bisnis untuk meningkatkan efisiensi operasional dan kualitas layanan. Salah satu penerapannya adalah pada bidang farmasi, khususnya dalam prediksi penjualan obat. Prediksi merupakan proses untuk memperkecil kesalahan, sehingga dapat meminimalisir antara kejadian yang sebenarnya. Prediksi penjualan yang akurat dapat membantu apotek dalam menentukan jumlah stok yang optimal, sehingga mengurangi risiko kekurangan atau kelebihan stok yang berdampak pada biaya operasional.

Prediksi adalah proses untuk meramalkan suatu variable di masa mendatang dengan berdasarkan pertimbangan data pada masa lampau. Data yang sering digunakan untuk melakukan prediksi adalah data yang berupa data kuantitatif. Prediksi tidak harus

memberikan jawaban secara pasti kejadian yang akan terjadi, melainkan berusaha untuk mencari jawaban sedekat mungkin yang akan terjadi. Istilah prediksi sama dengan ramalan atau perkiraan (forecast)[1]. Prediksi dapat menentukan kebutuhan dalam kurun waktu minggu, bulan, bahkan tahun hal ini berkaitan dengan dukungan jumlah data atau prediksi waktu yang dianalisis sehingga dapat diperhitungkan dalam memprediksi data pada tahun mendatang[2].

Dalam era globalisasi dan digitalisasi, teknologi informasi telah menjadi bagian penting dalam berbagai sektor, termasuk sektor Kesehatan. Salah satu bidang yang sangat dipengaruhi oleh perkembangan teknologi adalah industri farmasi. Apotik sebagai bagian dari industri farmasi kini tidak hanya berfungsi sebagai tempat penjualan obat, tetapi juga harus mampu mengelola dan mengatur stok obat dengan tepat, mengingat peran pentingnya dalam memastikan ketersediaan obat bagi Masyarakat.

Pengelolaan stok yang tidak optimal dapat mengakibatkan beberapa masalah, seperti kekurangan stok (stockout) atau kelebihan stok (overstock), yang pada akhirnya berdampak pada kualitas pelayanan dan efisiensi operasional apotek. Permasalahan tersebut umumnya terjadi karena pengelolaan data penjualan obat belum dilakukan secara optimal. Sebagian besar apotek masih melakukan pencatatan dan analisis secara manual, sehingga sulit untuk memprediksi jumlah kebutuhan obat di masa mendatang. Akibatnya, pengambilan keputusan terkait pengadaan obat seringkali tidak tepat sasaran. Apotik adalah suatu tempat tertentu, tempat dilakukannya pekerjaan kefarmasian dan penyaluran sediaan farmasi. Perbekalan kesehatan lainnya kepada masyarakat[3].

Apotik Difana sebagai salah satu apotek lokal untuk menghadapi tantangan dalam memprediksi kebutuhan stok obat harus dilakukan secara tepat, jika stok obat berlebih akan menyebabkan obat tersebut menumpuk hal ini bisa menyebabkan kerugian bagi apotik dikarenakan obat tersebut memiliki masa kadaluarsanya. Penerapan metode data mining menjadi penting. Data mining adalah proses ekstraksi informasi berharga dari data yang tersedia dan dapat digunakan untuk membuat prediksi. Salah satu metode yang umum dalam prediksi adalah K-Nearest Neighbor (KNN), yang sederhana namun efektif dalam menangkap pola dari data historis, menghadapi tantangan dalam memprediksi jumlah permintaan obat dari waktu ke waktu. Permintaan obat seringkali bersifat fluktuatif dan di pengaruhi oleh banyak faktor, seperti musim, promosi, perubahan tren Kesehatan, dan kebutuhan pasien. Dalam situasi ini, metode konvensional untuk mengelolah persediaan seringkali kurang efektif, karena tidak mampu menangkap dan memanfaatkan data historis secara optimal. Oleh karena itu, diperlukan pendekatan yang lebih canggih dan berbasis data untuk membantu apotik dalam membuat Keputusan terkait manajemen stok obat.

Data mining suatu istilah yang digunakan untuk menemukan pengetahuan yang tersembunyi di dalam database[4]. Data mining, sebagai bagian dari teknologi kecerdasan buatan, menawarkan berbagai metode analisis data yang mampu menggali informasi berharga dari data yang ada. Salah satu Teknik yang dapat digunakan untuk memprediksi penjualan adalah metode K-Nearest Neighbor (KNN) suatu metode klasifikasi dan prediksi yang bekerja berdasarkan kesamaan atau kedekatan data baru dengan data yang sudah ada. Metode ini sederhana namun efektif, terutama Ketika di terapkan pada data penjualan yang memiliki pola tertentu.

Penerapan metode K-Nearest Neighbor pada data penjualan obat di apotek Difana diharapkan dapat memberikan Solusi yang tepat untuk memprediksi kebutuhan stok di masa mendatang. Dengan memanfaatkan data penjualan historis, KNN dapat membantu apotik dalam menentukan jenis obat yang kemungkinan besar akan laku terjual dan dalam

jumlah berapa. Ini tidak hanya membantu mengoptimalkan pengelolaan stok, tetapi juga dapat meningkatkan kepuasan pelanggan karena obat-obatan yang dibutuhkan tersedia tepat waktu. K-Nearest Neighbors (K-NN) adalah metode klasifikasi yang menggunakan data pelatihan untuk menentukan kelas objek berdasarkan jarak terdekat. Proses ini melibatkan transformasi data pelatihan ke ruang multidimensi, dimana setiap dimensi mewakili karakteristik data. Algoritma K-NN relatif sederhana dan bekerja dengan menghitung jarak terdekat antara objek query dan sample pelatihan untuk menentukan jumlah tetangga yang optimal[5].

Selain itu, memprediksi penjualan obat juga dapat memberikan keuntungan finansial bagi apotik. Dengan stok yang di kelolah secara efisien, apotik dapat mengurangi pemborosan akibat obat-obatan yang tidak terjual atau mendekati kadaluarsa. Di sisi lain, risiko kehilangan penjualan karena kekurangan obat juga dapat diminimalkan. Oleh karena itu, penelitian ini di fokuskan pada penerapan metode K-Nearest Neighbor untuk memprediksi penjualan obat di Apotik Difana. Penelitian ini diharapkan dapat memberikan strategi manajemen stok yang lebih efektif dan berbasis data, sehingga Apotik Difana dapat menghadapi tantangan permintaan pasar dengan lebih baik dan lebih efisien. Dengan memanfaatkan teknologi berbasis website, prediksi penjualan dapat di akses dengan mudah dan kapan saja. Dalam penelitian ini, dilakukan pengembangan sistem prediksi penjualan obat di Apotik Difana menggunakan metode KNN berbasis website. Penelitian diharapkan dapat membantu apotek dalam mengambil keputusan stok obat yang lebih baik, sehingga meningkatkan efisiensi dan kualitas layanan kepada pelanggan.

2. LANDASAN TEORI

2.1 Sistem

Sistem adalah unsur-unsur yang bekerja secara berkelompok dan saling berhubungan erat sehingga dapat bekerja sama dalam menjalankan fungsi untuk mencapai tujuan. Secara umum, suatu sistem terdiri dari input, proses, dan output. menyajikan sistem sebagai sekelompok komponen yang saling berhubungan yang dirakit untuk mencapai tujuan tertentu[6]

2.2 Undefined Modeling Language (UML)

Unified Modelin Language (UML) adalah sebuah bahasa pemodelan perangkat lunak yang telah distandardisasi sebagai media penulisan cetak biru (blueprints) perangkat lunak (Pressman). UML bisa saja digunakan untuk visualisasi, spesifikasi, konstruksi dan dokumentasi beberapa bagian-bagian dari sistem yang ada dalam perangkat lunak.[7]

2.3 Konsep Dasar Data Mining

Data mining adalah proses analitis yang dirancang untuk mengeksplorasi sejumlah besar data untuk pengetahuan yang berharga, konsisten, dan tersembunyi. Salah satu metode yang digunakan dalam data mining adalah klasifikasi[8]

2.4 Metode K-Nearest Neighbor

Metode K-NN merupakan metode imputasi yang paling populer untuk menyelesaikan masalah *Missing Value*. Metode ini sederhana dan efektif dalam masalah *imputasi missing value*, karena paling banyak digunakan untuk menyelesaikan banyak masalah prediksi. Keunggulan metode imputasi dengan KNN dapat digunakan untuk meramalkan 2 jenis data yaitu data diskrit (nilai modus) dan data kontinu (nilai rata-rata)[16]

Berikut rumus k-nearest neighbor

$$d(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_n - y_n)^2} \quad (1)$$

Dimana :

$X = (x_1, x_2, \dots, x_n)$ adalah vektor fitur dari data yang diprediksi

$y = (y_1, y_2, \dots, y_n)$ adalah vektor fitur dari salah satu data pelatihan.

2.5 Website

Website adalah Kumpulan halaman-halaman yang digunakan untuk menampilkan informasi teks, gambar, diam atau gerak, animasi, suara, dan atau gabungan dari semuanya baik yang bersifat statis maupun dinamis yang membentuk suatu rangkaian bangunan yang saling terkait, yang masing-masing dihubungkan dengan jaringan-jaringan halaman.[9]

2.6 MySQL

MySQL merupakan DBMS yang memiliki kemampuan manajemen data pengguna, keamanan data, pencadangan dan pemulihan data serta banyak kemampuan lain terkait pemrosesan data. *MySQL* menggunakan Bahasa *SQL* untuk menghubungkan antara database server dengan perangkat lunak pengguna. *SQL* merupakan bahasa komputer standar yang digunakan untuk melakukan *query* manajemen (pengaksesan dan manipulasi) sistem basis data. *SQL* memiliki lima bagian dalam pemrosesan data, yaitu data *retrieving*, data *definition*, data *manipulation*, data *control*, dan data *transaction language*. [10]

2.7 XAMPP

XAMPP adalah perangkat lunak bebas, yang mendukung banyak sistem operasi, merupakan kompilasi dari beberapa program. Fungsinya adalah sebagai server yang berdiri sendiri (*localhost*), yang terdiri atas program *Apache HTTP Server*, *MySQL database*, dan penerjemah bahasa yang ditulis dengan bahasa pemrograman *PHP* dan *Perl*. Nama *XAMPP* merupakan singkatan dari X (empat sistem operasi apapun), *Apache*, *MySQL*, *PHP* dan *Perl*. Program ini tersebut dalam *GNU (General Public License)* dan bebas, merupakan *web server* yang mudah digunakan yang dapat melayani tampilan halaman *web* yang dinamis [11]

2.8 PHP

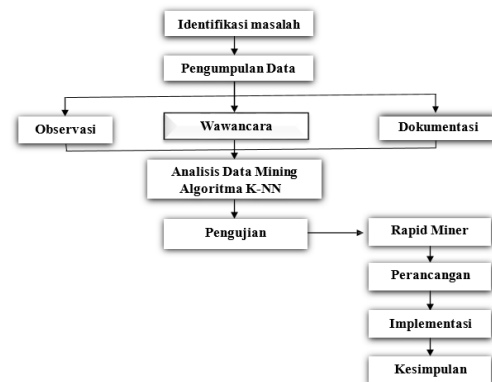
PHP adalah bahasa pemrograman yang biasa digunakan untuk *server-side scripting*. *PHP* merupakan bahasa pemrograman yang simpel namun *powerfull* dan tepat untuk digunakan pada *web server*. *PHP* dapat berjalan pada dapat berjalan pada semua sistem operasi yang sering digunakan, seperti *UNIX*, *Windows*, dan *MacOS*.

2.9 Rapidminer

Rapid Miner merupakan perangkat lunak yang digunakan untuk pengolahan data dan menjadi sebuah solusi untuk melakukan analisis terhadap *data mining*, *text mining* dan analisis prediksi. Aplikasi ini memudahkan *user* melakukan perhitungan data yang sangat banyak dengan menggunakan operator-operator yang berfungsi untuk melakukan modifikasi data. Data dihubungkan melalui node-node pada operator kemudian dihubungkan ke node hasil untuk menampilkan hasil akhir sehingga hasil dapat ditampilkan secara visual dengan grafik [12]

3. METODOLOGI PENELITIAN

kerangka penelitian menunjukkan langkah-langkah yang penulis lakukan pada Penelitian ini. Pada penelitian ini dilakukan penulis dengan berbagai macam langkah proses pada penelitian. Berikut ini adalah kerangka penelitian:



Gambar 1. Kerangka Penelitian

Penjelasan dari tabel diatas adalah sebagai berikut :

a. Identifikasi masalah

Pada penelitian ini, identifikasi masalah dilakukan untuk mengetahui tantangan yang dihadapi oleh Apotek Difana dalam mengelola penjualan obat, serta bagaimana data mining, khususnya metode K-Nearest Neighbor (K-NN), dapat memberikan solusi dalam memprediksi penjualan obat secara lebih akurat

b. Pengumpulan data

Pengumpulan data merupakan tahap awal yang sangat penting dalam penelitian ini, pengumpulan data peneliti memiliki beberapa metode salah satunya observasi, wawancara, serta studi literatur, karena kualitas prediksi yang dihasilkan oleh metode K-Nearest Neighbor sangat bergantung pada kualitas dan kelengkapan data yang digunakan. Data yang digunakan dalam penelitian ini diperoleh langsung

dari Apotik Difana, yang mencatat seluruh transaksi penjualan obat dalam periode tertentu. Proses pengumpulan data dilakukan dengan cara mengambil data historis penjualan obat, yang mencakup informasi tentang jenis obat, jumlah obat yang terjual, tanggal transaksi, harga, serta faktor-faktor yang berpengaruh pada penjualan seperti cuaca, promosi, atau perubahan harga. Data ini dikumpulkan dari sistem manajemen penjualan yang digunakan oleh Apotik Difana. Selain data transaksi penjualan, data tambahan juga dikumpulkan terkait dengan karakteristik produk, seperti kategori obat, dan data terkait lain yang dapat mempengaruhi pola penjualan. Seluruh data yang terkumpul kemudian dilakukan pembersihan (data cleaning) untuk memastikan bahwa data yang digunakan valid, konsisten, dan tidak mengandung kesalahan atau nilai yang hilang (missing values).

Setelah tahap pembersihan, data kemudian disusun dalam bentuk yang dapat dianalisis lebih lanjut menggunakan metode K-Nearest Neighbor, untuk memprediksi penjualan obat di masa mendatang. Dalam penelitian ini, dataset yang digunakan terdiri dari data transaksi penjualan dalam periode 1 tahun terakhir, yang diharapkan dapat memberikan gambaran yang representatif tentang pola penjualan obat di Apotik Difana

c. Analisis data menggunakan metode K-NN

Setelah data yang diperlukan terkumpul dan dilakukan pembersihan, langkah selanjutnya adalah melakukan analisis data menggunakan metode K-Nearest Neighbor (K-NN). Metode K-NN dipilih dalam penelitian ini karena kemampuannya dalam memprediksi variabel kontinu seperti jumlah penjualan obat berdasarkan kemiripan data sebelumnya

d. Perancangan sistem

Perancangan sistem ini bertujuan untuk menciptakan suatu model yang dapat memprediksi penjualan obat di Apotik Difana dengan menggunakan metode K-Nearest Neighbor (K-NN). Sistem yang dirancang akan mengolah data penjualan historis, serta faktor-faktor eksternal yang memengaruhi penjualan, untuk menghasilkan prediksi yang akurat mengenai jumlah penjualan obat pada periode mendatang. Dalam perancangan sistem ini, terdapat beberapa tahapan yang perlu dilalui, mulai dari perancangan arsitektur sistem hingga implementasi algoritma K-NN.

e. Mengimplementasikan dalam bentuk program

Pada tahap implementasi sistem prediksi penjualan obat menggunakan metode K-Nearest Neighbor, pemrograman dilakukan dengan menggunakan bahasa PHP. PHP dipilih karena fleksibilitasnya dalam mengelola data dan membangun aplikasi berbasis web. Dalam implementasi ini, PHP akan digunakan untuk memproses data penjualan obat, menjalankan algoritma K-NN, serta menghasilkan prediksi

penjualan yang dapat digunakan oleh pengelola apotik dalam mengambil keputusan.

f. Pengujian hasil

Pada tahap pengujian hasil, dilakukan evaluasi terhadap model yang telah dibangun dengan menggunakan data uji yang sebelumnya dipisahkan. Pengujian ini bertujuan untuk mengetahui seberapa baik model K-Nearest Neighbor (K-NN) dalam memprediksi penjualan obat di Apotik Difana berdasarkan data historis yang ada. Proses pengujian mencakup perbandingan antara nilai prediksi yang dihasilkan oleh model dengan nilai aktual dari data uji. Dengan demikian, pengujian ini dapat mengukur seberapa akurat prediksi yang diberikan oleh sistem.

g. Kesimpulan

Pada bab 3 ini, telah dijelaskan secara rinci tentang tahap-tahap yang dilakukan dalam penerapan metode data mining untuk memprediksi penjualan obat di Apotik Difana menggunakan algoritma K-Nearest Neighbor (K-NN). Beberapa langkah utama yang telah dilakukan meliputi pengumpulan data, identifikasi masalah, study literatur, pengumpulan data, analisis data menggunakan metode K-NN, perancangan sistem, mengimplementasikan dalam bentuk program, serta pengujian.

4. HASIL DAN PEMBAHASAN

4.1 Analisis sistem

Dalam tahap ini, peneliti menggunakan data transaksi barang sebagai bahan untuk menyelesaikan penelitian yang diambil dari Apotek Difana selama tahun 2024 yang dilakukan dengan teknik wawancara dan dokumentasi.

Data sample item Obat yang digunakan:

Tabel 4. 1 Data Item Obat

No	Nama Obat
1	talegi
2	obh surya
3	komik herbal sachet
4	komik obh anak
5	wasin kapsul
6	Antangin
7	Antangin permen
8	alkohol 70% medika
.....	
267	jesscool
268	lactor B
269	luvisma
270	hotin cream sachet
271	TJ joybee gummy
272	cavit d3

4.2 Pembersihan Data

Tahap awal yang dilakukan dalam proses data mining adalah pembersihan data (data cleaning). Pada tahap ini, data transaksi penjualan obat yang telah

diperoleh dari sumber data dilakukan pemeriksaan dan penyaringan untuk memastikan hanya data yang relevan dan sesuai kebutuhan analisis yang diproses lebih lanjut.

Dalam penelitian ini, data transaksi penjualan obat terdiri dari beberapa atribut, yaitu Nama Obat, Jumlah, Tanggal Penjualan, dan Jenis Obat. Namun, setelah dilakukan evaluasi terhadap kebutuhan analisis, atribut Jenis Obat diputuskan untuk dihapus karena tidak digunakan dalam proses analisis data mining. Sehingga, atribut yang digunakan dalam penelitian ini hanya terdiri dari Nama Obat, Jumlah, dan Tanggal Penjualan. Selain itu, proses pembersihan data juga mencakup pemilihan data berdasarkan sampel yang telah ditentukan. Data yang digunakan mencakup total sebanyak 3.600 lebih entri transaksi penjualan obat dengan rentang tanggal penjualan dari bulan Januari hingga Desember 2024.

Tabel 4. 2 Data Penjualan

No	Nama Obat	Jumlah	Tanggal Penjualan
1	transpulmin anak	19	2024-12-31
2	becefort	18	2024-12-31
3	OB herbal 100ml	11	2024-12-31
4	balsem aktif 20gr	8	2024-12-31
5	balsem otot geliga 10gr	17	2024-12-31
6	madu rasa	8	2024-12-31
7	amlodipine 10mg	17	2024-12-31
8	rohto cool	13	2024-12-31
9	tolak angin	9	2024-12-31
.....			
3597	Antangin	19	2024-01-01
3598	rapet wangi	12	2024-01-01
3599	paramex	13	2024-01-01
3600	IPI vitamin B kompleks	15	2024-01-01

4.3 Transformasi Data

Transformasi data dilakukan untuk menyesuaikan data penjualan obat agar dapat digunakan oleh algoritma *K-Nearest Neighbors Regressor* (KNNR) dalam memprediksi jumlah penjualan. Data awal yang terdiri dari id_obat, jumlah, dan tanggal, diolah dengan mengekstraksi dua fitur baru, yaitu hari (0–6 untuk Senin–Minggu) dan bulan (1–12), sehingga membentuk data dengan struktur yang sesuai untuk proses prediksi berbasis waktu.

Tabel 4. 3 Contoh Data Yang Sudah Di Transformasi

id_obat	nama_obat	jumlah	tanggal	hari	bulan
1	paracetamol	8	2024-11-23	5	11
2	cetirizine	12	2024-05-20	0	5
3	aspirin	8	2024-03-12	1	3
4	paracetamol	18	2024-02-14	2	2
5	cetirizine	7	2024-01-16	1	1
6	aspirin	5	2024-04-29	0	4
7	cetirizine	18	2024-04-11	3	4
8	paracetamol	13	2024-08-17	5	8
9	loratadine	18	2024-03-22	4	3
10	cetirizine	9	2024-05-22	2	5

4.4 Penerapan Algoritma KNN

a. Menentukan Parameter K (jumlah tetangga terdekat)

Nilai K pada algoritma KNN menunjukkan jumlah tetangga terdekat yang digunakan untuk prediksi. Jika K terlalu kecil, model menjadi sensitif terhadap noise. Sebaliknya, jika terlalu besar, hasil prediksi bisa terlalu umum. Oleh karena itu, pemilihan nilai K yang tepat sangat penting untuk akurasi model. Pada penelitian ini akan menggunakan 3 sebagai nilai K nya ($K=3$).

b. Menentukan Tetangga Terdekat (*K Neighbors*)

Menentukan tetangga terdekat (*K Neighbors*) dilakukan dengan menghitung jarak antara data yang ingin diprediksi dan seluruh data dalam dataset. Kemudian, data diurutkan berdasarkan jarak terkecil, dan diambil sebanyak K data terdekat. Tetangga inilah yang digunakan sebagai dasar untuk melakukan prediksi dalam algoritma KNN. Berikut contoh dengan prediksi pada salah satu obat yaitu talegin dengan data penjualan sebagai berikut :

Tabel 4. 4 Data Penjualan Obat talegin

NO	NAMA OBAT	JUMLAH	TANGGAL
1	talegi	11	2024-12-19
2	talegi	6	2024-10-23
3	talegi	6	2024-10-02
4	talegi	18	2024-09-23
5	talegi	8	2024-09-19
6	talegi	14	2024-07-23
7	talegi	12	2024-05-27
8	talegi	15	2024-05-07
9	talegi	16	2024-05-06
10	talegi	10	2024-02-29
11	talegi	16	2024-02-01
12	talegi	10	2024-02-01
13	talegi	6	2024-01-20
14	talegi	15	2024-01-10
15	talegi	6	2024-01-02

Sebelum data di gunakan data akan di Transformasi dengan mengekstraksi dua fitur baru, yaitu hari (0–6 untuk Senin–Minggu) dan bulan (1–12), sehingga membentuk data dengan struktur yang sesuai untuk proses prediksi berbasis waktu. Yang bentuk teranformsinya seperti berikut ini :

Tabel 4. 5 Data penjualan yang sudah di tranformasi

Tanggal	Hari	Bulan	Jumlah
2024-12-19	4	12	11
2024-10-23	3	10	6
2024-10-02	3	10	6
2024-09-23	1	9	18
2024-09-19	4	9	8
2024-07-23	2	7	14
2024-05-27	1	5	12
2024-05-07	2	5	15

2024-05-07	2	5	15
2024-05-06	1	5	16
2024-02-29	4	2	10
2024-02-01	4	2	10
2024-02-01	4	2	16
2024-01-20	6	1	6
2024-01-10	3	1	15
2024-01-02	2	1	6

Setelah data diolah dan disusun sebagaimana ditampilkan pada Tabel 4.5, data tersebut siap digunakan untuk melakukan prediksi penjualan di masa mendatang.

c. Menentukan Jarak Euclidean

Salah satu penerapan prediksi ini dilakukan pada tanggal 1 Agustus 2025 dengan menggunakan algoritma *K-Nearest Neighbors* (KNN) dan parameter $K = 3$. Pada tahap awal, dilakukan proses ekstraksi fitur dari tanggal prediksi. Tanggal 2025-08-01 jatuh pada hari Jumat, yang dalam *representasi* fungsi *weekday()* bernilai 4 (dengan ketentuan 0 = Senin, 6 = Minggu), dan berada pada bulan Agustus, yang direpresentasikan sebagai 8. Maka, fitur yang digunakan untuk prediksi adalah [hari = 4, bulan = 8]. Nilai fitur ini akan dibandingkan dengan data historis untuk menentukan tiga tetangga terdekat yang digunakan sebagai dasar perhitungan estimasi jumlah penjualan obat, dalam hal ini untuk obat *Talegi*. Berikut adalah penentuan k3 atau 3 tetangan terdekat: rumus jarak Euclidean

$$\text{Jarak} = \sqrt{(h_i - h_t)^2 + (b_i - b_t)^2}$$

Keterangan:

h_i = hari ke-i dari data historis

h_t = hari dari tanggal target

b_i = bulan dari data historis

b_t = bulan dari tanggal target

Karena untuk prediksi 1 Agustus 2025 sudah didapatkan nilai fitur nya maka akan digunakan rumus jarak Euclidean antara data historis (hari history ,bulan history) dengan target (4,8):

$$\text{Jarak} = \sqrt{(\text{hari}_i - 4)^2 + (\text{bulan}_i - 8)^2}$$

Tabel 4. 6 Menghitung Jarak Euclidean

Tabel 4. 7 Menghitung Jarak Euclidean

Hari (h_i)	Bulan (b_i)	Jumlah	Rumus Perhitungan	Jarak
3	12	11	$\sqrt{((3-4)^2 + (12-8)^2)} = \sqrt{(1^2 + 4^2)} = \sqrt{17}$	4.12
2	10	6	$\sqrt{((2-4)^2 + (10-8)^2)} = \sqrt{(2^2 + 2^2)} = \sqrt{8}$	2.83
2	10	6	$\sqrt{((2-4)^2 + (10-8)^2)} = \sqrt{8}$	2.83
0	9	18	$\sqrt{((0-4)^2 + (9-8)^2)} = \sqrt{(4^2 + 1^2)} = \sqrt{17}$	4.12
3	9	8	$\sqrt{((3-4)^2 + (9-8)^2)} = \sqrt{(1^2 + 1^2)} = \sqrt{2}$	1.41
1	7	14	$\sqrt{((1-4)^2 + (7-8)^2)} = \sqrt{(3^2 + 1^2)} = \sqrt{10}$	3.16
0	5	12	$\sqrt{((0-4)^2 + (5-8)^2)} = \sqrt{(4^2 + 3^2)} = \sqrt{25}$	5.00

0	5	12	$\sqrt{((0-4)^2 + (5-8)^2)} = \sqrt{(4^2 + 3^2)} = \sqrt{25}$	5.00
1	5	15	$\sqrt{((1-4)^2 + (5-8)^2)} = \sqrt{(3^2 + 3^2)} = \sqrt{18}$	4.24
0	5	16	$\sqrt{((0-4)^2 + (5-8)^2)} = \sqrt{25}$	5.00
3	2	10	$\sqrt{((3-4)^2 + (2-8)^2)} = \sqrt{(1^2 + 6^2)} = \sqrt{37}$	6.08
3	2	10	$\sqrt{((3-4)^2 + (2-8)^2)} = \sqrt{(1^2 + 6^2)} = \sqrt{(1+36)} = \sqrt{37}$	6.08
3	2	16	$\sqrt{((3-4)^2 + (2-8)^2)} = \sqrt{(1^2 + 6^2)} = \sqrt{(1+36)} = \sqrt{37}$	6.08
5	1	6	$\sqrt{((5-4)^2 + (1-8)^2)} = \sqrt{(1^2 + 7^2)} = \sqrt{50}$	7.07
2	1	15	$\sqrt{((2-4)^2 + (1-8)^2)} = \sqrt{(2^2 + 7^2)} = \sqrt{53}$	7.28
1	1	6	$\sqrt{((1-4)^2 + (1-8)^2)} = \sqrt{(3^2 + 7^2)} = \sqrt{58}$	7.62

d. Mengambila k = 3 Data Terdekat

Maka untuk 3 data terdekat yang bisa diambil yaitu seperti ini :

Hari (h_i)	Bulan (b_i)	Jumlah	Rumus Perhitungan	Jarak
2	10	6	$\sqrt{((2-4)^2 + (10-8)^2)} = \sqrt{(2^2 + 2^2)} = \sqrt{8}$	2.83
2	10	6	$\sqrt{((2-4)^2 + (10-8)^2)} = \sqrt{8}$	2.83
3	9	8	$\sqrt{((3-4)^2 + (9-8)^2)} = \sqrt{(1^2 + 1^2)} = \sqrt{2}$	1.41

Tabel 4. 8 K 3 Tetanga Terdekat

Tiga tetangga yang memiliki jarak terdekat inilah yang kemudian digunakan sebagai dasar dalam menentukan estimasi jumlah penjualan. Nilai-nilai penjualan dari ketiga tetangga tersebut dirata-ratakan untuk menghasilkan nilai prediksi akhir.

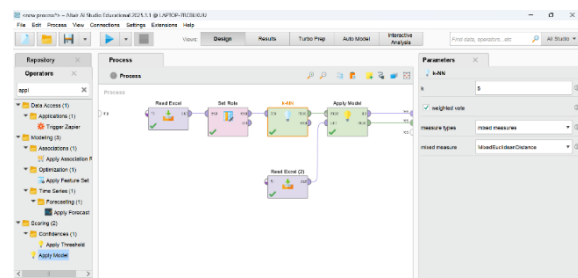
Untuk prediksi, seperti jumlah penjualan, algoritma KNN menggunakan rumus rata-rata dari nilai target tetangga terdekat:

$$\hat{y} = \frac{1}{K} \sum_{i=1}^K y_i$$

Hitung Rata-Rata Jumlah Penjualan :

$$\hat{y} = \frac{1}{3}(8 + 6 + 6) = \frac{20}{3} \approx 6,67 \approx 7 \text{ unit}$$

Dengan demikian, hasil prediksi jumlah penjualan untuk obat Talegi pada tanggal 1 Agustus 2025 adalah 7 unit. Nilai ini kemudian dapat digunakan sebagai dasar untuk perencanaan stok dan kebutuhan logistik, termasuk dalam menghitung stok awal optimal.



e. Menghitung Safety Stock

Safety Stock (stok pengaman) dibutuhkan untuk mengantisipasi ketidakpastian dalam proses penjualan, pembelian, atau distribusi barang. Tujuannya adalah agar tidak kehabisan stok ketika permintaan meningkat tiba-tiba atau terjadi keterlambatan pasokan.

Pada penelitian ini akan digunakan nilai 30% untuk safety stock karena nilai tersebut dianggap sebagai estimasi konservatif yang umum digunakan dalam praktik manajemen persediaan. Persentase ini cukup untuk mengantisipasi fluktuasi permintaan dan potensi keterlambatan pasokan, tanpa menyebabkan kelebihan stok yang berlebihan. Dengan demikian, tingkat 30% dipilih agar tetap efisien namun aman dalam menjaga ketersediaan barang

$$\text{Safety Stock} = 30\% \times 7 = 2.1 \approx 2$$

$$\text{Stok Awal Optimal} = 7 + 2 = 9 \text{ unit}$$

Maka untuk hasil akhirnya kana seperti ini:

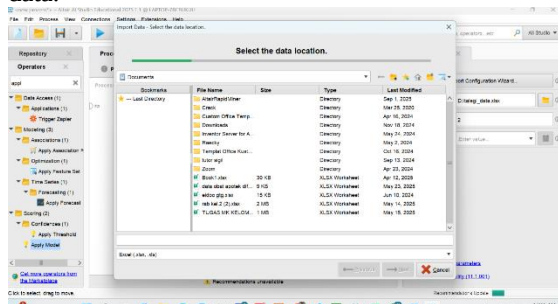
Tabel 4. 9 Hasil Prediksi

Tanggal	Obat	Jumlah Prediksi	Safety Stock	Stok Awal Optimal
2025-08-01	talegi	7	2	9

5. IMPLEMENTASI

a. Import data transaksi

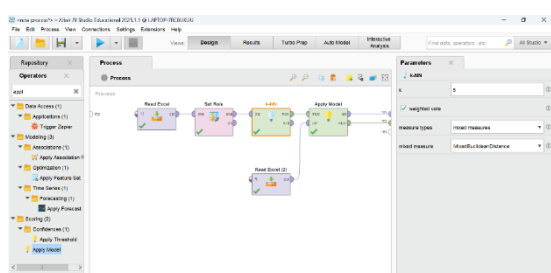
Untuk memulai proses implementasi, data yang sudah dirubah menjadi tabular kemudian dimasukkan ke *software* Rapidminer menggunakan fungsi import data.



Gambar 4.1 Import data transaksi

b. Proses algoritma KNN

Pada bagian *design* di *software* Rapidminer merupakan tempat dimana modul-modul pengolahan data bekerja



Gambar 4.2 Proses algoritma KNN

c. Hasil Prediksi

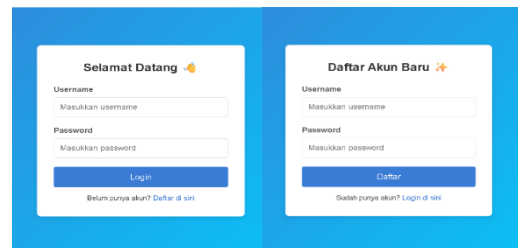
Berdasarkan hasil pengujian yang dilakukan dengan aplikasi RapidMiner terdapat 6 peredesi, berikut hasil pengujian menggunakan aplikasi RapidMiner

Row No.	prediction	Mari (h)	Bulan (h)	Jumlah
1	10.478	1	9	0
2	9.984	2	9	0
3	9.807	3	9	0
4	8.797	4	9	0
5	8.869	5	9	0
6	8.908	6	9	0

Gambar 4.3 Hasil prediksi

d. Antar muka sistem

Antarmuka sistem merupakan titik interaksi pengguna dan aplikasi atau layanan yang diakses melalui *browser Web*. Antarmuka ini dibangun menggunakan teknologi seperti *HTML*, *CSS* dan *JavaScript* Serta *Python* untuk algoritma KNN, yang memungkinkan tampilan yang responsif dan interaktif Serta Perediksi Yang Akurat. Berikut tampilan antarmuka system ini:



Gambar 4.17 Halaman Login Dan Daftar Akun

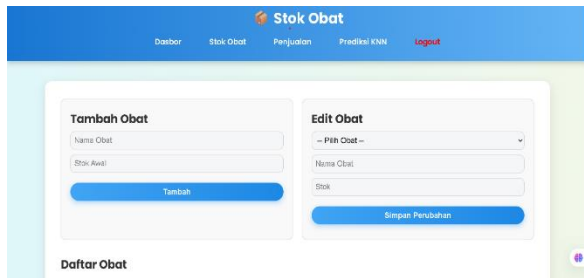
pada Gambar 4.17 adalah gambaran login yang bernuansa biru dengan gradasi biru mudah dengan tampilan sederhana dan segar di pandang mata



Gambar 4.18 Halaman Dasbor

Halaman ini menampilkan navbar biru dengan menu navigasi utama dan tombol Logout merah. Di bawahnya, terdapat kotak konten putih berisi sambutan dan penjelasan singkat tentang algoritma

KNN dalam 4 langkah. Footer biru di bagian bawah menampilkan teks © 2025 Sistem Prediksi KNN Obat.



Gambar 4.19 Halaman Stok Obat

Halaman pada Gambar 4.19 menampilkan fitur manajemen Stok Obat dengan tampilan rapi. Terdapat navbar biru dengan menu utama dan tombol *Logout* merah. Dua formulir utama disediakan: "Tambah Obat" untuk input nama dan stok awal, serta "Edit Obat" untuk mengubah data melalui *dropdown* dan *input*. Daftar obat ditampilkan di bawah untuk memudahkan pengelolaan stok secara efisien

6. KESIMPULAN

Berdasarkan hasil perancangan, implementasi, dan pengujian sistem yang telah dilakukan, dapat disimpulkan beberapa hal sebagai berikut:

Sistem prediksi penjualan obat telah berhasil dibangun menggunakan algoritma K-Nearest Neighbor (KNN) dengan berbagai fungsionalitas utama, seperti manajemen data obat, pencatatan penjualan, serta prediksi penjualan berdasarkan data historis. Penerapan algoritma KNN dalam sistem ini mampu menghasilkan prediksi yang cukup akurat dalam memperkirakan jumlah penjualan obat pada periode tertentu. Kemampuan algoritma ini dalam mengelompokkan data berdasarkan kemiripan pola penjualan sebelumnya menjadikannya efektif dalam mendukung proses pengambilan keputusan, khususnya dalam menentukan kebutuhan pengadaan dan pengelolaan stok obat secara optimal.

Selain itu, sistem dirancang dengan antarmuka yang sederhana dan terstruktur untuk memastikan kemudahan dalam penggunaan. Setiap fitur, seperti penambahan data obat, pencatatan transaksi penjualan, hingga prediksi penjualan, dapat diakses dengan mudah oleh pengguna. Dengan adanya sistem ini, proses pengelolaan data penjualan menjadi lebih efisien dan terukur, serta mampu meningkatkan ketepatan dalam perencanaan stok obat. Secara keseluruhan, sistem ini memberikan kontribusi nyata terhadap efisiensi operasional dan pengambilan keputusan yang berbasis data aktual.

REFERENCES

- [1] Siswo Adiguno, Yohanni Syahra, and Milfa Yetri, "Prediksi Peningkatan Omset Penjualan Menggunakan Metode Regresi Linier Berganda," Jun. 2022.
- [2] M. Faisal, W. S. Utami, and S. Parmica, "BY-SA) Implementasi Algoritma K-Nearest Neighbor (KNN) Dalam Memprediksi Indeks Kemiskinan".
- [3] R. K. Rizal Dzulkarnaen1, "APLIKASI PENGELOLAAN DATA PASIEN DI APOTEK BERBASIS ANDROID," Oct. 2019.
- [4] A. Nugraha, O. Nurdiawan, and G. Dwilestari, "PENERAPAN DATA MINING METODE K-MEANS CLUSTERING UNTUK ANALISA PENJUALAN PADA TOKO YANA SPORT," 2022.
- [5] A. Naufal Hilmi et al., "Implementasi Algoritma K-Nearest Neighbor (KNN) untuk Identifikasi Penyakit pada Tanaman Jeruk Berdasarkan Citra Daun," no. 2, pp. 107–117, 2024, doi: 10.62951/router.v2i2.78.
- [6] I. Hussain Alamyar and E. Nurmiati, "PENGARUH KECANGGIHAN TEKNOLOGI INFORMASI TERHADAP MANAJEMEN PENGETAHUAN," Jurnal Teknologi dan Sistem Informasi (JTSI), vol. 3, no. 1, pp. 64–70, 2022, [Online]. Available: <http://jim.teknokrat.ac.id/index.php/JTSI>
- [7] M. Sumiati, R. Abdillah, and A. Cahyo, "Pemodelan UML untuk Sistem Informasi Persewaan Alat Pesta".
- [8] B. I. Nugroho, Z. Ma'arif, and Z. Arif, "Tinjauan Pustaka Sistematis: Penerapan Data Mining Metode Klasifikasi Untuk Menganalisa Penyalahgunaan Sosial Media," 2022. [Online]. Available: www.jurnal.peradaban.ac.id
- [9] W. Sudrajat and I. Cholid, "K-NEAREST NEIGHBOR (K-NN) UNTUK PENANGANAN MISSING VALUE PADA DATA UMKM," 2023.
- [10] A. Suheri, S. Widaningsih, and H. Refiyana, "Sistem Informasi Pariwisata Berbasis Website Studi Kasus Sindangbarang Cianjur Selatan," Jurnal Interkom: Jurnal Publikasi Ilmiah Bidang Teknologi Informasi dan Komunikasi, vol. 17, no. 4, pp. 175–184, Jan. 2023, doi: 10.35969/interkom.v17i4.278.
- [11] A. Ni'amillah, A. Arif Alfin, and I. Kurniasari, "Siklus Hidup Pengembangan Sistem Basis Data Pada Sistem Informasi Buku Tamu di Badan Pusat Statistik Kabupaten Kediri Menggunakan MySQL," Jurnal Nasional Komputasi dan Teknologi Informasi, vol. 6, no. 1, 2023.
- [12] Baiq Ayu Indah Wardaningsih, Amri Muliawan Nur, and Imam Fathurrahman, "Aplikasi Penyedia Lowongan Kerja Di Wilayah NTB Berbasis Web Menggunakan Framework Laravel Dan Mysql," Jan. 2022.