

## Implementasi Data Mining Regresi Linear Berganda Pada Sistem Prediksi Penjualan Obat Pada Apotek XYZ

Fredy Suwanto<sup>1</sup>, Iis Pradesan<sup>2</sup>

<sup>1,2</sup> Sistem Informasi, Universitas Multi Data Palembang

Email: <sup>1</sup>fredy.suwanto56@gmail.com, <sup>2</sup>iis@mdp.ac.id,  
Email Penulis Korespondensi: [fredy.suwanto56@gmail.com](mailto:fredy.suwanto56@gmail.com)

### Abstrak

Persaingan di dunia bisnis penjualan seperti apotek sekarang menuntut para stakeholder apotek untuk menemukan strategi pengelolaan stok maupun pemasaran untuk memprediksi pangsa pasar agar dapat meningkatkan penjualan obat dan memberikan profitabilitas bagi apotek XYZ. Apotek XYZ belum menerapkan analisis secara manual maupun menggunakan sistem untuk memprediksi penjualan obat, sehingga sering mengalami kekosongan maupun penumpukan stok. Untuk menganalisis diperlukan suatu metode atau algoritma data mining yang mampu mengolah transaksi penjualan obat menjadi suatu informasi. Penelitian ini bertujuan mengembangkan sistem informasi prediksi penjualan obat di apotek XYZ yang nantinya diimplementasikan atau diintegrasikan sebuah model prediksi Regresi Linear Berganda yang diimplementasikan dalam sistem berbasis website, dengan mengikuti tahapan CRISP-DM untuk membantu apotek XYZ dalam memprediksi penjualan obat sehingga dapat mencegah kehabisan atau kelebihan stok dan dapat mendukung penjualan pada apotek. Variabel harga dan curah hujan digunakan sebagai prediktor terhadap jumlah penjualan. Berdasarkan hasil implementasi terhadap obat Arkavit, perhitungan manual mendapatkan hasil prediksi sebesar 3,87768 dan menggunakan sistem website yang telah dibuat mendapatkan hasil prediksi sebesar 3,88 obat. Hasil penelitian menunjukkan bahwa sistem mampu mereplikasi perhitungan manual secara akurat, dan beberapa obat menunjukkan performa model yang baik, meskipun sebagian lainnya masih memerlukan penyempurnaan. Secara keseluruhan, model dan sistem prediksi yang dibangun dapat membantu Apotek XYZ dalam memprediksi penjualan obat, dapat dijadikan dasar dalam pengambilan keputusan terkait pengelolaan stok obat secara lebih efisien.

**Kata Kunci:** Data Mining, Regresi Linear Berganda, Prediksi Penjualan, CRISP-DM, Apotek

### Abstract

*Competition in the sales business world such as pharmacies now requires pharmacy stakeholders to find stock management and marketing strategies to predict market share in order to increase drug sales and provide profitability for XYZ pharmacy. XYZ pharmacy has not implemented manual analysis or used a system to predict drug sales, so it often experiences empty or stock buildup. To analyze it, a data mining method or algorithm is needed that is able to process drug sales transactions into information. This study aims to develop a drug sales prediction information system at XYZ pharmacy which will later be implemented or integrated with a Multiple Linear Regression prediction model implemented in a website-based system, by following the CRISP-DM stages to help XYZ pharmacy in predicting drug sales so that it can prevent stockouts or excess stock and can support sales at the pharmacy. Price and rainfall variables are used as predictors of the number of sales. Based on the results of the implementation of Arkavit drugs, manual calculations obtained a prediction result of 3,87768 and using the website system that has been created obtained a prediction result of 3,88 drugs. The results showed that the system was able to accurately replicate manual calculations, and some drugs demonstrated good model performance, although others still needed improvement. Overall, the developed model and prediction system can assist XYZ Pharmacy in predicting drug sales and can be used as a basis for more efficient decision-making regarding drug inventory management.*

**Keywords:** Data Mining, Multiple Linear Regression, Sales Prediction, CRISP-DM, Pharmacy

## 1. PENDAHULUAN

Perkembangan teknologi informasi saat ini ditandai dengan peningkatan pesat di berbagai bidang termasuk di dunia kesehatan, salah satunya adalah apotek. Apotek merupakan salah satu tempat atau usaha yang menjual obat-obatan, alat kesehatan dan lainnya. Salah satu aspek penting untuk terjadinya sebuah transaksi jual beli pada apotek yaitu tersedianya obat-obatan. Persaingan di dunia bisnis penjualan seperti apotek sekarang menuntut para *stakeholder* apotek untuk menemukan strategi pengelolaan stok maupun pemasaran untuk memprediksi pangsa pasar agar dapat meningkatkan penjualan obat dan memberikan pelayanan yang maksimal terhadap konsumen. Persediaan obat yang efektif dapat memastikan tersedianya obat yang cukup dan dapat mencegah kekurangan stok yang berakibat menghambat pelayanan, serta mengoptimalkan pengeluaran apotek [1]. Salah satu cara untuk memprediksi penjualan atau kemungkinan yang akan terjadi di masa depan adalah dengan memanfaatkan dengan data transaksi penjualan obat pada apotek untuk dianalisis. Untuk menganalisis dibutuhkan suatu metode atau algoritma *data mining* yang mampu mengolah transaksi penjualan obat menjadi suatu informasi, hal ini dapat membantu para *stakeholder* sebagai dasar pengambilan suatu keputusan yang efektif dalam kelangsungan usaha.

*Data mining* merupakan rangkaian proses untuk mengetahui informasi dari suatu kumpulan data berupa pengetahuan menggunakan algoritma dan teknik dari ilmu statistik, mesin pembelajaran, dan sistem manajemen *database*

[2]. Salah satu teknik analisis *data mining* adalah Prediksi atau *Forecasting*. Proses memperkirakan kejadian pada waktu berikutnya, jadi data yang telah diolah menjadi informasi dan menghasilkan pengetahuan baru disebut prediksi [3]. Penelitian ini akan menggunakan regresi linear berganda untuk memprediksi bagian daripada penjualan untuk pengoptimalan persediaan pada apotek. Regresi linear berganda adalah bagian dari prediksi dalam *data mining* yang proses pengerjaannya berdasarkan banyaknya variabel yang digunakan untuk membentuk aturan ataupun rumus untuk melihat terdapat hubungan yang searah atau tidak dari hubungan variabel *independent* dan variabel *dependen* [4].

Pemilihan regresi linear berganda dalam penelitian ini didasarkan pada beberapa perbandingan. Pada penelitian yang dilakukan Novrizal Nazeriandy Siregar dan tim mengatakan bahwa analisis Regresi Linier Berganda banyak digunakan karena mampu mengukur ikatan atau dampak satu atau lebih variabel *independen* dan *dependen* [5]. Penelitian yang dilakukan Hilman Winnos dan timnya membandingkan Regresi Linear Berganda dan Autoregressive Integrated Moving Average (ARIMA) untuk prediksi saham PT. BSI, TbK. Hasilnya, perbandingan metode regresi linear berganda dan ARIMA menunjukkan akurasi tertinggi didapat oleh metode regresi linear berganda yaitu 98.9%, sedangkan metode ARIMA menghasilkan akurasi 97.64% [6]. Pada penelitian yang dilakukan Reza Syahputra Sinambela dan timnya, membandingkan metode Support Vector Reggresion dan metode Regresi Linear Berganda dalam memprediksi harga emas. Hasil akurasi yang didapat metode Regresi Linear Berganda 99,72% dan metode Support Vector Reggresion (SVR) 98.07% [7]. Dari penelitian yang telah disampaikan sebelumnya, dapat dikatakan metode regresi linear berganda unggul dan efektif dalam memprediksi sesuatu di masa depan. Dengan algoritma ini, diharapkan apotek dapat memprediksi penjualan sehingga apotek akan lebih efektif dalam membuat keputusan serta mendapatkan profitabilitas yang lebih.

Menurut hasil wawancara yang telah dilakukan, pihak apotek mengatakan bahwa curah hujan mempengaruhi penjualan obat, saat ini apotek belum pernah melakukan analisis data penjualan atau transaksi sebelumnya untuk memprediksi penjualan pada masa depan secara manual maupun menggunakan sistem, hal ini membuat apotek pernah kehabisan obat dan membuat obat pada apotek menumpuk atau tidak laku, sehingga diperlukan analisis yang akurat guna meningkatkan penjualan dan pengelolaan stok agar dapat bersaing pada pasar atau bisnis apotek.

Dalam melakukan penelitian ini, ada beberapa penelitian-penelitian sebelumnya yang dijadikan penulis sebagai referensi terkait topik yang di angkat.

Judul “*Comparison of the Performance of Multiple Linear Regression and Multi-Layer Perceptron Neural Network Algorithms in Predicting Drug Sales at Pharmacy XYZ*” yaitu tentang memprediksi penjualan analgesik secara agregat, tetapi tidak melakukan validasi apakah model yang sama cocok untuk semua jenis/kategori obat [8].

Judul “Algoritma Regresi Linier Berganda untuk Analisis Efisiensi Stok Produk di PT. Madu Pramuka Batang” yaitu tentang analisis efisiensi stok produk menggunakan regresi linear berganda, tetapi evaluasi atau akurasi yang digunakan hanya RMSE sehingga sulit untuk menilai kinerja model [3].

Judul “Implementasi Data Mining untuk Prediksi Persediaan Obat pada Puskesmas Kertapati menggunakan Regresi Linier Berganda” yaitu tentang memprediksi persediaan obat di Puskesmas. Penelitian ini tidak menjelaskan secara rinci bagaimana proses pemilihan data, pembersihan dilakukan [9].

Judul “Penerapan Metode Regresi Linear Berganda di RSUD Pratama Kerang untuk Prediksi Penggunaan Obat” yaitu tentang memprediksi persediaan obat di RSUD. Penelitian ini hanya melibatkan data internal saja dan tidak mempertimbangkan untuk memasukkan faktor” eksternal yang dapat mempengaruhi persediaan obat [10]

Judul “Prediksi Penjualan Crude Palm Oil (CPO) Menggunakan Metode Regresi Linear Berganda”, penelitian yang dilakukan tidak membagi data training dan data testing, koefisien  $b_2$  dan  $b_3 = 0$ ,  $X_2$  dan  $X_3$  tidak akan mempengaruhi rumus atau model karena hasilnya akan 0 sehingga rumus yang dipakai hanya  $Y = 171,635444271 + -0,000094297 * X_1$ . Penulis tidak menjelaskan mengapa dan apa dampaknya jika koefisien  $b_2$  dan  $b_3 = 0$  kepada pembaca. Penelitian ini juga tidak mengukur akurasi yang telah dilakukan [11].

Judul “Penerapan Metode Regresi Linier Sederhana untuk Prediksi Persediaan Obat Jenis Tablet” yaitu tentang memprediksi persediaan obat jenis ibuprofen tablet. Penelitian ini mengabaikan faktor-faktor penting lain yang mempengaruhi persediaan obat. Prediksi yang dilakukan hanya satu jenis obat (ibuprofen) dan tidak mewakili persediaan yang lain karena dihitung manual sehingga tidak bisa memprediksi jenis lainnya, sehingga penulis menyarankan untuk dibuat sebuah *software* untuk mempermudah pihak farmasi memprediksi jenis obat yang lain [12].

Berdasarkan penelitian terdahulu sebelumnya dapat disimpulkan bahwa sebagian besar penelitian menggunakan faktor-faktor internal untuk memprediksi kebutuhan dan penjualan obat tanpa mempertimbangkan faktor eksternal. Penerapan variabel eksternal curah hujan juga jarang ditemukan pada beberapa penelitian sebelumnya yang menunjukkan potensi untuk memasukkan curah hujan sebagai variabel penting untuk meningkatkan akurasi prediksi, terutama dalam konteks variabilitas penjualan obat berdasarkan cuaca. Selain curah hujan, variabel lain yang akan digunakan adalah harga. Pada penelitian yang dilakukan Risa Mustika dan Andini Setyo Anggraeni yaitu menganalisis faktor-faktor yang mempengaruhi keputusan pembelian pelanggan, salah satu faktor yang dianalisis yaitu harga dan hasil dari penelitian ini bahwa harga memiliki pengaruh negatif terhadap keputusan pembeli, jika harga obat naik 1 poin (sementara variabel lain tetap), maka keputusan pembelian pelanggan akan menurun sebesar 0,178 poin [13].

Penelitian ini akan menerapkan algoritma regresi linear berganda dalam menganalisis data transaksi historis penjualan pada apotek dengan tujuan dapat membuat sistem informasi yang nantinya diimplementasikan atau diintegrasikan sebuah model prediksi untuk memprediksi penjualan obat yang dapat dijadikan dasar dalam pengambilan keputusan dan

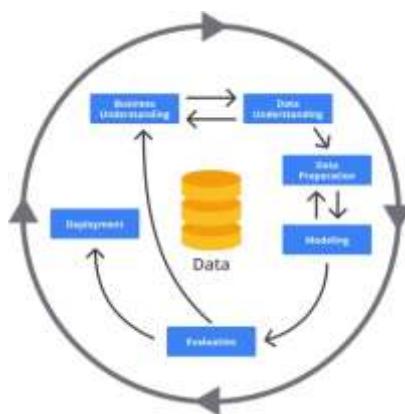
mengetahui tingkat keakuratan yang diperoleh dari algoritma regresi linear berganda. Selain itu, hasil analisis diharapkan dapat membantu Apotek XYZ untuk mengurangi risiko kehabisan atau menumpuknya obat, serta mendukung penjualan di apotek. Kiranya metode atau algoritma dalam ilmu data mining bisa digunakan untuk mendukung seorang *stakeholder* Apotek XYZ mendapatkan informasi yang diperlukan dalam merencanakan pembelian persediaan barang maupun strategi pemasaran untuk penjualan Apotek. Manfaat penelitian ini nantinya tidak hanya dirasakan oleh Apotek XYZ saja dan diharapkan Apotek XYZ dapat mengambil keputusan yang baik sehingga apotek dapat mengelola stok obat secara lebih efisien, mencegah terjadinya kekosongan maupun penumpukan obat sebagai upaya meningkatkan penjualan pada Apotek XYZ.

Berdasarkan yang telah disampaikan sebelumnya, maka penulis akan melakukan penelitian dengan judul “Implementasi Data Mining Regresi Linear Berganda pada Sistem Prediksi Penjualan Obat pada Apotek XYZ”.

## 2. METODOLOGI PENELITIAN

### 2.1 Tahapan Penelitian

Dalam penelitian ini penulis memakai metode penelitian Cross Industry Standard Process for Data Mining (CRISP-DM) yang dapat dilihat pada Gambar 1.



Gambar 1. Tahapan pada CRISP-DM [14]

Metode penelitian ini terbagi menjadi enam tahap, adapun penjelasan enam tahap tersebut sebagai berikut [15]:

a. *Business Understanding*

Tahapan ini merupakan langkah pertama yang bertujuan untuk mengenali dan merumuskan permasalahan yang perlu diselesaikan agar tujuan yang diinginkan tercapai.

b. *Data Understanding*

Pada tahap ini dilakukan persiapan awal terhadap data yang akan diproses dengan memastikan bahwa data dalam kondisi normal, lengkap dan konsisten sehingga layak digunakan untuk pemodelan sesuai dengan metode data mining yang akan digunakan.

c. *Data Preparation*

Tahap ini mencakup seluruh kegiatan pengolahan untuk membentuk dataset akhir. Data-data yang didapat biasanya memiliki kualitas rendah seperti nilai yang hilang, kesalahan input dan ketidakkonsistenan. Oleh karena itu, data perlu diproses terlebih dahulu melalui proses penghapusan data duplikasi, pengisian atau perbaikan data yang hilang, penyelarasan data yang tidak konsisten, dan memperbaiki kesalahan pengetikan.

d. *Modelling*

Pada tahap ini dilakukan pemilihan dan penerapan teknik *data mining* dengan menentukan algoritma yang digunakan. Dalam tahap ini, algoritma regresi linear berganda dipilih kemudian diterapkan ke dataset yang sudah disiapkan untuk mengatasi kebutuhan bisnis tertentu. Tahap pembuatan model juga mencakup penilaian dan analisa komparatif dari berbagai model yang dibangun

e. *Evaluation*

Setelah tahap modelling selesai dilakukan, model tersebut harus dievaluasi untuk melihat kualitas dan efektivitas sebelum disebarluaskan untuk digunakan. Pada tahap ini juga ditentukan apakah model dapat mencapai tujuan yang telah ditetapkan pada tahap pertama, apakah terdapat permasalahan penting dari penelitian yang tidak tertangani dengan baik, dan mengambil keputusan berkaitan dengan penggunaan hasil dari data mining

f. *Deployment*

Pada tahap akhir, penelitian yang telah dilakukan dibuat kesimpulan dan saran kemudian akan disampaikan sebagai rekomendasi ke pihak berkepentingan.

## 2.2 Teknik Pengumpulan Data

Teknik atau proses pengumpulan data dijelaskan sebagai berikut:

a. Studi Literatur

Penulis mencari berbagai sumber referensi yang relevan sebagai dasar teori atau acuan penelitian yang mendukung pada permasalahan kasus penelitian. Data tersebut akan mendukung penulis untuk melalukan dan mendukung penelitian.

b. Wawancara

Pengumpulan data dengan teknik wawancara dilakukan dengan tanya jawab dengan narasumber untuk mendapatkan informasi untuk mengetahui proses-proses pada objek yang hasilnya nanti akan dipakai dalam memutuskan fitur sistem prediksi yang nantinya akan dibuat.

Setelah studi literatur dan wawancara dilakukan terdapat beberapa dataset yang didapat untuk digunakan dalam penelitian ini yaitu dataset penjualan dan dataset curah hujan.

## 2.3 Data Mining

Data mining merupakan bidang keilmuan yang menggabungkan teknik dari *machine learning*, kecerdasan buatan, statistika, matematika, *database*, visualisasi untuk menemukan informasi yang berguna dan bermanfaat yang tersimpan dari basis data yang besar [16]. Data mining sering digunakan oleh peneliti atau perusahaan dalam menganalisa sebuah penjualan, yang nantinya dapat menciptakan sebuah strategi pemasaran, yang memungkinkan perusahaan bisa memperkirakan atau memprediksi sebuah keuntungan maupun kerugian [17]. Perusahaan dapat memakai teknik *data mining* untuk menganalisis penjualan dan mengira-ngira kemungkinan yang akan dipesan oleh *customer* di masa yang akan datang [18].

*Data mining* mempunyai beberapa fungsi dasar yang bisa digunakan untuk memperoleh pengetahuan atau informasi baru, yaitu

1. Prediksi (*prediction*)
2. Estimasi
3. Klasifikasi (*classification*)
4. Klastering (*Clustering*)
5. Asosiasi(*Association Rule*)

*Data mining* dikenal sebagai *Knowledge Discovery in Database* (KDD). KDD adalah sebuah proses model yang menggabungkan ilmu seperti matematika, statistika, database, kecerdasan buatan, sistem informasi manajemen dan berbasis pengetahuan [19].

### 2.3.1 Prediksi

Prediksi adalah upaya memperkirakan kejadian di masa mendatang berdasarkan data dari masa lalu (historis) melalui pendekatan ilmiah [20]. Data yang sering digunakan untuk melakukan prediksi adalah data yang berupa data kuantitatif. Prediksi tidak harus memberikan jawaban secara pasti apa yang akan terjadi, tetapi mencoba untuk mencari jawaban sedekat mungkin apa yang akan terjadi [16].

*Forecasting* akan mengambil kumpulan angka atau nilai-nilai yang tercatat dari waktu ke waktu kemudian teknik *Forecasting* ini akan memperkirakan nilai masa depan dengan menggunakan berbagai teknik *machine learning* dan teknik statistik yang berkaitan dengan musim, trend pada data [21]. *Forecasting* memiliki algoritma-algoritma untuk menyelesaikan sebuah permasalahan seperti *Linear Regression*, *Support Vector Machine*, *Time Series Forecasting*, *Neural Network* [22].

## 2.4 Algoritma Regresi Linear Berganda

Regresi linear berganda merupakan metode *statistic* dengan tujuan menganalisis kaitan antara variabel *dependen* dengan dua atau lebih variabel *independen* [23]. Dalam konteks prediksi penjualan, regresi linier berganda dipakai untuk menentukan dan memodelkan dampak variabel *independen* terhadap omset penjualan [24]. Regresi Linear Berganda adalah perkembangan dari Regresi Linear Sederhana. Dikatakan regresi linear sederhana jika jumlah variabel *independen* hanya ada satu saja, tetapi jika variabel *independen* nya lebih dari satu maka disebut dengan regresi linear berganda (*multiple regresi linear*) [25]. Algoritma cocok digunakan terhadap penelitian yang memiliki variabel *independen* lebih dari 1

Persamaan regresi linear berganda dapat dilihat pada persamaan 1:

$$Y = a + b_1X_1 + b_2X_2 \dots b_nX_n \quad (1)$$

Keterangan:

Y = Variabel terikat (*dependen*)

X = Variabel tidak terikat (*independen*)

a = Konstanta (*intercept*)

b = Koefisien regresi

## 2.5 Evaluasi Model Prediksi

### 2.5.1 MSE

MSE adalah metrik evaluasi yang menghitung rata-rata kesalahan dari selisih kuadrat antara data asli dan hasil prediksi. Nilai MSE yang rendah menunjukkan bahwa model memiliki kinerja yang lebih baik. Rumus MSE dapat dilihat pada persamaan 2:

$$MSE = \frac{1}{n} \sum_{i=1}^n [y_i - \hat{y}_i]^2 \quad (2)$$

Keterangan:

$\sum$  = penjumlahan terhadap kumpulan data

n = jumlah data.

$y_i$  = nilai aktual.

$\hat{y}_i$  = nilai prediksi

### 2.5.2 RMSE

RMSE (*Root Mean Squared Error*) adalah salah satu evaluasi yang digunakan untuk menilai tingkat akurasi sebuah model dalam prediksi dan mengukur seberapa baik model melakukan prediksi. Semakin rendah nilai RMSE atau semakin mendekati 0, maka hasil prediksi model dianggap semakin mendekati data sebenarnya. Walaupun RMSE selalu bernilai positif, nilai yang kecil menandakan performa model yang lebih baik. Rumus RMSE dapat dilihat pada persamaan 3:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (3)$$

Keterangan:

$\sum$  = penjumlahan terhadap kumpulan data

n = jumlah data.

$y_i$  = nilai aktual.

$\hat{y}_i$  = nilai prediksi

### 2.5.3 MAPE

MAPE adalah metrik untuk menilai atau mengukur akurasi prediksi yang dilakukan. Rumus MAPE dapat dilihat pada persamaan 4:

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \bar{y}}{\bar{y}} \right| \quad (4)$$

Keterangan:

$\sum$  = penjumlahan terhadap kumpulan data

n = jumlah data.

$\frac{y_i - \bar{y}}{\bar{y}}$  = Persentase kesalahan antara nilai aktual dan nilai yang diprediksi pada periode atau data ke-i.

Semakin sedikit nilai atau hasil MAPE maka tingkat akurasi model semakin baik. Persentase keakurasaian nilai MAPE dapat dilihat pada Tabel 1

Tabel 1. Persentase Keakurasaian MAPE [26]

Range MAPE	Arti Nilai
<10%	Sangat Baik
10-20%	Baik
20-50%	Layak
>50%	Buruk

## 2.6 Koefisien Determinasi ( $R^2$ )

Koefisien determinasi ( $R^2$ ) adalah statistik untuk menilai variabel *independent* dalam regresi dapat mengartikan variasi dalam variabel *dependen*.

Secara matematis, koefisien determinasi yang telah kita hitung dapat dirumuskan pada persamaan 5:

$$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2} \quad (5)$$

Keterangan:

$y_i$  = nilai aktual.

$\hat{y}_i$  = nilai prediksi

$\bar{y}$  = nilai rata-rata aktual.

Nilai  $R^2$  yang mendekati 1, menandakan variabel bebas mampu menjelaskan variabel terikat dengan sangat baik, interpretasi koefisien determinasi dapat dilihat pada Tabel 2.

Tabel 2. Interpretasi Koefisien Determinasi [15]

Interval Koefisien	Tingkat Hubungan

0,80 – 1,000	Sangat Kuat
0,60 – 0,799	Kuat
0,40 – 0,599	Cukup Kuat
0,20 – 0,399	Rendah
0,00 – 0,199	Cukup Rendah

### 3. HASIL DAN PEMBAHASAN

#### 3.1 Tahapan CRISP-DM

##### a. Business Understanding

Pada analisis ini dilakukan pemahaman mendalam tentang sistem yang ada untuk memahami alur proses. Dari hasil analisis sistem berjalan, diperoleh bahwa sistem yang saat ini berjalan untuk pengambilan keputusan mengenai pembelian atau pemesanan obat dilakukan berdasarkan perkiraan penjualan sebelumnya. Sistem yang berjalan saat ini memiliki keterbatasan karena belum mampu menyajikan gambaran akurat mengenai tren penjualan obat bebas. Hal ini sering menyebabkan ketidakakuratan dalam menentukan jumlah stok yang akan dibeli, baik kelebihan maupun kekurangan.

##### b. Data Understanding

Pada tahap *data understanding*, penulis akan melakukan pemahaman data yang diterima dalam bentuk *excel*. Data yang digunakan adalah data penjualan 68.931 baris yang berisi 8 atribut dan data curah hujan berisi 2 atribut. Data ini didapat melalui wawancara dan studi pustaka yang dilakukan. Dataset penjualan obat dan dataset curah hujan dapat dilihat pada Tabel 3 dan Tabel 4.

Tabel 3. Dataset Penjualan Obat

No	vc_nota	vc_tgl_nota	vc_k_obat	vc_n_obat	vc_n_satuan	dc_rupiah	dc_qty	dc_total
1	520240101002	01/01/2024	000151	Elkana tablet	TABLET	1205.22	10	12052.2
2	520240101002	01/01/2024	001049	VITAMIN B12 50mg	TABLET	189.82	20	3796.4
3	520240101002	01/01/2024	001051	VITAMIN B6 10mg	FLS	179.22	20	3584.4
...	...	...	...	...	...	....	...	...
689	520250930101	30/09/2025	001346	NELLCO SPESIAL 100 ML	FLS	44400	1	44400
31								

Sumber: Apotek XYZ

Keterangan:

vc\_nota: nomor nota transaksi

vc\_tgl\_nota = tanggal nota transaksi

vc\_k\_obat = kode obat

vc\_n\_obat = nama obat

vc\_n\_satuan = satuan atau kemasan obat

dc\_rupiah = harga satuan obat

dc\_qty = jumlah pembelian obat

dc\_total = jumlah harga pembelian obat

Tabel 4. Dataset Curah Hujan

Tanggal	Curahhujan
01-01-2024	0
02-01-2024	0
03-01-2024	0
...	...
30-09-2025	0.6

Sumber: BMKG [27]

Keterangan:

Tanggal: periode terjadinya pengukuran curah hujan

Curahhujan: jumlah air hujan yang jatuh

Setelah didapat dataset penjualan, dataset curah hujan dan melakukan pemahaman data, maka penentuan atribut yang akan digunakan adalah tanggal, vc\_n\_obat, vc\_n\_satuan, dc\_rupiah, curah hujan, dc\_qty. Berikut variabel yang dijadikan dasar, komponen atau faktor dalam memprediksi penjualan obat, dapat dilihat pada Tabel 5

Tabel 5. Variabel yang digunakan

No	Variabel	Nama Variabel	Keterangan
1	Y	Penjualan	Obat bebas yang terjual
2	X1	Harga	Nilai jual obat
3	X2	Cuaca	Tingkat curah hujan (mm)

c. *Data Preparation*

Pada Data Preparation akan melewati tahap-tahap select data, data integration, transformasi data dan cleaning data. Pertama, dilakukan select data untuk memilih data yang dibutuhkan atau digunakan untuk melakukan prediksi yang dapat dilihat pada Tabel 6 dan Tabel 7

Tabel 6. Atribut Data Penjualan

No	vc_tgl_nota	vc_n_obat	vc_n_satuan	dc_rupiah (X1)	dc_qty (Y)
1	01/01/2024	Elkana tablet	TABLET	1205.22	10
2	01/01/2024	VITAMIN B12 50mg	TABLET	189.82	20
3	01/01/2024	VITAMIN B6 10mg	FLS	179.22	20
...	...	...	...	....	...
689	31/09/2025	NELLCO SPESIAL	FLS	44400	1
31		100 ML			

Sumber: Apotek XYZ

Tabel 7. Atribut Data Curah Hujan

Tanggal	Curahhujan
01-01-2024	0
02-01-2024	0
03-01-2024	0
...	...
30-09-2025	0.6

Sumber: BMKG [27]

Selanjutnya, dilakukan proses *data integration* dengan menggabungkan dua sumber data, yaitu data penjualan dan data curah hujan. Penggabungan dilakukan berdasarkan atribut tanggal sehingga nilai curah hujan dapat diintegrasikan ke dalam dataset penjualan sebagai variabel independen X2. Hasil join atau integrasi nya dapat dilihat pada Tabel 8

Tabel 8. Dataset Merge (*Integration*)

no	vc_tgl_nota	vc_n_obat	vc_n_satuan	dc_rupiah (X1)	Curahhujan (X2)	dc_qty (Y)
1	01/01/2024	Elkana tablet	TABLET	1205.22	0	10
2	01/01/2024	VITAMIN B12 50mg	TABLET	189.82	0	20
3	01/01/2024	VITAMIN B6 10mg	FLS	179.22	0	20
...	...	...	...	....	...	...
689	31/09/2025	NELLCO SPESIAL	FLS	44400	0.6	1
31		100 ML				

Sumber: Apotek XYZ & BMKG[27]

Selanjutnya melakukan transformasi data. Pada tahap ini dilakukan untuk menyeragamkan atau menyamakan tipe data dari atribut yang digunakan. Perubahan akan dilakukan pada vc\_tgl\_nota dan tanggal. Pada atribut vc\_tgl\_nota dan tanggal akan diubah menjadi tipe data Date. Transformasi data dapat dilihat pada Tabel 9

Tabel 9. Transformasi Data

Atribut	Sebelum	Sesudah
vc_tgl_nota	Tipe Data String	Tipe Data Date
curahhujan	Tipe Data String	Tipe Data Date

Setelah itu dilakukan cleaning data untuk mencari atau memeriksa data yang tidak valid, tidak lengkap dan tidak konsisten pada masing” atribut. Data atau kolom yang bernilai kosong dan berisi “-“ akan menampilkan error

(missing data), atribut dc\_rupiah yang bernilai string (bukan angka) akan menghasilkan error karena tidak sesuai tipe data (invalid), pada atribut dc\_rupiah nilai atau data yang mengandung tulisan “Rp” di konversi menjadi decimal (10,2) contoh Rp10.000,50 akan diubah menjadi 10000,50, kemudian pada atribut curah hujan yang bernilai 8888, 9999, - akan disebut outlier yang kemudian akan dikonversikan menjadi angka (0) agar bisa di analisis. Setelah melakukan cleaning data, data yang awalnya berjumlah 68931 penjualan setelah dilakukan cleaning, data yang akan diproses berjumlah 68916 penjualan. Proses lengkap *cleaning data* dapat dilihat pada Tabel 10

Tabel 10. Cleaning Data

Atribut	Sebelum	Sesudah
vc_tgl_nota	p30-12-2024	Baris X : Format tanggal tidak valid
	-	Baris X : Kolom vc_tgl_nota berisi '-' (missing data)
		Baris X : Kolom vc_tgl_nota kosong
vc_n_obat	-	Baris X : Kolom vc_n_obat berisi '-' (missing data)
		Baris X : Kolom vc_n_obat kosong
vc_n_satuan	-	Baris X : Kolom vc_n_satuan berisi '-' (missing data)
		Baris X : Kolom vc_n_satuan kosong
dc_rupiah	-	Baris X : Kolom dc_rupiah berisi '-' (missing data)
	p21312	Baris X : Kolom dc_rupiah harus berupa angka, ditemukan 'p'
	Rp21,312	21312
dc_qty	10-	Pada Baris X: Kolom dc_qty harus berupa angka, ditemukan '10-'
	-	Baris X : Kolom dc_qty berisi '-' (missing data)
		Baris X : Kolom dc_rupiah kosong
curahhujan	8888	0
	9999	0
	-	0

Setelah melewati tahap” sebelumnya, hasil data preparation dapat dilihat pada Tabel 11.

Tabel 11. Dataset Akhir Yang Digunakan

no	vc_tgl_nota	vc_n_obat	vc_n_satuan	dc_rupiah (X1)	Curahhujan (X2)	dc_qty (Y)
1	01/01/2024	Elkana tablet	TABLET	1205.22	0	10
2	01/01/2024	VITAMIN B12 50mg	TABLET	189.82	0	20
3	01/01/2024	VITAMIN B6 10mg	FLS	179.22	0	20
...	...	...	...	....	...	...
689	31/09/2025	NELLCO SPESIAL	FLS	44400	0.6	1
16		100 ML				

Sumber: Apotek XYZ & BMKG [27]

d. *Modelling*

Pada tahap ini dilakukan implementasi menggunakan algoritma regresi linear berganda. Data yang telah disiapkan akan dilakukan pembagian *data training* dan *data testing*. Pembagian data yang dilakukan adalah 80% untuk *data training* dan 20% untuk *data testing*. Berikut merupakan *data training* obat Arkavit tablet sebagai sampel untuk dilakukan perhitungan yang dapat dilihat pada Tabel 12.

Tabel 12. Data Training Obat Arkavit Tablet

no	vc_tgl_nota	vc_n_obat	vc_n_satuan	dc_rupiah (X1)	Curahhujan (X2)	dc_qty (Y)
1	01/01/2024	Arkavit Tablet	Tablet	366.26	0	4
2	01/01/2024	Arkavit Tablet	Tablet	366.26	0	3
3	01/01/2024	Arkavit Tablet	Tablet	366.26	0	3
...	...	...	...	...	...	...
1839	10/06/2025	Arkavit Tablet	Tablet	366.26	0	4

Sumber: Apotek XYZ & BMKG [27]

$$n = 1839$$

$$\sum x_1 = 672774,02$$

$$\sum x_2 = 18511,7$$

$$\begin{aligned} \sum y &= 7210 \\ \sum x_1y &= 2632480,12 \\ \sum x_2y &= 72534,6 \\ \sum x_1^2 &= 246186932,5 \\ \sum x_2^2 &= 815843,63 \\ \sum x_1x_2 &= 3=6778210,506 \\ &\quad \begin{array}{ccccc} 1839 & 672774,02 & 18511,7 & b_0 & 85 \\ A = [ & 672774,02 & 246186932,5 & 6778210,506 ] & [b_1] = [20386,4] \\ & 18511,7 & 6778210,506 & 815843,63 & b_2 & 2665 \\ A = 70997434415649,8 & & & & \\ & 7210 & 672774,02 & 18511,7 & \\ A_0 = [2632480 & 246186932,5 & 6778210,51] & & & \\ & 72534,6 & 6778210,506 & 815843,63 & \\ A_0 = 2481572083036760 & & & & \\ & 1839 & 7210 & 18511,7 & \\ A_1 = [672774,02 & 2632480 & 6778210,51] & & & \\ & 18511,7 & 72534,6 & 815843,63 & \\ A_1 = -6023842253860,27 & & & & \\ & 1839 & 672774,02 & 7210 & \\ A_2 = [672774,02 & 246186932,5 & 2632480] & & & \\ & 18511,7 & 6778210,506 & 72534,6 & \\ A_2 = 52119748253,3778 & & & & \\ b_0 = \det A_0 / \det A & = 2481572083036760 / 70997434415649,8 & = 34,95298251 & & \\ b_1 = \det A_1 / \det A & = -6023842253860,27 / 70997434415649,8 & = -0,084845915 & & \\ b_2 = \det A_2 / \det A & = 52119748253,3778 / 70997434415649,8 & = 0,000734107 & & \end{array} \end{aligned}$$

Dari hasil perhitungan sebelumnya diperoleh:

$$Y = 34,95298 + -0,084845X_1 + 0,000734X_2$$

Diketahui:

$$\text{Nilai } X_1 = 366,26$$

$$\text{Nilai } X_2 = 0$$

Kemudian masukkan nilai  $X_1$  dan  $X_2$  tersebut ke dalam persamaan

$$Y = 34,95298 + -0,084845X_1 + 0,000734X_2$$

$$Y = 34,95298 + (-0,084845 * 366,26) + (0,000734 * 0)$$

$$Y = 3,87768 \text{ Tablet}$$

Secara umum dapat disimpulkan bahwa:

1. Model menunjukkan bahwa nilai awal variabel terikat ( $Y$ ) adalah 34,953 ketika seluruh variabel bebas bernilai nol

2. Variabel  $X_1$  berdampak negatif terhadap  $Y$ , yang ditunjukkan oleh koefisien regresi dengan nilai  $-0,084845$ . Artinya, setiap peningkatan  $X_1$  akan menurunkan nilai  $Y$ .

3. Variabel  $X_2$  berdampak positif terhadap  $Y$ , dengan koefisien regresi dengan nilai  $0,000734$ . Artinya, setiap peningkatan  $X_2$  akan meningkatkan nilai  $Y$ .

Dengan demikian, model regresi ini menggambarkan bahwa  $X_2$  menurunkan nilai  $Y$ , sedangkan  $X_2$  meningkatkan nilai  $Y$ , meskipun pengaruh  $X_2$  relatif lebih kecil.

e. *Evaluation*

Setelah didapat rumus regresi linear berganda, setelah itu penulis melakukan evaluasi untuk memberikan perbedaan atau perbandingan hasil prediksi dengan data yang sebenarnya. Evaluasi dilakukan dengan mencari nilai MSE, RMSE, MAPE dan R2. Tabel *Data Testing* dapat dilihat pada tabel 13

Tabel 13. *Data Testing*

no	Y	Yprediksi	error	absolute error	squared error	Absolute percentage error
1	4	3,858587711	0,1414	0,1414	0,01999	3,5353
2	3	3,858587711	-0,85859	0,85859	0,73717	28,6195
3	3	3,858587711	-0,85859	0,85859	0,73717	28,6195
...	...	...	...	...	...	...
460	3	3,859028111	-0,85903	0,85859	0,73792	28,6342

Dari perhitungan yang dilakukan yang dapat dilihat pada tabel 9 diperoleh nilai MSE, RMSE, MAPE dan R2.

$$\text{MSE} = 4,1862733$$

RMSE = 2,0460384

MAPE = 26,5825017

R<sup>2</sup> = -0,1036562

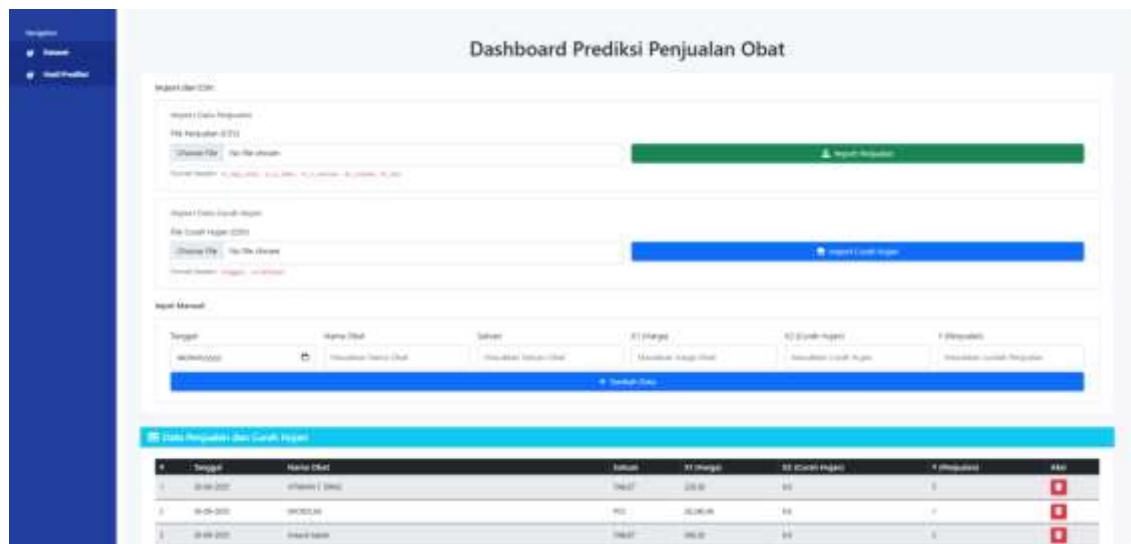
Dapat disimpulkan bahwa nilai RMSE memperlihatkan bahwa rata-rata kesalahan prediksi model adalah kira-kira 2 dari nilai actual, kemudian Nilai MAPE sebesar 26,58% menunjukkan bahwa akurasi dari prediksi yang dilakukan masih termasuk layak. Sementara itu, nilai R<sup>2</sup> yang bernilai negatif menandakan model tidak mampu mengartikan variasi data dengan baik.

## 3.2 Implementasi Sistem

Berikut ini adalah hasil sistem prediksi yang terdiri dari halaman dataset, halaman prediksi.

### a. Halaman Dataset

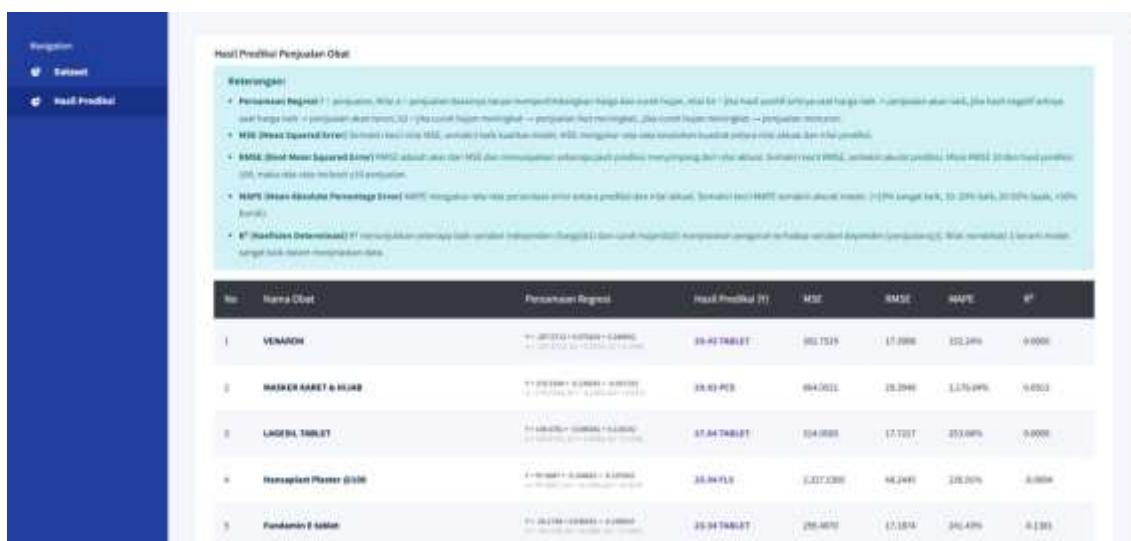
Fitur *dataset* ini digunakan pengguna untuk memasukkan data yang diperlukan yang nantinya berguna untuk memprediksi penjualan serta membuat modelling. Fitur dataset ini digunakan dengan cara mengimport file csv dan bisa menginput data secara manual. Halaman ini juga menampilkan dataset yang telah diimport atau diinput secara manual, halaman ini dapat dilihat pada gambar 2.



Gambar 2. Halaman Dataset

### b. Halaman Prediksi Penjualan Obat

Fitur prediksi penjualan obat ini digunakan pengguna untuk melihat hasil sebuah prediksi. Hasil prediksi penjualan akan dibuat per obat. Halaman ini dapat dilihat pada gambar 3.



Gambar 3. Halaman Hasil Prediksi

Berdasarkan penelitian yang sudah dilakukan, analisa dengan menerapkan Algoritma Regresi Linear Berganda. Pengujian yang dilakukan secara manual mendapatkan hasil prediksi sebesar 3,87768 tablet untuk Arkavit Tablet.

Pengujian menggunakan sistem website yang telah dibuat mendapatkan hasil prediksi sebesar 3,88 tablet untuk penjualan obat Arkavit Tablet di Apotek XYZ.

Perhitungan yang dilakukan secara manual dan menggunakan sistem website menghasilkan nilai yang sama. Dengan demikian, algoritma regresi linear berganda yang diimplementasikan dalam sistem telah berjalan sesuai dengan konsep dan perhitungan matematis yang sebenarnya, sehingga sistem mampu mereplikasi perhitungan manual secara akurat dan konsisten.

#### 4. KESIMPULAN

Berdasarkan penelitian yang dilakukan yang dilakukan, penerapan metode regresi linear berganda dengan tahapan CRISP-DM untuk prediksi penjualan obat di Apotek XYZ menghasilkan kesimpulan bahwa model prediksi telah berhasil dibuat dapat diimplementasikan atau diintegrasikan ke dalam sistem berbasis website. Hasil penelitian menunjukkan bahwa sistem yang dikembangkan mampu mereplikasi perhitungan regresi linear berganda secara akurat dan konsisten. Analisis dilakukan secara terpisah untuk setiap jenis obat, sehingga setiap obat memiliki model prediksi dan tingkat akurasi masing-masing. Beberapa obat menunjukkan hasil model prediksi yang baik dan mampu memprediksi penjualan obat secara lebih akurat, sementara sebagian lainnya memiliki performa model prediksi yang kurang optimal dan masih memerlukan penyempurnaan model. Meskipun demikian, implementasi algoritma regresi dalam sistem telah berjalan dengan benar, terbukti dari kesesuaian hasil prediksi antara perhitungan manual dan hasil sistem. Secara keseluruhan, sistem prediksi yang dibangun mampu melakukan proses perhitungan dengan konsisten serta bisa digunakan sebagai alat bantu untuk memprediksi penjualan obat di Apotek XYZ, meskipun masih ada ruang untuk meningkatkan kualitas model tetap diperlukan kedepannya melalui penambahan variabel *predictor* dan eksplorasi algoritma yang lainnya.

#### UCAPAN TERIMAKASIH

Terima kasih disampaikan kepada Bapak Iis Pradesan dan pihak-pihak yang telah mendukung dalam proses penyelesaian penelitian ini yang tidak dapat disebutkan satu persatu. Kiranya penelitian ini bisa memberikan manfaat bagi pembaca dan dapat meningkatkan kualitas penelitian selanjutnya.

#### DAFTAR PUSTAKA

- [1] A. Ardiansyah, A. T. Zy, dan A. Nugroho, “Implementasi data mining algoritma apriori pada sistem persediaan obat (studi kasus klinik pratama keluarga kesehatan),” *Journal of Information System, Applied, Management, Accounting and Research*, vol. 7, no. 3, hlm. 2598–8700, 2023, doi: 10.52362/jisamar.v7i3.1163.
- [2] F. Desryani Br Ginting, M. Ramadhan, R. mahyuni, S. Informasi, dan S. Triguna Dharma, “Penerapan Data Mining Dalam Pengelompokan Wilayah Toko Penjualan Air Mineral Kemasan Menggunakan Algoritma K-Means Clustering,” vol. 4, no. 5, hlm. 1243–1257, 2025, [Daring]. Tersedia pada: <https://ojs.trigunadharma.ac.id/index.php/jsi>
- [3] T. Hidayat, R. Darnis, dan D. Hidayatussa’adah, “Algoritma regresi linier berganda untuk analisis efisiensi stok produk di pt. Madu pramuka batang,” *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 12, no. 3, Agu 2024, doi: 10.23960/jitet.v12i3.4899.
- [4] Y. F. Wijaya dan A. Triayudi, “Penerapan data mining pada prediksi harga emas dengan menggunakan algoritma regresi linear berganda dan arima,” *Journal of Computer System and Informatics (JoSYC)*, vol. 5, no. 1, hlm. 73–81, Nov 2023, doi: 10.47065/josyc.v5i1.4615.
- [5] N. N. Siregar, Y. Syahra, dan M. Syaifudin, “Penerapan data mining untuk memprediksi penggunaan daya listrik pada pt.pln (persero) rayon medan selatan dengan menggunakan metode regresi linier berganda,” *Jurnal Sains Manajemen Informatika dan Komputer*, vol. 20, no. 1, hlm. 20–27, 2021, [Daring]. Tersedia pada: <https://ojs.trigunadharma.ac.id/>
- [6] H. Winnos, R. Septima, dan H. Gemasih, “Perbandingan metode regresi linier berganda dan autoregressive integrated moving average (arima) untuk prediksi saham pt. Bsi, tbk,” 2022.
- [7] R. S. Sinambela, M. Ula, dan A. F. Ulva, “Prediksi harga emas menggunakan algoritma regresi linear berganda dan support vector machine (svm),” *Jurnal Sistem dan Teknologi Informasi (JustIN)*, vol. 12, no. 2, hlm. 253, Apr 2024, doi: 10.26418/justin.v12i2.73386.
- [8] D. Arifuddin, K. Kusrini, dan K. Kusnawi, “Comparison of the performance of multiple linear regression and multi-layer perceptron neural network algorithms in predicting drug sales at pharmacy xyz,” *JURNAL SISFOTEK GLOBAL*, vol. 15, no. 1, hlm. 9, Mar 2025, doi: 10.38101/sisfotek.v15i1.15822.
- [9] Dahlia dan Andri, “Implementasi data mining untuk prediksi persediaan obat pada puskesmas kertapati menggunakan regresi linier berganda,” *Jurnal Sistem dan Informatika (JSI)*, vol. 15, no. 2, hlm. 95–103, Mei 2021, doi: 10.30864/jsi.v15i2.331.

# JURNAL SISTEM INFORMASI TGD

Volume 5, Nomor 1, Januari 2026, Hal 83-94

P-ISSN : 2828-1004 ; E-ISSN : 2828-2566

<https://ojs.trigunadharma.ac.id/index.php/jsi>



- [10] R. Rakhamadinnur, F. D. Marleny, dan Windarsyah, "Penerapan metode regresi linear berganda di rsud pratama kerang untuk prediksi penggunaan obat," *INTECH*, vol. 6, no. 1, hlm. 152–161, Mei 2025, doi: 10.54895/intech.v6i1.3099.
- [11] E. Lette, M. Zunaidi, dan W. Rista Maya, "Prediksi penjualan crude palm oil (CPO) menggunakan metode regresi linear berganda," *JURNAL SISTEM INFORMASI TGD*, vol. 1, hlm. 128–138, 2022, [Daring]. Tersedia pada: <https://ojs.trigunadharma.ac.id/index.php/jsi>
- [12] Harsiti, Z. Muttaqin, dan E. Srihartini, "Penerapan metode regresi linier sederhana untuk prediksi persediaan obat jenis tablet," *Sistem Informasi*, vol. 9, no. 1, hlm. 12–16, 2022.
- [13] R. Mustika dan A. S. Anggraeni, "Pengaruh harga, pelayanan dan kelengkapan obat terhadap keputusan pembeli apotek," *Jurnal Sintak*, vol. 1, no. 2, 2023, [Daring]. Tersedia pada: <https://doi.org/>
- [14] M. N. Alwi, "CRISP-DM: Tahapan, Studi Kasus, Kelebihan, dan Kekurangan," Dicoding. Diakses: 14 November 2025. [Daring]. Tersedia pada: <https://www.dicoding.com/blog/crisp-dm-tahapan-studi-kasus-kelebihan-dan-kekurangan/>
- [15] A. Novebrian Mahardja, I. Maulana, dan B. Arif Dermawan, "Penerapan metode regresi linear berganda untuk prediksi kerugian negara berdasarkan kasus tindak pidana korupsi," 2021. [Daring]. Tersedia pada: <http://jurnal.polibatam.ac.id/index.php/JAIC>
- [16] R. D. T. Wulandari dan N. Sari, "Prediksi peningkatan omset penjualan toko snack menggunakan metode regresi linier berganda," 2023.
- [17] A. Sumirso Wahyudi dan R. Zulfiandry, "Penerapan Metode Regresi Linear Berganda Dalam Prediksi Jumlah Pemakaian Obat Pada Rumah Sakit Rafflesia Kota Bengkulu," *Tahun 525 Jl. Meranti Raya*, vol. 20, no. 2, hlm. 341139, 2024.
- [18] R. Mulyawan, "Mengenal Pengertian Data Mining: Apa itu Penambangan Data? Menurut Ahli, Sejarah, Fungsi, Jenis, Proses dan Tahapannya!", rifqimulyawan. Diakses: 14 November 2025. [Daring]. Tersedia pada: <https://rifqimulyawan.com/blog/pengertian-data-mining/>
- [19] D. Elisa Sinaga dkk., "Analisis data mining algoritma decision tree pada prediksi persediaan obat (studi kasus : apotek franch farma)," vol. 2, no. 4, hlm. 123–131, 2022, [Daring]. Tersedia pada: <https://djournals.com/klik>
- [20] A. Damayanti, F. D. Marleny, dan A. A. Ningrum, "Implementasi regresi linear berganda untuk prediksi penjualan pada pt trimandiri sarana propetindo banjarmasin," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 13, no. 3, Jul 2025, doi: 10.23960/jitet.v13i3.6679.
- [21] NEX Datacenter, "Apa Saja Metode Dalam Data Mining?", nexdatacenter. Diakses: 14 November 2025. [Daring]. Tersedia pada: <https://www.nexdatacenter.com/id/apa-saja-metode-dalam-data-mining/>
- [22] R. Tineges, "Wajib Tahu, 3 Tipe Algoritma Machine Learning," dqlab. Diakses: 14 November 2025. [Daring]. Tersedia pada: <https://dqlab.id/3-tipe-algoritma-machine-learning>
- [23] D. F. S. Anjani, R. Astuti, W. Prihartono, dan R. Herdiana, "Prediksi penjualan handphone samsung di shopee menggunakan algoritma regresi linear berganda," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 13, no. 2, Apr 2025, doi: 10.23960/jitet.v13i2.6461.
- [24] F. Karina dan D. A. Dermawan, "Prediksi peningkatan omzet penjualan dengan menggunakan metode regresi linier berganda (studi kasus: ub makmur surabaya)," 2024.
- [25] D. M. Khairina, R. O. Shapanara, S. Maharani, dan H. Rahmania Hatta, "Data mining untuk estimasi sidang perkara narkotika menggunakan metode regresi linier berganda," 2022. [Daring]. Tersedia pada: <http://jurnal.polibatam.ac.id/index.php/JAIC>
- [26] R. Maulid, "Kriteria Jenis Teknik Analisis Data dalam Forecasting," dqlab. Diakses: 14 November 2025. [Daring]. Tersedia pada: <https://dqlab.id/kriteria-jenis-teknik-analisis-data-dalam-forecasting>
- [27] BMKG – Direktorat Data dan Komputasi, "Data harian – data iklim BMKG." Diakses: 30 Oktober 2025. [Daring]. Tersedia pada: <https://dataonline.bmkg.go.id/data-harian>