

Comparison of the Performance of Multiple Linear Regression Algorithms and Multi Layer Perceptron Neural Networks in Predicting Drug Sales

Perbandingan Performansi Algoritma Multiple Linear Regression dan Multi Layer Perceptron Neural Network dalam Memprediksi Penjualan Obat

Danang Arifuddin^{1*}, Kusrini², Kusnawi³

^{1,2,3}Program Studi Teknik Informatika, Fakultas Ilmu Komputer,
Universitas Amikom Yogyakarta, Indonesia

E-Mail: ¹danang22551202@students.amikom.ac.id,
²kusrini@amikom.ac.id, ³kusnawi@amikom.ac.id

Received Jan 31th 2025; Revised Apr 2nd 2025; Accepted Apr 16th 2025; Available Online Apr 20th 2025, Published Apr 20th 2025

Corresponding Author: Danang Arifuddin

Copyright © 2025 by Authors, Published by Institut Riset dan Publikasi Indonesia (IRPI)

Abstract

This study evaluates the attribute selection of internal (number of sales) and external (weather, commodity prices, inflation) variables using the correlation method, and compares the performance of Multiple Linear Regression (MLR) and Multi-Layer Perceptron Neural Network with backpropagation (MLPNN-b) algorithms in predicting the sales of analgesic drugs at "XYZ Pharmacy". Mean Squared Error (MSE) and Mean Absolute Percentage Error (MAPE) evaluation metrics were used to measure the prediction accuracy. The results show that the internal attribute "h-7" has the highest correlation (0.35) to daily sales, while external variables such as daily temperature, onion price, and interest rate also contribute. The MLPNN-b algorithm with certain parameters achieved a MAPE of 22.3% and MSE of 19,588 on single attributes, while MLR performed more evenly on combined attributes with a MAPE of 25.6% and MSE of 22,768. However, both models still suffer from underfitting with a fairly high level of prediction error. This study concludes that although MLPNN is superior in capturing non-linear relationships compared to MLR, the prediction accuracy is still not optimal. Therefore, exploration of hybrid models as well as integration of more external variables are recommended to improve sales prediction and support a more accurate pharmaceutical stock management system.

Keyword: Machine Learning, Multiple Linear Regression, Pharmacy, Sales Prediction

Abstrak

Penelitian ini mengevaluasi pemilihan atribut dari variabel internal (jumlah penjualan) dan eksternal (cuaca, harga komoditas, inflasi) menggunakan metode korelasi, serta membandingkan performansi algoritma Multiple Linear Regression (MLR) dan Multi-Layer Perceptron Neural Network dengan backpropagation (MLPNN-b) dalam memprediksi penjualan obat analgesik di "Apotek XYZ". Metrik evaluasi Mean Squared Error (MSE) dan Mean Absolute Percentage Error (MAPE) digunakan untuk mengukur akurasi prediksi. Hasil menunjukkan bahwa atribut internal "h-7" memiliki korelasi tertinggi (0,35) terhadap penjualan harian, sementara variabel eksternal seperti suhu harian, harga bawang merah, dan suku bunga juga memberikan kontribusi. Algoritma MLPNN-b dengan parameter tertentu mencapai MAPE 22,3% dan MSE 19.588 pada atribut tunggal, sedangkan MLR memiliki kinerja lebih merata pada atribut kombinasi dengan MAPE 25,6% dan MSE 22.768. Namun, kedua model masih mengalami underfitting dengan tingkat kesalahan prediksi yang cukup tinggi. Penelitian ini menyimpulkan bahwa meskipun MLPNN lebih unggul dalam menangkap hubungan non-linear dibandingkan MLR, akurasi prediksi masih belum optimal. Oleh karena itu, eksplorasi model hybrid serta integrasi lebih banyak variabel eksternal direkomendasikan untuk meningkatkan prediksi penjualan dan mendukung sistem manajemen stok farmasi yang lebih akurat.

Kata Kunci: Farmasi, Machine Learning, Multiple Linear Regression, Sales Prediction



1. PENDAHULUAN

Kekurangan obat masih menjadi masalah signifikan dalam sistem layanan kesehatan global, memengaruhi akses dan pasokan obat yang vital bagi kebutuhan pengobatan pasien [1], [2]. Tantangan ini diperburuk oleh biaya tambahan seperti penggantian, pengiriman, dan penyimpanan yang dapat meningkatkan harga obat, membuatnya sulit diakses [3]. Upaya prediksi dan pengelolaan stok obat yang efisien diperlukan untuk mengurangi dampak kekurangan ini, namun penelitian mendalam mengenai solusi prediksi tingkat konsumsi obat masih terbatas [2].

Penggunaan algoritma machine learning telah berkembang untuk mengoptimalkan prediksi stok obat, termasuk algoritma neural network seperti *Multilayer Perceptron Neural Network* (MLPNN) dengan backpropagation. Algoritma ini unggul dalam menangani pola data non-linear, sementara pendekatan hybrid seperti MLPNN dikombinasikan dengan *Autoregressive Integrated Moving Average* (ARIMA) atau *genetic algorithms* juga telah digunakan untuk meningkatkan akurasi prediksi [4], [5]. Penelitian lain membandingkan shallow dan deep neural networks, dengan shallow network menunjukkan performa terbaik dalam beberapa kasus [6]. Namun, metode sederhana seperti *Multiple Linear Regression* (MLR) tetap relevan sebagai alternatif praktis dan hemat biaya [7].

Apotek memiliki peran penting sebagai titik akhir rantai distribusi obat, menyediakan data permintaan real-time dan memastikan keberlanjutan pasokan obat secara rasional [8]. Namun, tantangan utama dalam pengelolaan stok obat adalah keterbatasan masa edar dan kebutuhan penyimpanan khusus, yang meningkatkan risiko obat kadaluwarsa atau kekurangan stok pada saat dibutuhkan. Ketidakmampuan apoteker dalam menganalisis pola permintaan secara akurat dapat menyebabkan inefisiensi dalam distribusi obat, berujung pada keterbatasan akses bagi pasien [9]. Oleh karena itu, diperlukan pendekatan yang lebih sistematis dan berbasis data untuk mengoptimalkan pengelolaan stok obat di tingkat apotek.

Pandemi Covid-19 semakin menyoroti pentingnya prediksi konsumsi obat yang akurat, terutama untuk analgesik yang mengalami lonjakan permintaan selama periode tersebut [3]. Obat-obatan seperti parasetamol, kodein, tramadol, dan morfin memerlukan pemantauan konsumsi yang lebih cermat, mengingat pengaruh berbagai faktor eksternal seperti perubahan pola kesehatan masyarakat dan kebijakan distribusi obat [10]. Dalam penelitian ini, machine learning menjadi solusi yang tepat karena kemampuannya dalam mengolah data historis dan mengidentifikasi pola yang kompleks, sehingga dapat menghasilkan prediksi yang lebih akurat dibandingkan metode konvensional. Dengan menerapkan model prediksi berbasis machine learning, penelitian ini bertujuan untuk memahami faktor-faktor internal dan eksternal yang memengaruhi tingkat konsumsi obat analgesik resep, guna mendukung pengelolaan stok yang lebih efektif, efisien, dan berkelanjutan.

Penelitian ini bertujuan untuk menganalisis korelasi antara variabel internal, yaitu volume penjualan obat, dan variabel eksternal seperti cuaca harian, harga komoditas (beras dan minyak goreng), inflasi, serta suku bunga kredit terhadap permintaan obat-obatan di daerah pedesaan dengan mayoritas masyarakat petani. Selain itu, penelitian ini juga membandingkan tingkat efektivitas prediksi permintaan obat-obatan menggunakan algoritma MLR dan MLPNN berdasarkan nilai Mean Squared Error (MSE) dan Mean Absolute Percentage Error (MAPE). Hasil yang ingin dicapai meliputi pemahaman terhadap hubungan antara faktor eksternal dengan permintaan obat serta evaluasi kinerja kedua algoritma dalam memprediksi penjualan obat di wilayah tersebut.

MLR telah digunakan secara luas dalam berbagai penelitian prediksi karena kemampuannya dalam memodelkan hubungan linear antara variabel independen dan dependen. Penelitian oleh Maulana (2024) menunjukkan bahwa MLR dapat digunakan untuk memprediksi permintaan obat dengan mempertimbangkan faktor sosial-ekonomi dan tren musiman [11]. Studi lain oleh Pratista (2023) mengembangkan model prediksi konsumsi obat menggunakan MLR dan menemukan bahwa model ini cukup akurat dalam skenario dengan data yang memiliki hubungan linear yang kuat [12]. Selain itu, penelitian oleh Agnes et al. (2023) dan Ginting et al. (2021) membandingkan MLR dengan model non-linear lainnya dalam memprediksi permintaan obat di apotek dan menunjukkan bahwa meskipun MLR lebih mudah diinterpretasikan, akurasinya dapat berkurang jika pola data bersifat non-linear [13], [14].

Di sisi lain, MLPNN telah diterapkan dalam berbagai studi prediksi karena kemampuannya menangani data non-linear dan mengenali pola kompleks. Sebagai contoh, penelitian oleh Artanti (2023) menunjukkan bahwa MLPNN dapat meningkatkan akurasi prediksi diabetes tahap awal dengan memanfaatkan data historis dan variabel lingkungan [15]. Penelitian lain oleh Yustisio (2022) menerapkan MLPNN untuk memprediksi jumlah pasien Covid-19 di rumah sakit dan menemukan bahwa model ini memiliki performa lebih baik dibandingkan metode statistik konvensional [16]. Selain itu, studi oleh Atianto (2023) mengembangkan pendekatan hybrid dengan mengombinasikan MLPNN dan algoritma optimasi untuk meningkatkan akurasi prediksi dalam pengelolaan stok obat [17]. Dari hasil penelitian ini, MLPNN terbukti lebih unggul dalam menangkap pola kompleks dibandingkan dengan metode linear seperti MLR, terutama ketika terdapat faktor eksternal yang berkontribusi dalam fluktuasi permintaan obat.

Penelitian ini mengatasi keterbatasan studi sebelumnya yang belum secara spesifik mempertimbangkan variabel eksternal yang relevan bagi masyarakat pedesaan, seperti akses layanan kesehatan, kondisi sosial-ekonomi, dan ketersediaan tenaga medis. Selain itu, masih jarang studi yang secara langsung membandingkan

performa algoritma MLR dan MLPNN dalam memprediksi permintaan obat. Oleh karena itu, penelitian ini menawarkan pendekatan yang lebih komprehensif dengan mengintegrasikan variabel eksternal dalam model prediksi serta mengevaluasi efektivitas metode linear dan non-linear. Hasil penelitian ini diharapkan dapat membantu manajemen apotek dalam perencanaan stok, memberikan pemahaman kepada peneliti mengenai kinerja algoritma prediksi, serta memberi informasi kepada produsen dan distributor obat mengenai jenis obat yang dibutuhkan masyarakat berdasarkan faktor eksternal seperti cuaca, harga bahan pokok, inflasi, dan suku bunga kredit.

2. LITERATUR REVIEW

2.1 Tinjauan Pustaka

Penelitian oleh Emmanuel et al. (2022) menunjukkan bahwa model time series linier kurang akurat dibandingkan neural network dalam memprediksi stok obat, meskipun neural network menghadapi tantangan dalam pembobotan dan evaluasi menyeluruh [1]. Nguyen et al. (2023) memanfaatkan data logistik untuk mencegah kekurangan stok obat di rumah sakit Prancis dengan indikator seperti rata-rata waktu pengiriman sebagai prediktor kondisi pasokan masa depan [2]. Holt-Winters Seasonal Additive dianggap lebih akurat daripada metode canggih seperti Genetic Algorithm, berkat nilai RMSE yang lebih kecil.

2.2. Data Mining

Menurut Kusriani dan Luthfi (2009)[18], data mining adalah proses menemukan pola dan hubungan signifikan dalam data menggunakan teknik statistik dan matematika. Larose (2005) [19] menambahkan bahwa campur tangan manusia diperlukan dalam setiap tahap data mining untuk menghindari kesalahan algoritma dan memastikan analisis yang tepat. Data mining tidak hanya menggali pengetahuan tetapi juga memprediksi tren masa depan [20].

2.3. Outlier dan Metode IQR

Outlier adalah nilai ekstrem yang dapat mengganggu analisis data [19]. Metode Inter-Quartile Range (IQR) digunakan untuk mendeteksi dan menghapus outlier, membagi data menjadi empat bagian dengan batasan Q1 (25%), median (50%), dan Q3 (75%). Pembersihan data ini penting untuk meningkatkan akurasi analisis [21].

2.4. Korelasi

Korelasi antara variabel dianalisis untuk memahami hubungan linier. Data dikumpulkan dari Farmacare.id, meliputi harga kebutuhan pokok (Siskaperbapo), cuaca (BMKG), inflasi (BPS Banyuwangi), dan suku bunga (OJK) selama periode Januari–Juni 2024. Harga pupuk dikecualikan karena nilainya tetap selama periode pengamatan.

2.5. Algoritma Multiple Linear Regression (MLR)

MLR memungkinkan analisis hubungan antara banyak variabel bebas dan satu variabel terikat. Menurut Kurniawan (2022) [22], MLR lebih efektif dibanding Simple Linear Regression (SLR) dalam menangani data dengan banyak faktor. Larose (2005) menyebut MLR sebagai metode penting dalam data mining untuk memanfaatkan variabel potensial dengan hubungan linier terhadap target [19]. Secara matematis MLR ditunjukkan pada persamaan 1.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + \epsilon \quad (1)$$

Dalam persamaan regresi linear, Y merepresentasikan variabel dependen atau target yang ingin diprediksi, sedangkan X_1, X_2, \dots, X_n merupakan variabel independen atau prediktor yang memengaruhi nilai dari Y . Nilai β_0 berfungsi sebagai intersep atau konstanta, yaitu nilai Y saat seluruh variabel X bernilai nol. Sementara itu, $\beta_1, \beta_2, \dots, \beta_n$ adalah koefisien regresi yang menggambarkan seberapa besar pengaruh masing-masing variabel independen terhadap variabel dependen. Terakhir, ϵ atau error term menunjukkan selisih antara nilai aktual dan nilai prediksi, yang mencerminkan variasi dalam Y yang tidak dapat dijelaskan oleh variabel X .

2.6. Algoritma Multi-Layer Perceptron Neural Network (MLPNN)

MLPNN adalah algoritma neural network dengan lapisan tersembunyi (hidden layers) yang memungkinkan prediksi data non-linear [19]. Struktur MLPNN terdiri dari input, hidden, dan output nodes yang terhubung penuh dengan bobot awal acak. Aktivasi seperti fungsi sigmoid umum digunakan karena kemampuannya menangkap pola kompleks [23]. Namun, semakin banyak hidden nodes, semakin kompleks dan lambat pengolahannya. Secara matematis MLPNN ditunjukkan pada persamaan 2 dan 3.

$$Z_j = f \left(\sum_{i=1}^n w_{ij} X_i + b_j \right) \quad (2)$$

dan outputnya dihitung sebagai:

$$Y_k = f \left(\sum_{j=1}^m v_{jk} Z_j + b_k \right) \quad (3)$$

Dalam jaringan saraf tiruan, setiap input ke jaringan dilambangkan dengan X_i , yang kemudian dikalikan dengan bobot koneksi w_{ij} antara input layer dan hidden layer. Hasil perkalian ini dijumlahkan dengan bias pada hidden layer, yang dilambangkan dengan b_j , lalu dilewatkan ke fungsi aktivasi $f(x)$ untuk menghasilkan output dari hidden layer yang disebut Z_j . Output dari hidden layer ini kemudian diteruskan ke output layer dengan bobot koneksi baru, yaitu v_{jk} , yang menghubungkan hidden layer ke output layer. Proses ini juga mencakup penambahan bias b_k pada output layer, sebelum akhirnya dilewatkan kembali ke fungsi aktivasi untuk menghasilkan output akhir jaringan yang disebut Y_k . Fungsi aktivasi yang digunakan pada jaringan ini dapat berupa sigmoid, ReLU, tanh, atau fungsi lainnya yang berfungsi untuk memperkenalkan non-linearitas ke dalam model.

2.7. Metode Backpropagation

Backpropagation adalah metode pembelajaran dalam MLPNN yang efektif untuk menangani masalah non-linear [18]. Menurut Lyu et al. (2022), backpropagation mampu belajar, beradaptasi, dan menghadapi error secara mandiri, menjadikannya metode yang fleksibel untuk berbagai aplikasi [24]. Backpropagation bekerja dengan menghitung error antara output prediksi dan nilai aktual, kemudian memperbarui bobot menggunakan metode perhitungan gradien dari fungsi loss. Persamaan umum untuk pembaruan bobot dalam backpropagation ditunjukkan pada persamaan 4.

$$w_{new} = w_{old} - \eta \frac{\partial E}{\partial w} \quad (4)$$

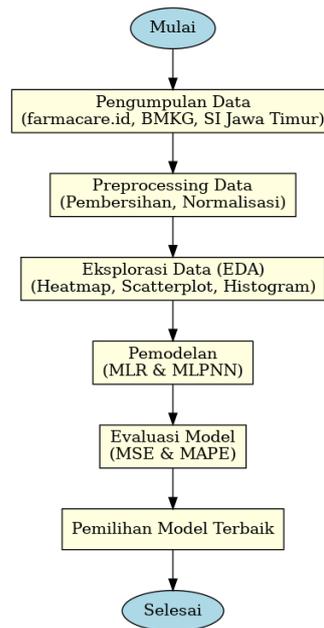
Dalam proses pelatihan jaringan saraf, pembaruan bobot dilakukan untuk meminimalkan error atau fungsi loss. Bobot baru w_{new} diperoleh dengan mengurangi hasil perkalian antara learning rate η dan turunan dari fungsi error E terhadap bobot $\frac{\partial E}{\partial w}$ dari bobot sebelumnya w_{old} . Learning rate η sendiri merupakan parameter yang menentukan seberapa besar langkah pembaruan bobot dilakukan dalam setiap iterasi. Sementara itu, turunan dari fungsi error terhadap bobot menunjukkan arah serta besarnya perubahan bobot yang diperlukan agar error dapat diminimalkan. Pendekatan ini dikenal sebagai metode gradient descent, yang merupakan dasar dari banyak algoritma optimasi dalam pembelajaran mesin. Penelitian ini mengintegrasikan pendekatan MLR dan MLPNN dengan backpropagation untuk mengevaluasi hubungan variabel internal dan eksternal terhadap prediksi stok obat, memastikan model yang efektif dalam mengelola pasokan farmasi.

3. METODOLOGI PENELITIAN

Penelitian ini merupakan studi kasus kuantitatif yang berfokus pada angka penjualan obat sebagai variabel target dan variabel prediktor (faktor internal dan eksternal). Penelitian bersifat eksploratif karena variabel dan algoritma yang digunakan dapat menjadi dasar bagi penelitian lanjutan. Algoritma dengan tingkat error terendah akan direkomendasikan untuk mendukung prediksi penjualan.

Data yang digunakan dalam penelitian ini diperoleh dari berbagai sumber, termasuk rekam medis dari farmacare.id, data cuaca dari BMKG, serta data harga komoditas dari Sistem Informasi Jawa Timur dalam rentang waktu 1 Januari–30 Juni 2024. Data yang dikumpulkan akan melalui proses Preprocessing Data, yang mencakup pembersihan dari nilai hilang dan outlier, serta normalisasi menggunakan metode Min-Max Scaling. Selanjutnya, dilakukan *Explanatory Data Analysis* (EDA) dengan visualisasi berupa heatmap, scatterplot, dan histogram untuk mengidentifikasi pola serta memilih atribut prediktor yang paling relevan.

Pemodelan dilakukan dengan menggunakan dua algoritma, yaitu *Multiple Linear Regression* (MLR) untuk menangkap hubungan linier antar variabel, serta *Multi-Layer Perceptron Neural Network* (MLPNN) untuk menangkap pola non-linear dalam data. Model ini dievaluasi berdasarkan Mean Squared Error (MSE) sebagai ukuran kesalahan prediksi absolut, serta *Mean Absolute Percentage Error* (MAPE) untuk mengevaluasi rata-rata kesalahan dalam bentuk persentase. Diagram alir metodologi penelitian ditampilkan pada Gambar X, yang menjelaskan alur mulai dari pengumpulan data, preprocessing, eksplorasi data, pemodelan, hingga evaluasi model guna mendapatkan algoritma dengan performa terbaik. Gambaran umum metodologi penelitian dapat ditunjukkan pada gambar 1.



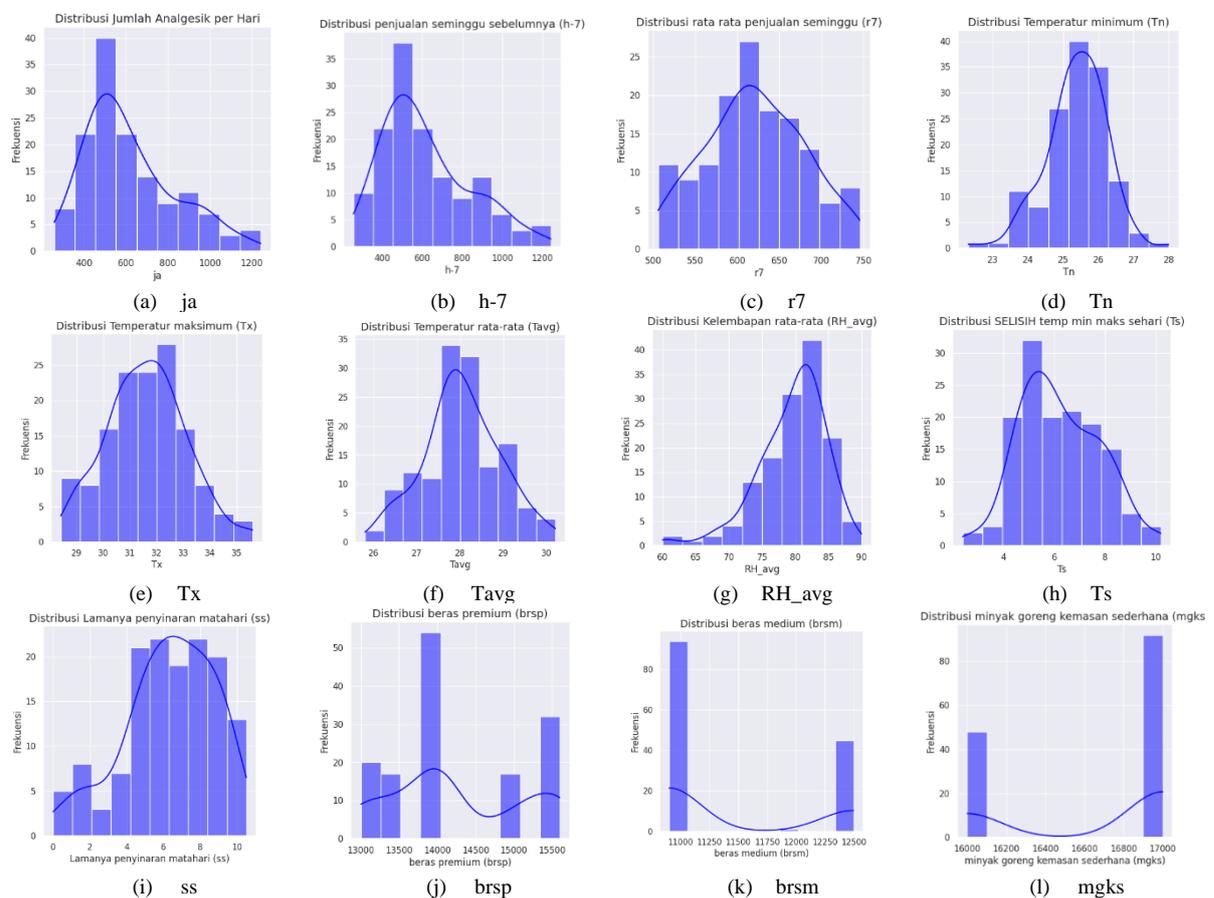
Gambar 1. Metodologi Penelitian

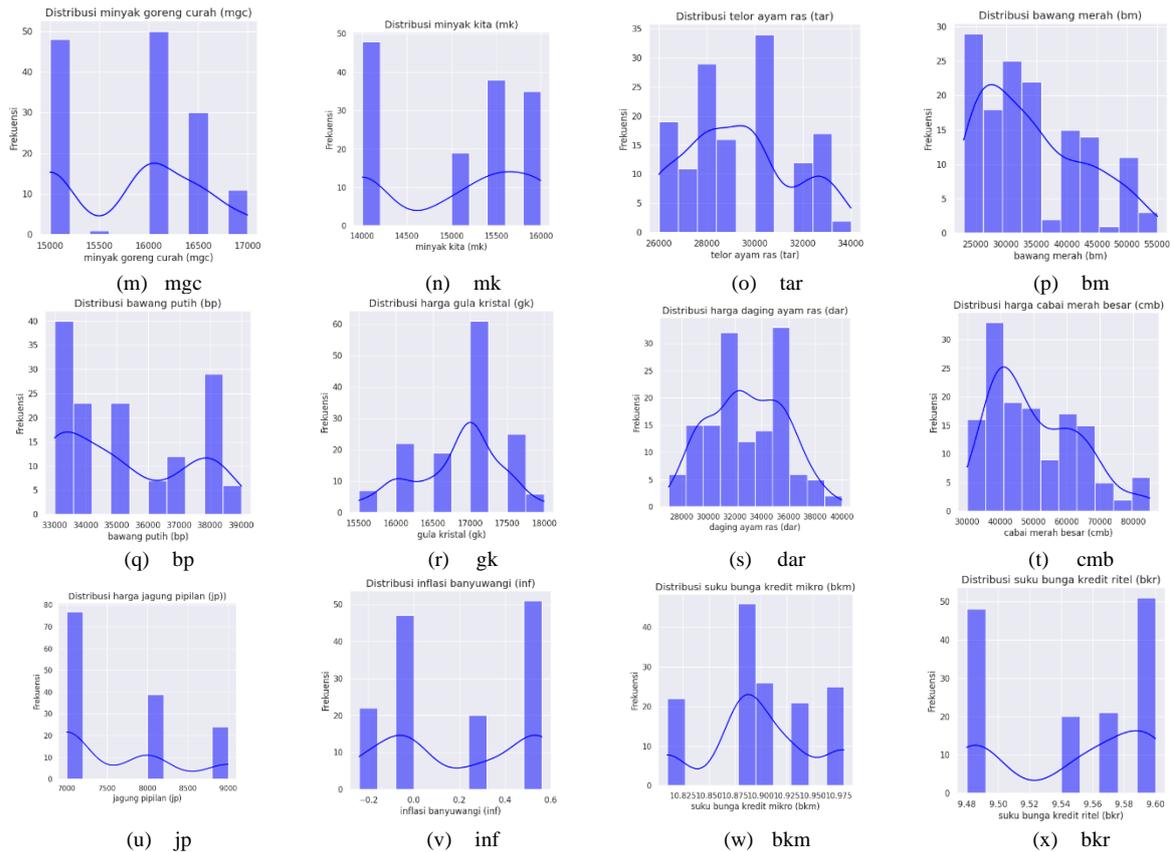
4. HASIL DAN PEMBAHASAN

4.1 Preprocessing (Pengolahan Awal Data)

4.1.1 Analisis Deskriptif Data

Penelitian ini menggunakan dataset yang mencakup atribut-atribut dari variabel internal dan eksternal. Analisis statistik deskriptif dilakukan untuk memahami pola distribusi data awal, sebagaimana ditampilkan dalam Tabel 1. Menurut Kusri dan Luthfi (2009) [18], analisis ini penting untuk memberikan gambaran awal sebelum tahap pengolahan data. Visualisasi data juga digunakan untuk mengidentifikasi outlier atau anomali yang dapat mengganggu harmonisasi data.





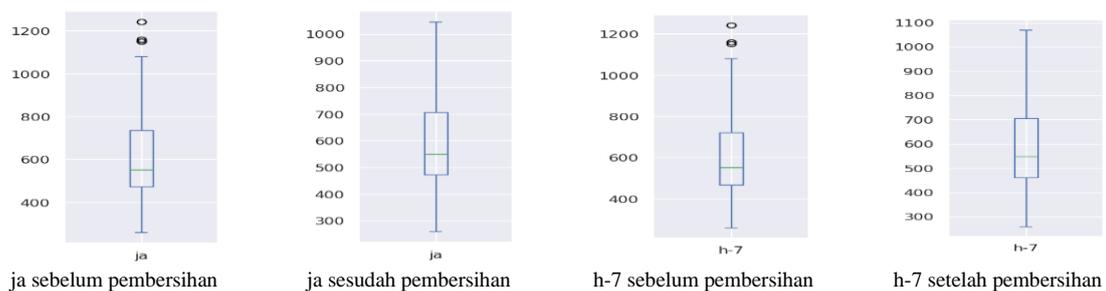
Gambar 2. Histogram semua atribut prediktor dan target

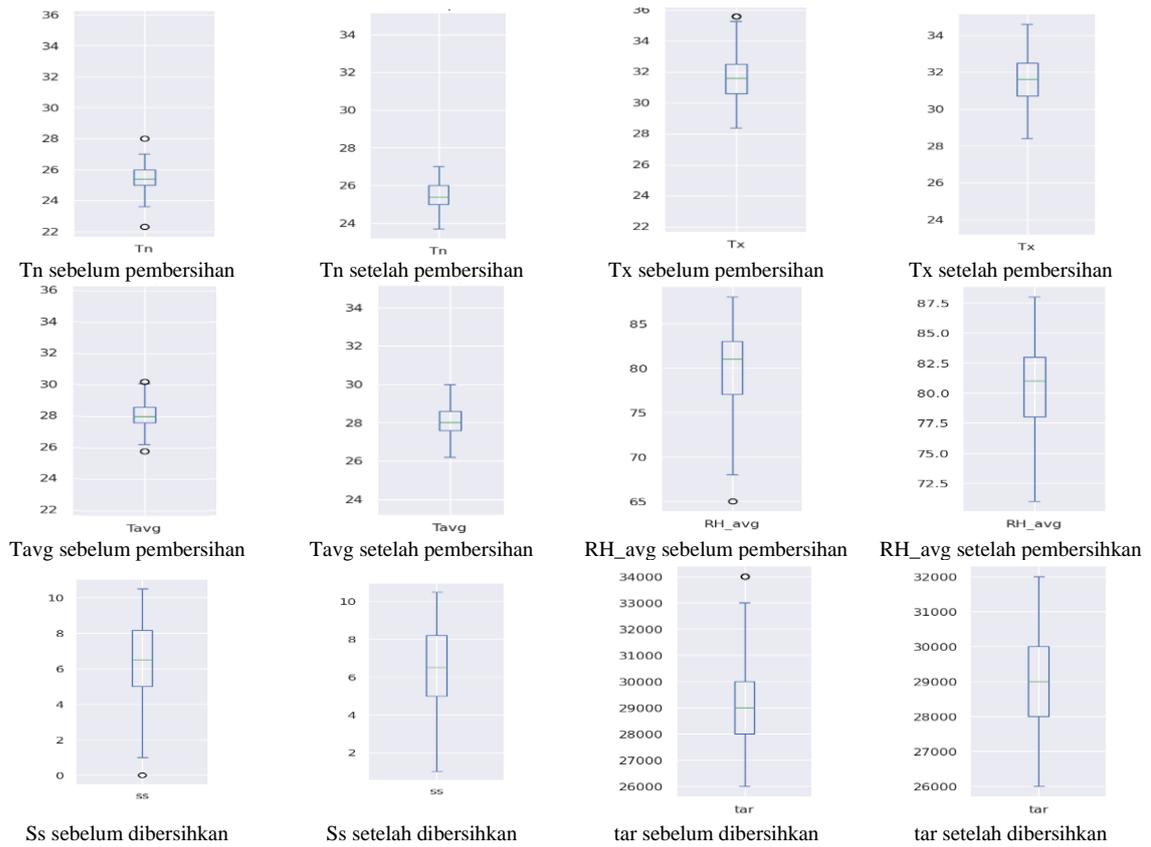
Pada gambar 2 dapat dilihat banyak grafik yang memiliki kemiringan (Skewness) baik ke kiri maupun ke kanan. Kemiringan distribusi data pada grafik menunjukkan adanya kemungkinan ketidakseimbangan dalam penyebaran data. Skewness ke kanan (positif) mengindikasikan bahwa sebagian besar nilai berada di sisi kiri dengan ekor panjang ke kanan, yang sering terjadi pada data dengan nilai ekstrem tinggi (outlier tinggi). Sebaliknya, skewness ke kiri (negatif) menunjukkan bahwa sebagian besar nilai berada di sisi kanan dengan ekor panjang ke kiri, yang bisa terjadi pada data dengan beberapa nilai sangat rendah (outlier rendah).

Adanya skewness ini penting untuk dianalisis karena dapat mempengaruhi pemodelan statistik dan machine learning. Sebagian besar model, seperti regresi linear dan neural network, bekerja lebih optimal pada data dengan distribusi normal, sehingga skewness yang signifikan dapat menyebabkan bias dalam estimasi parameter model. Selain itu, skewness yang berlebihan juga bisa menandakan adanya outlier atau anomali dalam data, yang mungkin berasal dari kesalahan pencatatan, kondisi ekonomi yang ekstrem, atau perubahan musiman yang tidak biasa. Untuk mengatasi masalah ini, teknik transformasi data seperti log transformation, Box-Cox transformation, atau square root transformation dapat digunakan untuk mendekati distribusi ke normal. Dengan memahami pola skewness dari atribut yang ditampilkan dalam grafik, langkah preprocessing seperti deteksi dan penanganan outlier dapat dilakukan sebelum melanjutkan ke tahap pemodelan data, sehingga hasil analisis menjadi lebih akurat dan dapat diandalkan.

4.1.2 Penanganan Outlier

Gambar 3 merupakan deretan perintah dari metode Inter Quatile Range (IQR) untuk mendrop outlier jumlah analgesic per hari per kolom dan mendrop baris dengan outlier.

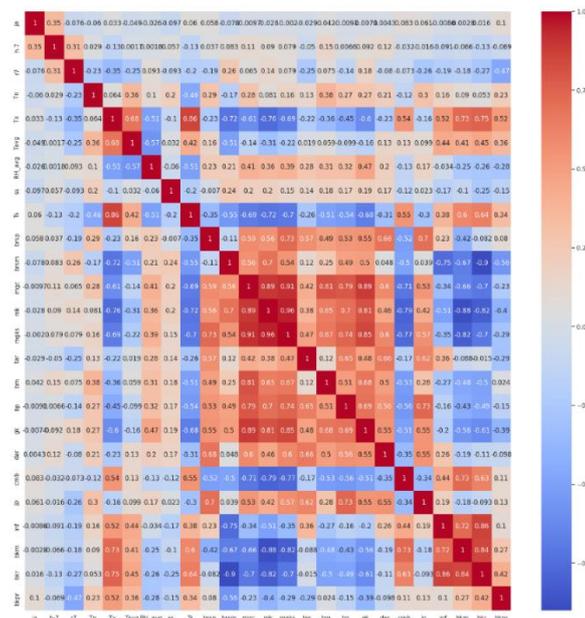




Gambar 3. Grafik boxplot atribut yang menderita outlier

4.2 Hubungan Antar Variabel (Korelasi)

Matriks korelasi antar atribut dengan atribut ‘jumlah analgesik per hari (ja) hasil preprocessing dapat ditunjukkan pada tabel 1. Pada gambar 4, heatmap korelasi antar atribut terlihat warna yang lebih gelap menunjukkan hubungan yang lebih kuat antara atribut. Pada gambar 7 terlihat bahwasanya hampir semua atribut terlihat memiliki korelasi yang lemah dengan atribut ‘jumlah analgesic per hari (ja)’. Hal ini akan menunjukkan semakin berat pemodelan tercapai khususnya pada model linear. Namun dalam penelitian ini akan lakukan 2 skenario pengujian. Pertama, akan mengambil nilai korelasi lebih dari sama dengan 0,25 sehingga didapatkan 1 atribut yaitu ‘h-7’. Sedangkan scenario pengujian kedua akan mencari pasangan atribut ‘h-7’ dengan cara melakukan pengamatan terhadap visualisasi grafik histogram.



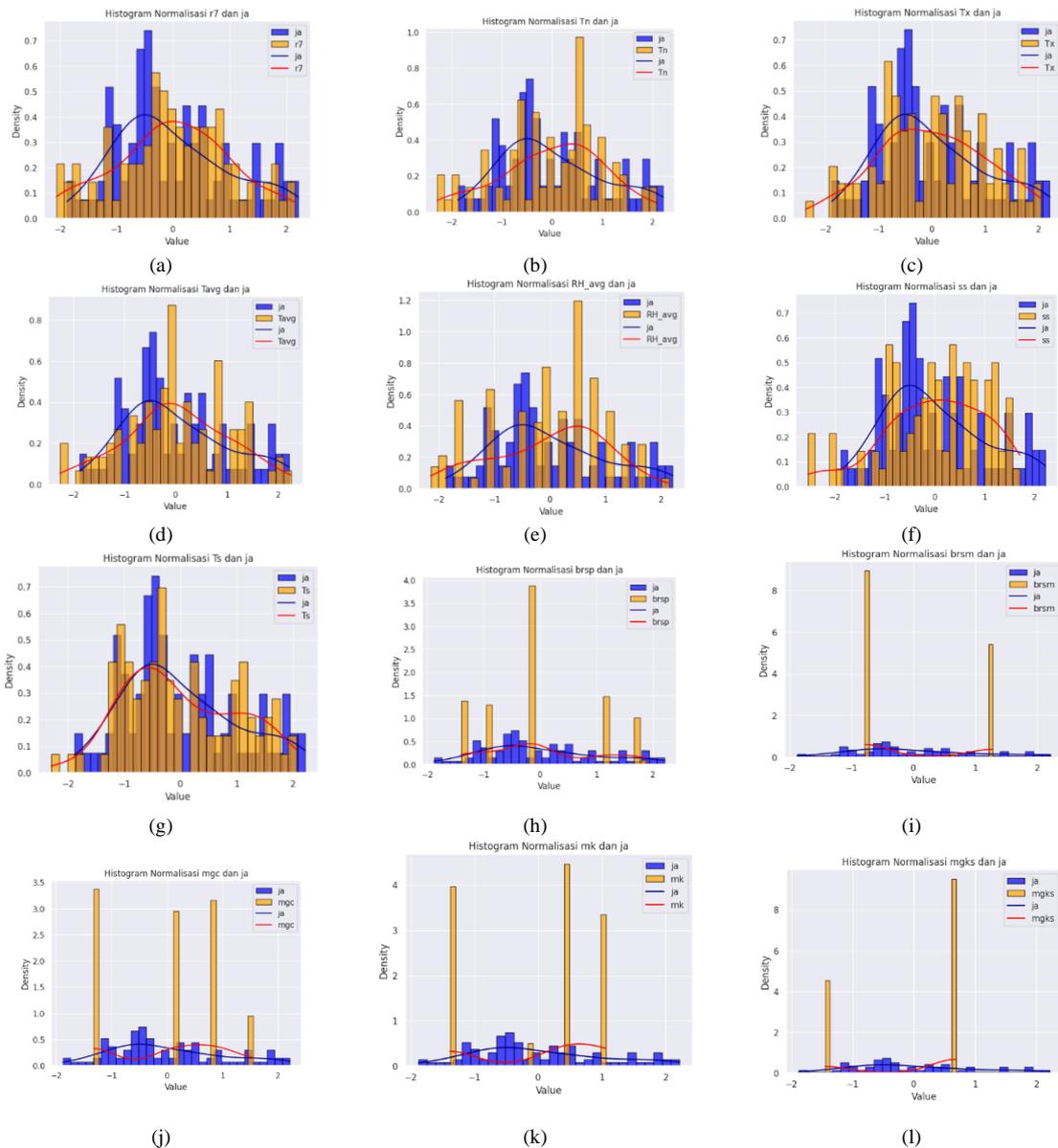
Gambar 4. Matriks korelasi antar atribut

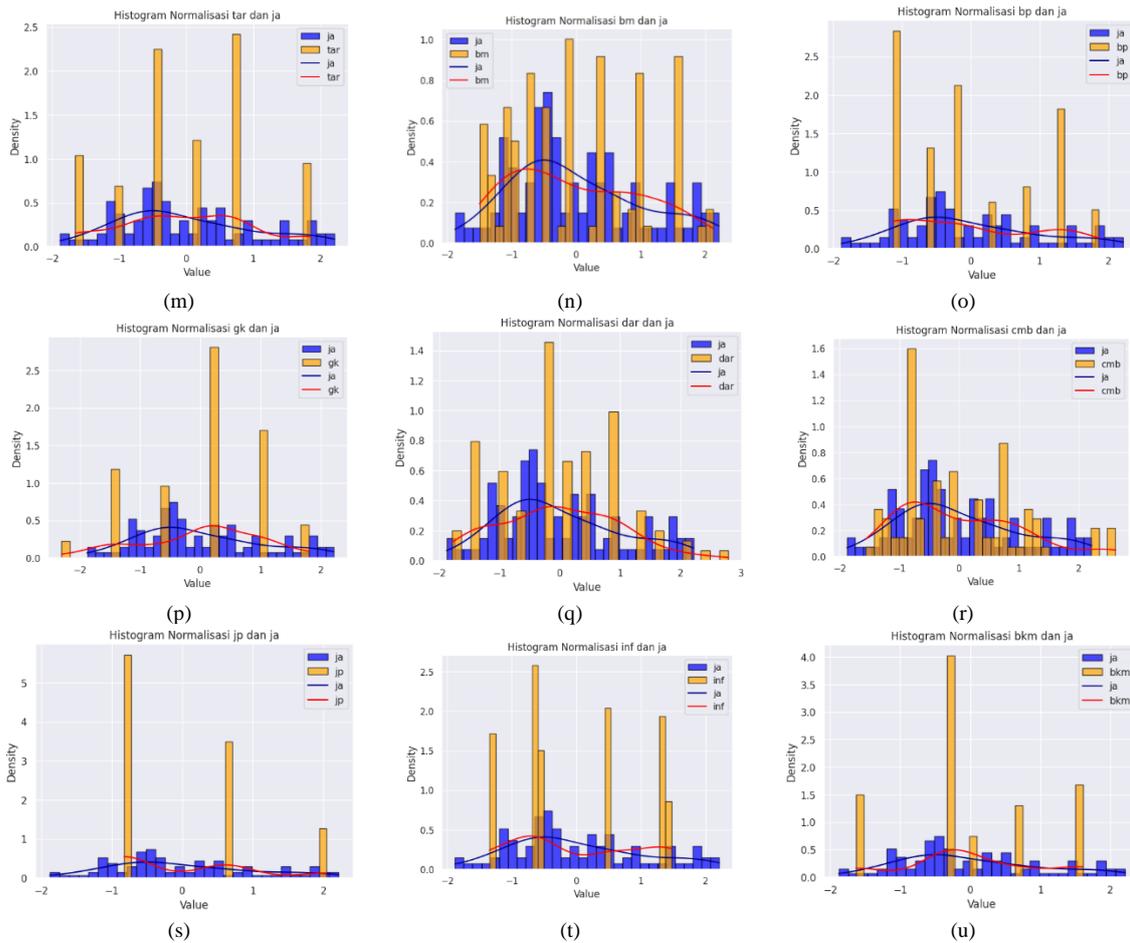
Tabel 1. Matriks korelasi antar atribut dengan atribut ‘jumlah analgesik per hari (ja)’

Atribut	Korelasi	Atribut	Korelasi	Atribut	Korelasi	Atribut	Korelasi
(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
ja	1,00	ss	-0,10	tar	-0,03	jp	0,06
h-7	0,35	Ts	0,06	bm	0,04	inf	-0,01
r7	-0,08	brsp	0,06	bp	-0,01	bkm	0,00
Tn	-0,06	brsm	-0,08	gk	-0,01	bkr	0,02
Tx	0,03	mgc	-0,01	dar	0,00	bkpr	0,10
Tavg	-0,05	mk	-0,03	cmb	0,08		
RH_avg	-0,03	mgks	0,00				

4.3 Pengamatan Visual grafik Plot dan Histogram

Selanjutnya mengamati visualisasi grafik histogram untuk mencermati kemungkinan adanya hubungan lain antara atribut prediktor dan atribut target pada gambar 5 untuk mencari pasangan atribut ‘h-7’.





Gambar 5. Grafik histogram ternormalisasi dengan StandardScaler() pada semua atribut prediktor yang akan dipasangkan dengan atribut 'h-7'

Dari gambar 8 tersebut bisa dicermati bahwasanya gambar 8 a, b, c, d, e, f, m, o, p, q, r, s, t, u, v tidak memiliki kecocokan pola dengan atribut 'ja'. Sedangkan gambar 8 h,i,j,k,l, terlihat data atribut prediktor memiliki jumlah yang terlalu kurang untuk mewakili hubungan dengan atribut 'ja'. Kemudian bila dicermati gambar 8 g, n dan x, dapat dilihat ada mayoritas kemiripan pola antara atribut prediktor dengan atribut 'ja'. Oleh karena itu penelitian ini juga menggunakan skenario kedua dengan membuat pasangan atribut atribut 'h-7' dengan 'Ts', 'h-7' dengan 'bm' dan 'h-7' dengan 'bkpr'.

4.4 Normalisasi Data

Normalisasi yang akan digunakan adalah normalisasi min-max. Dimana normalisasi ini bekerja berdasarkan jarak antara nilai terukur dengan nilai minimum dibagi rentang nilai. Menurut Larose (2005) [25], jaringan syaraf tiruan seperti MLPNN memerlukan input yang sudah dikodekan dalam rentang 0 hingga 1, Dimana masalah akan data yang bersifat kontinu bisa diatasi dengan normalisasi min-max karena jaringan syaraf terkenal akan kemampuan mentoleransi sedikit pelanggaran atas batas nilai minimum dan maksimum.

4.5 Model Prediksi dengan MLR dan MLPNN

4.5.1 Multiple Linear Regression

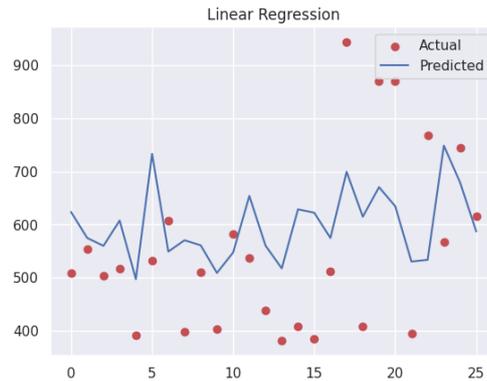
Multiple Linear regression merupakan algoritma yang memproses data dependent atau target berdasarkan data data independet (predictor) dengan memanfaatkan persamaan matematika linear [26]. Adapun kode program yang digunakan yaitu membagi Data ke dalam Data Training dan Data Test, dengan proporsi 80:20 dan diacak sebanyak 42 kali, Melihat dimensi array dan kerangka data. Shape akan memberikan informasi dalam bentuk tuple yang menunjukkan jumlah baris dan kolom, Membuat model regresi linier, Melatih model, Mendapatkan angka intercept, Mendapatkan angka angka koefisien persamaan linear, Mendapatkan nilai prediksi, Mendapatkan nilai evaluasi.

4.5.1.1 Atribut predictor ‘h-7’

Pada scenario pertama dengan atribut predictor hanya jumlah penjualan 7 hari sebelumnya (h-7), multiple linear regression mendapatkan dari python model.intercept = 405,215 dan model.coeficient = 0,328 maka bisa menerapkannya pada persamaan 1 sebagai berikut:

$$y = 405,215 + 0,328x_1$$

Dengan keterangan bahwa y merupakan nilai prediksi atau target atau dependent atribut, dan x₁ independent atribut ‘h-7’. Penjelasan dari model diatas dapat dijabarkan bahwa (a) Intercept (b₀)=405,215 merupakan nilai awal y ketika nilai x₁ bernilai nol, dan (b) Koefisien (b₁)= 0,328 berarti setiap kenaikan 1 unit pada x₁ akan menaikkan nilai prediksi y sebesar 0,328. Grafik nilai y actual dan nilai y prediksi ditunjukkan pada gambar 6.



Gambar 6. Grafik nilai y actual dan nilai y prediksi hasil model multiple linear regression dengan predictor yaitu ‘h-7’

Gambar 6 menunjukkan bahwa prediksi model (garis biru) tidak sepenuhnya sesuai dengan nilai aktual (titik merah). Beberapa prediksi jauh dari nilai aktual, menunjukkan model mengalami underfitting atau tidak mampu menangkap pola dengan baik. Penyebaran titik merah secara acak di sekitar garis biru mencerminkan variasi antara nilai aktual dan prediksi. Semakin dekat titik merah ke garis biru, semakin baik performa model. Namun, perbedaan besar di beberapa titik menunjukkan adanya error atau bias dalam model.

Nilai Evaluasi Model dapat dilihat pada penjelasan berikut:

1. Mean Squared Error (MSE): 22.544,477
MSE adalah rata-rata kuadrat dari error. Nilai besar ini menunjukkan prediksi sering meleset jauh dari nilai aktual, dengan RMSE sebesar 150,148, setara dengan rata-rata deviasi 150-unit analgesik per hari.
2. Mean Absolute Percentage Error (MAPE): 25,5%
MAPE menunjukkan rata-rata error prediksi sekitar 25,5% dari nilai aktual. Nilai ini tergolong tinggi, menandakan performa model masih perlu ditingkatkan.

4.5.1.2 Atribut predictor ‘h-7’ dan ‘suku bunga KPR (bkpr)’

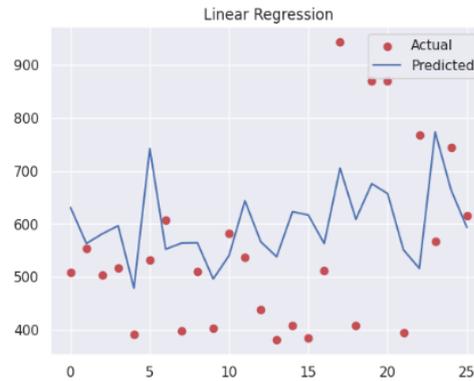
Pada scenario kedua dengan atribut predictor adalah jumlah penjualan 7 hari sebelumnya (h-7) dan ‘suku bunga KPR (kpr)’, multiple linear regression mendapatkan dari python model.intercept = -1904,898 dan model.coeficient “h-7” = 0,334 dan “bkpr” = 255,614 maka bisa menerapkannya pada persamaan 1 sebagai berikut

$$y = -1904,898 + 0,334(h - 7) + 255,614(bkpr)$$

Dengan keterangan y merupakan nilai prediksi atau target atau dependent atribut, x₁ = h-7, x₂ = bkpr, b₁ = koefisien x₁, b₂ = koefisien x₂, yang dapat dijelaskan seperti berikut:

1. Intercept (b₀)= -1904,898 merupakan nilai awal y ketika nilai x₁, x₂ bernilai nol.
2. Koefisien x₁ (b₁) = 0,334 berarti setiap kenaikan 1 unit pada x₁ akan menaikkan nilai prediksi y sebesar 0,335, dengan asumsi atribut lainnya tetap
3. Koefisien x₂ (b₂) = 255,614 berarti setiap kenaikan 1 unit pada x₂ akan menaikkan nilai prediksi y sebesar 255,614, dengan asumsi atribut lainnya tetap

Gambar 7 grafik nilai y actual dan nilai y prediksi hasil model multiple linear regression dengan predictor yaitu ‘h-6’ dan ‘suku bunga KPR (bkpr)’.



Gambar 7. grafik nilai y actual dan nilai y prediksi hasil model multiple linear regression dengan predictor yaitu 'h-7' berpasangan 'bkpr'

Grafik pada gambar 7 menunjukkan bahwa prediksi model (garis biru) tidak sepenuhnya sesuai dengan nilai sebenarnya (titik merah). Dalam beberapa bagian, model memberikan prediksi yang jauh dari nilai actual, menunjukkan kemungkinan model tidak mampu menangkap pola dengan baik atau ada underfitting. Titik merah yang tersebar secara acak di sekitar garis biru menunjukkan variasi nilai actual terhadap prediksi. Semakin dekat titik ke garis biru, semakin baik performa model pada data tersebut. Perbedaan yang besar antara nilai actual dan prediksi di beberapa titik menunjukkan adanya error atau bias dalam model. Adapun nilai-nilai evaluasi dapat dijelaskan sebagai berikut

1. Mean Squared Error (MSE): 22768,427

MSE adalah rata-rata kuadrat dari error (selisih antara nilai actual dan nilai prediksi). Dengan nilai MSE = 22768,427 ini berarti error cukup besar dalam prediksi. Besarnya MSE juga tergantung pada skala target, dan dalam hal ini, eror sekitar 22768,427 maka RMSE sebesar 150,89 menunjukkan bahwa prediksi sering meleset jauh dari nilai actual yaitu dikisaran 151-unit analgesic per hari

2. Mean Absolute Percentage Error (MAPE): 0.256 (atau 25,6%)

MAPE mengukur rata-rata persentase kesalahan prediksi terhadap nilai actual. Nilai MAPE 0.256 (atau 25,6%) berarti rata-rata eror model adalah sekitar 25,6% dari nilai actual yang dalam prakteknya 25,6% dianggap cukup tinggi.

4.5.1.3 Atribut predictor 'h-7' dan 'suku bunga Selisih Temperature maksimum-minimum per hari (Ts)'

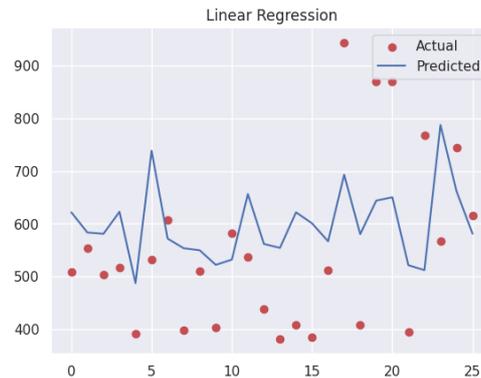
Pada scenario kedua dengan atribut predictor adalah jumlah penjuluan 7 hari sebelumnya (h-7) dan 'Selisih Temperature maksimum-minimum per hari (Ts)', multiple linear regression mendapatkan dari python model.intercept = 339,893 dan model.coeficient "h - 7" = 0,335 dan "Ts" = 9,765 maka bisa menerapkannya pada persamaan 1 sebagai berikut:

$$y = 339,893 + 0,335(h - 7) + 9,765(Ts)$$

Dengan keterangan y merupakan nilai prediksi atau target atau dependent atribut, x_1 merupakan atribut 'h-7', x_2 merupakan atribut 'Ts', b_1 merupakan koefisien x_1 , b_2 merupakan koefisien x_2 , yang dapat dijelaskan sebagai berikut:

1. Intercept (b_0)= 339,893 merupakan nilai awal y ketika nilai x_1 , x_2 bernilai nol.
2. Koefisien x_1 (b_1) = 0,335 berarti setiap kenaikan 1 unit pada x_1 akan menaikkan nilai prediksi y sebesar 0,335, dengan asumsi atribut lainnya tetap
3. Koefisien x_2 (b_2) = 9,765 berarti setiap kenaikan 1 unit pada x_2 akan menaikkan nilai prediksi y sebesar 9,765, dengan asumsi atribut lainnya tetap

Gambar 8 grafik nilai y actual dan nilai y prediksi hasil model multiple linear regression dengan predictor yaitu 'h-7' dan 'bm'.



Gambar 8. Grafik nilai y actual dan nilai y prediksi hasil model multiple linear regression dengan predictor yaitu 'h-7' berpasangan 'Ts'

Grafik pada gambar 8 menunjukkan bahwa prediksi model (garis biru) tidak sepenuhnya sesuai dengan nilai sebenarnya (titik merah). Dalam beberapa bagian, model memberikan prediksi yang jauh dari nilai aktual, menunjukkan kemungkinan model tidak mampu menangkap pola dengan baik atau ada underfitting. Titik merah yang tersebar secara acak di sekitar garis biru menunjukkan variasi nilai aktual terhadap prediksi. Semakin dekat titik ke garis biru, semakin baik performa model pada data tersebut. Perbedaan yang besar antara nilai aktual dan prediksi di beberapa titik menunjukkan adanya error atau bias dalam model. Adapun nilai-nilai evaluasi dapat dijelaskan sebagai berikut:

1. Mean Squared Error (MSE): 23.355

MSE adalah rata-rata kuadrat dari error (selisih antara nilai aktual dan nilai prediksi). Dengan nilai MSE = 23.355 ini berarti error cukup besar dalam prediksi. Besarnya MSE juga tergantung pada skala target, dan dalam hal ini, eror sekitar 23.355 maka RMSE sebesar 152,8 menunjukkan bahwa prediksi sering meleset jauh dari nilai actual yaitu dikisaran 153 unit analgesic per hari

2. Mean Absolute Percentage Error (MAPE): 0.258 (atau 25,8%)

MAPE mengukur rata-rata persentase kesalahan prediksi terhadap nilai aktual. Nilai MAPE 0.258 (atau 25,8%) berarti rata-rata eror model adalah sekitar 25,8% dari nilai actual yang dalam prakteknya 25,8% dianggap cukup tinggi.

4.5.1.4 Atribut predictor 'h-7' dan 'harga bawang merah (bm)'

Pada skenario kedua dengan atribut predictor adalah jumlah penjualan 7 hari sebelumnya (h-7) dan 'harga bawang merah (bm)', multiple linear regression mendapatkan dari python model.intercept = 392,5605 dan model.coeficient "h - 7" = 0,3256 dan "bm" = 0,00040475. maka bisa menerapkannya pada persamaan 1 sebagai berikut:

$$y = 392,5605 + 0,3256 (h - 7) + 0,0004(bm)$$

Dengan keterangan y merupakan nilai prediksi atau target atau dependent atribut, x_1 merupakan atribut 'h-7', x_2 merupakan atribut, 'bm', b_1 merupakan koefisien x_1 , b_2 merupakan koefisien x_2 , dengan penjelasan: Intercept (b_0)= 392,8861 merupakan nilai awal y ketika nilai x_1 , x_2 bernilai nol. Koefisien x_1 (b_1) = 0,326 berarti setiap kenaikan 1 unit pada x_1 akan menaikkan nilai prediksi y sebesar 0,326, dengan asumsi atribut lainnya tetap. Koefisien x_2 (b_2) = 0,0004 berarti setiap kenaikan 1 unit pada x_1 akan menaikkan nilai prediksi y sebesar 0,0004, dengan asumsi atribut lainnya tetap. Gambar 9 merupakan grafik nilai y actual dan nilai y prediksi hasil model multiple linear regression dengan predictor yaitu 'h-7' dan 'bm'.

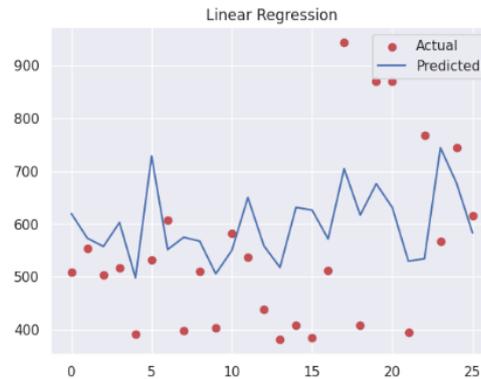
Grafik pada gambar 9 menunjukkan bahwa prediksi model (garis biru) tidak sepenuhnya sesuai dengan nilai sebenarnya (titik merah). Dalam beberapa bagian, model memberikan prediksi yang jauh dari nilai aktual, menunjukkan kemungkinan model tidak mampu menangkap pola dengan baik atau ada underfitting. Titik merah yang tersebar secara acak di sekitar garis biru menunjukkan variasi nilai aktual terhadap prediksi. Semakin dekat titik ke garis biru, semakin baik performa model pada data tersebut. Perbedaan yang besar antara nilai aktual dan prediksi di beberapa titik menunjukkan adanya error atau bias dalam model. Adapun nilai-nilai evaluasi dapat dijelaskan sebagai berikut:

1. Mean Squared Error (MSE): 22346,129

MSE adalah rata-rata kuadrat dari error (selisih antara nilai aktual dan nilai prediksi). Dengan nilai MSE = 22346,129 ini berarti error cukup besar dalam prediksi. Besarnya MSE juga tergantung pada skala target, dan dalam hal ini, eror sekitar 22346,129 maka RMSE sebesar 149,486 menunjukkan bahwa prediksi sering meleset jauh dari nilai actual yaitu dikisaran 150 unit analgesic per hari.

2. Mean Absolute Percentage Error (MAPE): 0.254 (atau 25,4%)

MAPE mengukur rata-rata persentase kesalahan prediksi terhadap nilai aktual. Nilai MAPE 0.254 (atau 25,4%) berarti rata-rata eror model adalah sekitar 25,4% dari nilai aktual yang dalam prakteknya 25,4% dianggap cukup tinggi.



Gambar 9. Grafik nilai y actual dan nilai y prediksi hasil model multiple linear regression dengan predictor yaitu 'h-7' berpasangan 'bm'

4.5.1.5 Kesimpulan dari 2 skenario pengujian model Regresi Linear

Model Regresi Linear bekerja dengan kurang memuaskan dengan hasil yang relatif sama pada kedua skenario sebagaimana resume pada table 9. Beberapa tantangan yang dapat memengaruhi kualitas dan akurasi pemodelan data antara lain adalah kurangnya jumlah data pelatihan, yang dapat menyebabkan model tidak mampu belajar pola yang representatif dari data. Selain itu, keberadaan outlier atau noise dalam data juga dapat mengganggu proses korelasi antaratribut dan menurunkan performa model. Atribut-atribut prediktor yang digunakan pun mungkin kurang informatif, sehingga tidak memberikan kontribusi signifikan terhadap proses prediksi. Misalnya, pada skenario kedua, atribut seperti 'bkpr', 'Ts', dan 'bm' menunjukkan korelasi yang terlalu lemah terhadap atribut target 'h-7', sehingga tidak cukup membantu dalam membangun model prediksi yang akurat. Hasil evaluasi linear regression dapat ditunjukkan pada tabel 2.

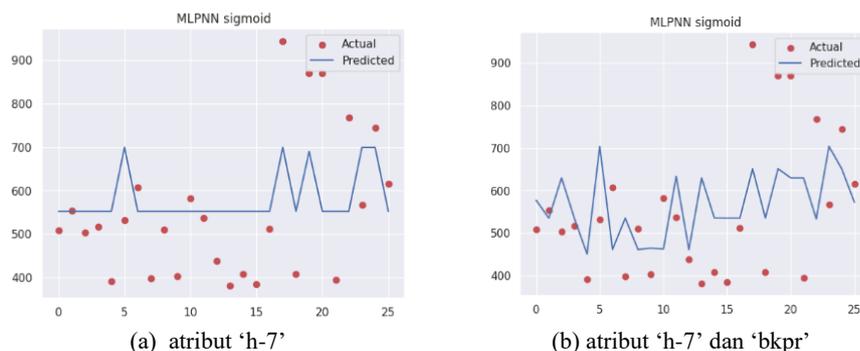
Tabel 2. Hasil evaluasi linear regression

Item	MLR "h-7", "bm"	MLR "h-7", "Ts"	MLR "h-7", "bkpr"	MLR "h-7"
MSE	22346,129	23355,282	22768,427	22544,477
MAPE	0,254	0,258	0,256	0,255

4.5.2 Multi-Layer Perceptron Neural Network (MLPNN)

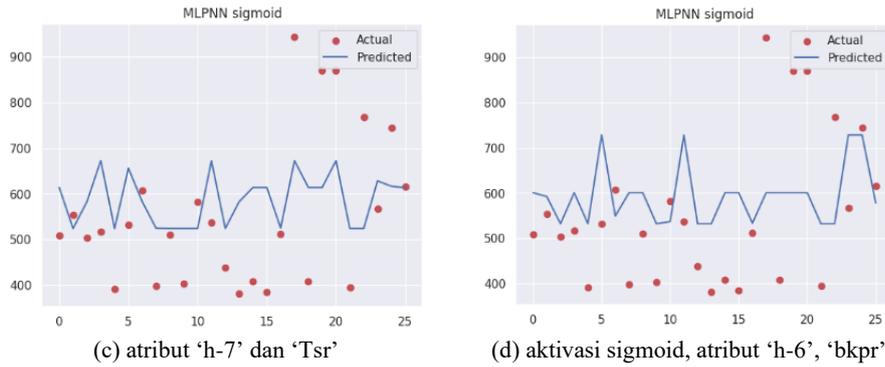
Menurut Desai & Shah (2021), MLPNN adalah algoritma jaringan syaraf tiruan dengan arsitektur feedforward yang meningkatkan akurasi melalui backward propagation, yaitu mekanisme pembelajaran dengan mereduksi error pada bobot. MLPNN terdiri dari input layer, satu atau lebih hidden layers, dan output layer, memungkinkan algoritma ini mempelajari pola data non-linear secara efektif untuk prediksi yang kompleks [29].

Dalam penelitian ini, MLPNN dengan backpropagation diterapkan menggunakan library TensorFlow, yang menyediakan alat sederhana untuk implementasi jaringan syaraf tiruan, termasuk backpropagation secara implisit [27]. Implementasi dilakukan dengan menggunakan Google Colab.# mengimpor library, modul dan kelas. Hasil Pengujian program python di atas dalam 4 skenario yang ditunjukkan pada gambar 10.



(a) atribut 'h-7'

(b) atribut 'h-7' dan 'bkpr'



Gambar 10. Grafik hasil pemodelan MLPNN backpropagation

Grafik pada gambar 10 menunjukkan bahwa prediksi model (garis biru) tidak sepenuhnya sesuai dengan nilai sebenarnya (titik merah). Dalam beberapa bagian, model memberikan prediksi yang jauh dari nilai actual. Hal ini menunjukkan kemungkinan model tidak mampu menangkap pola dengan baik atau ada underfitting. Titik merah yang tersebar secara acak di sekitar garis biru menunjukkan variasi nilai actual terhadap prediksi. Semakin dekat titik ke garis biru, semakin baik performa model pada data tersebut. Perbedaan yang besar antara nilai actual dan prediksi di beberapa titik menunjukkan adanya error atau bias dalam model. Evaluasi Model dari percobaan tersebut dapat dilihat pada tabel 3 dan dijelaskan sebagai berikut:

1. MSE (19,000–27,000): Prediksi model menunjukkan error besar dengan nilai terkecil 19,588, setara dengan RMSE 140, berarti meleset hingga 140-unit analgesik/hari.
2. MAPE (>22%): Rata-rata error model mencapai 22,3%, yang dianggap cukup tinggi untuk akurasi prediksi.

Tabel 3. Hasil evaluasi model MLPNN

	'h-7'	'h-7' dan 'bkpr'
MSE MLPNN	19588,936	21893,260
MAPE MLPNN	0,223	0,233
	'h-7', 'Ts'	'h-7', 'bm'
MSE MLPNN	24.999	26821,830
MAPE MLPNN	0,287	0,269

4.5.3 Perbandingan antara model MLR dengan MLPNN Backpropagation

Berdasarkan table 4 yang menampilkan perbandingan hasil evaluasi antara MLPNN backpropagation dan MLR, terlihat tidak ada model yang benar benar bisa menghasilkan model yang baik. Semua model menghasilkan MSE di kisaran lebih dari 19.000 yang berarti prediksi bisa meleset lebih dari 137-an butir atau rentang kesalahan prediksi 22,3% .

Tabel 4. Perbandingan antara model MLPNN dan MLR

Item	MSE	MAPE
MLPNN-sigmoid-linear "h-7"	19588,936	0,223
MLPNN-sigmoid-linear "h-7", "bkpr"	21893,260	0,233
MLR "h-7", "bm"	22346,129	0,254
MLR "h-7"	22544,477	0,255
MLR "h-7", "bkpr"	22768,427	0,256
MLR "h-7", "Ts"	23355,282	0,258
MLPNN-sigmoid-linear "h-7", "Ts"	24999,000	0,287
MLPNN-sigmoid-linear "h-7", "bm"	26821,830	0,264

4.5.4 Pengaruh Atribut Independen terhadap Model

Hasil evaluasi pada Tabel 11 menunjukkan bahwa tidak ada atribut independen yang memberikan pengaruh signifikan pada model. Regresi Linear gagal mendeteksi hubungan linear, sedangkan MLPNN Backpropagation juga tidak berhasil meskipun mampu menangkap hubungan non-linear [28]. Satu-satunya atribut yang menunjukkan potensi adalah "h-7," yang merepresentasikan penjualan 7 hari sebelumnya dengan korelasi cukup tinggi.

4.5.5 Validasi Temuan dengan Literatur

Temuan ini sejalan dengan Altamevia & Wijaya (2023) [29] yang menyoroti keunggulan neural network dalam prediksi stok farmasi. Tampati et al. (2022) [4] juga menegaskan pentingnya model hybrid untuk

meningkatkan akurasi. Penelitian ini mengusulkan eksplorasi atribut internal dan eksternal lainnya serta pengumpulan data dari wilayah yang lebih luas dengan sistem terintegrasi untuk mendapatkan data yang lebih komprehensif.

5 KESIMPULAN

Penelitian ini membandingkan performa Multiple Linear Regression (MLR) dan Multi-Layer Perceptron Neural Network (MLPNN) dalam memprediksi penjualan obat analgesik di Apotek XYZ. Hasil menunjukkan bahwa MLPNN dengan 1 hidden layer ($8 \times n$) +1 hidden node dan aktivasi sigmoid-linear lebih unggul dalam menangani pola data non-linear dibandingkan MLR, terlihat dari nilai MSE dan MAPE yang lebih rendah. Namun, kedua model masih memiliki keterbatasan dalam menghasilkan prediksi yang akurat, dengan nilai MSE di atas 19.000 dan rentang kesalahan prediksi mencapai 22,3%, yang mengindikasikan adanya underfitting.

Keterbatasan penelitian ini terletak pada kurangnya atribut independen yang memiliki pengaruh signifikan terhadap model, sebagaimana ditunjukkan dalam analisis korelasi. Meskipun MLPNN mampu menangkap hubungan non-linear, model ini tetap tidak dapat memberikan prediksi yang optimal. Selain itu, keterbatasan dalam cakupan data, seperti hanya menggunakan data dari satu apotek tanpa mempertimbangkan faktor eksternal seperti tren musiman atau kebijakan harga dari supplier, turut memengaruhi hasil prediksi. Untuk penelitian selanjutnya, direkomendasikan eksplorasi model hybrid yang menggabungkan teknik statistik dan deep learning untuk meningkatkan akurasi prediksi, sebagaimana disarankan oleh penelitian sebelumnya. Selain itu, pengumpulan data yang lebih luas dengan mempertimbangkan faktor eksternal seperti cuaca, harga bahan pokok, dan tren pasar dapat membantu memperbaiki kualitas model prediktif. Implementasi sistem database farmasi terintegrasi juga diperlukan untuk memastikan ketersediaan data historis yang lebih kaya dan mendukung prediksi yang lebih akurat dalam manajemen stok obat.

REFERENSI

- [1] G. Emmanuel, A. N. Hidayanto, and H. Prabowo, "A Literature Review: Drug Supply Chain and Neural Networks," in *2022 6th International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE)*, IEEE, 2022, pp. 559–564. <http://dx.doi.org/10.21203/rs.3.rs-3932079/v1>
- [2] A. Nguyen, O. Bougacha, B. Lekens, S. Lamouri, R. Pellerin, and C. Couvreur, "On the use of logistics data to anticipate drugs shortages through data mining," *Procedia Comput Sci*, vol. 219, pp. 949–956, 2023. <https://doi.org/10.1016/j.procs.2023.01.371>
- [3] European Association of Hospital Pharmacists, "EAHP Medicines Shortages Surveys," European Association of Hospital Pharmacists. <https://doi.org/10.1136/ejhpharm-2024-004090>
- [4] N. Tampati, "Prediksi Stok Obat pada Apotik Total Life Clinic Menggunakan Model Kombinasi Artificial Neural Network dan ARIMA," *JATISI (Jurnal Teknik Informatika Dan Sistem Informasi)*, vol. 9, no. 1, pp. 49–58, 2022. <https://doi.org/10.35957/jatisi.v9i1.1373>
- [5] M. Du, J. Luo, S. Wang, and S. Liu, "Genetic algorithm combined with BP neural network in hospital drug inventory management system," *Neural Comput Appl*, vol. 32, pp. 1981–1994, 2020. <https://link.springer.com/article/10.1007/s00521-019-04379-3>
- [6] R. Rathipriya, A. A. Abdul Rahman, S. Dhamodharavadhani, A. Meero, and G. Yoganandan, "Demand forecasting model for time-series pharmaceutical data using shallow and deep neural network model," *Neural Comput Appl*, vol. 35, no. 2, pp. 1945–1957, 2023. <https://doi.org/10.1007/s00521-022-07889-9>
- [7] S. Sahoo and M. K. Jha, "Groundwater-level prediction using multiple linear regression and artificial neural network techniques: a comparative assessment," *Hydrogeol J*, vol. 21, no. 8, p. 1865, 2013. <http://dx.doi.org/10.1007/s10040-013-1029-5>
- [8] S. Romano, H. Galante, D. Figueira, Z. Mendes, and A. T. Rodrigues, "Time-trend analysis of medicine sales and shortages during COVID-19 outbreak: Data from community pharmacies," *Research in Social and Administrative Pharmacy*, vol. 17, no. 1, pp. 1876–1881, 2021. <http://dx.doi.org/10.1016/j.sapharm.2020.05.024>
- [9] D. Saputra, "Analisis Pengendalian Internal Dalam Pengelolaan Persediaan Obat Pada Apotek Di Pekanbaru," *Jurnal Akuntansi dan Ekonomika*, vol. 12, no. 1, pp. 23–33, 2022. <https://doi.org/10.37859/jae.v12i1.3189>
- [10] V. Reinstadler *et al.*, "Monitoring drug consumption in Innsbruck during coronavirus disease 2019 (COVID-19) lockdown by wastewater analysis," *Science of the Total Environment*, vol. 757, p. 144006, 2021. <https://doi.org/10.1016/j.scitotenv.2020.144006>
- [11] D. A. Maulana, "Penggunaan Multiple Linear Regression Untuk Memprediksi Jumlah Kasus Campak Di Indonesia," Universitas Islam Sultan Agung Semarang, 2024.

- [12] S. Pratista, "Perbandingan Teknik Prediksi Pemakaian Obat Menggunakan Algoritma Simple Linear Regression dan Support Vector Regression," *Journal Building of Informatics, Technology and Science (BITS)*, vol. 5, no. 2, pp. 456–465, 2023. <https://doi.org/10.47065/bits.v5i2.4260>.
- [13] L. M. Ginting, M. M. T. Sigiyo, E. D. Manurung, and J. J. P. Sinurat, "Perbandingan Metode Algoritma Support Vector Regression dan Multiple Linear Regression Untuk Memprediksi Stok Obat," *J. Appl. Technol. Informatics Indones*, vol. 1, no. 2, pp. 29–34, 2021. <https://doi.org/10.47065/comforch.v3i1.1063>.
- [14] S. Agnesti, A. Nazir, I. Iskandar, E. Budianita, and I. Afrianty, "Comparison of Triple Exponential Smoothing and Support Vector Regression Algorithms in Predicting Drug Usage at Puskesmas," *Sistemasi: Jurnal Sistem Informasi*, vol. 12, no. 3, pp. 996–1006, 2023. <https://doi.org/10.32520/stmsi.v12i3.3499>.
- [15] P. H. Artanti, "Penerapan Neural Network dengan optimasi Ant Colony Optimization dan Backpropagation untuk membangun model prediksi diabetes tahap awal," Universitas Islam Negeri Maulana Malik Ibrahim, 2023. <https://doi.org/10.47065/josh.v6i1.6018>.
- [16] D. Yustisio, "Perbandingan Metode Extreme Learning Machine (ELM) dan Metode Multilayer Perceptron (MLP) Dalam Prediksi Jumlah Pasien Covid-19 Kota Semarang". <https://eprints.walisongo.ac.id/id/eprint/18407>.
- [17] D. B. Arianto, "Pengembangan Model Named Entity Recognition Untuk Pengenalan Entitas Pada Data Obat Indonesia," 2023.
- [18] E. T. L. Kusriani and E. Taufiq, "Algoritma data mining," *Yogyakarta: Andi Offset*, 2009.
- [19] D. T. Larose, "An introduction to data mining," *Traduction et adaptation de Thierry Vallaud*, p. 23, 2005.
- [20] X. Long and W. Chen, "Construction framework of smart tourism big data mining model driven by blockchain technology," *Heliyon*, vol. 10, no. 14, 2024. <https://doi.org/10.1016/j.heliyon.2024.e34159>.
- [21] A. Alabrah, "An improved CCF detector to handle the problem of class imbalance with outlier normalization using IQR method," *Sensors*, vol. 23, no. 9, p. 4406, 2023. <https://doi.org/10.3390/s23094406>.
- [22] D. Kurniawan, *Pengenalan machine learning dengan python*. Elex Media Komputindo, 2022.
- [23] M.-C. Popescu, V. E. Balas, L. Perescu-Popescu, and N. Mastorakis, "Multilayer perceptron and neural networks," *WSEAS Transactions on Circuits and Systems*, vol. 8, no. 7, pp. 579–588, 2009.
- [24] Z. Lyu *et al.*, "Back-propagation neural network optimized by K-fold cross-validation for prediction of torsional strength of reinforced Concrete beam," *Materials*, vol. 15, no. 4, p. 1477, 2022. <https://doi.org/10.3390/ma15041477>.
- [25] Larose, "Discovering Knowledge in Data: An Introduction to Data Mining," 2005.
- [26] S. Rath, A. Tripathy, and A. R. Tripathy, "Prediction of new active cases of coronavirus disease (COVID-19) pandemic using multiple linear regression model," *Diabetes & metabolic syndrome: clinical research & reviews*, vol. 14, no. 5, pp. 1467–1474, 2020. <https://doi.org/10.1016/j.dsx.2020.07.045>.
- [27] T. Filimonova, H. Samoylenko, A. Selivanova, Y. Yurchenko, and A. Parashchak, "Construction Of A Neural Network For Handwritten Digits Recognition Based On Tensorflow Library Applying An Error Backpropagation Algorithm.," *Eastern-European Journal of Enterprise Technologies*, vol. 126, no. 2, 2023. <https://doi.org/10.15587/1729-4061.2023.293682>.
- [28] M. Desai and M. Shah, "An anatomization on breast cancer detection and diagnosis employing multi-layer perceptron neural network (MLP) and Convolutional neural network (CNN)," *Clinical eHealth*, vol. 4, pp. 1–11, 2021. <https://doi.org/10.1016/j.ceh.2020.11.002>.
- [29] F. Altamevia, H. O. L. Wijaya, and E. Elmayati, "Analisis Pola Penjualan Obat di Apotek Srikandi Menggunakan Algoritma Supervised Learning," *Kesatria: Jurnal Penerapan Sistem Informasi (Komputer dan Manajemen)*, vol. 4, no. 1, pp. 170–176, 2023. <https://doi.org/10.30645/kesatria.v4i3.202>.