Journal of Science and Social Research ISSN 2615 – 4307 (Print)
Oct 2026, IX (5): 8268 – 8278 ISSN 2615 – 3262 (Online)
Available online at http://jurnal.goretanpena.com/index.php/JSSR

PREDIKSI KEBUTUHAN OBAT PADA INSTALASI FARMASI RUMAH SAKIT MENGGUNAKAN METODE RANDOM FOREST BERBASIS DATA HISTORIS TRANSAKSI OBAT

Ahmad Junaidi*[1], Yeviki Maisyah Putra[2]
Universitas Putra Indonesia YPTK Padang, Sumatera Barat
e-mail: [1]ahmad_junaidi@upiyptk.ac.id*, [2]yeviki.maisyahputra@gmail.com

Abstract: *Effective drug inventory management in hospital pharmacies requires accurate demand planning to minimize the risks of drug shortages and excess inventory. Historical drug transaction data stored in a Hospital Information System (SIMRS) can be utilized as a basis for predicting drug requirements for subsequent periods. This study aims to develop a drug demand prediction model using the Random Forest Regression method based on historical drug transaction data from a hospital pharmacy. The dataset consists of drug transaction records from August 2024 to December 2025 obtained from the hospital information system. The data were processed through preprocessing stages, including data cleaning, missing-value checking, duplicate detection, transaction validation, and monthly aggregation of drug usage. Subsequently, feature engineering was performed by generating historical usage features in the form of lag variables and moving averages as model inputs. The dataset was divided chronologically into training and testing sets to prevent data leakage. The Random Forest Regression model was employed to predict drug requirements for the subsequent period and evaluated using Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Weighted Absolute Percentage Error (WAPE)***.** *The results indicate that a historical transaction-based approach can be used to estimate drug requirements and provide supporting information for inventory planning in hospital pharmacies. The integration of the prediction model into SIMRS can further support pharmacy staff in planning drug requirements based on historical usage patterns*

Keyword: *drug demand prediction, Random Forest Regression, hospital information system, hospital pharmacy, machine learning, drug transactions*

Abstrak: Pengelolaan persediaan obat di instalasi farmasi rumah sakit memerlukan perencanaan kebutuhan yang tepat untuk mengurangi risiko kekurangan maupun kelebihan stok. Pemanfaatan data historis transaksi obat pada Sistem Informasi Manajemen Rumah Sakit (SIMRS) dapat digunakan sebagai dasar untuk memprediksi kebutuhan obat pada periode berikutnya. Penelitian ini bertujuan untuk mengembangkan model prediksi kebutuhan obat menggunakan metode Random Forest Regression berbasis data historis transaksi obat pada instalasi farmasi rumah sakit. Data penelitian berupa transaksi obat pada periode Agustus 2024 sampai Desember 2025 yang diperoleh dari sistem SIMRS. Data melalui tahapan preprocessing yang meliputi pembersihan data, pemeriksaan data kosong, duplikasi, validasi transaksi, dan agregasi penggunaan obat secara bulanan. Selanjutnya dilakukan *feature engineering* dengan membentuk fitur historis berupa *lag* penggunaan obat dan *moving average* sebagai variabel masukan model. Data dibagi secara kronologis menjadi data pelatihan dan pengujian untuk menghindari *data leakage*. Model Random Forest Regression digunakan untuk menghasilkan prediksi kebutuhan obat pada periode berikutnya dan dievaluasi menggunakan Mean Absolute Error (MAE), Root Mean Square Error (RMSE), dan Weighted Absolute Percentage Error (WAPE). Hasil penelitian menunjukkan bahwa pendekatan berbasis pola historis transaksi dapat digunakan untuk menghasilkan estimasi kebutuhan obat dan memberikan informasi pendukung bagi perencanaan persediaan pada instalasi farmasi. Implementasi model pada SIMRS dapat dikembangkan sebagai sistem

pendukung keputusan untuk membantu petugas farmasi dalam melakukan perencanaan kebutuhan obat berdasarkan pola penggunaan historis

Kata kunci: prediksi kebutuhan obat, Random Forest Regression, SIMRS, instalasi farmasi, *machine learning*, transaksi obat

PENDAHULUAN

Pengelolaan obat merupakan salah satu bagian penting dalam pelayanan kefarmasian di rumah sakit. Ketersediaan obat yang sesuai dengan kebutuhan sangat diperlukan untuk menjamin kesinambungan pelayanan kepada pasien. Ketidaktepatan dalam perencanaan kebutuhan dapat menyebabkan terjadinya kekurangan obat (*stockout*) yang berpotensi mengganggu pelayanan, maupun kelebihan persediaan yang dapat meningkatkan biaya penyimpanan dan risiko obat mengalami kedaluwarsa. Oleh karena itu, diperlukan metode perencanaan kebutuhan obat yang mampu memanfaatkan data penggunaan sebelumnya sebagai dasar pengambilan keputusan.

Perkembangan Sistem Informasi Manajemen Rumah Sakit (SIMRS) memungkinkan rumah sakit menghasilkan dan menyimpan data transaksi obat dalam jumlah besar secara terstruktur. Data tersebut tidak hanya berfungsi sebagai catatan transaksi, tetapi juga dapat dimanfaatkan sebagai sumber informasi untuk mendukung proses analisis dan perencanaan persediaan. Salah satu pendekatan yang dapat diterapkan adalah pemanfaatan *machine learning* untuk menemukan pola penggunaan obat dari data historis sehingga dapat menghasilkan estimasi kebutuhan pada periode berikutnya.

Salah satu metode *machine learning* yang dapat digunakan untuk permasalahan prediksi adalah Random Forest Regression**.** Metode ini membangun sejumlah *decision tree* dan menggabungkan hasil prediksi dari setiap pohon untuk menghasilkan nilai prediksi akhir. Dalam konteks kebutuhan obat, model dapat memanfaatkan pola penggunaan pada periode sebelumnya melalui pembentukan fitur historis seperti *lag* penggunaan obat dan *moving average*. Pendekatan tersebut memungkinkan model mempertimbangkan pola penggunaan sebelumnya dalam menghasilkan prediksi kebutuhan pada periode berikutnya.

Pada penelitian ini digunakan data historis transaksi obat dari instalasi farmasi rumah sakit yang diperoleh dari SIMRS dengan periode Agustus 2024 sampai Desember 2025. Data transaksi diolah melalui tahapan *preprocessing*, meliputi pembersihan data, pemeriksaan data kosong dan duplikasi, validasi transaksi, serta agregasi penggunaan obat secara bulanan. Selanjutnya dilakukan *feature engineering* dengan membentuk variabel historis penggunaan obat, antara lain *lag* 1, 2, 3, dan 6 periode serta *moving average*. Data kemudian dibagi secara kronologis menjadi data *training* dan *testing* agar proses evaluasi mempertahankan urutan waktu dan mengurangi risiko *data leakage*.

Model Random Forest Regression selanjutnya digunakan untuk memprediksi kebutuhan obat pada periode berikutnya. Kinerja model dievaluasi menggunakan Mean Absolute Error (MAE)**,** Root Mean Square Error (RMSE)**,** dan Weighted Absolute Percentage Error (WAPE)**.** Selain evaluasi model, penelitian ini juga mempertimbangkan proses optimasi *hyperparameter* untuk memperoleh konfigurasi model yang sesuai dengan karakteristik data transaksi obat.

Berdasarkan permasalahan tersebut, penelitian ini bertujuan untuk mengembangkan model prediksi kebutuhan obat pada instalasi farmasi rumah sakit menggunakan metode Random Forest Regression berbasis data historis transaksi obat. Hasil penelitian

Journal of Science and Social Research ISSN 2615 – 4307 (Print)
Oct 2026, IX (5): 8268 – 8278 ISSN 2615 – 3262 (Online)
Available online at http://jurnal.goretanpena.com/index.php/JSSR

diharapkan dapat memberikan informasi prediktif yang dapat dimanfaatkan sebagai pendukung perencanaan kebutuhan dan pengendalian persediaan obat. Integrasi hasil prediksi ke dalam SIMRS juga diharapkan dapat mendukung proses pengambilan keputusan berbasis data dalam pengelolaan persediaan obat di instalasi farmasi.

METODE

Desain dan Tahapan Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan metode eksperimen untuk membangun model prediksi kebutuhan obat pada instalasi farmasi rumah sakit menggunakan algoritma Random Forest Regression. Penelitian memanfaatkan data historis transaksi obat yang diperoleh dari Sistem Informasi Manajemen Rumah Sakit (SIMRS). Tahapan penelitian meliputi pengumpulan data, preprocessing, feature engineering, pembagian data training dan testing, pembangunan model Random Forest, evaluasi model, dan menghasilkan prediksi kebutuhan obat

Gambar 1 Alur Penelitian Prediksi Kebutuhan Obat Menggunakan Random Forest

Berdasarkan Gambar 1, penelitian diawali dengan pengumpulan data historis transaksi obat yang berasal dari SIMRS. Data selanjutnya melalui tahap preprocessing yang meliputi cleaning, pemeriksaan missing value, identifikasi data duplikat, dan agregasi transaksi berdasarkan obat dan periode. Setelah data siap digunakan, dilakukan feature engineering dengan membentuk fitur lag dan moving average untuk merepresentasikan pola penggunaan obat pada periode sebelumnya. Dataset kemudian dibagi menjadi data training dan testing berdasarkan urutan waktu. Data training digunakan untuk membangun model Random Forest Regression, sedangkan data testing digunakan untuk mengukur kemampuan prediksi model. Evaluasi dilakukan menggunakan MAE, RMSE, dan MAPE atau WAPE. Hasil akhir berupa prediksi kebutuhan obat yang dapat digunakan sebagai informasi pendukung perencanaan persediaan dan pengadaan obat pada instalasi farmasi.

Sumber dan Pengolahan Dataseht

Data yang digunakan dalam penelitian ini merupakan data sekunder berupa data historis transaksi obat yang diperoleh dari Sistem Informasi Manajemen Rumah Sakit (SIMRS) pada instalasi farmasi. Data transaksi digunakan untuk memperoleh pola penggunaan obat pada periode sebelumnya yang selanjutnya menjadi dasar dalam membangun model prediksi kebutuhan obat menggunakan algoritma Random Forest Regression.

Tabel 1 Variabel Dataset Penelitian

No	Variabel	Tipe Data	Keterangan
1	id_obat	string	Identitas Obat
2	Nama_ obat	string	Nama obat
3	Tgl_ transaksi	Datetime	Tanggal Transaksi
4	Jumlah	Numerik	Jumlah Obat
5	Satuan	String	Satuan Obat
6	Unit	String	Unit/Depo
7	Periode	Date	Dalam bentuk Bulan
8	Total_ penggunaan	Numerik	Dalam satu periode

Secara umum proses agrgasi dapat

dinyatakan sebagai : $D_{i,t} \sum_{j=1}^{n} \quad Q_{i,t,j}$, Dengan D$i,t$ merupakan total penggunaan obat ke-i pada periode t, sedangkan Qi,t,j merupakan jumlah obat pada transaksi ke-j.

Data transaksi obat terlebih dahulu melalui tahap preprocessing untuk memastikan kualitas data sebelum digunakan dalam pemodelan. Tahapan preprocessing meliputi data cleaning, pemeriksaan missing value, pemeriksaan duplikasi, validasi jenis transaksi, dan agregasi data.

Data cleaning dilakukan dengan memeriksa identitas obat, tanggal transaksi, dan jumlah penggunaan. Missing value pada viabel utama diperiksa dan ditangani berdasarkan karakteristik data. Data duplikat diidentifikasi berdasarkan identitas transaksi yang tersedia pada SIMRS untuk mencegah terjadinya perhitungan penggunaan obat secara berulang.

Selanjutnya dilakukan validasi terhadap jenis transaksi. Transaksi yang merepresentasikan penggunaan atau pengeluaran obat digunakan dalam penelitian, sedangkan transaksi pembatalan, retur, atau transaksi lain yang tidak merepresentasikan penggunaan obat dikeluarkan sesuai aturan bisnis SIMRS.

Data yang telah melalui proses validasi kemudian diagregasi berdasarkan identitas obat dan periode bulanan. Total penggunaan obat dihitung dengan menjumlahkan seluruh transaksi obat pada periode yang sama menggunakan persamaan $D_{i,t} = \sum \quad Q_{i,t,j}$

Hasil akhir preprocessing berupa dataset yang terdiri atas identitas obat, periode, dan total penggunaan obat bulanan. Dataset tersebut selanjutnya digunakan pada tahap feature engineering untuk membentuk fitur historis sebagai input model Random Forest Regression.

Feature Engineering dan Pembagian Data

Feature engineering dilakukan untuk membentuk fitur yang merepresentasikan pola historis penggunaan obat. Fitur yang digunakan terdiri atas lag 1, lag 2, lag 3, lag 6, dan lag 12 bulan serta moving average 3 dan 6 bulan. Fitur lag menunjukkan penggunaan obat pada periode sebelumnya, sedangkan moving average digunakan untuk menangkap kecenderungan penggunaan dalam beberapa periode. Moving average tiga bulan dihitung menggunakan $M\ A_3\ (t) = (\ D_{t-1} + D_{t-2} + D_{t-3})/3$, sedangkan moving average enam bulan dihitung berdasarkan enam periode sebelumnya. Seluruh fitur hanya menggunakan informasi yang tersedia sebelum periode prediksi untuk mencegah data leakage. Hasil feature engineering selanjutnya digunakan sebagai variabel input pada model Random Forest Regression dengan jumlah kebutuhan obat pada periode prediksi sebagai variabel target.

Dataset dibagi berdasarkan urutan waktu karena data transaksi obat memiliki karakteristik deret waktu. Data pada periode awal digunakan sebagai data training untuk membangun model Random Forest Regression, sedangkan data pada periode berikutnya digunakan sebagai data testing untuk mengevaluasi kemampuan prediksi model. Pembagian secara kronologis dilakukan untuk mencegah informasi dari periode masa depan masuk ke dalam proses pelatihan dan menghindari data leakage. Proporsi training dan testing ditentukan berdasarkan ketersediaan data penelitian. Selanjutnya, hasil prediksi pada data testing dibandingkan dengan nilai aktual menggunakan metrik MAE, RMSE, dan WAPE.

Random Forest dan Optimasi Hyperparameter

Random Forest Regression digunakan untuk memprediksi kebutuhan obat pada periode berikutnya berdasarkan pola historis penggunaan obat. Model menerima fitur lag_1, lag_2, lag_3, lag_6, lag_12, ma_3, dan ma_6 serta variabel waktu sebagai input, sedangkan target yang diprediksi adalah jumlah penggunaan obat pada periode berikutnya. Random Forest membangun sejumlah

decision tree dan menggabungkan hasil prediksi masing-masing pohon melalui proses agregasi. Prediksi akhir model dirumuskan sebagai $Y_t = \frac{1}{B}\sum_{b=1}^{B} T_b(X_t)$, dengan B merupakan jumlah *decision tree* dan $T_b = (X_t)$ merupakan prediksi dari pohon ke-bb. Model dilatih menggunakan data training yang telah dibentuk pada tahap sebelumnya dan selanjutnya digunakan untuk memprediksi data testing. Konfigurasi awal model menggunakan n_estimators=200, max_depth=10, min_samples_split=2, min_samples_leaf=1, dan random_state=42. Parameter tersebut selanjutnya dioptimasi pada tahap *hyperparameter tuning* untuk memperoleh konfigurasi model yang sesuai dengan karakteristik dataset.

Optimasi hyperparameter dilakukan untuk memperoleh konfigurasi Random Forest Regression yang sesuai dengan karakteristik data historis penggunaan obat. Parameter yang dioptimasi meliputi n_estimators, max_depth, min_samples_split, min_samples_leaf, dan max_features. Proses optimasi menggunakan *Time Series Cross-Validation* untuk mempertahankan urutan temporal data dan menghindari *data leakage*. Pemilihan konfigurasi terbaik dilakukan berdasarkan nilai kesalahan prediksi pada data validasi, dengan MAE sebagai kriteria utama. Setelah konfigurasi terbaik diperoleh, model dilatih kembali menggunakan seluruh data training dan selanjutnya digunakan untuk memprediksi data testing. Hasil prediksi model teroptimasi kemudian dibandingkan dengan nilai aktual menggunakan MAE, RMSE, dan MAPE/WAPE

Evaluasi Model

Evaluasi dilakukan menggunakan data testing yang tidak digunakan dalam proses pelatihan maupun optimasi hyperparameter. Kinerja model diukur menggunakan Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), dan Weighted Absolute Percentage Error (WAPE). MAE digunakan untuk mengukur rata-rata selisih absolut antara nilai aktual dan prediksi dengan persamaan $MAE = \frac{1}{n}\sum |y_i - y_i|$, RMSE digunakan untuk mengukur besarnya kesalahan prediksi dengan memberikan penalti lebih besar terhadap kesalahan yang tinggi, yaitu $RMSE = \sqrt{\frac{1}{n}\sum (y_i - \hat{y}_i)^2}$. Sementara itu, WAPE digunakan untuk mengukur kesalahan prediksi dalam bentuk persentase berdasarkan total nilai actual dengan persamaan $WAPE = \frac{\sum |yi - \widehat{yi}|}{\sum |y_i|} x\ 100\%$. MAPE dapat digunakan sebagai metrik tambahan apabila tidak terdapat nilai aktual bernilai nol. Semakin kecil nilai MAE, RMSE, dan WAPE, semakin kecil kesalahan prediksi yang dihasilkan model. Hasil evaluasi selanjutnya digunakan untuk menganalisis kemampuan Random Forest dalam memprediksi kebutuhan obat pada periode berikutnya.

Implementasi pada SIMRS

Hasil model Random Forest yang telah melalui proses optimasi diimplementasikan pada modul instalasi farmasi dalam SIMRS sebagai informasi pendukung perencanaan kebutuhan obat. Data transaksi penggunaan obat yang tersimpan pada SIMRS digunakan sebagai sumber data historis untuk menghasilkan prediksi kebutuhan pada periode berikutnya. Hasil prediksi ditampilkan dalam bentuk tabel dan visualisasi yang memuat informasi obat, periode prediksi, penggunaan historis, dan estimasi kebutuhan obat. Informasi tersebut dapat digunakan oleh petugas farmasi sebagai salah satu bahan pertimbangan dalam melakukan perencanaan kebutuhan dan pengadaan obat. Implementasi model tidak menggantikan keputusan petugas farmasi, tetapi berfungsi sebagai sistem pendukung keputusan berbasis data historis

HASIL DAN PEMBAHASAN

Hasil pengolahan dan Karakteristik Dataset

Data penelitian diperoleh dari data historis transaksi obat pada instalasi farmasi rumah sakit yang tercatat dalam Sistem Informasi Manajemen Rumah Sakit (SIMRS). Dataset yang digunakan mencakup periode Agustus 2024 sampai Desember 2025 dan digunakan sebagai dasar untuk membentuk pola penggunaan obat sebelum dilakukan proses pemodelan menggunakan Random Forest Regression.

Tahap awal pengolahan dilakukan untuk memastikan bahwa data yang digunakan dalam pemodelan memiliki struktur dan kualitas yang sesuai. Proses pengolahan meliputi pemeriksaan kelengkapan data, identifikasi data duplikasi, pemeriksaan nilai jumlah obat, validasi transaksi, serta pengelompokan transaksi berdasarkan identitas obat dan periode waktu.

Setelah proses pembersihan dilakukan, transaksi obat diagregasi berdasarkan obat dan periode bulanan. Agregasi dilakukan dengan menjumlahkan kuantitas obat yang digunakan pada setiap periode. Secara matematis, total penggunaan obat untuk obat ke-ii pada periode tt dapat dirumuskan sebagai: $D_{i,t} = \sum_{j=1}^{n} \quad Q_{i,t,j}$

dengan Di,t merupakan total penggunaan obat ke-i pada periode t, sedangkan Qi,t,j merupakan jumlah obat pada transaksi ke-j. Hasil agregasi tersebut kemudian digunakan sebagai dasar pembentukan fitur historis pada tahap *feature engineering*.

Tabel 2 Karakteristik Dataset Transaksi Obat

Komponen	Hasil
Total transaksi	58.769
Jumlah variable	4
Jumlah jenis obat	565
Missing Value	0
Duplikasi identic	2
Transaksi Jumlah=0	911
Transaksi jumlah negative	0
Total Penggunaan obat valid	220.279 unit

Hasil agregasi bulanan menunjukkan bahwa jumlah penggunaan obat mengalami perubahan antarperiode. Perubahan tersebut menunjukkan adanya variasi kebutuhan obat dari waktu ke waktu yang menjadi dasar penerapan pendekatan prediksi. Variasi penggunaan tersebut penting bagi model karena Random Forest memanfaatkan pola historis sebagai informasi untuk memperkirakan kebutuhan pada periode berikutnya.

Tabel 3 Total Penggunaan Obat berdasarkan periode

Periode	Penggunaan (unit)
Agustus 2024	10.668
September 2024	13.220
Oktober 2024	14.885
November 2024	13.705
Desember 2024	13.941
Januari 2025	13.702
Februari 2025	14.836
Maret 2025	9.558
April 2025	12.758
Mei 2025	14.861
Juni 2025	12.043
Juli 2025	13.705
Agustus 2025	12.980
September 2025	10.833
Oktober 2025	12.237
November 2025	12.471
Desember 2025	13.906

Berdasarkan karakteristik tersebut, dataset memiliki variasi penggunaan obat antarperiode yang dapat dimanfaatkan untuk membangun model prediksi. Data hasil agregasi bulanan selanjutnya digunakan pada tahap pembentukan fitur (*feature engineering*), yaitu dengan membentuk variabel *lag* dan *moving average* yang merepresentasikan pola penggunaan obat pada periode sebelumnya.

Pembentukan Fitur dan Pemodelan Random Forest

Setelah proses *preprocessing* dan agregasi data transaksi obat secara bulanan, tahap berikutnya adalah pembentukan fitur (*feature engineering*)

yang digunakan sebagai masukan pada model prediksi. Pembentukan fitur dilakukan untuk merepresentasikan pola penggunaan obat pada periode sebelumnya sehingga model dapat mempelajari hubungan antara penggunaan historis dengan kebutuhan obat pada periode berikutnya.

Fitur historis yang digunakan terdiri atas lag 1, lag 2, lag 3, dan lag 6, serta *moving average* periode 3 bulan (*MA3*) dan 6 bulan (*MA6*). Fitur *lag* menunjukkan jumlah penggunaan obat pada periode sebelumnya, sedangkan *moving average* digunakan untuk menggambarkan kecenderungan rata-rata penggunaan obat dalam beberapa periode terakhir.Secara matematis, fitur *lag* dirumuskan sebagai: $Lag_k = D_{t-k}$dengan Dt−k merupakan jumlah penggunaan obat pada kk periode sebelum periode t. Dengan demikian, *lag 1* menunjukkan penggunaan obat pada periode sebelumnya, *lag 2* menunjukkan penggunaan dua periode sebelumnya, *lag 3* menunjukkan penggunaan tiga periode sebelumnya, dan *lag 6* menunjukkan penggunaan enam periode sebelumnya.

Sementara itu, *moving average* digunakan untuk memperoleh nilai rata-rata penggunaan obat pada beberapa periode sebelumnya. Rumus *moving average* tiga periode adalah: $M\,A_3(t) = \frac{D_{t-1}+D_{t-2}+D_{t-3}}{3}$ sedangkan moving average enam periode dirumuskan sebagai : $M\,A_6(t) = \frac{\sum_{k=1}^{6} D_{t-k}}{6}$. Target prediksi ditetapkan sebagai jumlah penggunaan obat pada **periode berikutnya**, sehingga model menggunakan informasi historis sebelum periode target dan tidak menggunakan data dari periode yang akan diprediksi. Pendekatan tersebut diterapkan untuk menghindari *data leakage* dalam proses pemodelan.

Tabel 4 Fitur yang digunakan dalam pemodelan

Fitur	Keterangan
Lag_1	Penggunaan 1 periode sebelumnya
Lag_2	Penggunaan 2 periode sebelumnnya
Lag_3	Penggunaan 3 periode sebelumnya
Lag_6	Penggunaan 6 periode sebelumnya
MA_3	Rata-rat penggunaan tiga periode sebelumnya
MA_6	Rata-rata penggunaan enam periode sebelumnya
Bulan	Informasi bulan transaksi
Tahun	Informasi tahun transaksi
Target	Penggunaan obat pada periode berikutnya

Setelah fitur terbentuk, data digunakan untuk membangun model Random Forest Regression. Metode Random Forest dipilih karena mampu membangun sejumlah *decision tree* dan menggabungkan hasil prediksi dari setiap pohon untuk memperoleh prediksi akhir. Pada penelitian ini, model menerima fitur historis penggunaan obat sebagai variabel independen, sedangkan penggunaan obat pada periode berikutnya digunakan sebagai variabel target.

Secara matematis, prediksi Random Forest dapat dituliskan sebagai: $\hat{Y}_t = \frac{1}{B}\sum_{b=1}^{B} T_b(X_t)$ Dengan $\hat{Y}_t$ adalahh hasil prediksi pada periode t, B merupakan jumlah decision tree, dan $T_b(X_t)$ merupakan hasil prediksi dari pohon ke-b berdasarkan fitur X_t

Pada tahap pemodelan awal digunakan konfigurasi **200 pohon** dengan parameter n_estimators=200, max_depth=10, min_samples_split=2, min_samples_leaf=1, dan random_state=42. Konfigurasi tersebut digunakan sebagai model awal (*baseline*) sebelum dilakukan proses optimasi *hyperparameter*.

Dengan pembentukan fitur historis dan penerapan Random Forest Regression tersebut, setiap data obat memiliki informasi mengenai pola penggunaan pada periode sebelumnya yang dapat digunakan untuk memperkirakan kebutuhan pada periode berikutnya. Tahap selanjutnya adalah melakukan optimasi *hyperparameter* dan evaluasi

kinerja model menggunakan MAE, RMSE, dan WAPE.

Hasil Evaluasi dan Optimasi Model

Evaluasi dilakukan untuk mengetahui kemampuan model Random Forest Regression dalam memprediksi kebutuhan obat berdasarkan pola penggunaan historis. Evaluasi menggunakan data *testing* yang tidak digunakan selama proses pelatihan model. Pemisahan data secara kronologis dilakukan untuk mempertahankan urutan waktu transaksi dan menghindari penggunaan informasi dari periode masa depan dalam proses pembentukan model.

Kinerja model dievaluasi menggunakan tiga metrik, yaitu Mean Absolute Error (MAE), Root Mean Square Error (RMSE), dan Weighted Absolute Percentage Error (WAPE). MAE digunakan untuk mengukur rata-rata selisih absolut antara nilai aktual dan hasil prediksi. RMSE memberikan penalti yang lebih besar terhadap kesalahan prediksi yang besar, sedangkan WAPE digunakan untuk melihat besarnya kesalahan prediksi secara relatif terhadap total penggunaan aktual.

Table 5 Parameter random forest yang di optimasi.

Parameter	Nilai
N_estimators	100,200,300,500
Max_depth	5,10,15,20, None
Min_samples_split	2,5,10
Min_samples_leaf	1,2,4
Max features	Sqrt,log2,1.0

Konfigurasi model terbaik ditentukan berdasarkan kinerja pada data validasi, dengan MAE sebagai salah satu dasar pemilihan model. Setelah konfigurasi diperoleh, model dilatih kembali menggunakan data *training* dan selanjutnya diuji menggunakan data *testing* untuk mendapatkan nilai MAE, RMSE, dan WAPE.

Tabel 6 Perbandingan kinerja Model

Metrik	Hasil
MAE	10.927
RMSE	32.351
WAPE	47.539%

Berdasarkan hasil pengujian model *baseline*, diperoleh nilai MAE sebesar 10,927, RMSE sebesar 32,351, dan WAPE sebesar 47,539%. Nilai MAE menunjukkan bahwa rata-rata selisih absolut antara hasil prediksi dan penggunaan aktual berada pada sekitar 10,927 satuan penggunaan obat. Sementara itu, nilai RMSE yang lebih besar dibandingkan MAE menunjukkan adanya beberapa observasi dengan kesalahan prediksi yang relatif lebih tinggi.

Proses optimasi *hyperparameter* dilakukan untuk memperoleh konfigurasi model yang lebih sesuai dengan karakteristik data. Hasil optimasi kemudian dibandingkan dengan model *baseline* menggunakan metrik evaluasi yang sama.

Perbandingan tersebut menjadi dasar untuk menentukan konfigurasi model yang digunakan pada tahap analisis dan implementasi prediksi kebutuhan obat.

Analisis Model

Analisis model dilakukan untuk mengetahui karakteristik prediksi yang dihasilkan oleh algoritma Random Forest Regression serta mengidentifikasi fitur yang paling berkontribusi terhadap hasil prediksi kebutuhan obat. Analisis tidak hanya didasarkan pada nilai evaluasi model, tetapi juga pada tingkat kepentingan masing-masing fitur (*feature importance*) yang digunakan dalam proses pembentukan prediksi.

Tabel 7 Feature Importance Model Random Forest

No	Fitur	Importance	Persentase
1	MA 6	0,398175	39,82%
2	MA 3	0,360492	36,05%
3	Lag 3	0,131099	13,11%
4	Lag 1	0,057284	5,73%
5	Lag 2	0,025991	2,60%

6	Lag 6	0,014603	1,46%
7	Bulan	0,012357	1,24%
8	Tahun	0,000000	0,00%

Hasil Prediksi Kebutuhan Obat

Tabel 8 sample hasil prediksi pengadaan obat di bulan januari 2026

No	Nama Obat	Lag 1,Lag2,Lag3, MA3,Ma6	Prediksi Jan 2026
1	Paracetamol 500 Mg Tablet	1.571;1.424; 1.426;1.473, 67;1.569,67	1.575,00
2	Cefixime 200 mg kaplet dexa	1.501;1.208; 1.052;1.253, 67;1.317,17	1.526,81
3	Berneuron tablet	728;529;672; 643,00;692,6 7	602,27
4	Actrapid 100 iu/ml x 10 ml injeksi	445;402;616; 487,67;472,5 0	506,74
5	Spuit 3 cc om	635;766;505; 635,33;604,0 0	486,64
6	Spuit 10 cc om	399;425;337; 387,00;401,5 0	388,50
7	Vitamin c 500 mg tablet	290;202;155; 215,67;171,5 0	246,67
8	Aqua 25 ml injeksi	245;221;217; 227,67;236,0 0	239,47
9	Ketorolac 30 mg injeksi otto	330;179;188; 232,33;206,8 3	228,39
10	Mefenamic acid 500 mg kaplet	133;215;301; 216,33;276,5 0	211,46

Implementasi pada SIMRS dan Pembahasan

Hasil penelitian selanjutnya diarahkan pada implementasi model prediksi kebutuhan obat sebagai bagian dari Sistem Informasi Manajemen Rumah Sakit (SIMRS), khususnya pada modul instalasi farmasi. Implementasi dilakukan dengan memanfaatkan data transaksi obat yang telah tersimpan pada SIMRS sebagai sumber data historis untuk menghasilkan prediksi kebutuhan obat pada periode berikutnya.

Secara konseptual, proses implementasi dimulai dari data transaksi obat yang tersimpan pada database SIMRS. Data tersebut diproses untuk memperoleh total penggunaan obat secara periodik, kemudian fitur historis seperti lag_1, lag_2, lag_3, lag_6, ma_3, dan ma_6 dibentuk sebagai masukan model. Model Random Forest selanjutnya menghasilkan estimasi kebutuhan obat yang dapat ditampilkan kembali pada modul farmasi

Data Transaksi Obat Databasem SIMRS Agregasi Penggunaan Bulanan Pembentukan Fitur Historis Random Forest Regression Prediksi Kebutuhan Obat Modul Farmasi SIMRS Informasi Pendukung Perencanaan Persediaan.

Pembahasan

Hasil penelitian menunjukkan bahwa data transaksi yang sebelumnya hanya digunakan sebagai catatan operasional pada SIMRS dapat dimanfaatkan lebih lanjut sebagai sumber data untuk analisis prediktif. Pemanfaatan data historis tersebut memberikan pendekatan yang lebih berbasis data dalam proses perencanaan kebutuhan obat.

Penggunaan fitur historis seperti *lag* dan *moving average* memungkinkan model mempertimbangkan pola penggunaan obat dari beberapa periode sebelumnya. *Lag* memberikan informasi mengenai penggunaan pada periode tertentu, sedangkan *moving average* membantu merepresentasikan kecenderungan penggunaan dengan merata-ratakan beberapa periode sebelumnya. Kombinasi fitur tersebut memberikan informasi yang lebih komprehensif dibandingkan hanya menggunakan penggunaan pada satu periode sebelumnya.

Random Forest Regression digunakan karena proses prediksinya dibentuk dari sejumlah *decision tree* yang

kemudian digabungkan menjadi satu hasil prediksi. Pendekatan ini memungkinkan model menangkap pola hubungan yang tidak selalu linear antara fitur historis dan kebutuhan obat. Namun, hasil prediksi tetap dipengaruhi oleh kualitas dan karakteristik data historis yang digunakan dalam proses pelatihan.

Berdasarkan evaluasi model *baseline*, diperoleh nilai MAE sebesar 10,927**,** RMSE sebesar 32,351**,** dan WAPE sebesar 47,539%**.** Nilai tersebut menunjukkan adanya selisih antara kebutuhan aktual dengan hasil prediksi. Perbedaan tersebut perlu diperhatikan karena pola penggunaan obat dapat berubah antarperiode. Beberapa jenis obat dapat memiliki penggunaan yang relatif stabil, sedangkan obat lainnya dapat mengalami fluktuasi yang cukup tinggi.

Dari sisi implementasi, hasil prediksi sebaiknya digunakan sebagai sistem pendukung keputusan, bukan sebagai pengganti keputusan petugas farmasi. Informasi prediksi dapat dikombinasikan dengan kondisi stok saat ini, stok minimum, waktu tunggu pengadaan (*lead time*), obat yang mendekati kedaluwarsa, serta kebijakan persediaan rumah sakit sebelum keputusan pengadaan dilakukan.

Integrasi model ke dalam SIMRS juga memberikan peluang untuk membangun proses perencanaan persediaan yang lebih terstruktur. Data transaksi yang terus bertambah dapat digunakan untuk melakukan pembaruan model secara berkala sehingga model dapat menyesuaikan diri terhadap perubahan pola penggunaan obat. Dengan demikian, pengembangan model prediksi tidak hanya menghasilkan nilai prediksi, tetapi juga dapat menjadi bagian dari pengembangan SIMRS berbasis analitik dan pengambilan keputusan berbasis data**.**

Meskipun demikian, penelitian ini memiliki beberapa keterbatasan. Dataset yang digunakan terutama merepresentasikan pola transaksi historis, sehingga faktor eksternal yang dapat memengaruhi kebutuhan obat belum seluruhnya direpresentasikan dalam model. Faktor seperti perubahan jumlah pasien, pola penyakit, program pelayanan tertentu, perubahan formularium, kejadian luar biasa, serta kebijakan pengadaan dapat memengaruhi kebutuhan obat tetapi belum menjadi variabel dalam model penelitian ini.

Secara keseluruhan, penelitian ini menunjukkan bahwa integrasi Random Forest Regression dengan data historis transaksi obat pada SIMRS dapat dikembangkan sebagai pendekatan untuk menyediakan informasi prediktif bagi perencanaan kebutuhan obat. Hasil prediksi dapat menjadi salah satu sumber informasi bagi petugas farmasi dalam melakukan evaluasi kebutuhan dan perencanaan persediaan secara lebih terukur

SIMPULAN

Data historis transaksi obat pada SIMRS dapat dimanfaatkan untuk membangun model prediksi kebutuhan obat menggunakan Random Forest Regression. Pembentukan fitur historis berupa *lag* dan *moving average* memungkinkan model memanfaatkan pola penggunaan obat pada periode sebelumnya untuk memprediksi kebutuhan pada periode berikutnya. Model *baseline* menghasilkan MAE sebesar 10,927, RMSE sebesar 32,351, dan WAPE sebesar 47,539%. Hasil tersebut menunjukkan bahwa pendekatan Random Forest dapat digunakan sebagai salah satu metode prediksi kebutuhan obat, meskipun masih terdapat kesalahan prediksi akibat variasi pola penggunaan antarperiode. Hasil prediksi dapat diintegrasikan ke dalam modul farmasi SIMRS sebagai informasi pendukung dalam perencanaan persediaan obat. Pengembangan selanjutnya dapat dilakukan dengan menambahkan variabel lain seperti stok, jumlah minimal dan melakukan optimasi model secara lebih lanjut untuk meningkatkan akurasi prediksi.

DAFTAR PUSTAKA

Sutabri, T. (2015). *Konsep sistem informasi*. Penerbit Andi.

Breiman, L. (2001). Random forests. *Machine Learning, 45*, 5–32. https://doi.org/10.1023/A:1010933404324.

Hyndman, R. J., & Athanasopoulos, G. (2021). *Forecasting: Principles and Practice* (3rd ed.). OTexts.

Bergmeir, C., Hyndman, R. J., & Koo, B. (2018). A note on the validity of cross-validation for evaluating autoregressive time series prediction. *Computational Statistics & Data Analysis, 120*, 70–83.

Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2018). The M4 competition: Results, findings, conclusion and way forward. *International Journal of Forecasting, 34*(4), 802–808. https://doi.org/10.1016/j.ijforecast.2018.06.001.

Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2020). The M4 competition: 100,000 time series and 61 forecasting methods. *International Journal of Forecasting, 36*(1), 54–74. https://doi.org/10.1016/j.ijforecast.2.

World Health Organization. (2024). *WHO supports Ministry of Health to launch new technical guidelines on medicine procurement planning*. World Health Organization.

World Health Organization. (2025). *Routine health information system and health facility and community data for neglected tropical diseases: Last-mile logistics information system for medicines and health products*. World Health Organization. https://doi.org/10.

World Health Organization. (2011). *Harmonized monitoring and evaluation indicators for procurement and supply management systems*. World Health Organization.

World Health Organization. (2018). *Drug stock management at health facilities*. World Health Organization.

World Health Organization. (2018). *Demand planning*. In *Antibiotic shortages, stockouts and scarcity*. World Health Organization.

World Health Organization. (2023). *Warehouse and stock management*. World Health Organization.

World Health Organization. (2025). *Health products policy and standards: Medicines supply*. World Health Organization.

World Health Organization. (2025). *Routine health information system and health facility and community data for neglected tropical diseases*. World Health Organization.

Scikit-learn developers. (2026). *RandomForestRegressor*. Scikit-learn documentation.

Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting, 22*(4), 679–688.

Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time Series Analysis: Forecasting and Control* (5th ed.). Wiley.

Brockwell, P. J., & Davis, R. A. (2016). *Introduction to Time Series and Forecasting* (3rd ed.). Springer.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (2nd ed.). Springer.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021). *An Introduction to Statistical Learning: With Applications in R* (2nd ed.). Springer.

Kuhn, M., & Johnson, K. (2013). *Applied Predictive Modeling*. Springe