

IMPLEMENTASI ALGORITMA LSTM PADA PERAMALAN STOK OBAT (STUDI KASUS: PUSKESMAS BEBER)

Fajri Yanti, Betha Nurina Sari, Sofi Defiyanti

Program Studi Informatika S1, Fakultas Ilmu Komputer, Universitas Singaperbangsa
Karawang Jl. HS. Ronggo Waluyo, Telukjambe Timur, Karawang, Indonesia
2010631170069@student.unsika.ac.id

ABSTRAK

Pengelolaan obat merupakan bagian penting dari operasional Pusat Kesehatan Masyarakat (Puskesmas). Kesalahan dalam perencanaan stok dapat menyebabkan kekurangan atau kelebihan stok, yang mengganggu layanan farmasi di Puskesmas. Puskesmas Beber menyimpan stok obat untuk satu bulan dan merencanakan permintaan obat untuk bulan berikutnya. Saat ini, perencanaan dilakukan secara manual oleh puskesmas, yang seringkali tidak akurat, menyebabkan kelebihan atau kekurangan stok obat. Penelitian ini bertujuan untuk memprediksi lima jenis obat di Puskesmas Beber. Penelitian ini menggunakan metodologi CRISP-DM (*Cross Industry Standard Process for Data Mining*) LSTM (*Long Short Term Memory*). Pada tahap *business understanding* difokuskan pada pemahaman masalah penelitian, terutama dalam hal lima obat yang diminta oleh puskesmas. Tahap *data understanding* melibatkan pemeriksaan dataset dari Puskesmas Beber serta pemilihan dataset. Tahap *data preparation* melibatkan langkah-langkah pra-pemrosesan data. Tahap *modeling* LSTM diterapkan dengan tuning *hyperparameter* menggunakan *GridSearchCV*. Setelah dilakukan fitting model pada 80% data training, dilakukan prediksi pada 20% data testing untuk masing-masing obat. Evaluasi dilakukan menggunakan MAPE. Hasil evaluasi menunjukkan bahwa nilai terendah MAPE adalah obat Vitamin B Komplek Tablet dengan nilai MAPE sebesar 4.82%. Pada tahap *deployment*, dilakukan penyusunan laporan hasil prediksi yang dapat digunakan sebagai pertimbangan dalam perencanaan persediaan obat di Puskesmas Beber.

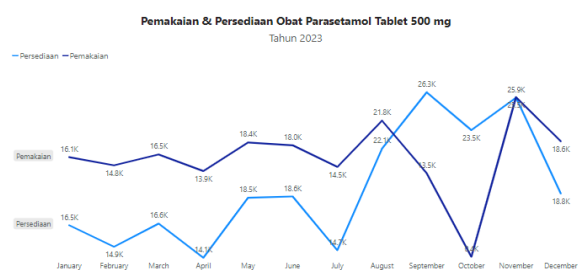
Kata kunci : Prediksi Stok Obat, LSTM, CRISP-DM, Deret Waktu, Puskesmas, MAPE

1. PENDAHULUAN

Pengelolaan obat merupakan aspek yang sangat penting dalam operasional Pusat Kesehatan Masyarakat (Puskesmas). Langkah-langkah dari perencanaan hingga evaluasi obat memiliki peran yang signifikan dalam menjamin ketersediaan obat yang aman, berkualitas, dan ekonomis. Namun, perencanaan stok yang tidak tepat dalam manajemen obat bisa menyebabkan kekurangan atau kelebihan stok, yang pada akhirnya akan mengganggu pelayanan farmasi di Puskesmas. Dampaknya termasuk ketidakpuasan pasien, penurunan kualitas layanan, dan potensi kerugian keuangan bagi Puskesmas. Kurang optimalnya pengelolaan obat tidak hanya mengancam kelangsungan layanan kesehatan, tetapi juga dapat menghambat pertumbuhan Puskesmas secara keseluruhan. Oleh karena itu, analisis mendalam terhadap sistem pengelolaan obat menjadi sangat penting untuk mengatasi tantangan ini dan memastikan efektivitas operasional Puskesmas secara keseluruhan.

Sebagai penyedia layanan kesehatan dasar, Puskesmas perlu memastikan ketersediaan obat yang memadai dan tepat waktu agar pasien mendapatkan pengobatan yang optimal. Puskesmas Beber, sebagai pusat pelayanan Kesehatan Masyarakat di daerah Beber, menyimpan persediaan obat untuk digunakan selama satu bulan dan membuat perencanaan permintaan obat untuk satu bulan berikutnya. Saat ini, perencanaan permintaan obat untuk bulan berikutnya masih dilakukan secara manual berdasarkan perkiraan

oleh petugas Puskesmas. Hal ini menyebabkan perencanaan persediaan obat menjadi tidak akurat karena ada obat yang memiliki kelebihan stok dan ada yang kekurangan. Oleh karena itu, diperlukan sistem yang dapat dipergunakan untuk memprediksi jumlah stok obat yang optimal untuk bulan berikutnya, sehingga dapat menjadi acuan dalam permintaan obat ke Dinas Kesehatan.



Gambar 1. Pemakaian parasetamol tablet 500 mg per bulan pada tahun 2022

Gambar 1 menunjukkan bahwa persediaan dan jumlah pakai obat parasetamol tablet 500 mg cenderung berlebih setiap bulan. Oleh karena itu diperlukan penelitian lebih mengenai prediksi stok obat pada Puskesmas Beber. Hasil prediksi berupa pola jumlah pemakaian obat sehingga bisa dijadikan bahan pertimbangan pihak Puskesmas dalam perencanaan stok obat bulan berikutnya.

Hasil penelitian yang dilakukan oleh [1] menunjukkan bahwa penggunaan Long

Short-term Memory (LSTM) untuk memprediksi penjualan produk farmasi di Apotik Suganda juga menunjukkan tingkat akurasi yang cukup baik, dengan tingkat kesalahan prediksi sekitar 4%. Penelitian yang dilakukan oleh [2] LSTM memiliki nilai RMSE lebih kecil daripada algoritma SVM pada prediksi kebutuhan obat mingguan. Berdasarkan latar belakang tersebut, penelitian ini berfokus pada implementasi algoritma LSTM pada prediksi stok obat Puskesmas Beber. Dalam studi [12],

Long Short Term Memory (LSTM) diimplementasikan pada dataset penjualan obat "X" di PT. Metiska Farma untuk memprediksi penjualan di masa depan. Dari dua parameter uji coba, percobaan kelima menunjukkan kinerja terbaik. Hasil penelitian menunjukkan bahwa metode optimal penggunaan LSTM menggunakan komposisi data pelatihan 90% dan data pengujian 10%, dengan rentang interval $[-1,1]$, dan epoch sebanyak 1500. Metode ini menghasilkan RMSE dalam bentuk rupiah sebesar 13,762,154.00, dan persentase rata-rata kesalahan model antara nilai prediksi dan nilai aktual harian terkecil dengan MAPE sebesar 12%. Penelitian lain [13],

Hasil evaluasi terhadap model LSTM dan ARIMA menunjukkan bahwa LSTM memiliki akurasi prediksi yang lebih tinggi untuk data penjualan obat di apotek, dengan selisih rata-rata nilai RMSE sebesar 13.643 dibandingkan ARIMA. Meskipun LSTM mengalami overfitting pada data pelatihan, hasil ini menekankan pentingnya mempertimbangkan generalisasi selain akurasi dalam machine learning. Dengan akurasi yang lebih baik dalam memprediksi nilai target, LSTM terbukti lebih efektif dalam menangani data penjualan obat dibandingkan metode tradisional seperti ARIMA.

2. TINJAUAN PUSTAKA

2.1. Peramalan (forecasting)

Peramalan (forecasting) merupakan ilmu yang digunakan untuk memperkirakan kejadian di masa depan dengan pengambilan data historis kemudian memproyeksikannya ke masa yang akan datang [3].

Memprediksi juga dapat didasarkan pada keahlian keputusan (judgement), yang pada gilirannya didasarkan pada historis dan pengalaman. Memprediksi kejadian masa depan ini bisa menguntungkan diberbagai keilmuan dari penjualan, cuaca, daya beli, kondisi iklim, dan lain sebagainya

Dalam melakukan peramalan, apapun jenisnya, terdapat lima langkah yang dapat diikuti berdasarkan Hanke & Wichern [4]:

- a. Formulasi masalah dan pengumpulan data
Untuk metode peramalan kuantitatif, penting untuk memiliki data yang relevan dan akurat. Jika data yang diperlukan tidak tersedia, mungkin perlu dilakukan evaluasi ulang terhadap perumusan masalah atau metode peramalan yang digunakan.

- b. Manipulasi dan pembersihan data
Beberapa data mungkin tidak relevan atau memiliki nilai yang hilang yang perlu diestimasi. Data mungkin perlu dikonversi ke unit yang berbeda atau diproses lebih lanjut, misalnya dengan mengkonsolidasikannya dari berbagai sumber. Beberapa data hanya relevan untuk periode historis tertentu, dan sering kali perlu diubah formatnya untuk memenuhi kebutuhan metode peramalan yang digunakan.
- c. Pembentukan dan evaluasi model
Langkah ini melibatkan menyesuaikan data yang terkumpul ke dalam model peramalan untuk mengurangi kesalahan peramalan.
- d. Implementasi model (peramalan aktual)
Setelah data yang relevan terkumpul dan model yang tepat dipilih, model tersebut digunakan untuk melakukan peramalan. Peramalan untuk periode saat ini, berdasarkan data historis, sering digunakan untuk memeriksa keakuratan proses.
- e. Evaluasi peramalan
Langkah ini mencakup penilaian terhadap model peramalan yang dibuat setelah data terkumpul dan model dipilih. Peramalan untuk periode sekarang, dengan menggunakan data historis yang diketahui, sering digunakan untuk mengevaluasi keakuratan proses. Implementasi model terdiri dari model peramalan aktual yang dibuat ketika data yang sesuai telah terkumpul dan terpilihnya model peramalan yang sesuai. Peramalan untuk periode sekarang dengan nilai historis aktual diketahui sering kali digunakan untuk mengecek keakuratan dari proses.

2.2. CRISP-DM

CRISP-DM adalah pendekatan yang umum digunakan oleh para ahli untuk menyelesaikan masalah menggunakan proses *data mining*. Pendekatan ini terdiri dari enam langkah, yakni Pemahaman Bisnis, Pemahaman Data, Persiapan Data, Pemodelan, Evaluasi, dan Implementasi [5]. Berikut 6 tahapan CRISP-DM berdasarkan [6]:

- a. Pemahaman Bisnis
Langkah pertama dalam proses ini dikenal sebagai pemahaman bisnis, di mana fokusnya adalah pada pengertian tujuan dari perspektif bisnis, mengidentifikasi permasalahan, dan merancang rencana strategi awal untuk mencapai tujuan pengolahan data.
- b. Pemahaman Data
Fase pemahaman data dimulai dengan pengumpulan data awal untuk memperdalam pemahaman tentang karakteristik data, menemukan potensi masalah dalam kualitas data, mendapatkan wawasan awal dari data, atau menemukan segmen data yang menarik untuk membentuk hipotesis tentang informasi tersembunyi. Proses ini melibatkan empat langkah, yakni pengumpulan data awal, deskripsi data, eksplorasi data, dan pengecekan kualitas data.

c. Persiapan Data

Langkah berikutnya adalah persiapan data, yang melibatkan serangkaian tindakan untuk mengorganisir dataset akhir atau data yang akan dimasukkan ke dalam alat pemodelan dari data mentah awal. Ini melibatkan pemilihan tabel, catatan, dan atribut yang relevan, serta melakukan transformasi dan pembersihan data untuk memastikan kesiapan data untuk proses pemodelan.

d. Pemodelan

Setelah tahap persiapan data selesai, langkah berikutnya adalah memilih teknik pemodelan yang paling sesuai. Pada tahap ini, parameter-parameter diperbaiki agar mencapai nilai optimal. Beberapa teknik pemodelan mungkin memiliki persyaratan khusus tergantung pada jenis data yang dimiliki, sehingga mungkin diperlukan revisi terhadap tahap persiapan data untuk menyesuaikan data dengan persyaratan teknik pemodelan tertentu.

e. Evaluasi

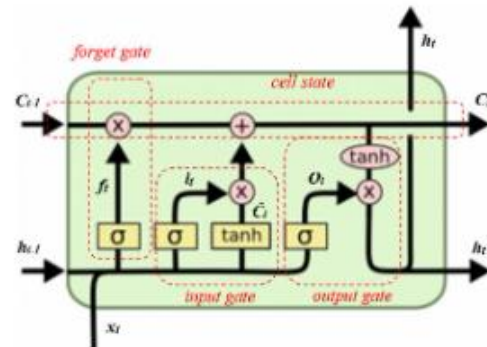
Fase evaluasi bertujuan untuk menilai kinerja model berdasarkan tujuan bisnis yang telah ditetapkan serta mengidentifikasi faktor-faktor bisnis yang mungkin menjadi penyebab ketidakberhasilan model. Evaluasi ini mencakup temuan yang relevan dengan tujuan bisnis awal serta dapat mengungkap masalah potensial, informasi tambahan, atau petunjuk untuk perbaikan di masa mendatang.

f. Implementasi

Langkah terakhir adalah implementasi, yang sering kali tidak berhenti pada pembuatan model. Fase ini bergantung pada kebutuhan pengguna, mulai dari penyusunan laporan hingga implementasi proses yang lebih kompleks di seluruh organisasi. Pengetahuan yang dihasilkan harus disajikan secara terstruktur dan relevan agar dapat dimanfaatkan secara efektif.

2.3. LSTM

Neural network (NN) sederhana hanya terdiri dari input, model network, dan output (*one to one*) untuk regresi linier dan klasifikasi. Pengembangan selanjutnya adalah model *one to many*, di mana output sebelumnya digunakan sebagai input untuk proses berikutnya. Karena data diproses secara beruntun, output digunakan sebagai input untuk iterasi berikutnya. Model ini kemudian berkembang menjadi *Recurrent Neural Network* (RNN), yang memproses input dengan melibatkan data sebelumnya. RNN menggunakan multi-layer dan perulangan dalam pemrosesan data runtun waktu, sehingga masuk dalam deep learning. Namun, RNN menghadapi masalah *vanishing gradient* pada urutan data panjang, sehingga hanya mampu mengelola data dengan ketergantungan pendek.[7]



Gambar 2. Struktur LSTM

Gambar 2 menunjukkan struktur LSTM. Untuk menangani data time series atau deret waktu, LSTM menghubungkan informasi sebelumnya dengan informasi berikutnya secara efektif dan mampu menyimpan data dalam jangka waktu yang panjang. Berikut adalah langkah-langkah dan aturan dalam metode LSTM[7]. Pertama, menghitung nilai sigmoid dan tanh. Kedua, mengubah data menjadi masalah supervised learning, yang mempelajari fungsi pemetaan dari input ke output. Data yang digunakan berbentuk time series, dengan output hari kemarin dilambangkan ($t-1$) dan hari ini (t). Ketiga, menerapkan min-max scaling untuk normalisasi. Keempat, menghapus variabel yang tidak diperlukan. Kelima, membagi data menjadi data latih dan data uji. Keenam, memodelkan LSTM dengan memilih jumlah neuron, hidden layer, dan epoch yang sesuai. Selanjutnya, menghitung nilai forget gate, input gate, memperbaiki memory cell, menghitung output gate, dan nilai output akhir. Terakhir, mengevaluasi model LSTM yang dikembangkan.[8]

2.4. Evaluasi Model

MAPE (Mean Absolute Percentage Error) merupakan salah satu metrik evaluasi yang umum digunakan dalam peramalan untuk mengukur akurasi suatu model peramalan dalam memprediksi nilai-nilai tertentu. MAPE mengukur rata-rata persentase kesalahan antara nilai aktual dan nilai prediksi dalam suatu dataset. MAPE memberikan gambaran tentang seberapa besar kesalahan rata-rata dalam prediksi sebagai persentase dari nilai aktual. Semakin kecil nilai MAPE, semakin baik model peramalan.

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{Y_i - \hat{Y}_i}{Y_i} \right| \times 100 \quad (1)$$

Keterangan:

n : Jumlah sampel dalam data

Y_i : Nilai aktual

$\hat{Y}_i$: Nilai prediksi

Tabel 1 menampilkan kategori kriteria nilai MAPE untuk suatu metode. Suatu metode dianggap memiliki kinerja yang sangat baik jika nilai MAPE kurang dari 10%. Semakin kecil nilai MAPE yang

diperoleh, semakin baik kinerja metode tersebut[11].

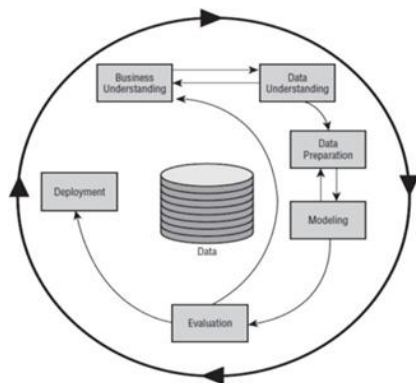
Table 1 Kategori Nilai MAPE

Nilai MAPE (%)	Kriteria
$\leq 10\%$	Sangat Baik
10% - 20%	Baik
20% - 50%	Cukup
$>50\%$	Buruk

Tabel 1 menunjukkan bahwa nilai MAPE $\leq 10\%$ maka kinerja model tersebut dikategorikan sebagai model dengan kinerja yang sangat baik. Nilai MAPE rentang 10% hingga 20% dikategorikan dengan kinerja baik, sedangkan nilai MAPE dengan rentang 20% hingga 50% dikategorikan sebagai model dengan kinerja cukup. Nilai MAPE $>50\%$ dikategorikan sebagai kinerja model buruk.

3. METODE PENELITIAN

Metodologi yang digunakan pada penelitian ini adalah CRISP-DM (CRoss-Industry Standard Process for Data Mining) yang terdiri dari enam tahap dalam pengolahan data, yaitu: *Business Understanding*, *Data Understanding*, *Data Preparation*, *Modeling*, *Evaluation*, dan *Deployment*.



Gambar 3. Flow CRISP-DM[10]

Gambar 3 menampilkan *flow* metodologi CRISP-DM, berikut merupakan detail untuk setiap *step* metodologi pada penelitian ini.

3.1. Pemahaman Bisnis (Business Understanding)

Pada tahapan ini akan dilakukan observasi ke tempat yang dijadikan objek penelitian yaitu Puskesmas Beber kemudian mencari kondisi bisnis Puskesmas Beber. Setelah mengetahui potensi datanya kemudian menentukan data yang akan diolah dengan menilai situasi yang akan dijadikan objek penelitian tersebut. Dalam hal ini dapat menentukan tujuan dari penelitian. Observasi ini akan dihasilkan informasi yang dibutuhkan, tujuan dari penelitian ini adalah untuk memprediksi stok obat untuk 12 bulan berikutnya.

3.2. Pemahaman Data (Data Understanding)

Pada tahapan ini proses pengumpulan data, kemudian data tersebut akan dipelajari dan dipahami lebih lanjut. Data yang dibutuhkan dalam penelitian ini adalah data stok optimum periode Januari 2019 – November 2023. Data yang digunakan terdiri dari tanggal, nama obat, satuan, stok awal, penerimaan, persediaan, stok akhir, dan total pemakaian.

3.3. Pengolahan Data (Data Preparation)

Pada tahap ini akan dilakukannya persiapan pada data yang diperoleh dari tahap data understanding data yang diperoleh yaitu Data yang digunakan terdiri dari Tanggal, Nama Obat, Satuan, Stok Awal, Penerimaan, Persediaan, Stok Akhir, dan Total Pemakaian.

Pada tahap ini perlu dilakukan prapemrosesan data terlebih dahulu. Langkah-langkah yang dilakukan yaitu *Data Selection* merupakan tahapan menyeleksi data-data yang relevan dengan permasalahan dalam penelitian ini. Setelah seleksi data tahapan berikutnya adalah *Data Cleaning*, tahapan pembersihan data atau attribute yang tidak lengkap atau tidak diperlukan.

3.4. Pemodelan (Modeling)

Pada tahap ini merupakan tahapan untuk melakukan pengujian dan pelatihan terhadap data yang sudah dikumpulkan dan diolah. Tahap ini menggunakan algoritma LSTM (Long Short Term Memory). Pemodelan dilakukan menggunakan *python* dan *jupyter notebook*.

3.5. Evaluasi (Evaluation)

Pada tahap ini akan dilakukan pengukuran serta validasi keakuratan hasil yang dicapai oleh model yang sudah digunakan. Perhitungan nilai eror model LSTM dilakukan untuk mengetahui perbandingan antara data asli dengan data hasil pembelajaran atau data training. Pengukuran dilakukan menggunakan MAPE.

3.6. Penyebaran (Deployment)

Pada tahap ini merupakan tahap akhir dari metodologi CRISP-DM (*Cross-Industry Standard Process for Data Mining*). Laporan akhir ini berisi pengetahuan hasil dari penelitian dan evaluasi yang telah dilakukan pada proses data mining. Presentasi hasil akan disusun dalam bentuk yang mudah dipahami dan laporan ini dapat dijadikan sumber untuk penelitian selanjutnya.

4. HASIL DAN PEMBAHASAN

4.1. Pemahaman Bisnis (Business Understanding)

Tahap awal *business understanding* yaitu melakukan analisis permasalahan stok persediaan Puskesmas Beber, berdasarkan latar belakang dan tujuan penelitian ini, jumlah stok obat pada Puskesmas Beber memiliki pengaruh pada manajemen persediaan obat serta pengelolaan sumber daya keuangan Puskesmas Beber. Selain itu, ketersediaan obat juga berkaitan erat dengan penilaian pelayanan masyarakat.

Penelitian ini berfokus pada jumlah pemakaian 5 jenis obat per hari di Puskesmas Beber. Jumlah pemakaian obat ini akan digunakan untuk prediksi stok obat Puskesmas Beber. Penelitian ini menggunakan algoritma LSTM.

4.2. Pemahaman Data (Data Understanding)

Tahapan pemahaman data dilakukan dengan memahami karakteristik data yang akan digunakan sehingga bisa memudahkan proses *cleaning data*. Data yang digunakan merupakan data harian, Gambar 4 menampilkan karakteristik data awal.

```
Data columns (total 39 columns):
# Column Non-Null Count Dtype
---
0 number 117 non-null object
1 nama_obat 117 non-null object
2 satuan 117 non-null object
3 26.0 21 non-null object
4 27.0 0 non-null float64
5 28.0 29 non-null float64
6 29.0 23 non-null float64
7 30.0 23 non-null float64
8 1.0 19 non-null float64
9 2.0 21 non-null float64
10 3.0 27 non-null float64
11 4.0 0 non-null float64
12 5.0 26 non-null float64
13 6.0 21 non-null float64
14 7.0 24 non-null float64
15 8.0 20 non-null float64
16 9.0 25 non-null float64
17 10.0 19 non-null float64
18 11.0 0 non-null object
19 12.0 27 non-null float64
20 13.0 24 non-null float64
21 14.0 21 non-null float64
22 15.0 23 non-null float64
23 16.0 14 non-null float64
24 17.0 26 non-null float64
25 18.0 0 non-null float64
26 19.0 22 non-null float64
27 20.0 20 non-null float64
28 21.0 20 non-null float64
29 22.0 26 non-null float64
30 23.0 20 non-null float64
31 24.0 20 non-null float64
32 25.0 0 non-null float64
```

Gambar 4. Informasi data awal

```
nan_count = data.isna().sum()
print(nan_count)
```

number	0
nama_obat	0
satuan	0
26.0	96
27.0	117
28.0	88
29.0	94
30.0	94
1.0	98
2.0	96
3.0	90
4.0	117
5.0	91
6.0	96
7.0	93
8.0	97
9.0	92
10.0	98
11.0	117
12.0	90
13.0	93
14.0	96
15.0	94
16.0	103
17.0	91
18.0	117
19.0	95
20.0	97
21.0	97
22.0	91
23.0	97
24.0	97
25.0	117

Gambar 5. Missing value

Gambar 4 merupakan deskripsi informasi umum dari data, untuk analisis lebih lanjut diperlukan tipe data yang sesuai dengan kebutuhan analisis.

Berdasarkan gambar 5, berapa kolom data terdapat *missing value*. Setelah tahapan pemahaman data selanjutnya adalah data preparation.

4.3. Pengolahan Data (Data Preparation)

Tahapan selanjutnya berupa *data preparation*, tahapan ini meliputi *data selection*, *data cleaning*, *data integration*, serta normalisasi data. Penelitian ini hanya membutuhkan data nama obat dan total pemakaian per hari selama 2019-2023. Objek penelitian ini berupa lima jenis obat yang ditentukan oleh pihak puskesmas. Setelah tahapan seleksi data, tahapan selanjutnya adalah mengatasi *NaN value* pada data. Pada gambar 4.15 menunjukkan terdapat *missing value*, pada penelitian ini *missing value* pada data dikarenakan tidak ada pemakaian pada tanggal tersebut. Oleh karena itu, pihak Puskesmas mengosongkan kolom apabila tidak ada pemakaian obat.

```
1 data = data.drop(columns=['satuan', 'total_pemakaian'])
2 data.head()
```

	nama_obat	26.0	27.0	28.0	29.0	30.0	1.0	2.0	3.0	...	15.0	16.0	17.0	18.0	19.0	20.0	21.0	22.0	23.0	24.0
37	Asam efenkamat (As. Ef.) Tablet 50 mg	36	NaN	48.0	25.0	56.0	48.0	22.0	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
112	Gilbert Clonylstat Tablet 100 mg	60	NaN	117.0	81.0	114.0	72.0	75.0	69.0	NaN	...	69.0	90.0	NaN	48.0	79.0	48.0	60.0	75.0	27.0
225	Parasetamol Tablet 500 mg	420	NaN	497.0	319.0	365.0	420.0	351.0	352.0	NaN	...	353.0	302.0	NaN	364.0	382.0	306.0	310.0	335.0	245.0
230	Prokolan HCl (Pro. HCl) Tablet 10 mg	40	NaN	43.0	24.0	24.0	18.0	24.0	12.0	NaN	...	NaN	12.0	NaN	42.0	48.0	12.0	24.0	12.0	10.0
505	Vitamin B Komplek Tablet	60	NaN	66.0	54.0	38.0	50.0	70.0	20.0	NaN	...	70.0	66.0	NaN	70.0	104.0	72.0	112.0	116.0	94.0

5 rows x 31 columns

Gambar 6. NaN value

Pada penelitian ini, untuk mengatasi *missing value* atau *NaN*, ditampilkan pada gambar 5 dan 6, pada data diganti dengan nilai nol. Hasil penanganan *missing value* ditunjukkan pada gambar 7.

```
data.head()
```

	number	nama_obat	satuan	26.0	27.0	28.0	29.0	30.0	1.0	2.0	...	22.0	23.0	24.0	25.0	nan	nan	nan	nan	total_pemakaian
1	1	Acetazol 500 mg (Acet. Azetifazolat)	Tablet	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0
2	2	Acetazol 500 mg	Tablet	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0
3	3	Acetazol 400 mg Tablet	Tablet	0	0	0	0	10.0	0	0	...	12.0	0	0	0	0	0	0	0	0
4	4	Acetazol Cream 5 %	Tablet	0	0	0	2.0	2.0	0	0	...	1.0	0	0	0	0	18	0	0	0
15	15	Abendazole 400 mg	Tablet	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0

Gambar 7. Penanganan missing value

Pada raw data, tanggal dijadikan sebagai baris, oleh karena itu perlu dilakukan transposisi data agar memudahkan proses integrasi data. Setelah dilakukan transposisi data, format tanggal pada data perlu dilakukan penyesuaian agar bisa dilakukan pemodelan, pada raw data kolom tanggal hanya tercatat tanggal, tidak berupa format tanggal lengkap. Kolom 'Date' dilakukan penyesuaian terhadap tanggal pencatatan pemakaian obat disesuaikan dengan aturan pencatatan laporan pemakaian Puskesmas Beber. Puskesmas Beber membuat pencatatan pemakaian sheet baru setiap tanggal 25 bulan sebelumnya hingga tanggal 24 bulan berjalan. Oleh karena itu, kolom Date dilakukan date formatting sesuai tanggal pemakaian yang sebenarnya, sehingga data yang digunakan tetap valid. Hasil transposed data dan penyesuaian tanggal ditunjukkan pada gambar 8.

```
df = []
index_to_keep = df.transposed.index[df.transposed['Date'] == '1-0'][0]
last_month = df.transposed.iloc[index_to_keep].copy()
current_month = df.transposed.iloc[index_to_keep].copy()
last_month['Date'] = '2022-12-1' + last_month['Date'].astype(str).str.replace('-', '')
current_month['Date'] = '2022-12-1' + current_month['Date'].astype(str).str.replace('-', '')
df.append(last_month)
df.append(current_month)

result_df = pd.concat(df, ignore_index=True)
result_df
```

	Date	Asam askorbat (Vit. C) Tablet 50 mg	Gilitinil Gaseykolet Tablet 100 mg	Parasetamol Tablet 500 mg	Peridaksin HCl (Vit. B6) Tablet 10 mg	Vitamin B Komplek Tablet
0	2022-11-26	82.0	190.0	596.0	12.0	190.0
1	2022-11-27	0.0	0.0	0.0	0.0	0.0
2	2022-11-28	64.0	174.0	438.0	12.0	80.0
3	2022-11-29	72.0	127.0	452.0	18.0	96.0
4	2022-11-30	36.0	90.0	296.0	24.0	36.0
5	2022-12-1	24.0	38.0	155.0	12.0	36.0

Gambar 8 Hasil *transposed data*

Tahapan selanjutnya adalah integrasi data, berikut merupakan hasil integrasi yang ditunjukkan pada tabel 2. Tabel 2 merupakan data total pemakaian lima jenis obat per bulan

Tabel 2. Dataset hasil *data preparation*

Date	Obat1	Obat2	Obat3	Obat4	Obat5
2019-01	24557	3936.3	5509	4030	2474
2019-02	25490	5352	4789	4033	2294
2019-03	26414	4834	6096	4135	2684
2019-04	23887	4213	6518	3734	2653
2019-05	17293	4828	3736	2092	1714
2019-06	11304	2388	3944	2250	1268
2019-07	19893	4570	5398	2435	2501
2019-08	19833	4566	4606	2054	2627
2019-09	20660	5731	3892	2120	2370
2019-10	24030	5232	3698	2588	2739
2019-11	21990	3809	3308	3379	2952
2019-12	22380	5047	3729	3225	3031
2020-01	26307	5707	3201	3605	3589
2020-02	26339	5118	3740	3325	2847
2020-03	23998	4840	2789	3840	2618
2020-04	13051	2445	2040	1719	1506
2020-05	9938	1408	1341	934	1819
2020-06	10924	2580	1866	928	1261
2020-07	10795	1652	1826	1158	1143
2020-08	9111	1673	2557	1031	1030
2020-09	7912	2201	1780	1212	1444
2020-10	9588	3329	2236	1005	1001
2020-11	10425	2246	1692	1538	1303
2020-12	7002	1481	1086	1585	1002
2021-01	2586	320	669	666	326
2021-02	2457	584	516	646	254
2021-03	6528	786	1020	893	406
2021-04	11235	2026	1722	2381	1349
2021-05	9827	2295	2088	2145	1694
2021-06	12732	3157	3234	2747	1670
2021-07	11156	1526	1456	0	707
2021-08	11976	1338	1454	0	1147
2021-09	15342	2145	1701	0	1508
2021-10	15812	3724	2428	374	1289
2021-11	19089	6209	3519	2410	1495
2021-12	16171	4902	3640	3131	1035
2022-01	14166	3176	3045	1393	871
2022-02	15906	4463	3442	1973	772
2022-03	14082	3549	3021	3821	861
2022-04	12415	2088	1372	1900	705
2022-05	9820	2421	1767	1807	1049
2022-06	13706	3332	2781	2092	1262
2022-07	13013	3924	2014	2676	1184
2022-08	18151	7069	2156	3844	1864
2022-09	19787	5035	1160	4272	1262
2022-10	19933	7301	1982	2795	1614

Date	Obat1	Obat2	Obat3	Obat4	Obat5
2022-11	19439	6448	2461	2892	1267
2022-12	16693	5160	2464	1983	1109
2023-01	14829	4093	1867	2389	1376
2023-02	15213	4226	1643	1899	1238
2023-03	17178	4138	2261	1892	1287
2023-04	13292	3065	1656	1499	1148
2023-05	19394	4484	1823	3034	1780
2023-06	15920	3157	2691	1782	1646
2023-07	16870	3937	2455	2377	1622
2023-08	22431	5854	3207	2406	1814
2023-09	25084	5471	2929	2589	1258
2023-10	24286	5648	3053	3360	1228
2023-11	23524	4918	2570	3586	1298

Tahapan selanjutnya yaitu normalisasi data, tahapan ini digunakan untuk mengubah rentang data dengan nilai yang sama dan lebih kecil. Normalisasi dilakukan dengan menggunakan skala (0,1)[9].

```

scalers = {}
scaled_df = pd.DataFrame()

for column in df.columns:
    scaler = MinMaxScaler(feature_range=(0, 1))
    scaled_data = scaler.fit_transform(df[[column]])
    scalers[column] = scaler
    scaled_df[column] = scaled_data.flatten()

```

Gambar 9. Normalisasi data

Setelah dilakukan normalisasi, dilakukan pemodelan menggunakan LSTM(Long Short Term Memory).

4.4. Pemodelan (Modeling)

Tahap selanjutnya adalah membagi dataset menjadi data latih dan data uji. Pembagian antara training set dan test set berdasarkan studi empiris menunjukkan bahwa model terbaik diperoleh jika menggunakan 20% data untuk testing dan 80% data untuk training[10]. Penelitian ini membagi data training dan data testing dengan perbandingan 80:20. Code yang ditampilkan pada gambar 10 merupakan tahapan dari split data.

```

train_size = int(len(scaled_df) * 0.8)
test_size = len(scaled_df) - train_size

train_scaler = scaled_df.iloc[:train_size]
test_scaler = scaled_df.iloc[train_size:]

```

Gambar 10. Split data

Model LSTM merupakan suatu struktur yang terdiri dari beberapa layer. Menurut penelitian[9], layer ini penting untuk memantau jumlah nilai pada setiap layer agar dapat mengatasi masalah *underfitting* atau *overfitting*.

Penelitian [10] menyebutkan hasil pengujian nilai layer berbanding lurus dengan nilai prediksinya, semakin nilai layer kecil semakin buruk nilai prediksi. Namun, model akan mengalami *overfitting* apabila nilai layer terlalu besar. Berikut rincian arsitektur jaringan LSTM yang digunakan pada penelitian ditunjukkan pada gambar 11:

- LSTM *Input Layer* dengan 100 neuron, dengan nilai input shape 1, 1, sesuai dengan jumlah *timesteps* dan *features*.
- LSTM *Hidden Layer* dengan 100 neuron
- Dropout Hidden Layer* sebanyak 0,2 persen dari jumlah neuron, layer ini bertujuan untuk mencegah *overfitting* dalam model, secara acak menjatuhkan atau mengabaikan sejumlah neuron tertentu.
- Layer terakhir adalah *Dense Output Layer* dengan 1 neuron.

```
def build_model(optimizer='adam'):
    model = Sequential()
    model.add(Input(shape=(1, len(df.columns))))
    model.add(LSTM(100, return_sequences=True))
    model.add(LSTM(100))
    model.add(Dropout(0.2))
    model.add(Dense(1))
    model.compile(loss='mse', optimizer=optimizer)
    return model

keras_regressor = KerasRegressor(model=build_model, verbose=1, shuffle=False)

parameters = {
    'batch_size': [16, 32, 64],
    'epochs': [10, 100],
    'optimizer': ['adam']
}
```

Gambar 11. Model LSTM

Penelitian ini menggunakan *GridSearchCV* untuk mendapatkan kombinasi parameter yang optimal sehingga menghasilkan *loss function* MSE (*Mean Square Error*) kecil. Parameter yang telah diuji dengan berbagai nilai untuk *batch_size* dan *epoch* menunjukkan hasil yang baik pada nilai-nilai berikut: *batch_size* dengan nilai (16, 32, 64), *epochs* (10, 100), dan *optimizer* adalah *adam*. parameter terbaik untuk pemodelan Piridoksin HCl (Vit. B6) Tablet 10 mg dengan epoch:100, batch_size: 16 dengan optimizer adam. Berdasarkan *output GridSearchCV* untuk pemodelan LSTM semua obat mempunyai nilai parameter yang sama. Kombinasi nilai parameter, Gliseril Guayakolat Tablet 100 mg, Asam askorbat (Vit. C) Tablet 50 mg, Vitamin B Komplek Tablet, Piridoksin HCl (Vit. B6) Tablet 10 mg, dan Paracetamol Tablet 500 mg yang mempunyai batch_size 16 dan epoch 100, dengan optimizer adam. Tahapan selanjutnya adalah melakukan prediksi terhadap data testing menggunakan algoritma LSTM dengan kombinasi parameter *GridSearchCV*. Code ditunjukkan pada Gambar 12.

```
lstm_model = GridSearchCV(estimator=keras_regressor, param_grid=parameters, cv=2)
lstm_model.fit(train_X, train_scaler[column_name])

lstm_model = lstm_model.best_estimator_
forecast_sc = lstm_model.predict(test_X)

forecast_sc = forecast_sc.reshape(-1, 1)
scaler = scalers[column_name]
forecast = scaler.inverse_transform(forecast_sc)
```

Gambar 12. Prediksi data testing

4.5. Evaluasi (Evaluation)

Penelitian ini menggunakan *Root Mean Square Error* (RMSE) sebagai metrik untuk mengevaluasi kinerja model. Tabel 3 menunjukkan hasil evaluasi kinerja model LSTM.

Tabel 3. Evaluasi model

Nama Obat	MAPE(%)
Obat1	8.71%
Obat2	5.21%
Obat3	14.49%
Obat4	4.82%
Obat5	8.01%

Berdasarkan Kategori nilai MAPE, nilai MAPE yang ditunjukkan pada Tabel 1, obat obat1, obat2, obat4, dan obat5 termasuk dalam kategori model dengan kinerja sangat baik, sedangkan obat3 termasuk dalam kategori baik.

4.6. Penyebaran (Deployment)

Tahapan berikutnya yaitu *deployment*, nilai prediksi untuk 6 bulan berikutnya ditampilkan pada Tabel 3. Tabel disajikan dalam laporan.

Tabel 3. Prediksi Stok 12 Bulan

Date	Obat1	Obat2	Obat3	Obat4	Obat5
2023-12	19230	4300	3115	3627	1562
2024-1	18775	3713	5168	1228	1607
2024-2	16229	3829	2344	3699	1361
2024-3	19305	3743	3365	1078	1658
2024-4	16244	3412	4560	3772	1443
2024-5	19686	3573	2399	931	1353
2024-6	20167	3466	3487	3845	1300
2024-7	19308	3429	4033	788	1202
2024-8	20242	3333	2480	3915	1211
2024-9	19944	3279	3485	650	1176
2024-10	21742	3262	3605	3981	1138
2024-11	22405	3213	2557	520	1104

5. KESIMPULAN DAN SARAN

Penelitian ini bertujuan memprediksi stok lima jenis obat di UPTD Puskesmas Beber untuk periode 2019-2023 menggunakan metode LSTM, yang diterapkan dengan penyesuaian hyperparameter melalui *GridSearchCV*. Evaluasi prediksi menggunakan MAPE menunjukkan bahwa metode LSTM sangat efektif, dengan prediksi untuk obat obat1, obat2, obat4, dan obat5 termasuk dalam kategori model dengan kinerja sangat baik, sedangkan obat3 termasuk dalam kategori baik. Kesimpulannya, metode LSTM efektif untuk peramalan stok obat di Puskesmas Beber dengan nilai MAPE terendah sebesar 4.82% pada obat4. Saran untuk penelitian selanjutnya yaitu mempertimbangkan faktor lain selain data penggunaan obat yang dapat mempengaruhi perencanaan stok obat di puskesmas dan menerapkan metode LSTM pada data multivariat untuk meningkatkan akurasi dan relevansi prediksi stok obat.

DAFTAR PUSTAKA

- M. I. Anshory, Y. Priyandari, dan Yuniaristanto, "Peramalan Penjualan Sediaan Farmasi Menggunakan Long Short-term Memory: Studi Kasus pada Apotik Suganda.", *Performa: Media Ilmiah Teknik Industri*, vol. 19, no. 2, pp. 159-

174. 2020. Berikut adalah contoh sitasi IEEE untuk tesis tersebut:
- [2] E. N. Alam, "Prediksi Jumlah Kebutuhan Obat dengan Support Vector Regression dan Long Short-Term Memory," M.S. thesis, Institut Teknologi Bandung, 2019.
- [3] E. S. Putri and M. Sadiki, "Prediksi Penjualan Produk Untuk Mengestimasi Kebutuhan Bahan Baku Menggunakan Perbandingan Algoritma LSTM dan ARIMA," *Jurnal Fakultas Teknik Informatika Format*, vol. 10, no. 2, Universitas Mercu Buana, Jakarta, 2021. DOI: <http://dx.doi.org/10.22441/format.2021.v10.i2.007>.
- [4] T. I. Pratiwi, "Analisis Reaksi Pasar Modal Negara ASEAN Terhadap Pengumuman COVID-19," *Indonesia Banking School (IBS)*, Jakarta, 2021.
- [5] C. Schröer, F. Kruse, and J. M. Gómez, "A systematic literature review on applying crisp-dm process model," *Procedia Computer Science*, vol. 181, pp. 526–534, 2021. [Online]. Available: <https://doi.org/10.1016/j.procs.2021.01.199>
- [6] A. Bayangkari Karno, "Analisis Data Time Series Menggunakan LSTM (Long Short Term Memory) Dan ARIMA (Autocorrelation Integrated Moving Average) Dalam Bahasa Python.", *Ultima InfoSys : Jurnal Ilmu Sistem Informasi*, vol. 11, no. 1, pp. 1-7, Jul. 2020.
- [7] N. Selle, N. Yudistira, dan C. Dewi, "Perbandingan Prediksi Penggunaan Listrik dengan Menggunakan Metode Long Short Term Memory (LSTM) dan Recurrent Neural Network (RNN)," *Jurnal Teknologi Informasi Dan Ilmu Komputer*, vol. 9, no. 1, pp. 155-162, 2022.
- [8] A. Lathifah, "Cross-Industry Standard Process For Data Mining (Crisp-DM) Untuk Menemukan Pola Asosiasi Pada Data Tracer Study Lulusan Perguruan Tinggi Keagamaan Islam," UIN Syarif Hidayatullah Jakarta, Jakarta, 2023.
- [9] M. Karakaya, "LSTM: Understanding the Number of Parameters," *Medium*, 2020. [Online]. Available: <https://medium.com/deep-learning-with-keras/lstm-understanding-the-number-of-parameters-c4e087575756>.
- [10] A. Tussifah, "Analisis Perbandingan Kinerja Model ARIMA, LSTM, dan GRU Pada Stock Price Forecasting," M.S. thesis, Universitas Islam Negeri Syarif Hidayatullah, 2022.
- [11] I. Listiowarni, N.P. Dewi, dan A.K.W. Hapantenda, "Perbandingan Metode Double Exponential Smoothing dan Double Moving Average untuk Peramalan Harga Beras Eceran di Kabupaten Pamekasan," *Jurnal Komputer Terapan*, vol. 6, no. 2, pp. 158–169, 2020.
- [12] L. Wiranda and M. Sadikin, "Penerapan Long Short Term Memory pada Data Time Series untuk Memprediksi Penjualan Produk PT. Metiska Farma," *Jurnal Nasional Pendidikan Teknik Informatika*, vol. 8, no. 3, pp. 184, Dec. 2019, ISSN: 2089-8673 (Print), ISSN: 2548-4265 (Online).
- [13] A. Kurniawati, M. Sabri Ahmad, M. Fhadli, dan S. Lutfi, "Analisis Perbandingan Metode Time Series Forecasting Untuk Prediksi Penjualan Obat Di Apotek (Studi Kasus: Kimia Farma Apotek Takoma)", *JATI*, vol. 3, no. 1, hlm. 96–106, Jul 2023.