

**PREDIKSI PENJUALAN OBAT MENGGUNAKAN MODEL *LSTM* DAN
ANALISIS *TIME SERIES* PADA DATA TRANSAKSI PASIEN BPJS**

SKRIPSI

Diajukan sebagai Salah Satu Syarat untuk Mendapatkan
Gelar Sarjana Komputer (S.Kom) Program Studi Informatika



MUH. YUSUF SYAHBUL BASYAR

105841112020

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS MUHAMMADIYAH MAKASSAR**

2025



MAJELIS PENDIDIKAN TINGGI PIMPINAN PUSAT MUHAMMADIYAH
UNIVERSITAS MUHAMMADIYAH MAKASSAR
FAKULTAS TEKNIK



PENGESAHAN

Skripsi atas nama **Muh. Yusuf Syahbul Basyar** dengan nomor induk Mahasiswa **105841112020**, dinyatakan diterima dan disahkan oleh Panitia Ujian Tugas Akhir/Skripsi sesuai dengan Surat Keputusan Dekan Fakultas Teknik Universitas Muhammadiyah Makassar Nomor : 0001/SK-Y/55202/091004/2025, sebagai salah satu syarat guna memperoleh gelar Sarjana Teknik pada Program Studi Teknik Informatika Fakultas Teknik Universitas Muhammadiyah Makassar pada hari Kamis, 27 Februari 2025.

Panitia Ujian :

1. Pengawas Umum

Makassar,

14 Dzulhijjah 1446 H
10 Juni 2025 M

a. Rektor Universitas Muhammadiyah Makassar

Dr. Ir. H. Abd. Rakhim Nanda, ST., MT., IPU

b. Dekan Fakultas Teknik Universitas Hasanuddin

Prof. Dr. Eng. Muhammad Isran Ramli, S.T., MT., ASEAN, Eng

2. Penguji

a. Ketua : Rizki Yuslana Bakti ST., MT.

b. Sekretaris : Chyquitha Danuputri, S.Kom., M.T.

3. Anggota

1. Lukman, S.Kom., MT.

2. Titin Wahyuni, S.Pd., MT.

3. Desi Anggreani, S.Kom., MT.

Mengetahui

Pembimbing I

Pembimbing II

Muhyiddin A.M. Hayat, S.Kom., M.T.

Fahrim Irfan Rahman, S.Kom., M.T.

Dekan



Muhammad Syafaat S. Kuba, S.T., M.T.

NBM : 975 288

Gedung Menara Iqra Lantai 3

Jl. Sultan Alauddin No. 259 Telp. (0411) 866 972 Fax (0411) 865 588 Makassar 90221

Web: <https://teknik.unismuh.ac.id/>, e-mail: teknik@unismuh.ac.id





HALAMAN PENGESAHAN

Tugas Akhir ini diajukan untuk memenuhi syarat ujian guna memperoleh gelar Sarjana Komputer (S.Kom) Program Studi Informatika Fakultas Teknik Universitas Muhammadiyah Makassar.

Judul Skripsi : PREDIKSI PENJUALAN OBAT MENGGUNAKAN MODEL LSTM DAN ANALISIS TIME SERIES PADA DATA TRANSAKSI PASIEN BPJS

Nama : MUH. YUSUF SYAHBUL BASYAR

Stambuk : 105 84 111 2020

Makassar, 10 Juni 2025

Telah Diperiksa dan Disetujui
Oleh Dosen Pembimbing;

Pembimbing I

Pembimbing II

Muhyiddin A. M. Hayat, S.Kom., M.T.

Fahrim Irfhamna Rahman, S.Kom., M.T.

Mengetahui,
Ketua Prodi Informatika



Muhyiddin A. M. Hayat, S.Kom., M.T.

NBM : 1504 577



ABSTRAK

MUH. YUSUF SYAHBUL BASYAR. Prediksi Penjualan Obat Menggunakan Model Lstm dan Analisis Time Series Pada Data Transaksi Pasien Bpjs (dibimbing oleh Muhyddin A.M. Hayat, S.Kom., MT dan Fahrim Irhamna Rachman, S.Kom., M.T).

Prediksi penjualan obat yang akurat diperlukan untuk meningkatkan efisiensi persediaan dan kualitas pelayanan kesehatan, terutama di PT Kimia Farma Apotek (KFA) yang melayani pasien BPJS Kesehatan. Permintaan obat yang fluktuatif dan pola penjualan yang kompleks menjadi tantangan dalam perencanaan persediaan. Penelitian ini bertujuan untuk memprediksi penjualan obat secara akurat dengan menerapkan teknik data mining menggunakan model Long Short-Term Memory (LSTM). Data dikumpulkan dari transaksi di PT Kimia Farma Apotek (KFA) dan diproses melalui pra-pemrosesan yang mencakup pembersihan data, transformasi menjadi time series, serta pembagian data menjadi 80% data latih dan 20% data uji. Model LSTM dilatih selama 50 epoch menggunakan fungsi aktivasi ReLU dan Softmax pada lapisan output untuk menghasilkan probabilitas prediksi. Hasil penelitian menunjukkan bahwa model LSTM memiliki Mean Squared Error (MSE) sebesar 0.0021 dan akurasi prediksi mencapai 99.79%, menunjukkan kemampuan model dalam menghasilkan prediksi yang mendekati nilai aktual. Namun, R^2 Score sebesar -0.14 menunjukkan keterbatasan dalam menangkap variasi data yang kompleks. Penelitian ini menunjukkan bahwa model LSTM memiliki potensi besar untuk diterapkan dalam prediksi penjualan obat dan membantu pengambilan keputusan strategis dalam manajemen persediaan. Selain itu, pendekatan ini berkontribusi positif dalam meningkatkan efisiensi perencanaan persediaan obat di PT Kimia Farma Apotek.

Kata kunci: long short-term memory (LSTM), data mining, prediksi penjualan obat, time series, bpjs kesehatan, kimia farma apotek.

ABSTRACT

MUH. YUSUF SYAHBUL BASYAR. *Drug Sales Prediction Using LSTM Model and Time Series Analysis on BPJS Patient Transaction Data (supervised by Muhyddin A.M. Hayat, S.Kom., MT and Fahrim Irhamna Rachman, S.Kom., M.T).*

Accurate drug sales predictions are essential for improving inventory efficiency and healthcare service quality, especially at PT Kimia Farma Apotek (KFA), which serves BPJS Health patients. Fluctuating drug demand and complex sales patterns pose challenges in inventory planning. This study aims to accurately predict drug sales by applying data mining techniques using the Long Short-Term Memory (LSTM) model. Data was collected from transactions at PT Kimia Farma Apotek (KFA) and underwent preprocessing, including data cleaning, transformation into time series, and splitting into 80% training data and 20% testing data. The LSTM model was trained for 50 epochs using ReLU and Softmax activation functions in the output layer to produce prediction probabilities. The results show that the LSTM model achieved a Mean Squared Error (MSE) of 0.0021 and a prediction accuracy of 99.79%, indicating the model's ability to generate predictions close to actual values. However, an R^2 Score of -0.14 reveals limitations in capturing complex data variations. This study demonstrates that the LSTM model has significant potential for implementation in drug sales prediction and can support strategic decision-making in inventory management. Additionally, this approach contributes positively to enhancing inventory planning efficiency at PT Kimia Farma Apotek.

Keywords: long short-term memory (LSTM), data mining, drug sales prediction, time series, bpjs health, kimia farma apotek.

KATA PENGANTAR

Segala puji bagi Allah SWT atas rahmat dan petunjuk-Nya kepada penulis. Semoga sholawat dan salam tercurah kepada Nabi Muhammad SAW, seorang revolusioner sejati yang menjadi teladan bagi seluruh umat, menyebarkan Islam hingga kita dapat merasakan berkahnya sebagai Muslim. Dengan izin-Nya, penulis berhasil menyelesaikan skripsi yang berjudul "Prediksi Penjualan Obat Menggunakan Model LSTM dan Analisis Time Series Pada Data Transaksi Pasien BPJS."

Penulisan skripsi ini disusun sebagai salah satu syarat untuk menyelesaikan Program Sarjana (S1) di Program Studi Informatika, Fakultas Teknik, Universitas Muhammadiyah Makassar. Penulis berharap skripsi ini dapat menjadi referensi tambahan bagi para pembaca, terutama mahasiswa informatika, dan masyarakat umum. Penulis menyadari bahwa banyak pihak yang terlibat dalam penyusunan skripsi ini. Oleh karena itu, penulis ingin mengucapkan terima kasih yang sebesar-besarnya kepada:

1. Kepada kedua Orang Tua yang senantiasa memberikan kasih sayang, bimbingan, dukungan materi, dan doa yang selalu dipanjatkan kepada Allah untuk kesuksesan penulis. Pengorbanan mereka yang tak ternilai telah menjadi motivasi utama yang memungkinkan penulis menyelesaikan perjalanan studi hingga meraih gelar sarjana.
2. Bapak Prof. Dr. H. Ambo Asse, M.Ag sebagai Rektor Universitas Muhammadiyah Makassar.
3. Ibu Dr. Hj. Ir. Nurnawaty, ST., MT sebagai Dekan Fakultas Teknik Universitas Muhammadiyah Makassar.
4. Bapak Muhyddin A.M. Hayat, S.Kom., MT sebagai Ketua Program Studi Informatika, Fakultas Teknik Universitas Muhammadiyah Makassar.
5. Bapak Muhyddin A.M. Hayat, S.Kom., MT sebagai Dosen Pembimbing I dan Bapak Fahrir Irhamna Rachman, S.Kom., M.T sebagai Dosen Pembimbing II yang selalu meluangkan waktu dan pemikirannya untuk membimbing dan mengarahkan penulis dalam penyusunan skripsi ini.

6. Seluruh Dosen dan Staf Fakultas Teknik Universitas Muhammadiyah Makassar.

Semoga Tuhan Yang Maha Esa memberikan pahala yang lebih besar kepada beliau sebagai penutup dari segala ucapan. Penulis berharap skripsi ini dapat memberikan manfaat bagi para pembaca pada umumnya dan khususnya bagi penulis sendiri. ***“Billahi Fii Sabilil Haq Fastabiqul Khaerat”***



Makassar 15 Juni 2024

Penulis

DAFTAR ISI

ABSTRAK.....	ii
ABSTRACT	iii
KATA PENGANTAR	iv
DAFTAR ISI	vi
DAFTAR GAMBAR	viii
DAFTAR TABLE	ix
DAFTAR ISTILAH.....	x
BAB I PENDAHULUAN	1
A. Latar Belakang	1
B. Rumusan Masalah	2
C. Tujuan Penelitian.....	3
D. Manfaat Penelitian	3
E. Ruang Lingkup Penelitian.....	3
F. Sistematika Penulisan.....	4
BAB II TINJAUAN PUSTAKA	5
A. Landasan Teori	5
B. Penelitian Terkait.....	10
C. Kerangka Pikir	14
BAB III METODE PENELITIAN.....	15
A. Tempat dan Waktu Penelitian	15
B. Alat dan Bahan	15
C. Perancangan Sistem	15
D. Teknik Pengujian Sistem.....	18
E. Teknik Analisis Data.....	19
BAB IV HASIL DAN PEMBAHASAN	21
A. Pengambilan Data	21
B. Preprocessing Data	22
C. Implementasi Model.....	26

D. Hasil Pengujian Data.....	34
BAB V PENUTUP	40
A. KESIMPULAN.....	40
B. SARAN	41
DAFTAR PUSTAKA	42
LAMPIRAN	44



DAFTAR GAMBAR

Gambar 1. Arsitektur LSTM.....	7
Gambar 2. Kerangka pikir	14
Gambar 3. Proses Penelitian.....	16
Gambar 4. Flowchart Model LSTM	17
Gambar 5. Pengambilan data transaksi pasien BPJS	21
Gambar 6. Data transaksi pasien BPJS.....	22
Gambar 7. Hasil pelatihan model LSTM.....	35
Gambar 8. Grafik Hasil Perbandingan Nilai Aktual dan Nilai Prediksi pada setiap Epoch....	38




DAFTAR TABLE

Table 1. Tahap Sebelum Cleaning.....	23
Table 2. Tahap Sesudah Cleaning	24
Table 3. Hasil Transformasi Data.....	25
Table 4. Hasil Evaluasi Model LSTM.....	36



DAFTAR ISTILAH



Data Mining	: Proses menggali informasi atau pola tersembunyi dari kumpulan data yang besar menggunakan teknik statistik, machine learning, atau algoritma lainnya untuk mendukung pengambilan keputusan.
BPJS Kesehatan	: Badan Penyelenggara Jaminan Sosial Kesehatan di Indonesia yang menyediakan layanan kesehatan untuk masyarakat.
Data Time Series	: Data yang dikumpulkan atau diukur dalam urutan waktu yang berurutan dan biasanya digunakan untuk analisis tren atau prediksi masa depan.
Epoch	: Satu putaran penuh melalui seluruh data pelatihan dalam proses pembelajaran model machine learning.
LSTM (Long Short-Term Memory)	: Model jaringan saraf buatan yang digunakan dalam deep learning untuk mengatasi masalah ketergantungan jangka panjang dalam data time series.
Mean Squared Error (MSE)	: Metode evaluasi model yang menghitung rata-rata kuadrat kesalahan antara nilai prediksi dan nilai aktual.
Model Prediksi	: Algoritma yang digunakan untuk memperkirakan nilai masa depan berdasarkan data historis.

Pra-pemrosesan Data	: Tahap awal dalam data mining yang mencakup pembersihan data, transformasi data, dan pembagian data untuk meningkatkan kualitas data sebelum digunakan dalam model.
ReLU (Rectified Linear Unit)	: Fungsi aktivasi yang digunakan dalam neural network untuk memperkenalkan non-linearitas pada model.
R ² Score	: Metrik evaluasi yang menunjukkan seberapa baik model memprediksi variasi dalam data target.
Softmax	: Fungsi aktivasi pada lapisan output yang menghasilkan probabilitas prediksi untuk beberapa kelas.
Time Series Analysis	: Teknik analisis yang digunakan untuk mengevaluasi data time series guna menemukan pola, tren musiman, dan perilaku data lainnya.
Flowchart	: Diagram yang menampilkan alur atau langkah-langkah dalam suatu program atau prosedur sistem secara logis dan terstruktur.
Python	: Bahasa pemrograman tingkat tinggi yang umum digunakan dalam pengembangan perangkat lunak dan pemrograman komputer.

BAB I

PENDAHULUAN

A. Latar Belakang

Dalam industri kesehatan, terutama di sektor farmasi, prediksi penjualan obat memiliki peran yang sangat penting dalam manajemen persediaan, perencanaan produksi, dan pelayanan kepada pasien. Salah satu kelompok yang memerlukan perhatian khusus adalah peserta BPJS (Badan Penyelenggara Jaminan Sosial) di Indonesia. Badan usaha farmasi yang merupakan badan usaha yang mencari laba dengan melakukan riset, mengembangkan, dan menjual obat-obatan khususnya dalam ruang lingkup kesehatan. Badan usaha farmasi mempunyai persaingan ketat yang terjadi karena dampak dari tingginya penawaran serta permintaan obat dikalangan masyarakat baik kalangan bawah, menengah serta juga kalangan atas. Dengan jumlah peserta yang sangat besar dan kebutuhan obat yang bervariasi, manajemen persediaan obat menjadi tantangan yang kompleks dan krusial. Menanggapi tantangan ini, PT Kimia Farma Apotek (KFA) menyediakan layanan kesehatan yang terintegrasi, meliputi layanan farmasi (apotek), klinik kesehatan, laboratorium klinik, dan optik, dengan konsep *One Stop Health Care Solution* (OSHcS), sehingga semakin memudahkan masyarakat mendapatkan layanan kesehatan berkualitas (Sulaeman & Nurcahyani, 2022).

Memahami pentingnya menganalisis data berbasis waktu dan memproyeksikan nilai-nilai masa depan dari data *time series* merupakan salah satu tantangan utama yang dihadapi oleh banyak analis di berbagai sektor. Sektor-sektor tersebut mencakup keuangan dan ekonomi, manajemen produksi, analisis kebijakan sosial dan politik, serta penelitian tentang dampak keputusan manusia dan kebijakan terhadap lingkungan (Wiratama et al., 2023). Penelitian ini akan menggunakan data transaksi pasien BPJS sebagai dasar untuk melakukan analisis *time series* dan implementasi model LSTM (Long Short Time Memory) dalam memprediksi penjualan obat.

Selain itu, analisis time series pada data transaksi pasien BPJS akan memberikan wawasan mendalam mengenai pola penjualan obat, tren musiman, dan faktor-faktor yang mempengaruhi permintaan obat. Teknik ini merupakan metode statistik yang digunakan untuk mengevaluasi data yang diukur atau diamati secara berkelanjutan selama periode waktu tertentu, sehingga memungkinkan pemahaman yang lebih baik tentang dinamika penjualan obat (Febriyana et al., 2023). Informasi ini sangat berharga bagi manajemen KFA dalam merumuskan strategi yang lebih efektif dan efisien dalam mengelola persediaan obat serta meningkatkan kualitas pelayanan kepada pasien.

Penelitian ini bertujuan untuk menganalisis pola-pola penjualan obat dari data transaksi pasien BPJS Kesehatan dan menerapkan model LSTM untuk menghasilkan prediksi penjualan obat berdasarkan informasi yang diperoleh dari data transaksi tersebut. LSTM adalah salah satu bentuk pengembangan neural network yang dirancang khusus untuk pemodelan data *time series*. Teknologi ini efektif dalam menangani ketergantungan jangka panjang pada data masukan (Angioni et al., 2021).

Penelitian ini diharapkan dapat menganalisis pola penjualan obat secara efektif dan menggunakan model LSTM untuk menghasilkan prediksi penjualan yang lebih akurat, yang akan meningkatkan efisiensi perencanaan persediaan obat dan mendukung pengambilan keputusan yang lebih baik dalam manajemen obat di fasilitas kesehatan.

B. Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, rumusan masalah dalam penelitian ini adalah: “Bagaimana model LSTM dapat diterapkan untuk memprediksi penjualan obat secara akurat berdasarkan data transaksi pasien BPJS Kesehatan?”.

C. Tujuan Penelitian

Tujuan yang ingin dicapai dalam penelitian ini adalah: Untuk memprediksi penjualan obat secara akurat dari data transaksi pasien BPJS Kesehatan dengan menerapkan model LSTM.

D. Manfaat Penelitian

Manfaat yang diharapkan dari penelitian ini adalah:

1. Bagi User

Diharapkan penelitian ini dapat membantu pengguna dalam mengoptimalkan persediaan obat, mengurangi pemborosan, dan memastikan ketersediaan obat sesuai kebutuhan.

2. Bagi Peneliti

- a. Menjadi salah satu persyaratan yang harus dipenuhi untuk menyelesaikan program S1.
- b. Memperkaya literatur dan dapat memberikan kontribusi pada pengembangan keilmuan di bidang *data science*, khususnya dalam penerapan model LSTM untuk analisis pola permintaan obat.

3. Bagi Universitas

- a. Sebagai referensi untuk penelitian yang akan dilakukan di masa depan.
- b. Sebagai bahan evaluasi universitas untuk meningkatkan reputasi universitas dalam bidang riset dan pendidikan, terutama dalam bidang teknologi informasi dan kesehatan.

E. Ruang Lingkup Penelitian

1. Penelitian ini terbatas pada prediksi penjualan obat di apotek yang melayani pasien BPJS Kesehatan, dengan fokus pada PT Kimia Farma Apotek (KFA).
2. Dataset yang digunakan mencakup data transaksi penjualan obat dari pasien BPJS Kesehatan pada PT Kimia Farma Apotek (KFA).
3. Melakukan pra-pemrosesan data untuk menghapus atau mengoreksi data yang hilang atau tidak valid

4. Model yang digunakan untuk analisis dan prediksi adalah Long Short-Term Memory (LSTM).
5. Model akan dievaluasi menggunakan metrik evaluasi seperti *Mean Squared Error* (MSE), *Mean Absolute Error* (MAE), dan *Root Mean Squared Error* (RMSE).
6. Mengimplementasikan model yang telah dilatih untuk menghasilkan prediksi penjualan obat di masa depan.
7. Menganalisis hasil prediksi untuk menentukan apakah model LSTM memberikan prediksi yang akurat dan andal

F. Sistematika Penulisan

Secara umum, penulisan ini dibagi menjadi beberapa bagian utama sebagai berikut:

BAB I PENDAHULUAN

Bab ini menjelaskan secara singkat dan jelas latar belakang, rumusan masalah, tujuan dan manfaat, batasan masalah, metodologi yang digunakan, serta sistematika penulisan.

BAB II TINJAUAN PUSTAKA

Bab ini mengulas teori-teori yang mendukung penulis dalam menyelesaikan skripsi.

BAB III METODE PENELITIAN

Bab ini membahas teknik penelitian serta sumber daya yang digunakan dalam pengembangan sistem.

BAB II

TINJAUAN PUSTAKA

A. Landasan Teori

1. Prediksi Penjualan Obat

Prediksi adalah proses memperkirakan jumlah sesuatu yang akan terjadi di masa depan dengan menganalisis data masa lalu menggunakan metode ilmiah, terutama metode statistika. Dalam hal ini, prediksi melibatkan penggunaan data historis untuk mengestimasi nilai yang akan datang. Proses prediksi sering kali bergantung pada beberapa variabel prediksi, yang umumnya berasal dari data deret waktu historis, untuk memperkirakan permintaan atau kejadian di masa depan (ZUDI ITA BELA, 2022). Proses ini dilakukan dengan mengambil data masa lalu dan memasukkannya ke dalam masa depan menggunakan model yang sistematis. Jadi, dapat disimpulkan bahwa prediksi dapat memberikan hasil untuk masa yang akan datang dengan menghitung kejadian di masa sebelumnya. Namun, prediksi pada dasarnya tidak memiliki keakuratan 100%. Kesalahan prediksi di masa lalu digunakan untuk mengoreksi prediksi mendatang dengan cara yang berlawanan dengan kesalahan tersebut (Sari et al., 2023).

Penjualan adalah aktivitas penting bagi sebuah perusahaan, baik dalam bentuk barang maupun jasa, yang melibatkan transaksi antara dua pihak atau lebih dengan menggunakan alat pembayaran yang sah. Penjualan ini merupakan salah satu sumber pendapatan utama, baik bagi individu maupun perusahaan yang terlibat dalam transaksi jual beli. Untuk memenuhi permintaan pembeli dan mencapai kepuasan pelanggan, perusahaan mengembangkan berbagai rencana strategis guna memperoleh laba melalui penjualan. Semakin besar volume penjualan, semakin besar pula pendapatan yang dihasilkan oleh Perusahaan (Wibowo, 2018).

Dari penjelasan diatas, dapat disimpulkan bahwa Prediksi penjualan adalah proses ilmiah menggunakan analisis data masa lalu, terutama metode statistika, untuk memperkirakan penjualan di masa mendatang (Anshory et al., 2020). Proses ini melibatkan pengumpulan data historis dan penerapan model sistematis untuk

menghasilkan estimasi yang akurat. Meskipun tidak 100% akurat, kesalahan prediksi sebelumnya dapat meningkatkan keakuratan prediksi berikutnya. Prediksi penjualan sangat penting bagi perusahaan, karena membantu memahami pola dan tren penjualan, mengembangkan strategi efektif, memenuhi permintaan pembeli, meningkatkan kepuasan pelanggan, dan memaksimalkan laba.

2. Data Mining

Data Mining adalah proses menemukan korelasi, pola, dan tren baru yang bermakna dengan menyaring sejumlah besar data yang tersimpan dalam repositori, menggunakan teknologi pengenalan pola serta teknik statistik dan matematika. *Data mining* melibatkan pengambilan data dari data *warehouse* berdasarkan prediksi variabel. Prediksi ini digunakan untuk menemukan data dari sekumpulan data yang ditentukan dalam berbagai domain, seperti kecerdasan buatan, basis data, dan lainnya. Penggunaan data mining memiliki dampak signifikan di berbagai bidang, termasuk kesehatan. Dalam sistem medis, data mining diperlukan untuk mengekstrak informasi dari database sehingga memungkinkan diagnosis penyakit yang lebih akurat (Setiawan et al., 2022).

3. Data Time Series

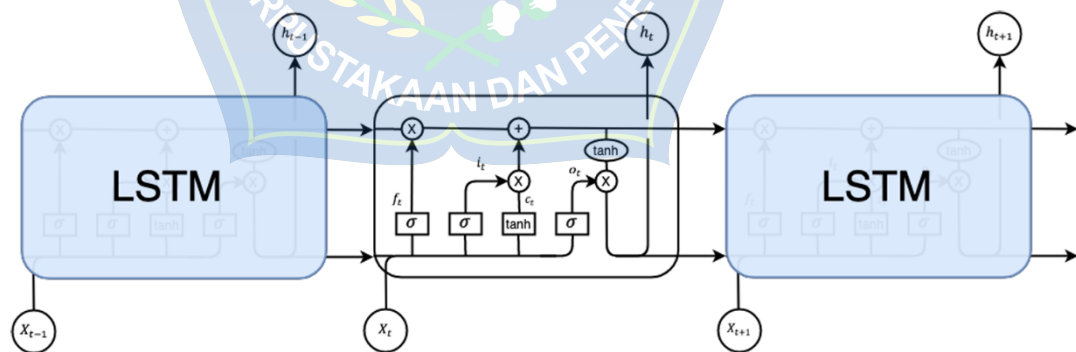
Time Series adalah metode kuantitatif yang melibatkan analisis banyak data berpola dari masa lalu, membuatnya sederhana dan efektif untuk perkiraan masa depan. *Data Time Series* adalah kumpulan data yang diukur atau diamati pada interval waktu yang konsisten, seperti harian, mingguan, bulanan, atau tahunan. Metode ini memiliki beberapa kelebihan, antara lain mudah digunakan untuk peramalan jangka pendek, dapat mewakili rentang waktu yang luas dari deret waktu jangka pendek secara fleksibel, dan menghasilkan perhitungan kuantitatif pada prosedur pengujian kesesuaian model. Hal ini menjadikannya acuan yang baik untuk prediksi penjualan obat (Syifa & Kusumawardani, 2023).

4. Long Short-Term Memory (LSTM)

Long Short-Term Memory (LSTM) adalah hasil pengembangan algoritma deep learning yang telah menghasilkan modifikasi dari metode RNN. Metode ini,

yang dikembangkan oleh Hochreiter dan Schmidhuber, telah diterapkan dalam berbagai bidang seperti peramalan (*forecasting*) dan pengenalan suara (*speech recognition*). LSTM dirancang untuk mengatasi kelemahan RNN (*Recurrent Neural Network*), yang kesulitan dalam membuat prediksi berdasarkan informasi masa lalu yang disimpan dalam periode waktu yang lama. Masalah utama pada RNN adalah seiring berjalannya waktu, informasi yang lebih lama cenderung dilupakan dan digantikan dengan memori baru, sehingga mengurangi akurasi prediksi. LSTM menyelesaikan masalah ini dengan kemampuan untuk mempertahankan informasi penting dalam jangka waktu yang lebih lama, meningkatkan akurasi dalam prediksi berbasis data urutan waktu. (Sundari, 2020).

LSTM mengatasi masalah ini dengan menggunakan memory cell dan gate unit yang mampu menyimpan informasi setiap masukan untuk jangka waktu panjang, bahkan ketika data yang digunakan sangat besar (Studi et al., 2020). Cell state memuat struktur yang disebut cell gate atau gerbang sel. Setiap gate terdiri dari empat bagian yaitu *input gate*, *forget gate*, *memory-cell state gate* dan *output gate*. Dalam membuat prediksi, LSTM menggunakan data historis (masa lampau) dan data saat ini. Berdasarkan struktur cell pada model LSTM, berikut gambar arsitektur LSTM (Kurniawati et al., 2023):



Gambar 1. Arsitektur LSTM

1. Foget Gate

Gate ini berperan dalam memproses informasi dari setiap data masukan serta menentukan data mana yang harus disimpan atau dibuang dalam sel memori. Langkah pertama dalam proses ini adalah menentukan informasi yang perlu dihilangkan dari keadaan sel. Lapisan sigmoid, yang dikenal sebagai "*forget gate layer*" bertugas mengambil keputusan tersebut. Forget gate layer menerima h_{t-1} dan x_t sebagai masukan, lalu menghasilkan output berupa nilai antara 0 dan 1 yang akan mempengaruhi keadaan sel sebelumnya, c_{t-1} .

2. Input Gate

Langkah berikutnya dalam LSTM adalah menentukan informasi baru yang akan disimpan dalam cell state. Proses ini melibatkan dua gerbang dalam input gate. Pertama, lapisan sigmoid yang disebut "*input gate layer*" berfungsi untuk menentukan nilai mana yang perlu diperbarui. Setelah itu, lapisan tanh menghasilkan vektor kandidat baru, \tilde{C}_t , yang dapat ditambahkan ke cell state. Kedua hasil ini kemudian digabungkan untuk memperbarui cell state. Berikut adalah rumus untuk *input gate layer* (i_t) dan vektor kandidat baru \tilde{C}_t .

3. Cell State

Pada tahapan ini, cell state lama (c_{t-1}) akan diperbaharui menjadi cell state baru c_t . Dengan mengalikan c_{t-1} lama dengan f_t , untuk menghapus informasi yang telah ditentukan sebelumnya pada tahapan forget gate layer. Selanjutnya tambahkan dengan $i_t * \tilde{C}_t$. Ini adalah nilai kandidat baru dan akan digunakan untuk memperbaharui cell state

4. Output Gate

Langkah terakhir dari LSTM bertujuan untuk memutuskan hasil keluaran. Output harus sesuai dengan cell state yang telah diproses sebelumnya. Pertama lapisan sigmoid dijalankan untuk menentukan

bagian cell state mana yang menjadi output. Kemudian, menempatkan cell state melalui tanh (untuk mendorong nilai antara -1 dan 1) dan mengalikannya dengan output dari sigmoid gate, sehingga hanya mengeluarkan sesuai dengan apa yang diputuskan sebelumnya

Dengan kemampuan ini, LSTM dapat memproses, memprediksi, dan mengklasifikasikan data dalam jangka waktu yang panjang, menjadikannya lebih efektif daripada RNN dalam berbagai aplikasi analisis data (Wiratama et al., 2023).

5. PT. Kimia Farma Apotek

PT. Kimia Farma Apotek adalah anak perusahaan PT. Kimia Farma (Persero) Tbk., salah satu perusahaan BUMN terbesar di Indonesia di bidang farmasi dan pelayanan kesehatan. Didirikan pada 4 Januari 2003, KFA menyediakan layanan kesehatan terintegrasi, termasuk apotek, klinik kesehatan, laboratorium klinik, dan optik, dengan konsep *One Stop Health Care Solution*. Sejak 2011, KFA telah memudahkan masyarakat mendapatkan layanan kesehatan berkualitas. Kimia Farma juga telah meningkatkan sistem ERP sebagai bagian dari kesiapan menghadapi era digital (Royyana, 2018).

6. BPJS Kesehatan Indonesia

Badan Penyelenggara Jaminan Sosial Kesehatan adalah badan hukum yang bertanggung jawab kepada Presiden, berfungsi menyelenggarakan program jaminan kesehatan di Indonesia. BPJS Kesehatan mengelola Jaminan Kesehatan Nasional untuk seluruh warga negara, mencakup layanan promotif, preventif, kuratif, dan rehabilitatif, termasuk obat dan bahan medis habis pakai. Dalam operasinya, BPJS Kesehatan bermitra dengan berbagai penyedia layanan kesehatan seperti rumah sakit, klinik, praktik dokter, apotek, dan optik (Septian, 2022).

Dengan demikian, BPJS Kesehatan Indonesia adalah lembaga yang mengelola program Jaminan Kesehatan Nasional, memastikan akses layanan kesehatan yang

komprehensif dan berkualitas bagi seluruh warga negara melalui kerja sama dengan berbagai penyedia layanan kesehatan.

7. Mean Squared Error (MSE)

Metode Mean Squared Error (MSE) adalah rata-rata dari kuadrat selisih antara nilai aktual dan nilai prediksi, digunakan untuk mengukur tingkat kesalahan dalam peramalan. MSE sangat penting dalam evaluasi model prediksi penjualan obat karena memberikan gambaran jelas tentang seberapa jauh prediksi model dari nilai aktual. Metrik ini membantu mengidentifikasi dan mengoreksi kesalahan prediksi, meningkatkan akurasi model (ZUDI ITA BELA, 2022). Semakin kecil nilai MSE, semakin baik model dalam meramalkan penjualan obat, karena menunjukkan perbedaan yang lebih kecil antara nilai prediksi dan nilai aktual. Selain itu, MSE memudahkan pemilihan model terbaik dari beberapa kandidat dengan membandingkan nilai MSE mereka.

8. Mean Absolute Error (MAE)

Metode Mean Absolute Error (MAE) adalah rata-rata dari nilai absolut kesalahan prediksi. Untuk mengetahui metode prediksi dengan akurasi tinggi, penting untuk menghitung tingkat kesalahan dalam prediksi tersebut. Semakin kecil tingkat kesalahan yang dihasilkan, semakin baik prediksinya. MAE adalah standar umum yang digunakan untuk mengukur kesalahan prediksi dan menilai akurasi model (ZUDI ITA BELA, 2022). Dengan menggunakan MAE, perusahaan dapat mengevaluasi performa berbagai model prediksi dan memilih model yang paling sesuai berdasarkan tingkat kesalahannya. Selain itu, MAE dapat digunakan untuk membandingkan kinerja model prediksi dari berbagai periode waktu atau berbagai jenis obat, memastikan strategi penjualan yang lebih baik dan pelayanan yang optimal kepada pelanggan.

B. Penelitian Terkait

1. Gunawan Abdi Putra Utama, 2021

Pada penelitian yang dilakukan oleh Gunawan Abdi Putra Utama dengan judul penelitian “Prediksi Transaksi Penjualan Produk Pada Data *Time Series*

Menggunakan LSTM Dalam Mengestimasi Persediaan Produk”. Penelitian ini bertujuan untuk mengevaluasi efektivitas model time series berbasis (LSTM) dalam memprediksi transaksi penjualan produk untuk estimasi persediaan produk. Dalam penelitian ini, digunakan algoritma LSTM untuk menganalisis data *time series* penjualan dan mengukur akurasi prediksi melalui nilai *Root Mean Square Error* pada data testing untuk lima produk secara berurutan. Hasil yang diperoleh menunjukkan nilai RMSE untuk kelima produk tersebut adalah 0.22, 0.23, 0.30, 0.09, dan 0.34. Berdasarkan hasil evaluasi, produk D menunjukkan performa terbaik dengan nilai RMSE terkecil yaitu 0.09. Temuan ini menunjukkan bahwa model LSTM yang digunakan dalam penelitian ini dapat menjadi alat yang bermanfaat bagi perusahaan untuk memperkirakan kebutuhan persediaan produk secara lebih akurat dan efektif.

2. Westlie Wiratama, Lutfi Aulia Alifah, Alta Gurusinga, dan Evta Indra, 2023

Pada penelitian yang dilakukan oleh Westlie Wiratama, Lutfi Aulia Alifah, Alta Gurusinga, dan Evta Indra dengan judul penelitian “Prediksi Turis Mancanegara ke Indonesia Menggunakan Metode EDA Time Series dan LSTM”. Penelitian ini menganalisis prediksi jumlah wisatawan mancanegara ke Indonesia menggunakan metode Exploratory Data Analysis (EDA) dan algoritma Long Short-Term Memory (LSTM) dalam konteks data time series. Berdasarkan hasil puncak jumlah wisatawan ke Indonesia terjadi pada tahun 2019 dengan lebih dari 15 juta wisatawan. Benua Asia adalah benua dengan jumlah wisatawan terbanyak, dengan Malaysia sebagai negara pengirim wisatawan terbanyak. Model LSTM menghasilkan nilai MAPE yang menunjukkan akurasi yang baik, yaitu 13,84% untuk Benua Asia dan 15,96% untuk Negara ASEAN. Prediksi menunjukkan jumlah wisatawan tertinggi akan terjadi pada Agustus 2024 dengan 580 ribu wisatawan dari Benua Asia dan pada Mei 2025 dengan 580 ribu wisatawan dari Negara ASEAN.

3. Efrike Sofyani Putri, 2020

Pada penelitian yang dilakukan oleh Efrike Sofyani Putri dengan judul penelitian “Prediksi Penjualan Produk PT Indonesia Saji Nusantara Untuk Mengestimasi Kebutuhan Bahan Baku Menggunakan Perbandingan Algoritma LSTM Dan ARIMA”. Penelitian ini mengevaluasi prediksi penjualan makanan frozen food di PT. Indonesia Saji Nusantara dengan menggunakan dua algoritma machine learning, yaitu Long Short-Term Memory (LSTM) dan AutoRegressive Integrated Moving Average (ARIMA). Hasil penelitian menunjukkan bahwa LSTM lebih akurat daripada ARIMA. Model LSTM menggunakan 63% data latih dan 37% data uji, dengan range interval $[-1,1]$ dan 100 epoch. Evaluasi menunjukkan bahwa LSTM memiliki nilai Root Mean Square Error (RMSE) rata-rata sebesar 0,22, sementara ARIMA sebesar 60,21. Nilai Mean Absolute Percentage Error (MAPE) untuk LSTM adalah 29,57%, lebih rendah dibandingkan dengan ARIMA yang sebesar 73%.

4. Linda Sundari, 2020

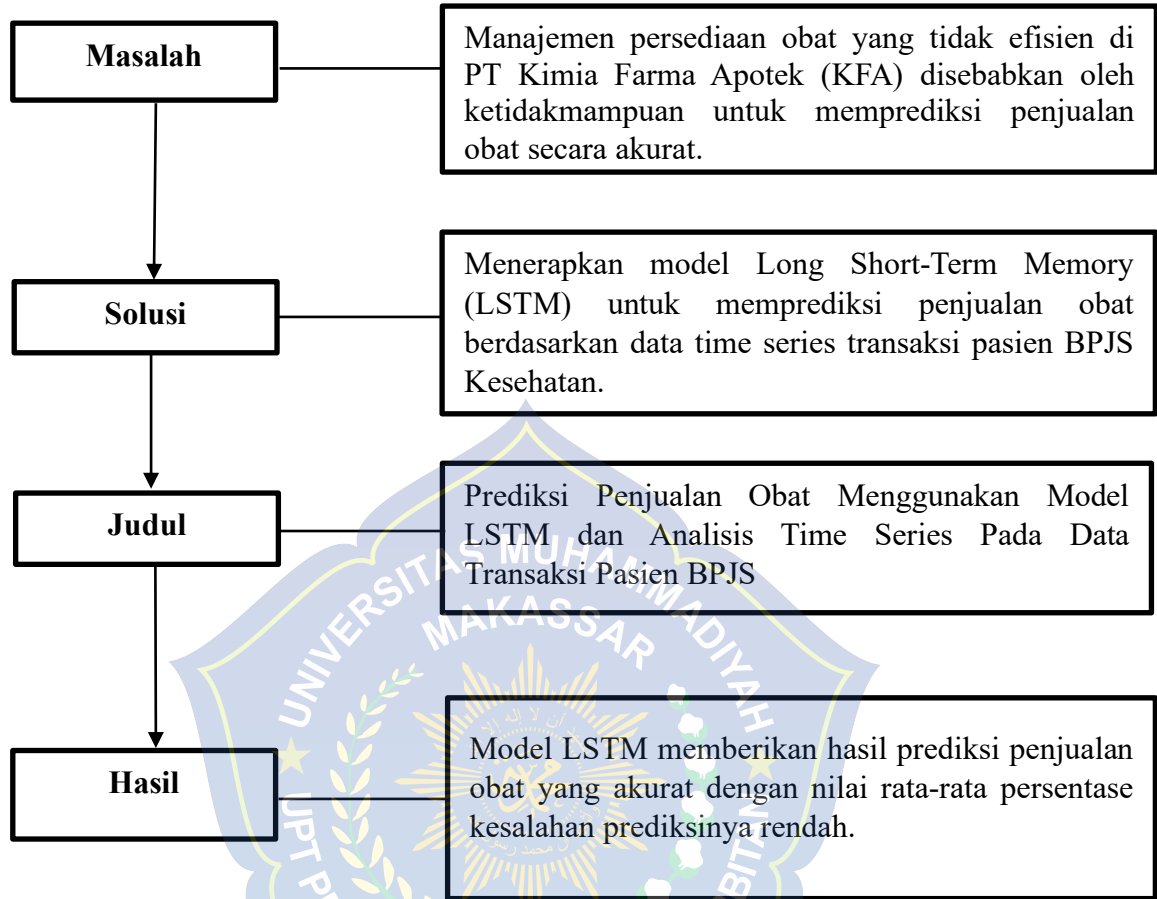
Pada penelitian yang dilakukan oleh Linda Sundari dengan judul penelitian “Pemodelan Time Series Untuk Peramalan Suhu Udara Menggunakan Metode Long Short Term Memory (LSTM) (Studi Kasus: Stasiun Klimatologi Lampung)”. Penelitian ini mengevaluasi peramalan suhu udara di Stasiun Klimatologi Lampung menggunakan model Long Short-Term Memory (LSTM). Hasil penelitian menunjukkan bahwa model LSTM terbaik terbentuk dari pembagian data training sebesar 80% dan data testing sebesar 20%. Model LSTM untuk peramalan Suhu Udara Minimum, Suhu Udara Maksimum, dan Suhu Udara Rata-Rata dibangun dari 25 neuron hidden, 4 batch size, dan 50 epoch. Evaluasi menunjukkan bahwa nilai Root Mean Square Error (RMSE) dan Mean Absolute Percentage Error (MAPE) untuk Suhu Udara Minimum adalah 0,89 dan 2,59%, untuk Suhu Udara Maksimum adalah 0,85 dan 2,32%, serta untuk Suhu Udara Rata-Rata adalah 0,88 dan 2,44%. Dengan nilai MAPE kurang dari 10%, model LSTM ini sangat baik untuk peramalan suhu udara. Peramalan Suhu Udara Minimum dan Suhu Udara Rata-Rata dari 01 November 2022 hingga 31 Januari

2023 menunjukkan tren penurunan, sementara Suhu Udara Maksimum menunjukkan tren peningkatan.

5. Maulana Ichwan Anshory, Yusuf Priyandari, dan Yuniaristanto, 2020

Pada penelitian yang dilakukan oleh Maulana Ichwan Anshory, Yusuf Priyandari, dan Yuniaristanto dengan judul penelitian “Peramalan Penjualan Sediaan Farmasi Menggunakan Long Short-term Memory: Studi Kasus pada Apotik Suganda”. Penelitian ini menganalisis akurasi dan efisiensi penggunaan Long Short-Term Memory (LSTM) untuk peramalan penjualan sediaan farmasi di Apotek Suganda. Hasil penelitian menunjukkan bahwa LSTM cukup akurat dalam melakukan peramalan dengan nilai rata-rata persentase kesalahan peramalan (MAPE) sekitar 4%. Tidak terdapat perbedaan signifikan antara nilai MAPE untuk peramalan permintaan produk secara harian (4,6541%) maupun mingguan (4,7109%). Parameter ideal yang digunakan dalam penelitian ini adalah n_steps sebesar 4 dan epoch sebesar 100. Namun, dalam implementasinya untuk pembuatan modul program peramalan yang akan diintegrasikan ke dalam aplikasi penjualan, perlu dipertimbangkan waktu yang dibutuhkan untuk proses pembentukan model LSTM menggunakan data training. Waktu yang dibutuhkan untuk training model cukup lama, yaitu sekitar 5,14 jam untuk data sebanyak 546 produk.

C. Kerangka Pikir



Gambar 2. Kerangka pikir

BAB III

METODE PENELITIAN

A. Tempat dan Waktu Penelitian

Penelitian ini dilakukan secara offline dengan mengambil beberapa data transaksi penjualan obat pada pasien BPJS. Adapun pelaksanaan penelitian ini dilakukan selama satu bulan pada bulan Juni-Juli 2024.

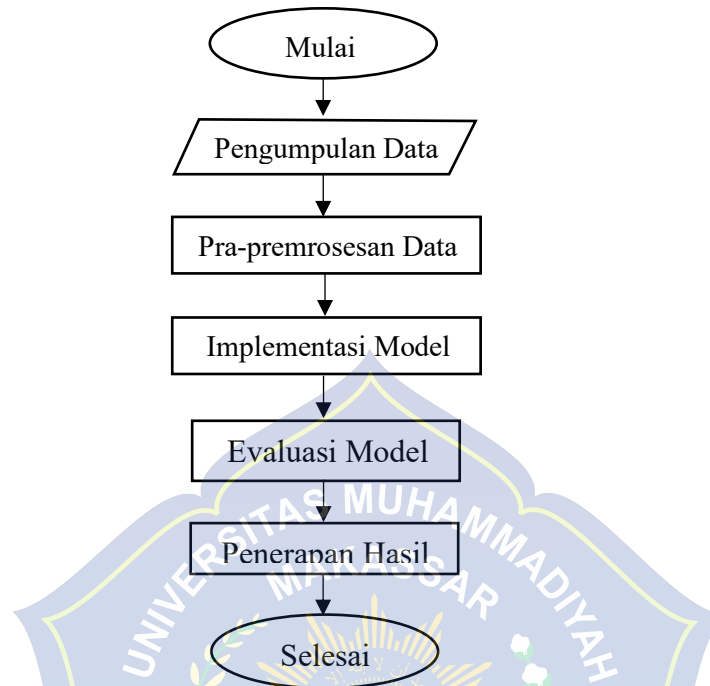
B. Alat dan Bahan

1. Kebutuhan Hardware (perangkat keras)
 - a. Laptop Acer Aspire 5
2. Kebutuhan Software (perangkat lunak)
 - a. Microsoft word
 - b. Google Colab
 - c. Excel
 - d. Python
 - e. Scikit-Learn

C. Perancangan Sistem

Flowchart atau diagram alur adalah representasi visual yang menggambarkan urutan langkah-langkah dan Keputusan yang diperlukan untuk menyelesaikan suatu proses dalam program. Setiap langkah digambarkan dalam bentuk diagram dan dihubungkan oleh garis atau panah yang menunjukkan arah aliran proses.

1. Flowchart Proses Penelitian

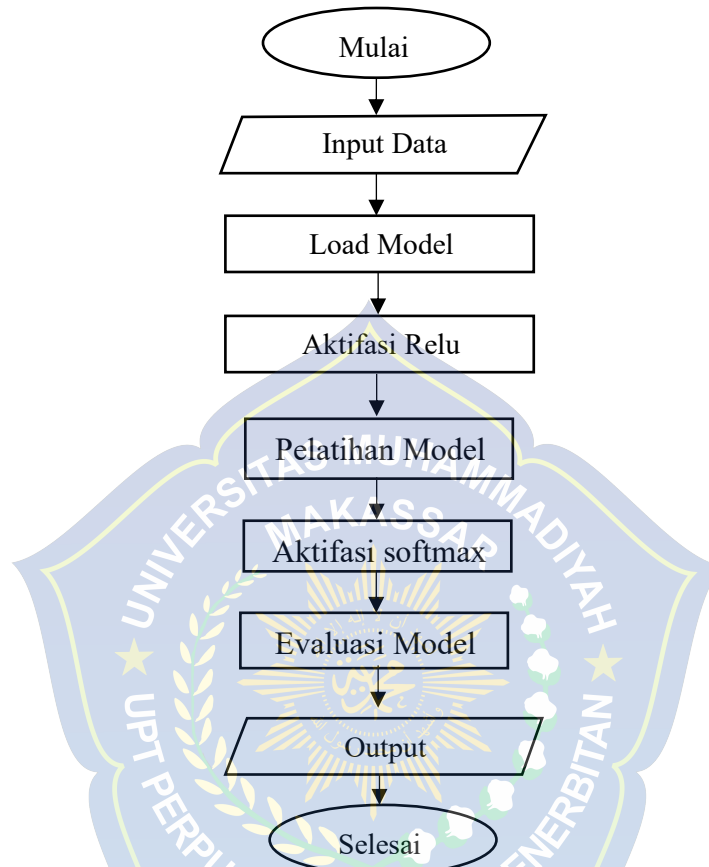


Gambar 3. Proses Penelitian

Pada gambar 3 diatas, proses penelitian dimulai dengan melakukan pengumpulan data, pengumpulan data dilakukan secara *offline* dari transaksi obat di PT Kimia Farma Apotek. Pada tahap pra-pemrosesan data, dilakukan langkah-langkah seperti pembersihan data: Menghapus data yang duplikat, menangani data yang hilang, dan memperbaiki inkonsistensi, transformasi data: Mengubah data transaksi menjadi format *time series* yang sesuai untuk model LSTM, dan pembagian data: data dibagi menjadi dua bagian, yaitu data latih dan data uji. Selanjutnya implementasi model untuk menentukan arsitektur jaringan LSTM seperti jumlah lapisan dan unit LSTM, kemudian menetapkan parameter seperti *learning rate*, *batch size*, dan *epoch* dan melatih model LSTM menggunakan data pelatihan. Selanjutnya evaluasi model dengan menghitung matrik untuk mengevaluasi akurasi prediksi model. Kemudian hasil yang dikeluarkan yaitu

hasil akurasi prediksi dari dari model LSTM. Terakhir penarikan kesimpulan dilakukan berdasarkan hasil analisis yang telah dilakukan.

2. Flowchart model LSTM



Gambar 4. Flowchart Model LSTM

Pada gambar 4 diatas, flowchart model LSTM dimulai dengan memasukkan data sebagai input, kemudian load model LSTM. Setelah itu, proses berlanjut ke fungsi aktivasi Relu untuk memperkenalkan non-linearitas ke dalam model. Selanjutnya melatih model menggunakan data latih untuk memprediksi, lalu menerapkan fungsi aktivasi Softmax pada layer output untuk menghasilkan probabilitas prediksi. Kemudian mengevaluasi model dengan data uji dan menghitung metrik performa seperti MSE, MAE, dan RMSE. Hasil akhir prediksi dari model LSTM diberikan sebagai output.

D. Teknik Pengujian Sistem

Dalam penelitian ini, teknik pengujian sistem dilakukan dengan menguji performa model LSTM dalam memprediksi penjualan obat berdasarkan data transaksi pasien BPJS. Proses pengujian mencakup beberapa tahap penting yang bertujuan untuk mengevaluasi seberapa efektif dan akurat model LSTM dalam melakukan prediksi penjualan.

Pengujian dilakukan dengan pembagian data transaksi obat menjadi dua bagian: 80% sebagai data latih dan 20% sebagai data uji. Model LSTM dilatih menggunakan data latih ini dengan parameter yang telah ditentukan seperti *learning rate*, *batch size*, dan *epoch*. Setelah model dilatih, prediksi dilakukan terhadap data uji. Untuk mengevaluasi akurasi dan efektivitas model LSTM, beberapa metrik evaluasi digunakan, termasuk *Mean Squared Error* (MSE), *Mean Absolute Error* (MAE), *Root Mean Squared Error* (RMSE).

Untuk mengevaluasi akurasi dan efektivitas model LSTM, beberapa metrik evaluasi digunakan, termasuk *Mean Squared Error* (MSE), *R2 Score*, dan *Accuracy*.

Rumus MSE:

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

- y_i = Nilai aktual
- \hat{y}_i = Nilai prediksi
- n = Jumlah data

Rumus R2 Score:

$$R2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (2)$$

- y_i = Nilai aktual
- \hat{y}_i = Nilai prediksi
- \bar{y} = Rata-rata nilai aktual
- n = Jumlah data

Rumus Accuracy:

$$\text{Accuracy} = \frac{\text{Jumlah Prediksi Benar}}{\text{Total Prediksi}} \times 100\% \quad (3)$$

Setelah evaluasi model LSTM, hasilnya berupa nilai kesalahan terkecil untuk memastikan bahwa model LSTM memberikan hasil yang optimal.

E. Teknik Analisis Data

Proses analisis data merupakan langkah sistematis untuk mengorganisir dan menginterpretasikan informasi yang diperoleh melalui wawancara, observasi, dan dokumentasi. Ini mencakup pengorganisasian data ke dalam kategori, pemecahan menjadi unit-unit kecil, sintesis, identifikasi pola, pemilihan informasi yang signifikan untuk dipelajari, serta pengambilan kesimpulan. Tujuan analisis data adalah untuk mempermudah pemahaman informasi, baik bagi peneliti maupun pihak lain yang terlibat atau membaca hasil analisis. Untuk mencapai hasil yang diinginkan, peneliti melakukan beberapa tahapan analisis berikut:

1. Pengumpulan Data:

Proses pengumpulan data melibatkan pengambilan informasi yang bersumber dari data historis dari transaksi obat di PT Kimia Farma Apotek (KFA) yang mencakup beberapa jenis data yaitu: Nama Pasien, No. Keanggotaan BPJS, No. SEP, No. Transaksi, No. Resep, Tanggal Pelayanan, Nama Obat, Kode SAP, Qty, HNA, Harga, dan Nama Penjamin.

2. Preprocessing:

Langkah selanjutnya adalah persiapan data untuk tahap pengolahan. Dalam penelitian ini, proses preprocessing melibatkan beberapa langkah penting seperti menghapus duplikasi, menangani data yang hilang, memperbaiki inkonsistensi data, Konversi data transaksi menjadi format time series, dan membagi data menjadi set pelatihan dan set pengujian.

3. Penyajian Data:

Peneliti menyajikan data yang telah direduksi secara sistematis dan terstruktur. Tujuan penyajian ini adalah untuk memudahkan pemahaman informasi yang terkandung dalam data.

4. Kesimpulan:

Peneliti menyajikan kesimpulan berdasarkan data yang telah dianalisis, dengan tujuan memberikan jawaban terhadap permasalahan yang diteliti.



BAB IV

HASIL DAN PEMBAHASAN

A. Pengambilan Data

Pengambilan data transaksi pasien BPJS yang digunakan dalam penelitian ini diperoleh melalui system manajemen Apotik Kimia Farma Cendrawasih menggunakan mekanisme ekstraksi langsung melalui SQL query yang dijalankan pada basis data internal Apotik. Data yang diambil mencakup informasi lengkap mengenai setiap transaksi penjualan obat dari pasien BPJS mulai dari nomor transaksi, nama pasien, tanggal transaksi, jenis obat, jumlah obat, hingga total biaya yang harus dibayarkan. Rentang waktu pengambilan data adalah 1 bulan dengan periode data selama 2 tahun dari Januari 2023 hingga Desember 2024 dengan tujuan untuk memberikan representasi tren yang cukup baik selama periode tersebut.

Pada gambar 4 menunjukkan proses pengambilan data melalui system manajemen Apotik Kimia Farma.

The screenshot displays the 'pos.kimiafarma.co.id/report/insurance.payment' web application. The interface includes a header with the 'POS kimia farma' logo and user information for 'KJ.Cendrawasih'. The main section is titled 'Laporan Apotek' and contains a form for 'Laporan Pembayaran dengan Jaminan'. This form has two columns of filters: 'Pilih Jangka Waktu' (with radio buttons for 'Semua', 'Satu bulan terakhir', and 'Satu minggu terakhir') and 'Pilih tanggal' (with 'Periode Awal' and 'Periode Akhir' date pickers). A 'FILTER' button is located below these filters. Below the filter section, there is a 'Laporan Pembayaran dengan Jaminan' section with a 'FILTER JAMINAN' button and a 'Cari' input field. At the bottom, a summary bar shows 'Total Pendapatan dari periode 05 Januari 2025 s/d 05 Januari 2025' with a value of 'Rp. 0'. A table header is visible at the very bottom, listing various transaction details like 'No. Transaksi', 'No. Resep', 'Tgl LPH', 'Nama Pasien', 'Kode SAP', 'Nama Obat', 'Qty', 'HNA', 'Harga/Obat', 'Embalase', 'Total Harga', 'Total Ditanggung', 'Poli RS', 'Nama Dokter', 'Nama Penjamin', 'No. Rujukan', 'No. Keanggotaan', 'No. SEP', and 'Status Bridging'. At the bottom right, there are buttons for 'UNDUH FILE EXCEL' and 'PREVIEW CETAK LAPORAN'.

Gambar 5. Pengambilan data transaksi pasien BPJS

Setelah melakukan pengambilan data transaksi pasien BPJS dari system manajemen apotik, kemudian hasilnya disimpan dalam format excel.

Berikut adalah hasil pengambilan data transaksi pasien BPJS yang telah disimpan ke dalam Excel:

Gambar 6. Data transaksi pasien BPJS

B. Preprocessing Data

Tahapan *Preprocessing Data* adalah langkah penting dalam menyiapkan dan membersihkan data mentah sebelum digunakan untuk analisis. Proses ini dilakukan secara manual dengan mencakup berbagai langkah untuk memastikan bahwa data dalam kondisi yang sesuai dan siap untuk diproses lebih lanjut.

Berikut Langkah-langkah dalam *Preprocessing Data*:

1. *Cleaning* (Pembersihan)

Tahap pembersihan data melibatkan proses pemeriksaan dan pengolahan data untuk memastikan konsistensi, kelengkapan, dan kualitasnya. Pada tahap pembersihan ini, data yang mengandung nilai kosong atau tidak relevan akan dihapus.

Table 1. Tahap Sebelum Cleaning

No	...	Nama obat	Jumlah obat	Harga/ obat	Total harga	Poli RS	Nama Penjamin	No Rujukan	...
1	...	Furosemide	30	Rp.145					...
	...	Candesartan	30	Rp.271	Rp.15.075	Umum	BPJS	-	...
	...	Bisoprolol	15	Rp.173					...
2	...	Candesartan	30	Rp.271					...
	...	Bisoprolol	30	Rp.173	Rp.16.230	Umum	BPJS	-	...
	...	Vitamin B Kompleks	30	Rp.97					...
...
19.413	...	Glimepiride	30	Rp.419					...
	...	Metformin	90	Rp.237	Rp.41.490	Umum	BPJS	-	...
	...	Simvastatin	30	Rp.253					...

Table 2. Tahap Sesudah Cleaning

No	Nama obat	Jumlah obat	Harga/ obat	Total harga
1	Furosemide	30	Rp.145	
	Candesartan	30	Rp.271	Rp.15.075
	Bisoprolol	15	Rp.173	
2	Candesartan	30	Rp.271	
	Bisoprolol	30	Rp.173	Rp.16.230
	Vitamin B Kompleks	30	Rp.97	

19.413	Glimepiride	30	Rp.419	
	Metformin	90	Rp.237	Rp.41.490
	Simvastatin	30	Rp.253	

2. Transformasi Data

Pada tahap Transformasi data merupakan proses mengubah data mentah menjadi format yang lebih sesuai untuk analisis atau pemodelan, khususnya dalam *Machine Learning* seperti LSTM. Dalam konteks penelitian ini, transformasi data mencakup penyusunan data time series yang akan digunakan sebagai input dan target dalam model prediksi.

Table 3. Hasil Transformasi Data

No	Nama obat	Harga	Jumlah obat Januari 2023	Total harga Januari 2023	Jumlah obat Februari 2023	Total harga Februari 2023	...
1	ACARBOSE 100MG TAB@100 GDX	Rp.1216	1200	Rp.1459200	1020	Rp.1240320	...
2	ACARBOSE 50 MG (DUS 100 TAB)	Rp.902	1980	Rp.1785960	0	Rp. 0	...
3	ADALAT OROS 30MG TAB@30 BPJS	Rp.6306	120	Rp.756720	210	Rp.1324260	...
...
204	VITAMIN B6 10MG TAB@200 GIMFAR BPJS	Rp.106	3811	Rp.403966	195	Rp.20670	...

	VITAMIN B6 HCL DUS@200					...
205	TAB (PEHA)	Rp.107	0	Rp.0	0	Rp.0

C. Implementasi Model

1. Implementasi Model LSTM

Model Long Short-Term Memory (LSTM) digunakan untuk memprediksi penjualan obat berdasarkan data historis. LSTM adalah jenis Recurrent Neural Network (RNN) yang dirancang untuk memproses dan menganalisis data sekuensial atau data time series. Model ini memiliki kemampuan untuk mengingat informasi dalam jangka waktu panjang, yang penting dalam analisis deret waktu seperti tren penjualan.

```
import pandas as pd
import numpy as np
from sklearn.preprocessing import MinMaxScaler
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import MinMaxScaler
from sklearn.model_selection import train_test_split
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import LSTM, Dense
import matplotlib.pyplot as plt
from sklearn.metrics import mean_squared_error, r2_score
```

Program ini memanfaatkan pustaka Python seperti pandas untuk membaca data, numpy untuk komputasi numerik, dan scikit-learn untuk normalisasi data dengan MinMaxScaler serta membagi data menjadi pelatihan dan pengujian menggunakan train_test_split. Model LSTM dibangun menggunakan TensorFlow, dengan layer LSTM untuk menangani data deret waktu dan layer Dense untuk menghasilkan prediksi. Evaluasi dilakukan

menggunakan metrik seperti MSE dan R^2 score, sementara visualisasi hasil dibantu oleh matplotlib.pyplot. Program ini bertujuan memprediksi tren penjualan obat berdasarkan data historis.

```
# Import data dari file Excel
data = pd.read_excel('Data penjualan obat pasien bpjs
(Edit).xlsx')
```

Perintah di atas digunakan untuk mengimpor data dari file Excel yang bernama 'Data penjualan obat pasien bpjs (Edit).xlsx' ke dalam sebuah DataFrame menggunakan pustaka pandas. Dengan menggunakan fungsi `pd.read_excel()`, file Excel tersebut dibaca dan isinya dimuat dalam variabel `data`, yang memungkinkan untuk dilakukan analisis atau manipulasi data lebih lanjut dalam format tabel.

```
# Hanya ambil kolom jumlah obat terjual (kolom genap
setelah harga obat)
jumlah_columns = [col for i, col in enumerate(data.columns)
if i > 2 and i % 2 == 0]
jumlah_data = data[jumlah_columns]
```

Perintah di atas memilih kolom-kolom yang berisi jumlah obat terjual, yang berada di posisi genap setelah kolom harga obat. Dengan menggunakan list comprehension, kolom-kolom tersebut dipilih berdasarkan indeks yang lebih besar dari 2 dan bernomor genap, lalu data dari kolom-kolom tersebut disimpan dalam variabel `jumlah_data`.

```
# Normalisasi data
scaler = MinMaxScaler()
normalized_data = scaler.fit_transform(jumlah_data)
```

Perintah di atas melakukan normalisasi data pada kolom jumlah obat terjual menggunakan `MinMaxScaler`. Fungsi `fit_transform()` digunakan untuk mengubah nilai dalam `jumlah_data` ke dalam rentang 0 hingga 1, lalu hasilnya disimpan dalam variabel `normalized_data`.

```
# Persiapan fitur dan target untuk LSTM
sequence_length = 3 # Panjang urutan waktu yang digunakan
sebagai input
X, y = [], []
```

Perintah di atas mempersiapkan data untuk model LSTM (Long Short-Term Memory) dengan menentukan panjang urutan waktu (`sequence_length`) yang akan digunakan sebagai input. Variabel `X` akan menyimpan fitur (input), sedangkan `y` akan menyimpan target (output) untuk pelatihan model. Namun, pada kode ini, `X` dan `y` masih berupa list kosong yang nantinya akan diisi dengan data yang dipersiapkan.

```
for i in range(len(normalized_data) - sequence_length):
    X.append(normalized_data[i:i + sequence_length])
    y.append(normalized_data[i + sequence_length])

X = np.array(X)
y = np.array(y)
```

Perintah di atas membangun urutan waktu untuk fitur (`X`) dan target (`y`) berdasarkan panjang urutan yang telah ditentukan. Dalam loop, data `normalized_data` dipotong menjadi sub-sekuens sepanjang `sequence_length` untuk fitur `X`, dan target `y` diambil sebagai elemen setelah urutan tersebut. Setelah loop selesai, `X` dan `y` diubah menjadi array NumPy untuk memudahkan pemrosesan lebih lanjut, seperti pelatihan model LSTM.

```
# Split data menjadi data latih dan data uji
X_train, X_test, y_train, y_test = train_test_split(X, y,
test_size=0.2, random_state=42)
```

Perintah di atas membagi data menjadi dua bagian: data latih (`X_train`, `y_train`) dan data uji (`X_test`, `y_test`) menggunakan fungsi `train_test_split`. 80% data digunakan untuk pelatihan, sementara 20% digunakan untuk pengujian (ditentukan dengan `test_size=0.2`). Proses ini juga dipastikan

konsisten dengan menggunakan `random_state=42` untuk mengontrol pembagian data yang acak.

```
print("Shape of X_train:", X_train.shape)
print("Shape of y_train:", y_train.shape)
print("Shape of X_test:", X_test.shape)
print("Shape of y_test:", y_test.shape)
```

Perintah di atas digunakan untuk menampilkan dimensi (shape) dari data latih (X_train, y_train) dan data uji (X_test, y_test). Fungsi shape akan menunjukkan jumlah baris (contoh data) dan kolom (fitur) dalam masing-masing array, sehingga kita dapat mengetahui ukuran data yang akan digunakan dalam pelatihan dan pengujian model.

```
# Membangun model LSTM
model = Sequential()
model.add(LSTM(50, activation='relu',
input_shape=(X_train.shape[1], X_train.shape[2])))
model.add(Dense(y_train.shape[1]))
model.compile(optimizer='adam', loss='mse')
```

Perintah di atas membangun sebuah model LSTM menggunakan Keras. Model ini dimulai dengan lapisan LSTM yang memiliki 50 unit dan fungsi aktivasi ReLU, yang menerima input berupa urutan data dengan bentuk yang ditentukan oleh dimensi data latih. Setelah itu, model dilengkapi dengan lapisan Dense yang menghasilkan output sesuai dengan jumlah kolom pada data target. Model kemudian dikompilasi dengan menggunakan optimizer adam dan fungsi loss mean squared error (mse), yang umum digunakan untuk masalah regresi.

```
# Untuk menyimpan hasil prediksi setelah setiap epoch
predictions_per_epoch = []
```


Perintah di atas mendefinisikan sebuah list kosong `predictions_per_epoch` yang digunakan untuk menyimpan hasil prediksi model setelah setiap epoch selama pelatihan. List ini memungkinkan untuk memantau perkembangan prediksi model seiring dengan berjalannya epoch.

```
# Fungsi callback untuk menyimpan prediksi setelah setiap epoch
def store_predictions(epoch, logs):
    y_pred_epoch = model.predict(X_test)
    predictions_per_epoch.append(y_pred_epoch.flatten())
```

Perintah di atas mendefinisikan sebuah fungsi callback `store_predictions` yang akan dipanggil setelah setiap epoch selama pelatihan. Fungsi ini melakukan prediksi pada data uji (`X_test`) menggunakan model yang telah dilatih, kemudian menyimpan hasil prediksi tersebut dalam list `predictions_per_epoch`. Prediksi tersebut diratakan (`flatten`) sebelum disimpan agar memiliki format yang sesuai untuk analisis lebih lanjut.

```
# Melatih model dengan callback
from tensorflow.keras.callbacks import LambdaCallback
```

Perintah di atas mengimpor `LambdaCallback` dari pustaka Keras. `LambdaCallback` memungkinkan pembuatan callback kustom yang dapat digunakan dalam proses pelatihan model. Callback ini bisa digunakan untuk berbagai tujuan, seperti menyimpan hasil prediksi setelah setiap epoch, memonitor metrik, atau melakukan tindakan lain sesuai kebutuhan saat pelatihan berlangsung.

```
history = model.fit(X_train, y_train, epochs=50,
                    batch_size=32, validation_data=(X_test, y_test),
                    verbose=1,
                    callbacks=[LambdaCallback(on_epoch_end=store_predictions)])
```

Perintah di atas melatih model LSTM dengan data latih (`X_train`, `y_train`) selama 50 epoch, menggunakan batch size 32. Selama pelatihan,

data uji (`X_test`, `y_test`) digunakan untuk validasi model. Fungsi `verbose=1` akan menampilkan informasi proses pelatihan. Callback `LambdaCallback(on_epoch_end=store_predictions)` digunakan untuk memanggil fungsi `store_predictions` setelah setiap epoch selesai, yang menyimpan hasil prediksi pada setiap epoch ke dalam list `predictions_per_epoch`.

```
# Menampilkan ringkasan model
model.summary()
```

Perintah di atas menampilkan ringkasan (summary) dari model yang telah dibangun. Fungsi `model.summary()` akan menunjukkan informasi tentang arsitektur model, termasuk jumlah lapisan, jenis lapisan, jumlah parameter yang dapat dilatih, serta bentuk input dan output setiap lapisan. Ini berguna untuk memverifikasi struktur model sebelum pelatihan dilakukan.

2. Evaluasi Model LSTM

```
# Evaluasi model
eval_loss = model.evaluate(X_test, y_test, verbose=0)
print("Evaluation Loss:", eval_loss)
```

Perintah di atas mengevaluasi model menggunakan data uji (`X_test`, `y_test`). Fungsi `model.evaluate()` menghitung nilai loss (kerugian) pada data uji berdasarkan fungsi loss yang telah ditentukan sebelumnya (dalam hal ini, mse). Nilai loss yang dihitung kemudian disimpan dalam variabel `eval_loss` dan ditampilkan dengan menggunakan `print()`. Dengan `verbose=0`, tidak ada informasi tambahan yang akan ditampilkan selama evaluasi.

```
# Prediksi pada data uji
y_pred = model.predict(X_test)
```

Perintah di atas digunakan untuk membuat prediksi menggunakan model pada data uji (`X_test`). Fungsi `model.predict()` akan menghasilkan prediksi berdasarkan input yang diberikan dan menyimpan hasilnya dalam

variabel `y_pred`. Prediksi ini dapat digunakan untuk menganalisis performa model atau membandingkannya dengan nilai target sebenarnya (`y_test`).

```
# Menghitung akurasi dengan metrik tambahan
mse = mean_squared_error(y_test, y_pred)
r2 = r2_score(y_test, y_pred)
accuracy = (1 - mse) * 100 # Estimasi persentase akurasi
sebagai inversi MSE
print("Mean Squared Error:", mse)
print("R2 Score:", r2)
print("Accuracy (%):", accuracy)
```

Perintah di atas digunakan untuk mengevaluasi performa model dengan tiga metrik utama. Pertama, **Mean Squared Error (MSE)** dihitung menggunakan `mean_squared_error()` untuk mengukur seberapa besar kesalahan rata-rata kuadrat antara nilai target (`y_test`) dan hasil prediksi (`y_pred`). Selanjutnya, **R2 Score** dihitung dengan `r2 score()` untuk menilai sejauh mana model dapat menjelaskan variansi dalam data. Terakhir, **Accuracy** dihitung sebagai inversi dari MSE, dengan rumus $(1 - mse) * 100$, yang memberikan estimasi persentase akurasi model. Hasil dari ketiga metrik ini kemudian ditampilkan untuk menilai kualitas model secara keseluruhan.

```
# Plot hasil prediksi vs nilai aktual
plt.figure(figsize=(10, 6))
plt.plot(y_test.flatten(), label='Actual', color='blue')
plt.plot(y_pred.flatten(), label='Predicted', color='red',
linestyle='dashed')
plt.title('Actual vs Predicted Values')
plt.xlabel('Sample Index')
plt.ylabel('Normalized Value')
plt.legend()
plt.show()
```

Perintah di atas digunakan untuk memvisualisasikan hasil prediksi model dibandingkan dengan nilai aktual. Menggunakan matplotlib, grafik ditampilkan dengan sumbu x mewakili indeks sampel, dan sumbu y mewakili nilai yang dinormalisasi. Dua kurva digambar: satu untuk nilai **aktual** (`y_test`, berwarna biru) dan satu untuk nilai **prediksi** (`y_pred`, berwarna merah dengan garis putus-putus). Judul, label sumbu, dan legenda ditambahkan untuk memperjelas grafik, yang memungkinkan perbandingan visual antara nilai yang diprediksi dan nilai yang sesungguhnya.

```
# Grafik: Prediksi vs Nilai Aktual dan Epoch
plt.figure(figsize=(10, 6))
```

Perintah di atas memulai pembuatan grafik dengan ukuran 10x6 inci menggunakan matplotlib. Ini menyediakan area untuk menampilkan visualisasi yang lebih besar dan jelas, yang dapat digunakan untuk membandingkan hasil prediksi terhadap nilai aktual seiring dengan perubahan pada epoch pelatihan model. Selanjutnya, elemen-elemen seperti data dan label akan ditambahkan untuk menyelesaikan grafik sesuai kebutuhan.

```
# Plot Actual (Data Uji)
plt.plot(y_test.flatten(), label='Actual', color='blue')

# Plot Prediksi untuk setiap epoch
for i, pred in enumerate(predictions_per_epoch):
    plt.plot(pred, label=f'Predicted at Epoch {i+1}',
             linestyle='dashed', alpha=0.5)

plt.title('Actual vs Predicted Values at Different Epochs')
plt.xlabel('Sample Index')
plt.ylabel('Normalized Value')
plt.legend()
plt.show()
```

Perintah di atas digunakan untuk memvisualisasikan perbandingan antara nilai aktual dan hasil prediksi model pada setiap epoch pelatihan. Grafik yang dihasilkan menampilkan kurva biru untuk nilai aktual dari data uji (y_{test}), sementara setiap prediksi pada epoch yang berbeda digambarkan dengan garis putus-putus berwarna yang berbeda dan transparansi untuk memberikan gambaran yang jelas mengenai perkembangan prediksi. Dengan menambahkan judul, label sumbu, dan legenda, grafik ini memungkinkan kita untuk memantau bagaimana model meningkatkan prediksi seiring berjalannya waktu dan seiring dengan setiap epoch pelatihan.

D. Hasil Pengujian Data

Pengujian data dalam penelitian ini dimulai dengan pembagian data menjadi dua bagian, yaitu data latih (training data) dan data uji (testing data), menggunakan metode train-test split. Data yang digunakan untuk pelatihan berjumlah 80%, sementara sisanya, yaitu 20%, digunakan sebagai data uji untuk mengevaluasi performa model. Data yang digunakan adalah data penjualan obat dengan beberapa fitur yang relevan, seperti jumlah obat terjual dan harga obat.

Setelah pembagian data, model LSTM dibangun dan dilatih menggunakan data latih. Model ini dirancang untuk memprediksi jumlah obat terjual berdasarkan urutan waktu. Dalam proses pelatihan, model mengalami 50 epoch, yang memungkinkan model untuk mempelajari hubungan temporal antara berbagai titik data yang ada.

```

Epoch 1/50
2/2 _____ 0s 201ms/step
6/6 _____ 3s 144ms/step - loss: 0.0106 - val_loss: 0.0020
Epoch 2/50
2/2 _____ 0s 5ms/step
6/6 _____ 1s 36ms/step - loss: 0.0094 - val_loss: 0.0020
Epoch 3/50
2/2 _____ 0s 4ms/step
6/6 _____ 0s 30ms/step - loss: 0.0083 - val_loss: 0.0020
Epoch 4/50
2/2 _____ 0s 9ms/step
6/6 _____ 0s 40ms/step - loss: 0.0054 - val_loss: 0.0020
Epoch 5/50
2/2 _____ 0s 9ms/step
6/6 _____ 0s 43ms/step - loss: 0.0075 - val_loss: 0.0020
Epoch 6/50
2/2 _____ 0s 8ms/step
6/6 _____ 0s 39ms/step - loss: 0.0076 - val_loss: 0.0020
Epoch 7/50
2/2 _____ 0s 8ms/step
6/6 _____ 0s 36ms/step - loss: 0.0055 - val_loss: 0.0020
Epoch 8/50
2/2 _____ 0s 6ms/step
6/6 _____ 0s 46ms/step - loss: 0.0051 - val_loss: 0.0020
Epoch 9/50
2/2 _____ 0s 5ms/step
6/6 _____ 0s 40ms/step - loss: 0.0062 - val_loss: 0.0022
Epoch 10/50
2/2 _____ 0s 5ms/step
6/6 _____ 0s 23ms/step - loss: 0.0067 - val_loss: 0.0023
Epoch 11/50
2/2 _____ 0s 6ms/step
6/6 _____ 0s 29ms/step - loss: 0.0071 - val_loss: 0.0021
Epoch 12/50
2/2 _____ 0s 6ms/step
6/6 _____ 0s 25ms/step - loss: 0.0082 - val_loss: 0.0021
Epoch 13/50
2/2 _____ 0s 5ms/step
6/6 _____ 0s 24ms/step - loss: 0.0064 - val_loss: 0.0021
Epoch 14/50
2/2 _____ 0s 4ms/step
6/6 _____ 0s 23ms/step - loss: 0.0090 - val_loss: 0.0021
Epoch 15/50
2/2 _____ 0s 7ms/step
6/6 _____ 0s 24ms/step - loss: 0.0062 - val_loss: 0.0022
Epoch 16/50
2/2 _____ 0s 4ms/step
6/6 _____ 0s 28ms/step - loss: 0.0086 - val_loss: 0.0022
Epoch 17/50
2/2 _____ 0s 5ms/step
6/6 _____ 0s 26ms/step - loss: 0.0087 - val_loss: 0.0021
Epoch 18/50
2/2 _____ 0s 5ms/step
6/6 _____ 0s 24ms/step - loss: 0.0057 - val_loss: 0.0021
Epoch 19/50
2/2 _____ 0s 4ms/step
6/6 _____ 0s 25ms/step - loss: 0.0085 - val_loss: 0.0021
Epoch 20/50
2/2 _____ 0s 6ms/step
6/6 _____ 0s 25ms/step - loss: 0.0061 - val_loss: 0.0021
Epoch 21/50
2/2 _____ 0s 6ms/step
6/6 _____ 0s 30ms/step - loss: 0.0066 - val_loss: 0.0021
Epoch 22/50
2/2 _____ 0s 6ms/step
6/6 _____ 0s 24ms/step - loss: 0.0123 - val_loss: 0.0021
Epoch 23/50
2/2 _____ 0s 5ms/step
6/6 _____ 0s 23ms/step - loss: 0.0118 - val_loss: 0.0021
Epoch 24/50
2/2 _____ 0s 5ms/step
6/6 _____ 0s 23ms/step - loss: 0.0077 - val_loss: 0.0021

```

Gambar 7. Hasil pelatihan model LSTM

Proses pelatihan model LSTM berlangsung selama 50 epoch, dengan setiap epoch menghasilkan informasi mengenai waktu yang dibutuhkan untuk setiap langkah pelatihan serta nilai loss dan val_loss. Loss menunjukkan kesalahan pada data latih, sedangkan val_loss menunjukkan kesalahan pada data uji (validation). Pada awal epoch, nilai loss dan val_loss cukup fluktuatif, namun seiring berjalannya waktu, keduanya cenderung menurun, menunjukkan bahwa model semakin baik dalam memprediksi data. Meskipun demikian, ada beberapa lonjakan kecil pada nilai loss, yang mungkin disebabkan oleh variabilitas dalam data atau pelatihan. Pada akhir epoch ke-50, model mencatatkan loss sebesar 0.0077 dan val_loss sebesar 0.0021, menunjukkan bahwa model cukup stabil dan memiliki performa yang baik dalam memprediksi data uji.

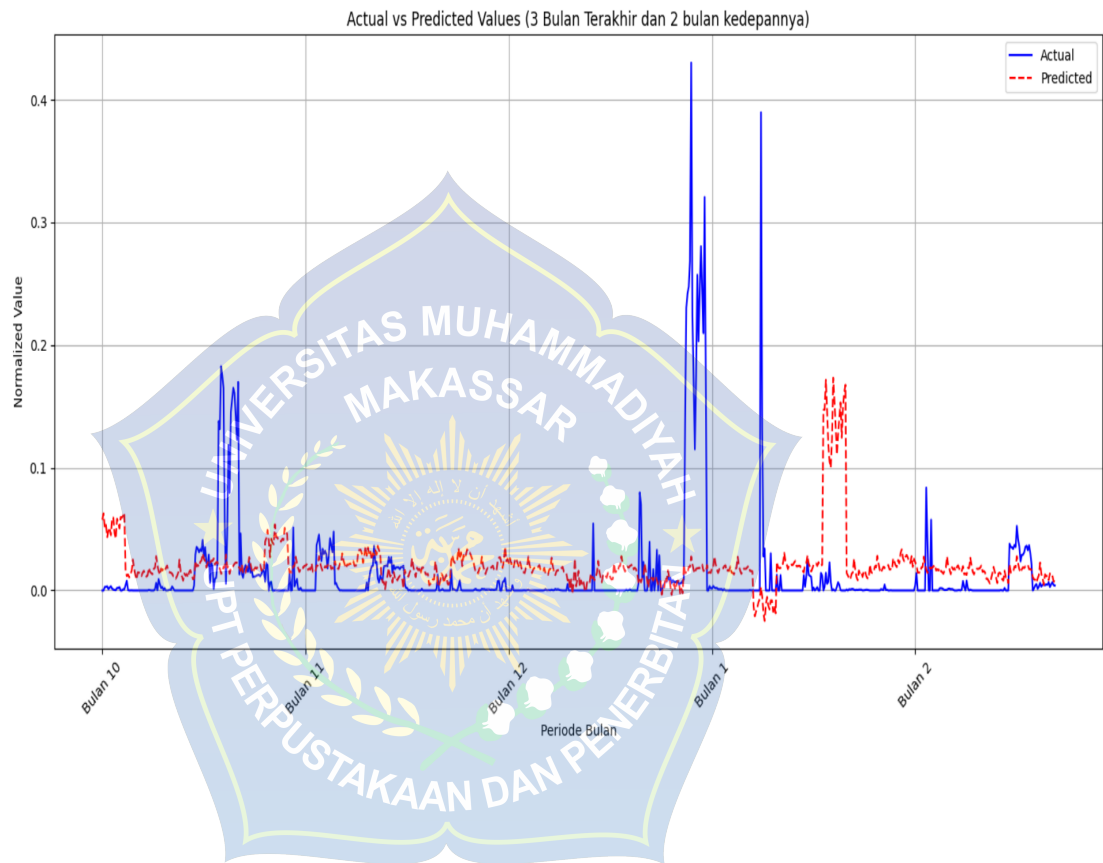
Setelah model selesai dilatih, dilakukan evaluasi terhadap data uji untuk mengukur kinerja model menggunakan beberapa metrik evaluasi.

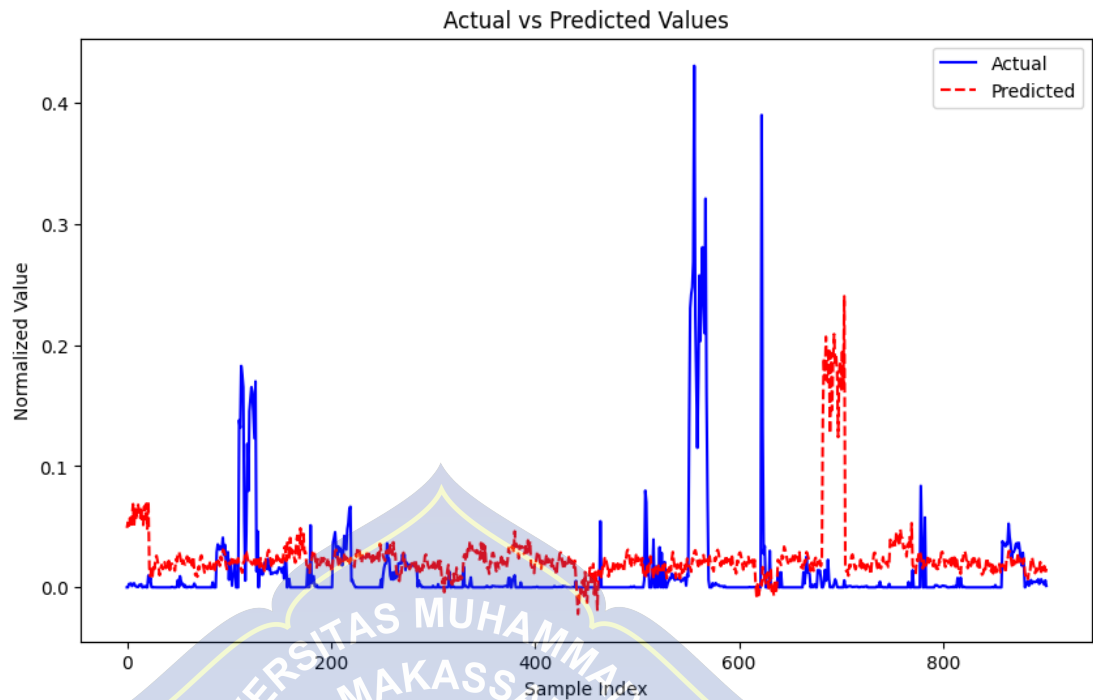
Table 4. Hasil Evaluasi Model LSTM

Mean Squared Error	: 0.0020862370230014444
R2 Score	: -0.14262983973904056
Accuracy (%)	: 99.79137629769986

Hasil evaluasi menunjukkan bahwa Mean Squared Error (MSE) mencapai nilai 0.0020862370230014444, yang mengindikasikan bahwa rata-rata kesalahan kuadrat antara nilai prediksi dan nilai aktual sangat kecil. Nilai R2 Score sebesar -0.14262983973904056 menunjukkan adanya peluang untuk lebih meningkatkan kemampuan model dalam menangkap variasi data secara menyeluruh. Meskipun begitu, model berhasil menghasilkan prediksi rata-rata yang mendekati nilai aktual dengan akurasi tinggi. Di sisi lain, akurasi yang

dihitung sebagai inversi dari MSE menghasilkan nilai 99.79137629769986%, yang menunjukkan bahwa sebagian besar prediksi sangat dekat dengan nilai target. Hasil ini memberikan indikasi positif bahwa model memiliki potensi yang baik sebagai dasar analisis dan dapat ditingkatkan lebih lanjut untuk menangkap pola dan tren yang lebih kompleks dalam data.





Gambar 8. Grafik Hasil Perbandingan Nilai Aktual dan Nilai Prediksi pada setiap Epoch

Grafik ini menampilkan perbandingan antara nilai aktual (garis biru) dan nilai prediksi model pada berbagai epoch selama proses pelatihan. Dengan demikian, setiap titik pada grafik merepresentasikan penjualan bulanan. Dari total data, 800 sampel (20% dari keseluruhan dataset) digunakan untuk pengujian, yang mewakili 5 bulan terakhir dari periode waktu dataset, sementara sisanya, yaitu 3200 sampel, digunakan untuk pelatihan model. Data disusun secara kronologis sehingga data uji mencakup 800 data terakhir sebelum waktu sekarang dalam urutan waktu. Karena terdapat 4000 data dalam 24 bulan, rata-rata terdapat sekitar 166,67 data per bulan. Ini menunjukkan bahwa lebih dari satu data dikumpulkan setiap bulan. Dengan 800 data uji, periode waktu yang tercakup adalah sekitar 4,8 bulan atau dibulatkan menjadi 5 bulan terakhir dalam dataset. Pada grafik, sumbu X

menunjukkan urutan bulan dalam data uji, dimulai dari bulan pertama hingga bulan terakhir dari periode pengujian. Setiap titik pada grafik merepresentasikan penjualan bulanan pada bulan tersebut. Grafik ini menunjukkan prediksi penjualan bulanan untuk 800 data terakhir dari 5 bulan terakhir secara kronologis, sehingga dapat dianalisis performa model dalam memprediksi penjualan bulanan dari waktu ke waktu. Secara umum, terlihat bahwa model menunjukkan peningkatan performa seiring bertambahnya epoch. Pada epoch awal, prediksi model masih cenderung fluktuatif dan kurang sesuai dengan nilai aktual. Namun, pada epoch-epoch berikutnya, prediksi model menjadi semakin konsisten dan mendekati pola nilai aktual, terutama untuk data dengan nilai yang lebih rendah. Hal ini menunjukkan bahwa model berhasil mempelajari pola dasar dalam data secara bertahap selama proses pelatihan. Meskipun terdapat beberapa tantangan dalam menangkap puncak-puncak yang tinggi, prediksi yang stabil pada sebagian besar sampel mengindikasikan bahwa model memiliki potensi untuk terus ditingkatkan dalam memahami tren data penjualan obat.

BAB V

PENUTUP

A. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa model Long Short-Term Memory (LSTM) memiliki potensi yang signifikan dalam memprediksi jumlah penjualan obat berdasarkan data historis. Berikut adalah poin-poin utama dari kesimpulan penelitian ini:

1. Evaluasi Performa Model

Model LSTM berhasil dievaluasi dalam memprediksi jumlah penjualan obat berdasarkan data historis dan data yang digunakan telah dinormalisasi untuk meningkatkan efektivitas pelatihan model.

2. Pembagian Data dan Pelatihan Model

Data dibagi menjadi 80% untuk pelatihan dan 20% untuk pengujian, Model dilatih selama 50 epoch untuk memahami pola temporal dalam data.

3. Hasil Pengujian Model

Dari hasil pengujian model tersebut menunjukkan nilai Mean Squared Error (MSE) yang rendah (0.0021) menunjukkan rata-rata kesalahan prediksi yang kecil sehingga Akurasi prediksi berdasarkan inversi MSE mencapai 99.79%, yang menunjukkan prediksi mendekati nilai actual. Nilai R^2 Score sebesar -0.14 mengindikasikan model masih memiliki keterbatasan dalam menangkap tren kompleks, terutama pada data dengan fluktuasi tinggi.

4. Analisis Grafik Prediksi

Model mampu memprediksi nilai rata-rata dengan baik tetapi mengalami kesulitan dalam memprediksi nilai puncak yang tinggi. Diperlukan perbaikan lebih lanjut, seperti tuning parameter, penambahan fitur, atau eksplorasi arsitektur jaringan yang lebih kompleks.

5. Implikasi dan Potensi Penerapan

Model LSTM memiliki potensi besar untuk diterapkan dalam prediksi penjualan obat. Optimalisasi lebih lanjut dapat meningkatkan akurasi model dan membantu pengambilan keputusan strategis dalam manajemen stok obat. Pendekatan ini berkontribusi positif bagi sektor kesehatan, terutama dalam efisiensi pengelolaan persediaan obat.

B. SARAN

Berdasarkan hasil penelitian, terdapat beberapa saran untuk pengembangan penelitian ini lebih lanjut yaitu:

1. Optimisasi model LSTM dapat dilakukan dengan mengeksplorasi parameter seperti jumlah neuron, ukuran batch, jumlah epoch, serta penerapan mekanisme dropout untuk mengurangi overfitting.
2. Menambahkan fitur yang relevan, seperti pola penyakit, hari libur, atau faktor musiman, dapat membantu meningkatkan akurasi model dalam memprediksi penjualan obat. Selain itu, penggunaan teknik rekayasa fitur, seperti lag features atau moving average, dapat memperkaya informasi dalam data.

DAFTAR PUSTAKA

- Angioni, S. A., Giansante, C., Ferri, N., Ballarin, L., Pampanin, D. M., Marin, M. G., Bargione, G., Vasapollo, C., Donato, F., Virgili, M., Petetta, A., Lucchetti, A., Cabuga Jr, C. C., Masendo, C. B. ., Hernando, B. J. ., Joseph, C. C. ., Velasco, J. P. ., Angco, M. K. ., Ayaton, M. A., ... Barile, N. B. (2021). PREDIKSI TRANSAKSI PENJUALAN PRODUK PADA DATA TIME SERIES MENGGUNAKAN LSTM DALAM MENGESTIMASI PERSEDIAAN PRODUK. *Fisheries Research*, 140(1), 6. [http://dspace.ucuenca.edu.ec/bitstream/123456789/35612/1/Trabajo de Titulacion.pdf%0Ahttps://educacion.gob.ec/wp-content/uploads/downloads/2019/01/GUIA-METODOLOGICA-EF.pdf%0Ahttp://dx.doi.org/10.1016/j.fishres.2013.04.005%0Ahttps://doi.org/10.1038/s41598-](http://dspace.ucuenca.edu.ec/bitstream/123456789/35612/1/Trabajo%20de%20Titulacion.pdf%0Ahttps://educacion.gob.ec/wp-content/uploads/downloads/2019/01/GUIA-METODOLOGICA-EF.pdf%0Ahttp://dx.doi.org/10.1016/j.fishres.2013.04.005%0Ahttps://doi.org/10.1038/s41598-)
- Anshory, M. I., Priyandari, Y., & Yuniaristanto, Y. (2020). Peramalan Penjualan Sediaan Farmasi Menggunakan Long Short-term Memory: Studi Kasus pada Apotik Suganda. *Performa: Media Ilmiah Teknik Industri*, 19(2), 159–174. <https://doi.org/10.20961/performa.19.2.45962>
- Febriyana, N., Ridho Abdillah, I., Lailatus, S., Oktavia, J., & Hidayati, C. (2023). Analisis Perbandingan Time Series Sektor Media Di Indonesia Perusahaan Media Yang Terdaftar Di Bursa Efek Indonesia Tahun 2020-2022. *Neraca Manajemen, Ekonomi*, 3(2).
- Kurniawati, A., Ahmad, M. S., Fhadli, M., & Lutfi, S. (2023). Analisis Perbandingan Metode Time Series Forecasting Untuk Prediksi Penjualan Obat Di Apotek (Studi Kasus : Kimia Farma Apotek Takoma) Comparative Analysis of Methodstime Series Forecasting for Prediction of Drug Sales in Pharmacy (Case Study : Chemical. *Jurnal Jaringan Dan ...*, 3(1), 96–106. <https://doi.org/00.0000/jati>
- Royyana, A. (2018). STRATEGI TRANSFORMASI DIGITAL PADA PT. KIMIA FARMA (PERSERO) TBK. *Journal of information system public health* 3(3). *Jurnal Sistem Informasi Kesehatan Masyarakat Journal of Information Systems for Public Health*, 3(3), 15–32.
- Sari, P. P., Hidayat, A. T., Wijaya, H. O. L., & Armanto. (2023). Prediksi Penjualan Obat Menggunakan Metode Forecasting Exponential Smoothing Models (Kasus Pada Apotek Simpang F). *BRAHMANA: Jurnal Penerapan Kecerdasan Buatan*, 4(2), 129–137.
- Septian, E. D. (2022). Kewajiban BPJS Kesehatan dalam Pemberian Pelayanan

- Telekonsultasi Klinis yang Dilakukan Antara Dokter dan Pasien BPJS. *Verdict: Journal of Law Science*, 1(1), 37–49. <https://www.ojs.wahanapublikasi.com/index.php/vjlaws/article/view/34>
- Setiawan, D., Surojudin, N., & Hadikristanto, W. (2022). Prediksi Penjualan Obat Dengan Algoritma Regresi Linear. *Prosiding Sains Dan Teknologi*, 1(1), 237–246.
- Studi, P., Informatika, T., Komputer, F. I., & Buana, U. M. (2020). <https://lib.mercubuana.ac.id/>.
- Sulaeman, R. S., & Nurcahyani, E. (2022). Analisis Altman Z-Score Untuk Menilai Tingkat Kesehatan Keuangan Perusahaan Manufaktur Sub Sektor Farmasi Yang Terdaftar Di Bursa Efek Indonesia Periode 2017-2020. *AKSELERASI: Jurnal Ilmiah Nasional*, 4(1), 77–92. <https://doi.org/10.54783/jin.v4i1.530>
- Sundari, L. (2020). *PEMODELAN TIME SERIES UNTUK PERAMALAN SUHU UDARA MENGGUNAKAN METODE LONG SHORT TERM MEMORY (LSTM) (Studi Kasus: Stasiun Klimatologi Lampung)*. July, 1–23.
- Syifa, M. K., & Kusumawardani, D. M. (2023). Implementasi Metode Time Series Dalam Forecasting Penggunaan Satusehat. *Jurnal Pengembangan Sistem Informasi Dan Informatika*, 4(4), 14–25. <https://doi.org/10.47747/jpsii.v4i4.1223>
- Wibowo, D. A. (2018). Prediksi Penjualan Obat Herbal Hp Pro Menggunakan Algoritma Neural Network. *Technologia: Jurnal Ilmiah*, 9(1), 33. <https://doi.org/10.31602/tji.v9i1.1100>
- Wiratama, W., Aulia Alifah, L., Gurusinga, A., Indra, E., Sistem Informasi, J., & Sains Dan Teknologi, F. (2023). Prediksi Turis Mancanegara ke Indonesia Menggunakan Metode EDA Time Series dan LSTM. *Jurnal Riset Sistem Informasi Dan Teknik Informatika (JURASIK)*, 8(2), 524–537. <https://tunasbangsa.ac.id/ejurnal/index.php/jurasik>
- ZUDI ITA BELA, Z. I. B. (2022). Sistem Prediksi Penjualan Obat Menggunakan Metode Single Moving Average (Studi Kasus: Apotek Wilujeng Kecamatan Panceng Kab.Gresik). *Indexia*, 4(1), 47. <https://doi.org/10.30587/indexia.v4i1.3638>

LAMPIRAN

Lampiran 1. Data Set transaksi pasien

No	...	<u>Nama obat</u>	<u>Jumlah</u> <u>obat</u>	<u>Harga/</u> <u>obat</u>	<u>Total</u> <u>harga</u>	<u>Poli</u> <u>RS</u>	<u>Nama</u> <u>Penjamin</u>	<u>No</u> <u>Rujukan</u>	...
1	...	Furosemide	30	Rp.145					...
	...	Candesartan	30	Rp.271	Rp.15.075	Umum	BPJS	-	...
	...	Bisoprolol	15	Rp.173					...
2	...	Candesartan	30	Rp.271					...
	...	Bisoprolol	30	Rp.173	Rp.16.230	Umum	BPJS	-	...
	...	Vitamin B Kompleks	30	Rp.97					...
...
19.413	...	Glimepiride	30	Rp.419					...
	...	Metformin	90	Rp.237	Rp.41.490	Umum	BPJS	-	...
	...	Simvastatin	30	Rp.253					...

No	Nama obat	Jumlah obat	Harga/ obat	Total harga
1	Furosemide	30	Rp.145	
	Candesartan	30	Rp.271	Rp.15.075
	Bisoprolol	15	Rp.173	
2	Candesartan	30	Rp.271	
	Bisoprolol	30	Rp.173	Rp.16.230
	Vitamin B Kompleks	30	Rp.97	
...
19.413	Glimepiride	30	Rp.419	
	Metformin	90	Rp.237	Rp.41.490
	Simvastatin	30	Rp.253	

No	Nama obat	Harga	Jumlah obat Januari 2023	Total harga Januari 2023	Jumlah obat Februari 2023	Total harga Februari 2023	...
1	ACARBOSE 100MG TAB@100 GDX	Rp.1216	1200	Rp.1459200	1020	Rp.1240320	...
2	ACARBOSE 50 MG (DUS 100 TAB)	Rp.902	1980	Rp.1785960	0	Rp. 0	...
3	ADALAT OROS 30MG TAB@30 BPJS	Rp.6306	120	Rp.756720	210	Rp.1324260	...
...
204	VITAMIN B6 10MG TAB@200 GIMFAR BPJS	Rp.106	3811	Rp.403966	195	Rp.20670	...
205	VITAMIN B6 HCL DUS@200 TAB (PEHA)	Rp.107	0	Rp.0	0	Rp.0	...

Lampiran 2. Source Code

```
import pandas as pd
import numpy as np
from sklearn.preprocessing import MinMaxScaler
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import MinMaxScaler
from sklearn.model_selection import train_test_split
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import LSTM, Dense
import matplotlib.pyplot as plt
from sklearn.metrics import mean_squared_error, r2_score

# Import data dari file Excel
```

```

data = pd.read_excel('Data penjualan obat pasien bpjs
(Edit).xlsx')

# Hanya ambil kolom jumlah obat terjual (kolom genap
setelah harga obat)
jumlah_columns = [col for i, col in enumerate(data.columns)
if i > 2 and i % 2 == 0]
jumlah_data = data[jumlah_columns]

# Normalisasi data
scaler = MinMaxScaler()
normalized_data = scaler.fit_transform(jumlah_data)

# Persiapan fitur dan target untuk LSTM
sequence_length = 3 # Panjang urutan waktu yang digunakan
sebagai input
X, y = [], []
for i in range(len(normalized_data) - sequence_length):
    X.append(normalized_data[i:i + sequence_length])
    y.append(normalized_data[i + sequence_length])

X = np.array(X)
y = np.array(y)
# Split data menjadi data latih dan data uji
X_train, X_test, y_train, y_test = train_test_split(X, y,
test_size=0.2, random_state=42)
print("Shape of X_train:", X_train.shape)
print("Shape of y_train:", y_train.shape)
print("Shape of X_test:", X_test.shape)
print("Shape of y_test:", y_test.shape)

# Membangun model LSTM
model = Sequential()
model.add(LSTM(50, activation='relu',
input_shape=(X_train.shape[1], X_train.shape[2])))
model.add(Dense(y_train.shape[1]))
model.compile(optimizer='adam', loss='mse')

# Untuk menyimpan hasil prediksi setelah setiap epoch
predictions_per_epoch = []

```

```

# Fungsi callback untuk menyimpan prediksi setelah setiap
epoch
def store_predictions(epoch, logs):
    y_pred_epoch = model.predict(X_test)
    predictions_per_epoch.append(y_pred_epoch.flatten())
# Melatih model dengan callback
from tensorflow.keras.callbacks import LambdaCallback
history = model.fit(X_train, y_train, epochs=50,
batch_size=32, validation_data=(X_test, y_test),
                        verbose=1,
callbacks=[LambdaCallback(on_epoch_end=store_predictions)])
# Menampilkan ringkasan model
model.summary()
# Evaluasi model
eval_loss = model.evaluate(X_test, y_test, verbose=0)
print("Evaluation Loss:", eval_loss)
# Prediksi pada data uji
y_pred = model.predict(X_test)
# Menghitung akurasi dengan metrik tambahan
mse = mean_squared_error(y_test, y_pred)
r2 = r2_score(y_test, y_pred)
accuracy = (1 - mse) * 100 # Estimasi persentase akurasi
sebagai inversi MSE
print("Mean Squared Error:", mse)
print("R2 Score:", r2)
print("Accuracy (%):", accuracy)
# Plot hasil prediksi vs nilai aktual
plt.figure(figsize=(10, 6))
plt.plot(y_test.flatten(), label='Actual', color='blue')
plt.plot(y_pred.flatten(), label='Predicted', color='red',
linestyle='dashed')
plt.title('Actual vs Predicted Values')
plt.xlabel('Sample Index')

```

```

plt.ylabel('Normalized Value')
plt.legend()
plt.show()

# Grafik: Prediksi vs Nilai Aktual dan Epoch
plt.figure(figsize=(10, 6))
# Plot Actual (Data Uji)
plt.plot(y_test.flatten(), label='Actual', color='blue')

# Plot Prediksi untuk setiap epoch
for i, pred in enumerate(predictions_per_epoch):
    plt.plot(pred, label=f'Predicted at Epoch {i+1}',
             linestyle='dashed', alpha=0.5)

plt.title('Actual vs Predicted Values at Different Epochs')
plt.xlabel('Sample Index')
plt.ylabel('Normalized Value')
plt.legend()
plt.show()

```

Lampiran 3. Hasil Plagiasi

 **MAJELIS PENDIDIKAN TINGGI PIMPINAN PUSAT MUHAMMADIYAH
UNIVERSITAS MUHAMMADIYAH MAKASSAR
UPT PERPUSTAKAAN DAN PENERBITAN**
Alamat kantor: Jl.Sultan Alauddin NO.259 Makassar 90221 Tlp.(0411) 866972,881593, Fax.(0411) 865588

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

SURAT KETERANGAN BEBAS PLAGIAT

**UPT Perpustakaan dan Penerbitan Universitas Muhammadiyah Makassar,
Menerangkan bahwa mahasiswa yang tersebut namanya di bawah ini;**

Nama : Muh. Yusuf Syahbul Basyar
Nim : 105841112020
Program Studi : Teknik Informatika
Dengan nilai:

No	Bab	Nilai	Ambang Batas
1	Bab 1	10 %	10 %
2	Bab 2	25 %	25 %
3	Bab 3	10 %	10 %
4	Bab 4	10 %	10 %
5	Bab 5	5 %	10 %

Dinyatakan telah lulus cek plagiat yang diadakan oleh UPT- Perpustakaan dan Penerbitan Universitas Muhammadiyah Makassar Menggunakan Aplikasi Turnitin.

Demikian surat keterangan ini diberikan kepada yang bersangkutan untuk dipergunakan seperlunya.

Makassar, 12 Februari 2025
Mengetahui
Kepala UPT- Perpustakaan dan Penerbitan,


Nuzuliah S. Huda, M.I.P.
NBM. 964 591

Jl. Sultan Alauddin no 259 makassar 90222
Telepon (0411)866972,881 593,fax (0411)865 588
Website: www.library.unismuh.ac.id
E-mail : perpustakaan@unismuh.ac.id

 Dipindai dengan CamScanner

MUH. YUSUF SYAHBUL BASYAR 105841112020 Bab I

ORIGINALITY REPORT

10%

SIMILARITY INDEX

9%

INTERNET SOURCES

2%

PUBLICATIONS

7%

STUDENT PAPERS

PRIMARY SOURCES

1

digilib.uin-suka.ac.id
Internet Source

2%

2

vdocuments.mx
Internet Source

2%

3

Submitted to Fakultas Ekonomi Universitas
Indonesia
Student Paper

2%

4

www.coursehero.com
Internet Source

2%

5

eprints.uny.ac.id
Internet Source

2%

Exclude quotes On

Exclude bibliography On

Exclude matches On



MUH. YUSUF SYAHBUL BASYAR 105841112020 Bab II

ORIGINALITY REPORT

25% SIMILARITY INDEX	23% INTERNET SOURCES	8% PUBLICATIONS	13% STUDENT PAPERS
--------------------------------	--------------------------------	---------------------------	------------------------------

PRIMARY SOURCES

1	tunasbangsa.ac.id Internet Source	3%
2	digilib.unila.ac.id Internet Source	3%
3	Submitted to Padjadjaran University Student Paper	2%
4	etheses.uin-malang.ac.id Internet Source	2%
5	ejurnal.umri.ac.id Internet Source	1%
6	ZUDI ITA BELA ZUDI ITA BELA. "SISTEM PREDIKSI PENJUALAN OBAT MENGGUNAKAN METODE SINGLE MOVING AVERAGE (STUDI KASUS: APOTEK WILUJENG KECAMATAN PANCENG KAB.GRESIK)", Indexia, 2022 Publication	1%
7	Submitted to fpptijateng Student Paper	1%
8	www.sttrcepu.ac.id Internet Source	1%
9	publikasi.mercubuana.ac.id Internet Source	1%
10	pengantarbisnisreza.blogspot.com Internet Source	1%
11	pdfcoffee.com Internet Source	1%

MUH. YUSUF SYAHBUL BASYAR 105841112020 Bab III

ORIGINALITY REPORT

10%

SIMILARITY INDEX

9%

INTERNET SOURCES

2%

PUBLICATIONS

5%

STUDENT PAPERS

PRIMARY SOURCES

1

repository.unsri.ac.id

Internet Source

2%

2

Mochamad Alfian Rosid, Gunawan Gunawan,
Edwin Pramana. "Centroid Based Classifier
With TF – IDF – ICF for Classification of
Student's Complaint at Appliation E-Complaint
in Muhammadiyah University of Sidoarjo"
Journal of Electrical and Electronic
Engineering-UMSIDA, 2016
Publication

2%

3

eprints.ummetro.ac.id

Internet Source

2%

4

repository.ipb.ac.id

Internet Source

2%

5

www.mdpi.com

Internet Source

2%

Exclude quotes On

Exclude bibliography On

Exclude matches < 2%

MUH. YUSUF SYAHBUL BASYAR 105841112020 Bab IV

ORIGINALITY REPORT

10%	9%	9%	10%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Universitas Sultan Ageng Tirtayasa Student Paper	1%
2	Submitted to Queen's University of Belfast Student Paper	1%
3	Submitted to University of Bucharest Student Paper	1%
4	gifadn.medium.com Internet Source	1%
5	stackoverflow.com Internet Source	1%
6	wenku.csdn.net Internet Source	1%
7	jurnal.polibatam.ac.id Internet Source	1%
8	Michael Yuichi, Yeremia Alfa Susetyo. "Klasifikasi Penyakit Migrain dengan Metode Naïve Bayes pada Dataset Kaggle", Jurnal Indonesia : Manajemen Informatika dan Komunikasi, 2025 Publication	1%
9	jsw6701.github.io Internet Source	1%
10	dokumen.pub Internet Source	

MUH. YUSUF SYAHBUL BASYAR 105841112020 Bab V

ORIGINALITY REPORT

5%

SIMILARITY INDEX

5%

INTERNET SOURCES

0%

PUBLICATIONS

0%

STUDENT PAPERS

PRIMARY SOURCES

1

eprints.umm.ac.id

Internet Source

2%

2

e-journal.unkhair.ac.id

Internet Source

2%

Exclude quotes

On

Exclude bibliography

On

Exclude matches

< 2%