

**ANALISIS PREDIKSI HARGA SAHAM HAPAG-LLOYD AG  
MENGUNAKAN METODE LONG SHORT-TERM MEMORY  
(LSTM)**

**SKRIPSI**

**Diajukan untuk Memenuhi Salah Satu Syarat Kelulusan Strata 1  
di Program Studi Informatika Universitas Widyatama**

**Oleh**

**NAMA : FRANSISCUS KRISTIAN SUSANTO**

**NPM : 40621100012**



**PROGRAM STUDI INFORMATIKA  
FAKULTAS TEKNIK UNIVERSITAS WIDYATAMA  
BANDUNG  
2025**

# **LEMBAR PENGESAHAN**

## **ANALISIS PREDIKSI HARGA SAHAM HAPAG-LLOYD AG MENGUNAKAN METODE LONG SHORT-TERM MEMORY (LSTM)**

### **SKRIPSI**

Program Studi Informatika

Fakultas Teknik

Universitas Widyatama

**Oleh:**

**Fransiscus Kristian Susanto**

**40621100012**

Telah disetujui dan disahkan di Bandung, \_\_\_\_\_ 2025

Menyetujui,

**Pembimbing,**

**Dr. Feri Sulianta, S.T., M.T.**

**NIDN. 0420027901**

Mengetahui,

**Ketua Program Studi Informatika,**

**Dekan Fakultas Teknik,**

**Ari Purno Wahyu Wibowo, S.Kom.,**

**M.Kom.**

**NIDN. 0415078402**

**Dr. Arief Rahmana, ST., MT.,**

**CIPMP., IPU., ASEAN Eng.**

**NIDN. 0429097401**

## SURAT PERNYATAAN

Saya yang bertanda tangan di bawah ini:

Nama : Fransiscus Kristian Susanto  
NPM : 40621100012  
Tempat, Tanggal Lahir : Bandung, 21 Juni 2003  
Alamat Asal : Bandung  
Alamat Bandung : Komplek Permata Biru Blok H No. 93, Cinunuk,  
Kec. Cileunyi

Dengan ini menyatakan bahwa skripsi dengan judul: **ANALISIS PREDIKSI HARGA SAHAM HAPAG-LLOYD AG MENGGUNAKAN METODE LONG SHORT-TERM MEMORY (LSTM)** adalah hasil pekerjaan saya dan seluruh ide, pendapat, atau materi dari sumber lain telah dikutip dengan cara penulisan referensi yang sesuai.

Pernyataan ini saya buat dengan sebenar-benarnya dan jika pernyataan ini tidak sesuai dengan kenyataan, maka saya bersedia menanggung sanksi yang akan dikenakan kepada saya termasuk pencabutan gelar Sarjana yang telah saya dapatkan.

Bandung, \_\_\_\_\_ 2025

Fransiscus Kristian Susanto

## ABSTRAK

Analisis prediksi harga saham merupakan hal yang penting dalam mempertimbangan perdagangan modal terhadap suatu perusahaan. Dengan menggunakan metode Long Short-Term Memory (LSTM), prediksi tersebut memberikan gambaran terkait mengenali pola terhadap data time series dan memperkirakan pergerakan harga saham dalam data historis. Analisis prediksi yang dikembangkan memanfaatkan algoritma LSTM sebagai sebuah jenis jaringan saraf berulang yang dirancang untuk mengatasi masalah dalam urutan data waktu panjang. Penelitian ini melibatkan pengumpulan data historis harga saham Hapag-Lloyd AG dan menjalankan data preprocessing terhadap dataset. Model LSTM kemudian dikembangkan dan dilatih menggunakan dataset tersebut hingga akhirnya memberikan evaluasi terhadap model tersebut. Hasil analisis diharapkan dapat memberikan wawasan tentang efektivitas metode LSTM dalam memprediksi harga saham dan menjadi salah satu bahan pertimbangan dalam menentukan keputusan pemegangan modal di suatu perusahaan.

**Kata Kunci:** Saham, Prediksi, LSTM, Harga.

## ***ABSTRACT***

*Stock price prediction analysis is important in considering capital trading for a company. By using the Long Short-Term Memory (LSTM) method, the prediction provides an overview of recognizing patterns in time series data and estimating stock price movements in historical data. The prediction analysis developed utilizes the LSTM algorithm as a type of recurrent neural network designed to solve problems in long-term data sequences. This study involves collecting historical Hapag-Lloyd AG stock price data and running data preprocessing on the dataset. The LSTM model is then developed and trained using the dataset until finally providing an evaluation of the model. The results of the analysis are expected to provide insight into the effectiveness of the LSTM method in predicting stock prices and become one of the considerations in determining capital holding decisions in a company.*

**Keywords:** *Stocks, Prediction, LSTM, Price.*

## KATA PENGANTAR

Puji dan syukur penulis panjatkan kepada Tuhan Yang Maha Esa karena dengan rahmat dan karunia-Nya, penulis dapat menyelesaikan tugas akhir yang berjudul **“Analisis Prediksi Harga Saham Hapag-Lloyd AG Menggunakan Metode Long Short-Term Memory (LSTM)”** tepat pada waktunya. Tugas akhir ini disusun dalam rangka memenuhi salah satu syarat untuk menyelesaikan studi Strata 1 pada Program Studi Informatika Universitas Widyatama.

Dalam menyelesaikan tugas akhir ini penulis mendapatkan banyak bantuan, masukan, bimbingan, dan dukungan dari berbagai pihak. Oleh karena itu, penulis ingin mengucapkan terima kasih kepada:

1. Orang tua dan kakak tercinta yang memberikan perhatian dan kasih sayang kepada penulis.
2. Dr. Feri Sulianta, S.T., M.T. selaku dosen pembimbing yang telah memberi masukan dan waktunya dalam membimbing penulis.
3. Seluruh dosen pengajar di Program Studi Informatika Universitas Widyatama yang telah membekali penulis dengan berbagai ilmu pengetahuan.
4. Teman-teman Program Studi Informatika Angkatan 2021 yang telah memberi dukungan dalam pengerjaan tugas akhir.
5. Semua teman dan pihak yang telah memberikan bantuan kepada penulis baik secara langsung maupun tidak langsung.

Penulis menyadari bahwa penyusunan tugas akhir ini masih banyak kekurangan sehingga kritik dan saran yang membangun sangat membantu penulis guna perbaikan tugas akhir ini. Penulis berharap tugas akhir ini dapat bermanfaat bagi semua pihak.

Bandung, \_\_\_\_\_ 2025

Fransiscus Kristian Susanto

## DAFTAR ISI

<b>LEMBAR PENGESAHAN .....</b>	<b>i</b>
<b>SURAT PERNYATAAN .....</b>	<b>ii</b>
<b>ABSTRAK .....</b>	<b>iii</b>
<b><i>ABSTRACT</i> .....</b>	<b>iv</b>
<b>KATA PENGANTAR.....</b>	<b>v</b>
<b>DAFTAR ISI.....</b>	<b>vi</b>
<b>DAFTAR GAMBAR .....</b>	<b>ix</b>
<b>DAFTAR TABEL .....</b>	<b>x</b>
<b>DAFTAR LAMPIRAN .....</b>	<b>xi</b>
<b>BAB I PENDAHULUAN.....</b>	<b>1</b>
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	3
1.3 Tujuan Penelitian .....	3
1.4 Batasan Penelitian .....	3
1.5 Manfaat Penelitian .....	4
<b>BAB II KAJIAN PUSTAKA .....</b>	<b>5</b>
2.1 Tinjauan Pustaka.....	5
2.2 Landasan Teori.....	11
2.2.1 Saham.....	11
2.2.2 Hapag-Lloyd AG .....	14
2.2.3 Pengembangan Sistem <i>Waterfall</i> .....	14
2.2.4 Persyaratan Sistem Perangkat Lunak.....	16
2.2.5 <i>Data Mining</i> .....	18
2.2.6 <i>Long Short-Term Memory</i> .....	20
2.2.7 <i>Cross-Industry Standard Process for Data Mining</i> .....	27
2.2.8 <i>Time Series Forecasting</i> .....	30
2.2.9 <i>Data Flow Diagram</i> .....	33
2.2.10 <i>NoSQL Database</i> .....	34
2.2.11 <i>MongoDB</i> .....	35
2.2.12 <i>Streamlit</i> .....	38
2.2.13 <i>System Usability Scale</i> .....	39

<b>BAB III METODE PENELITIAN .....</b>	<b>42</b>
3.1 Pengembangan Sistem <i>Waterfall</i> .....	42
3.1.1 <i>Planning</i> .....	42
3.1.2 <i>Analysis</i> .....	43
3.1.3 <i>Design</i> .....	44
3.1.4 <i>Implementation</i> .....	45
3.1.5 <i>Use</i> .....	46
3.2 <i>Cross-Industry Standard Process for Data Mining</i> .....	47
3.2.1 <i>Business Understanding</i> .....	47
3.2.2 <i>Data Understanding</i> .....	47
3.2.3 <i>Data Preparation</i> .....	48
3.3.4 <i>Modeling</i> .....	48
3.3.5 <i>Evaluation</i> .....	49
3.3.6 <i>Deployment</i> .....	50
3.3 Kerangka Penelitian .....	51
<b>BAB IV HASIL PENELITIAN DAN PEMBAHASAN .....</b>	<b>52</b>
4.1 Analisis Kebutuhan .....	52
4.1.1 Kebutuhan Fungsional.....	52
4.1.2 Kebutuhan Sistem .....	52
4.1.3 Kebutuhan Non-Fungsional.....	52
4.2 Perancangan Sistem.....	52
4.2.1 Perancangan Perangkat Lunak .....	52
4.2.2 Perancangan Basis Data .....	52
4.2.3 Perancangan Antarmuka .....	52
4.3 Implementasi Sistem .....	59
4.3.1 Batasan Implementasi.....	59
4.3.2 Implementasi Perangkat Keras .....	60
4.3.3 Implementasi Perangkat Lunak .....	60
4.3.4 Implementasi Antarmuka.....	60
4.4 Pengujian Sistem .....	60
4.4.1 Pengujian Algoritma .....	60
4.4.2 <i>Black Box Testing</i> .....	61
4.4.3 <i>System Usability Scale</i> .....	61



<b>BAB V PENUTUP .....</b>	<b>62</b>
5.1    Kesimpulan .....	62
5.2    Saran .....	62
<b>DAFTAR PUSTAKA.....</b>	<b>63</b>
<b>LAMPIRAN.....</b>	<b>68</b>

## **DAFTAR GAMBAR**

## **DAFTAR TABEL**

## **DAFTAR LAMPIRAN**

<b>Lampiran 1. Riwayat Hidup .....</b>	<b>68</b>
<b>Lampiran 2. Kartu Bimbingan Tugas Akhir .....</b>	<b>69</b>

# **BAB I**

## **PENDAHULUAN**

### **1.1 Latar Belakang**

Investasi merupakan salah satu cara penting bagi individu maupun perusahaan untuk meningkatkan nilai asetnya. Dalam dunia investasi terdapat berbagai jenis instrumen yang dapat dipilih oleh investor di mana salah satunya adalah saham. Investasi saham memberikan kesempatan bagi investor untuk memperoleh keuntungan yang menarik. Oleh karena itu, saham menjadi pilihan yang menarik untuk investasi jangka panjang meskipun terdapat risiko yang perlu dipertimbangkan.

Investasi saham umumnya dilakukan pada perusahaan terbuka di mana terdapat banyak perusahaan yang sahamnya diperdagangkan di pasar modal dan dapat dibeli oleh masyarakat umum. Salah satu sektor yang menarik perhatian investor adalah sektor logistik yang memiliki peran penting dalam rantai pasokan barang dan distribusi global. Di sektor logistik terdapat perusahaan-perusahaan transportasi peti kemas yang berperan penting dalam mendukung kegiatan perdagangan internasional. Salah satu perusahaan yang memiliki pengaruh besar dalam sektor ini adalah Hapag-Lloyd AG, sebuah perusahaan transportasi peti kemas asal Jerman yang terkenal secara global.

Namun, investasi saham tidaklah tanpa tantangan. Salah satu tantangan terbesar adalah memprediksi harga saham yang sering kali dipengaruhi oleh berbagai faktor ekonomi, politik, dan sosial baik dari satu negara tersebut maupun antar negara. Maka karena itu analisis prediksi harga saham menjadi sangat penting bagi para investor dalam membuat keputusan yang lebih baik dan lebih menguntungkan.

Dalam hal ini, teknik analisis prediksi harga saham dapat menggunakan metode *data mining*. *Data mining* atau penambangan data merupakan proses pencarian pola atau informasi berharga dari data yang besar dan dapat digunakan untuk menganalisis berbagai aspek yang mempengaruhi harga saham. Salah satu metodologi yang sering digunakan dalam proyek penambangan data adalah CRISP-DM (*Cross-Industry Standard Process for Data Mining*). CRISP-DM memberikan

panduan dalam melakukan analisis data dengan langkah-langkah yang sistematis yang mulai dari pemahaman masalah hingga penerapan model.

Penelitian terkait analisis prediksi harga saham dengan menggunakan metodologi CRISP-DM telah dilakukan. Penelitian Saepulrohman dkk. (2025) mengembangkan analisis prediksi harga saham PT Aneka Tambang Tbk. dengan mengimplementasikan CRISP-DM. Penelitian ini menarik kesimpulan bahwa dengan mengikuti tahapan yang terstruktur berdasarkan metodologi CRISP-DM maka model yang dibangun mampu menunjukkan performa yang stabil. Hasil tersebut memberikan kontribusi penting bagi para pelaku pasar dan pembaca sebagai dasar untuk mengembangkan penelitian lebih lanjut dalam mengembangkan model prediksi harga saham yang lebih akurat dan andal (Saepulrohman dkk., 2025).

Salah satu metode yang populer dalam mengantisipasi hal di atas untuk analisis prediksi harga saham adalah model *Long Short-Term Memory* (LSTM), yang merupakan jenis jaringan saraf tiruan berbasis *Recurrent Neural Networks* (RNN). LSTM sangat efektif dalam memproses data deret waktu (*time series*), seperti data harga saham yang memiliki ketergantungan waktu. Model ini memiliki kemampuan untuk mengingat pola jangka panjang dalam data yang sangat berguna dalam memprediksi pergerakan harga saham.

Menurut penelitian yang dilakukan oleh Gourav Bathla (2020), model LSTM dibandingkan dengan metode *Support Vector Regression* (SVR) dalam menganalisis berbagai data indeks saham, termasuk S&P 500, NYSE, NSE, BSE, NASDAQ, dan Dow Jones Industrial Average. Hasil dari analisis tersebut menunjukkan bahwa LSTM mampu menghasilkan tingkat akurasi yang lebih tinggi dibandingkan dengan SVR (Bathla, 2020).

Penelitian ini bertujuan untuk menganalisis prediksi harga saham Hapag-Lloyd AG menggunakan metode LSTM dengan menerapkan metodologi CRISP-DM dalam proses analisis data. Melalui penelitian ini, diharapkan dapat memberikan wawasan lebih dalam tentang cara kerja LSTM dalam memprediksi harga saham serta memberikan informasi yang berguna bagi investor dalam membuat keputusan investasi.

## 1.2 Rumusan Masalah

Adapun rumusan masalah dalam tugas akhir ini adalah sebagai berikut.

1. Bagaimana cara memprediksi harga saham Hapag-Lloyd AG menggunakan metode LSTM?
2. Bagaimana implementasi metodologi CRISP-DM dalam kasus analisis prediksi harga saham menggunakan LSTM?
3. Bagaimana cara menerapkan dan mengevaluasi penggunaan aplikasi analisis prediksi harga saham Hapag-Lloyd AG?

## 1.3 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah dijabarkan, adapun tujuan penelitian adalah sebagai berikut.

1. Mengembangkan model prediksi harga saham Hapag-Lloyd AG menggunakan metode LSTM dengan menggunakan data historis yang tersedia.
2. Menyusun pelaksanaan metodologi CRISP-DM dalam proses analisis prediksi harga saham, yang melibatkan tahapan-tahapan seperti pemahaman masalah, pembacaan awal data, pemrosesan data, pemodelan, evaluasi, dan penerapan aplikasi.
3. Menganalisis dan mengevaluasi penggunaan aplikasi terkait analisis prediksi harga saham.

## 1.4 Batasan Penelitian

Adapun batasan masalah dalam tugas akhir ini adalah sebagai berikut.

1. Data yang digunakan dalam penelitian ini hanya mencakup harga saham Hapag-Lloyd AG selama periode tertentu yang tersedia di Yahoo Finance sebagai sumber data. Data tersebut mencakup harga penutupan harian yang digunakan untuk model prediksi harga saham.
2. Penelitian ini hanya menggunakan metode LSTM untuk menganalisis prediksi harga saham Hapag-Lloyd AG tanpa membandingkannya dengan metode lain.

3. Penelitian ini terbatas pada implementasi metodologi CRISP-DM dalam analisis prediksi harga saham dengan menggunakan variabel harga penutupan tanpa mencakup analisis fundamental lainnya yang berkaitan dengan saham.

### **1.5 Manfaat Penelitian**

Adapun manfaat dalam penelitian ini adalah sebagai berikut.

1. Penelitian ini dapat memberikan wawasan yang lebih baik dalam memprediksi harga saham Hapag-Lloyd AG, yang dapat digunakan untuk pengambilan keputusan investasi yang lebih informatif.
2. Hasil penelitian ini dapat menjadi referensi untuk penelitian lebih lanjut dalam penggunaan metode LSTM untuk analisis prediksi harga saham, serta penerapan metodologi CRISP-DM dalam proyek penambangan data lainnya.
3. Penelitian ini dapat memberikan contoh implementasi praktis dalam pengolahan dan pemodelan data urutan waktu menggunakan metode LSTM untuk prediksi harga saham.
4. Penelitian ini dapat memberikan pemahaman lebih luas mengenai pentingnya investasi saham, khususnya dalam sektor logistik, serta memberikan gambaran tentang cara kerja teknik prediksi harga saham menggunakan pendekatan *data mining*.



## BAB II

### KAJIAN PUSTAKA

#### 2.1 Tinjauan Pustaka

Bagian ini membahas sejumlah penelitian terdahulu yang memiliki keterkaitan dengan topik skripsi ini. Studi-studi sebelumnya memberikan cara pandang mengenai konsep dasar, penerapan, serta temuan yang diperoleh dari penelitian sejenis. Pemanfaatan penelitian terdahulu sebagai landasan dalam tinjauan pustaka digunakan untuk mendukung analisis prediksi harga saham dengan menggunakan metode LSTM.

Salah satu penelitian yang mendukung kajian ini dilakukan oleh Nadia Afrin Ritu dkk dengan menggunakan data dari Bursa Saham Dhaka yang mencakup empat bank, seperti Bank Asia, Brac Bank, Dhaka Bank, dan Islami Bank. Data yang dianalisis berasal dari periode 15 Juli 2018 hingga 25 Maret 2020 dengan fokus pada variabel harga penutupan. Data dibagi dengan proporsi 80% untuk pelatihan dan 20% untuk pengujian model. Sebelum diterapkan dalam model LSTM, seluruh data terlebih dahulu melalui proses standarisasi. Evaluasi performa model dilakukan menggunakan metrik *Root Mean Square Error* (RMSE) dengan nilai untuk masing-masing adalah 4,5124 (Brac Bank), 1,1898 (Bank Asia), 0,84084 (Islami Bank), dan 0,33193 (Dhaka Bank) yang menunjukkan variasi performa antar saham yang diuji (Ritu dkk., 2025).

Dalam penelitian yang dilakukan oleh Firhan Abdillah Mahbubi dkk, model LSTM dimanfaatkan untuk meramalkan harga saham nikel dari PT Resource Alam Indonesia Tbk. Data historis yang digunakan mencakup periode Januari 2021 hingga Mei 2024 yang kemudian dibagi menjadi dua bagian, seperti 80% untuk pelatihan dan 20% untuk pengujian. Berbagai hyperparameter disesuaikan untuk memaksimalkan akurasi model selama proses pelatihan. Hasil terbaik diperoleh saat menggunakan optimasi Adam dengan jumlah 150 epoch. Evaluasi performa model menunjukkan metrik prediksi yang cukup rendah dengan nilai *Mean Absolute Error* (MAE) sebesar 33,15, RMSE sebesar 48,14, *Mean Squared Error* (MSE) sebesar 2317,33, dan *Mean Absolute Percentage Error* (MAPE) sebesar 7,39 (Mahbubi dkk., 2025).

Penelitian yang dilakukan oleh Rendy Saputra dkk menggunakan data historis saham Tesla Inc. yang mencakup periode dari 21 Mei 2020 hingga 21 Mei 2024. Data yang dianalisis meliputi harga pembukaan, harga tertinggi, harga terendah, dan harga penutupan harian. Seluruh variabel masukan terlebih dahulu dinormalisasi dengan menggunakan teknik StandardScaler. Data kemudian dibagi menjadi 80% untuk pelatihan dan 20% untuk pengujian. Model prediksi yang dikembangkan mengandalkan dua lapisan LSTM yang masing-masing berisi 50 neuron dan diakhiri dengan satu lapisan dense. Proses pelatihan berlangsung selama 100 epoch dengan batch size sebesar 4 dan tanpa menerapkan teknik *shuffling*. Hasil evaluasi menunjukkan performa model yang cukup baik dengan nilai RMSE sebesar 0,042 dan MAPE sebesar 18,7% (Saputra dkk., 2024).

Dalam penelitian yang dilakukan oleh Muhammad Sahroni dkk, data saham PT Bank Syariah Indonesia yang diperoleh dari situs Yahoo Finance dengan fokus data pada harga penutupan harian yang telah disesuaikan (*adjusted close*) selama periode 31 Mei 2019 hingga 31 Mei 2024. Data tersebut dibagi menjadi dua bagian, yakni 80% untuk pelatihan dan 20% untuk pengujian. Sebelum pemodelan, data terlebih dahulu dinormalisasi menggunakan metode MinMaxScaler. Model yang digunakan adalah LSTM dengan parameter optimasi Adam. Struktur jaringan terdiri atas dua lapisan tersembunyi LSTM dengan fungsi aktivasi tanh dan sigmoid serta ditambahkan mekanisme dropout. Pengujian dilakukan dengan berbagai kombinasi parameter seperti jumlah neuron sebesar 50, ukuran batch masing-masing sebesar 4, 16, 64, 128, dan jumlah epoch masing-masing sebesar 50, 100, 150, 200. Pemilihan parameter terbaik dilakukan dengan pendekatan Grid Search dan cross-validation. Evaluasi model menggunakan metrik RMSE dan MAPE menunjukkan bahwa kombinasi batch size 4 dan epoch 200 memberikan hasil terbaik dengan nilai MAPE 1,72% dan RMSE 53,49. Secara keseluruhan model mampu memprediksi harga saham dengan tingkat akurasi sebesar 98,28% (Sahroni dkk., 2024).

Dalam penelitian yang dilakukan oleh Bayu Dwi Handika dan Sugianto, metode LSTM diterapkan untuk melakukan prediksi harga saham dari dua perusahaan perbankan, yaitu PT Bank BTPN Tbk (BTPN) dan PT Bank Rakyat Indonesia (Persero) Tbk (BBRI). Proses pelatihan model dilakukan dengan

membagi data historis menjadi 80% untuk pelatihan dan 20% untuk pengujian. Data yang digunakan terlebih dahulu dinormalisasi menggunakan teknik Min-Max Scaler dan jumlah epoch yang digunakan adalah 66 untuk model saham BTPN serta 55 untuk model saham BBRI. Hasil dari pemodelan ini menunjukkan tingkat akurasi yang baik dengan nilai Mean Absolute Error (MAE) masing-masing sebesar 0,01444 untuk saham BTPN dan 0,01406 untuk saham BBRI (Handika & Sugianto, 2024).

Alrafiful Rahman dkk dalam penelitiannya menganalisis data historis pergerakan harga saham harian BNI selama periode 16 April 2001 hingga 6 Januari 2023. Fokus utama penelitian ini adalah pada harga penutupan saham di mana data dibagi menjadi dua bagian, seperti 80% digunakan untuk pelatihan model dan 20% untuk pengujian. Model LSTM diterapkan dengan optimasi Adam dalam beberapa konfigurasi epoch, yaitu pada 25, 50, 75, dan 100. Hasil terbaik diperoleh saat model dijalankan dengan 100 epoch dan ukuran batch 64 yang menghasilkan nilai koefisien determinasi ( $R^2$ ) sebesar 0,9928 dan MAPE sebesar 1,53% (Rahman dkk., 2024).

Dalam penelitian yang dilakukan oleh Edwin Setiawan Nugraha dkk dengan memanfaatkan data harga penutupan saham PT Astra International Tbk yang diambil dari situs Yahoo Finance dalam periode antara 2 Januari 2015 hingga 30 Desember 2020. Proses awal pengolahan data mencakup pembersihan data, pembagian dataset menjadi 80% untuk pelatihan dan 20% untuk pengujian, serta normalisasi menggunakan metode Min-Max Scaling. Proses pemodelan dilakukan dengan bantuan algoritma optimasi Adam dan menggunakan fungsi *mean squared error*. Sebanyak dua belas model dievaluasi dengan berbagai kombinasi hyperparameter termasuk variasi jumlah epoch (50, 100, 200, dan 400) dan ukuran batch (4, 16, dan 64). Dari seluruh model yang diuji, model dengan konfigurasi batch size 4 dan 50 epoch terbukti paling optimal. Performa model tersebut diukur menggunakan metrik RMSE sebesar 151.910, MSE sebesar 23.076.561, MAE sebesar 118.128, dan MAPE sebesar 2,3% (Nugraha dkk., 2024).

Tatas Handharu Sworo dan Arief Hermawan melakukan penelitian yang berfokus pada prediksi harga saham BRI dengan memanfaatkan data historis harga penutupan dari Januari 2006 hingga Desember 2023. Model LSTM digunakan

untuk mempelajari pola pergerakan harga saham dari data historis tersebut. Setelah proses pelatihan selesai, performa model dievaluasi menggunakan sejumlah metrik evaluasi dilakukan dan hasil pengujian model memiliki kinerja yang baik yang ditunjukkan dengan nilai MSE sebesar 0,000279, MAE sebesar 0,0133, dan RMSE sebesar 0,0167 pada data pelatihan (Handharu Sworo & Hermawan, 2024).

Ni Putu Noviyanti Kusuma dalam penelitiannya memanfaatkan data harga penutupan saham yang telah disesuaikan dari 17 emiten *blue chip* yang tergabung dalam indeks IDX30 pada periode 2019 hingga 2022. Penerapan teknik normalisasi min-max scaling digunakan untuk menyesuaikan skala data. Setelah melalui proses preprocessing, data dibagi menjadi dua bagian, yakni 80% untuk pelatihan dan 20% untuk pengujian. Model prediksi kemudian dikembangkan menggunakan algoritma LSTM. Akurasi hasil prediksi dievaluasi menggunakan metode MAPE dengan hasil rata-rata sebesar 2,6% yang menandakan bahwa model ini memiliki performa yang sangat baik dalam memprediksi harga saham (Kusuma, 2024).

Arnes Anandita dan Tri Wahyuningsih dalam penelitiannya memanfaatkan data harga penutupan dari indeks saham syariah JII70 yang dikumpulkan selama periode 1 Januari 2020 hingga 30 Januari 2024. Data ini kemudian dipisahkan menjadi dua bagian, yaitu 90% digunakan untuk pelatihan dan 10% untuk pengujian. Sebelum pelatihan dimulai, data distandarisasi untuk meningkatkan stabilitas dan akurasi model. Model prediksi yang dibangun menggunakan arsitektur LSTM dengan konfigurasi satu masukan dan satu keluaran serta dilengkapi dengan 50 *hidden layer units*. Proses pelatihan berlangsung selama 500 epoch dengan menggunakan optimisasi Adam serta menerapkan batas gradien sebesar 1 dan learning rate sebesar 0,005. Akurasi model diuji menggunakan metrik RMSE dan hasilnya menunjukkan nilai RMSE sebesar 6,8993 yang menandakan bahwa model mampu memprediksi pergerakan indeks saham syariah secara cukup akurat (Anandita & Wahyuningsih, 2024).

Abdul Rosyd dkk menggunakan data historis saham BCA yang diambil dari situs Yahoo Finance dengan mencakup periode 1 Januari 2020 hingga 30 Oktober 2023. Proses awal dimulai dengan tahap preprocessing data dengan fokus analisis deret waktu ditujukan pada feature *Low*. Tahapan berikutnya adalah pembagian 80% data sebagai data latih dan normalisasi data menggunakan metode Min-Max

Scaling. Setelah transformasi selesai, maka dilakukan tahap pemodelan data dengan membangun arsitektur model LSTM Sequential yang terdiri dari dua lapisan LSTM masing-masing berisi 50 neuron serta dua lapisan dense dengan 25 dan 1 neuron. Proses uji model menggunakan optimasi Adam dan *mean squared error* sebagai fungsi loss model. Selama pelatihan, model disesuaikan dengan iterasi terhadap data latih, pengaturan jumlah epoch dan batch size, serta penyesuaian bobot berdasarkan hasil prediksi. Model yang telah dilatih kemudian digunakan untuk melakukan prediksi terhadap data baru dan performanya dievaluasi menggunakan metrik RMSE, MAPE, dan MSE. Hasil evaluasi menunjukkan bahwa model ini bekerja dengan baik dengan nilai RMSE sebesar 40,85, MAPE sebesar 0,71%, dan MSE sebesar 6.662,76 (Rosyd dkk., 2024).

Penelitian yang dilakukan oleh Zhenglin Li dkk menggunakan data historis harga penutupan saham dari Apple, Google, Microsoft, dan Amazon yang diambil dari Yahoo Finance sejak tanggal 1 Januari 2012. Data ini kemudian dibagi menjadi dua bagian, seperti 95% digunakan untuk melatih model dan 5% sisanya untuk menguji performanya. Model yang digunakan berbasis LSTM dengan arsitektur dua lapisan LSTM yang masing-masing memiliki 128 dan 64 unit serta dilanjutkan dengan dua lapisan dense berukuran 25 dan 1 unit. Model ini dibangun dengan menggunakan algoritma optimisasi Adam dan fungsi loss MSE. Proses pelatihan dijalankan selama satu epoch dengan ukuran batch sebesar satu. Untuk mengevaluasi hasil prediksi model pada data pengujian menggunakan metrik RMSE yang menghasilkan nilai 18,89 yang mencerminkan tingkat akurasi data uji model dalam memprediksi harga saham (Z. Li dkk., 2023).

Penelitian yang dilakukan oleh Eka Patriya dkk menggunakan data harga penutupan IHSG dari 10 Juli 2008 hingga 1 Maret 2023. Sebelum dimasukkan ke dalam model LSTM, data menjalani proses preprocessing seperti normalisasi, penentuan jumlah sampel berdasarkan time step sebesar 50, serta pembagian data menjadi 70% untuk pelatihan dan 30% untuk pengujian. Model LSTM dibangun dengan pendekatan sequential yang terdiri atas beberapa lapisan LSTM yang dilanjutkan dengan dense layer dan output layer untuk menghasilkan satu nilai prediksi. Evaluasi model dilakukan dengan menggunakan metrik MSE, sedangkan proses pelatihannya mengandalkan algoritma optimisasi Adam. Model diuji dengan

berbagai konfigurasi hyperparameter, seperti jumlah epoch (30, 35, 40, 45, dan 50) dan ukuran batch sebesar 32. Selain itu, performa model juga dianalisis melalui beberapa rasio pembagian data, yaitu 7:3, 8:2, dan 9:1 dengan pengukuran kinerja menggunakan RMSE. Hasil terbaik diperoleh saat model dilatih dengan 50 epoch dan rasio data 7:3 yang menghasilkan nilai RMSE sebesar 63,672 (Patriya dkk., 2023).

Dalam penelitian yang dilakukan oleh Juan Syidi Prasetyo dkk, data yang digunakan berasal dari harga penutupan harian saham ANTM dan MDKA diambil dari situs Yahoo Finance dengan periode waktu sejak 1 Januari 2017. Model prediksi LSTM terdiri dari beberapa lapisan yang diawali dengan lapisan pertama berisi 50 neuron dan diikuti oleh lapisan kedua serta dilanjutkan dengan lapisan dense ketiga yang memiliki 25 neuron. Lapisan keluaran model terdiri dari satu neuron untuk menghasilkan prediksi akhir. Model ini menggunakan parameter optimasi Adam. Dari hasil evaluasi kinerja model diperoleh nilai MAPE sebesar 3,08% untuk saham MDKA dan 2,17% untuk saham ANTM yang menunjukkan tingkat kesalahan prediksi yang relatif rendah (Syidi dkk., 2023).

I Ketut Agung Enriko dkk dalam penelitiannya menggunakan data saham TLKM dengan periode waktu dari 1 Januari 2019 hingga 11 Januari 2023. Data tersebut diperoleh melalui platform Yahoo Finance. Fokus utama dari penelitian ini adalah harga saham pembukaan. Dataset dibagi dengan proporsi 75% untuk pelatihan dan 25% untuk pengujian. Penerapan algoritma LSTM dilakukan dalam beberapa tahap yang mencakup penggunaan optimasi Adam, SGD, dan RMSprop dengan variasi jumlah epoch yang berbeda, yakni 25, 50, 75, dan 100. Pengukuran akurasi menggunakan metrik MAPE dan akurasi tertinggi tercatat pada optimasi Adam dengan 100 epoch dengan menghasilkan akurasi sebesar 98,59%. Sebaliknya, akurasi terendah tercatat pada optimasi RMSprop dengan 25 epoch yang hanya mencapai 97,61%. Pada optimasi SGD, akurasi tertinggi tercatat pada 50 epoch, sementara akurasi tertinggi tercatat pada 100 epoch pada optimasi RMSprop. Selanjutnya model menggunakan fungsi loss MSE. Nilai loss terendah ditemukan pada optimasi Adam dan RMSprop dengan 100 epoch yang menghasilkan nilai loss 0,0011. Sementara itu, nilai loss tertinggi sebesar 0,003656 ditemukan pada optimasi SGD dengan 25 epoch (Enriko dkk., 2023).

Berdasarkan berbagai penelitian sebelumnya menunjukkan bahwa metode LSTM telah menunjukkan performa yang baik dalam menganalisis dan memprediksi data time series termasuk harga saham. Namun, setiap penelitian memiliki perbedaan dalam hal pendekatan dan objek yang diteliti. Oleh karena itu, penelitian ini berupaya memberikan kontribusi ilmu dan pengetahuan terkait penerapan metode LSTM kepada penulis terhadap penelitian pada saham Hapag-Lloyd AG. Penelitian diharapkan dapat memperluas pemahaman dan memberikan wawasan baru terhadap penerapan LSTM dalam analisis data *time series*.

## **2.2 Landasan Teori**

Bagian ini membahas sejumlah konsep yang mendasari topik skripsi ini. Landasan teori memberikan wawasan mengenai berbagai pengetahuan untuk mendalami topik analisis prediksi harga saham menggunakan LSTM. Kegunaan teori yang telah dihimpun dalam landasan teori digunakan untuk merumuskan proses analisis prediksi harga saham dengan menggunakan metode LSTM.

### **2.2.1 Saham**

Di dalam dunia investasi terdapat banyak instrumen yang dapat dipilih oleh investor perorangan, seperti deposito, valuta asing, reksadana, tanah, logam, dan saham. Semua instrumen tersebut mempunyai kelebihan dan kekurangan pada masing-masing perspektif investor. Tetapi, salah satu instrumen investasi menarik yang dapat dilakukan di berbagai negara ialah saham.

Saham merupakan surat berharga yang menjadi tanda kepemilikan seseorang atau badan terhadap suatu perusahaan. Biasanya perusahaan dapat mengeluarkan dua jenis saham yang terdiri atas saham biasa dan preferen. Saham biasa adalah pemilik asli dari perusahaan yang menanggung risiko dan mendapatkan profit. Pemegang saham biasa mempunyai hak suara dalam RUPS (Rapat Umum Pemegang Saham) dan turut serta dalam menentukan kebijakan perusahaan. Apabila perusahaan dibubarkan, maka pemegang saham biasa akan membagi sisa aset perusahaan setelah dikurangi bagian pemegang saham preferen.

Saham preferen memperoleh hak istimewa dalam pembayaran dividen yang tetap setiap tahunnya dibandingkan saham biasa. Hak dividen pemegang saham preferen akan dikumpulkan apabila perusahaan pada tahun tersebut tidak mampu membagikan dividen. Pemegang saham preferen juga akan mendapatkan pembayaran dari sisa-sisa aset perusahaan sebelum pemegang saham biasa jika perusahaan dibubarkan. Sebagai konsekuensinya, pemegang saham preferen mempunyai hak suara terbatas dalam menentukan kebijakan perusahaan.

Bursa saham terpengaruh pada faktor-faktor ekonomi dan non-ekonomi dari dalam dan luar negeri. Perdagangan saham antara pembeli dan penjual saham hanya dapat dilakukan di bursa efek melalui perusahaan sekuritas. Harga saham dapat bergerak secara dinamis dan tidak ada jaminan apabila suatu perusahaan memiliki laba terkontrol atau pembagian dividen yang besar, maka harga sahamnya akan stabil naik. Adapun karena suatu hal, harga saham pada perusahaan dapat terus-menerus menurun meskipun labanya besar (Hin, 2008).

Investor mempunyai karakteristik dalam menentukan kategori saham untuk memperhitungkan profit yang berpotensi diraih. Kategori *blue chip* merupakan saham perusahaan yang mempunyai reputasi baik, catatan pertumbuhan profit stabil dari tahun ke tahun, dan dapat memberikan dividen secara stabil. Beberapa perusahaan tersebut telah diakui dalam pengelolaan secara profesional dalam menghasilkan produk/jasa berkualitas. Kelompok *growth stock* dapat diartikan kepada perusahaan yang menampilkan pertumbuhan pendapatan dan laba yang lebih tinggi dari rata-rata sektornya. Perusahaan *growth stock* cenderung menggunakan labanya untuk ekspansi dan sedikit bahkan tidak mengalokasikan pada dividen.

Terdapat saham lain yang termasuk dalam *defensive stock* di mana perusahaan dapat bertahan dalam keadaan ekonomi yang tidak stabil. Ketika kondisi perekonomian sudah membaik, perusahaan tersebut dapat mempertahankan labanya. Kategori *cyclical stock* mendapatkan momentum pada musim-musim tertentu. Beberapa perusahaan dapat memiliki musim berbeda yang bergantung pada agenda yang dilakukan oleh beberapa



kelompok masyarakat, seperti keluarga yang memiliki anak dalam memperlengkap kebutuhan tahun ajaran baru, persiapan hari-hari raya keagamaan, atau kebijakan terhadap naik atau turunnya suku bunga (Tambunan, 2007).

Ciri khas yang dimiliki saham biasanya dilakukan berdasarkan sifat kepemilikan atas nama dan untuk investasi jangka panjang. Saham dengan mudah dapat dipindahtangankan melalui proses jual beli pada bursa efek di suatu negara. Potensi keuntungan yang dapat diraih investor saham salah satunya dividen sebagai sejumlah pengembalian yang dilandaskan pada jumlah kepemilikan atas saham perusahaan bagi pemegang saham. Keuntungan saham lainnya dapat diperoleh dari *capital gain*, yaitu hasil penjualan saham dari selisih nilai jual yang lebih tinggi dari nilai belinya. Keunikan pada keuntungan saham juga dapat dijaminkan ke bank sebagai agunan untuk memperoleh kredit.

Berinvestasi saham tidak terhindarkan dari risiko yang bernama *capital loss* sebagai kerugian hasil penjualan dari selisih nilai jual yang lebih rendah dari nilai belinya. Risiko likuidasi dapat terjadi jika perusahaan bubar dan pemegang saham memiliki klaim terakhir terhadap aset perusahaan setelah semua kewajiban perusahaan telah dibayar. Saham juga dapat di-*delisting* dari pencatatan bursa sehingga tidak dapat diperdagangkan lagi.

Pasar pada bursa terbagi menjadi tiga, seperti pasar reguler, pasar negosiasi, dan pasar tunai. Pasar reguler merupakan pasar dengan pembentukan harga saham dilakukan dari tawar menawar terus-menerus secara lelang. Untuk pembentukan harga dengan negosiasi langsung antara sekuritas pembeli dan penjual dilakukan pada pasar negosiasi. Dan pasar tunai terjadi dengan pembentukan harga pada sekuritas yang tidak melunasi kewajiban penyelesaian transaksi di pasar reguler dan negosiasi.

Tempat melakukan pembelian saham dapat dilakukan pada pasar perdana di mana saham suatu perusahaan ditawarkan kepada masyarakat untuk pertama kalinya. Investor hanya bisa membeli saham pada mekanisme yang bernama *Initial Public Offering* (IPO) dan menjualnya di pasar sekunder. Pasar sekunder ialah tempat jual beli saham yang dapat dilakukan

di hari perdagangan setelah perusahaan dikatakan tercatat di bursa. Sekuritas akan menjadi jembatan dalam pembelian dan penjualan saham (Situmorang dkk., 2010).

### 2.2.2 Hapag-Lloyd AG

Hapag-Lloyd Aktiengesellschaft (AG) beroperasi pada aktivitas pengiriman global dengan menyediakan berbagai jenis kontainer untuk memenuhi berbagai kebutuhan kargo. Pada tahun 2025, Hapag-Lloyd AG menawarkan armada dengan kapasitas kapal sebanyak 2,4 juta TEU (*Twenty-foot Equivalent Unit*) dan kontainer sebanyak 3,7 juta TEU yang menjadikan salah satu armada kontainer berpendingin terbesar di dunia.

Perjalanan Hapag-Lloyd dimulai dari abad ke-19 pada saat perusahaan pendirinya yang bernama Hamburg-Amerikanische Packetfahrt-Actien-Gesellschaft (Hapag) and North German Lloyd berlayar untuk pertama kalinya membawa kargo umum dan penumpang ke New York. Dengan perubahan yang sangat dinamis, berkembanglah perusahaan pelayaran peti kemas ternama dan aktif seperti saat ini.

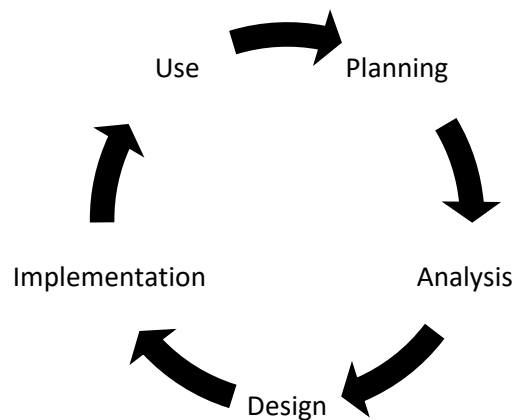
“*We care. We move. We deliver.*” merupakan nilai-nilai perusahaan Hapag-Lloyd AG yang menjadi inti dari semua hal yang perusahaan lakukan. Nilai-nilai ini menceritakan kisah tentang asal-usul, pendefinisian, dan bimbingan dalam perjalanan perusahaan ke depan. Nilai-nilai tersebut juga mengarahkan cara perusahaan dalam berbisnis, inspirasi, motivasi, dan membantu perusahaan menemukan jalan yang benar (Hapag-Lloyd AG, 2025).

### 2.2.3 Pengembangan Sistem *Waterfall*

Rekayasa perangkat lunak menawarkan dokumentasi dan pemeliharaan terhadap produksi model dan proses pada perangkat lunak tersebut. Memperhitungkan lamanya waktu dan pengupayaan pengembangan dapat dijadikan modal untuk menimbang perjalanan prosedur pengembangan perangkat lunak. *Life cycle model* pada suatu pengembangan perangkat lunak umumnya dapat berdampak pada pengembangan perangkat lunak dan

biasanya siklus tersebut dinamakan *software development life cycle* (SDLC) (Aggarwal & Singh, 2005).

SDLC merupakan pelaksanaan dengan suatu cara yang disarankan untuk mengurai masalah dalam pengembangan sistem informasi. Para pengembang memahami beberapa tahap pada upaya pengembangan yang mesti dilakukan dalam urutan tertentu. Aliran kegiatan satu arah yang mengarah pada penuntasan proyek dikenal dengan pendekatan *waterfall*.



**Gambar 2.1** Siklus Hidup Pengembangan Sistem *Waterfall*  
(Sumber: McLeod dan Schell, 2007)

Proyek direncanakan dengan menulis sumber daya yang dibutuhkan untuk dikumpulkan dalam melakukan pengembangan. Penjabaran masalah diwujudkan pada tahap perencanaan dan analisis. Analisis sistem dijalankan untuk mendalami masalah dan memilih persyaratan fungsional sistem yang baru.

Penjabaran dan evaluasi dalam pemilihan solusi berada pada tahap desain. Selanjutnya solusi tersebut dikembangkan dan diimplementasikan pada suatu sistem. Sistem digunakan dalam jangka waktu yang lama dengan mengoleksi umpan balik untuk mengetahui pemecahan masalah yang dikerjakan pada sistem tersebut. Pada waktu suatu sistem tidak lagi berguna dan mesti diganti, siklus hidup dimulai kembali kepada tahap perencanaan (McLeod & Schell, 2007).

Struktur model *waterfall* mampu mengendalikan proyek pengembangan perangkat lunak dikarenakan dokumentasi yang teratur. *Waterfall* mementingkan urutan langkah sebelum menjalani fase lainnya sehingga dapat menghemat waktu pengembangan dan menemukan desain dan kesalahan persyaratan yang lebih khusus. *Waterfall* sesuai untuk perancangan sistem yang memerlukan lebih sedikit sumber daya pada ruang lingkup proyek yang kecil (Aroral, 2021).

#### **2.2.4 Persyaratan Sistem Perangkat Lunak**

Persyaratan sistem perangkat lunak (*software system requirements*) merupakan pemaparan mendetail berkenaan dengan fungsi, layanan, dan batasan operasional yang dilakukan sistem perangkat lunak. Persyaratan sistem tersebut memberikan informasi khusus berkaitan dengan pelaksanaan sistem yang mendukung proses bisnis suatu perusahaan. Persyaratan sistem dapat dirumuskan sebelum pengambilan keputusan untuk melangsungkan pengembangan sistem.

Persyaratan sistem tidak berdiri sendiri dan satu persyaratan umumnya menghasilkan/membatasi persyaratan lainnya. Persyaratan tersebut menentukan kegunaan yang dibutuhkan untuk meyakinkan bahwa layanan/fitur dapat dilangsung secara efektif. Persyaratan sistem perangkat lunak dapat dikategorikan menjadi persyaratan fungsional atau non-fungsional.

Persyaratan fungsional merupakan pernyataan layanan yang mesti disiapkan sistem, seperti menanggapi masukan pengguna dan perbuatan sistem pada situasi tertentu. Persyaratan fungsional dapat dinyatakan secara tegas terkait tindakan-tindakan yang dilakukan oleh sistem. Persyaratan non-fungsional ialah batasan layanan/fungsi yang disajikan oleh sistem dan berlangsung secara menyeluruh.

Persyaratan fungsional berpegang pada jenis pengembangan perangkat lunak, pengguna, dan pendekatan pada penulisan persyaratan tersebut. Persyaratan fungsional yang ditulis sebagai persyaratan pengguna mesti ditulis dengan bahasa yang dimengerti oleh pengguna dan pengelola sistem.

Persyaratan ini memuat hal-hal yang sebaiknya dilakukan oleh sistem secara spesifik untuk menggambarkan cara kerja yang ada di suatu organisasi. Spesifikasi persyaratan fungsional sistem mengandung semua layanan dan informasi yang dibutuhkan oleh pengguna dan persyaratan tersebut tidak boleh bertentangan di antara pengembang, pengguna, dan pengelola sistem.

Persyaratan non-fungsional tidak secara langsung berhubungan dengan layanan spesifik yang diberikan sistem kepada penggunanya. Persyaratan ini dapat memutuskan batasan pada eksekusi sistem, seperti kemampuan perangkat masukan dan keluaran atau gambaran data yang digunakan dalam antarmuka dengan sistem lain.

Pelaksanaan persyaratan non-fungsional dapat menjalar pada seluruh sistem dikarenakan dapat memengaruhi keseluruhan arsitektur sistem daripada komponen individual dan mendatangkan beberapa persyaratan fungsional dalam memilih layanan sistem baru yang dibutuhkan bilamana persyaratan non-fungsional tersebut dijalankan. Persyaratan non-fungsional dapat berasal dari kebutuhan karakteristik perangkat lunak (persyaratan produk), organisasi pengembangan perangkat lunak (persyaratan organisasi), atau sumber eksternal.

Persyaratan produk (*product requirements*) memutuskan perilaku waktu berjalan perangkat lunak. Persyaratan produk meliputi persyaratan kegunaan (*usability*), efisiensi (*efficiency*), keandalan (*dependability*), dan keamanan (*security*).

Persyaratan organisasi (*organizational requirements*) merupakan persyaratan sistem menyeluruh yang berasal dari prosedur dalam organisasi pengembang dan pengguna. Bagian persyaratan organisasi terdiri dari persyaratan lingkungan (*environmental*), operasional (*operational*), dan pengembangan (*development*).

Persyaratan eksternal (*external requirements*) berasal dari komponen-komponen eksternal terhadap sistem dan proses pengembangannya. Persyaratan eksternal memuat persyaratan peraturan (*regulatory*), etika (*ethical*), dan legislatif (*legislative*) (Sommerville, 2016).

### 2.2.5 Data Mining

*Data mining* atau *Knowledge Discovery in Databases* (KDD) adalah kumpulan teknik eksplorasi yang dilandaskan pada metode dan alat analisis untuk menangani sejumlah besar informasi. Teknik-teknik tersebut dapat menemukan pola-pola baru yang dapat membantu perusahaan dalam memahami bisnis dan mengambil keputusan perusahaan. Teknik *data mining* (penambangan data) dapat digunakan untuk jumlah yang kecil, tetapi peluang menemukan sesuatu yang baru dan menarik lebih besar apabila jumlah datanya besar.

Penambangan data semakin menarik dikarenakan ada banyak pertumbuhan data dan kapasitas penyimpanan data, penurunan biaya pemrosesan, lingkungan yang bersaing, dan perangkat lunak penambangan data yang mudah diakses. Penambangan data pada suatu perusahaan menggunakan penyimpanan data terpisah di mana basis data terpisah tersebut berupa gudang data yang menyimpan informasi untuk pengambilan keputusan perusahaan tersebut dengan mencakup data terkini dan historis.

Proses penambangan data pada umumnya mempunyai langkah-langkah yang dimulai dengan *requirements analysis*. Para pengambil keputusan perusahaan membentuk tujuan yang dicapai dari proses penambangan data. Penggunaan pada penambangan data terjadi karena ide dan teknik yang jelas pada masalah yang ingin dipecahkan. Pemilihan dan

pengumpulan data dapat dijabarkan dari pencarian basis data menggunakan gudang data sampai pemilihan sampel datanya. Pembersihan dan persiapan data dilaksanakan untuk mencari dan menangani data yang hilang, konflik data, dan ketidakpastian pada suatu data.

Eksplorasi dan validasi penambangan data menerapkan sejumlah teknik yang dapat diartikan perusahaan. Penggunaan beberapa teknik untuk eksplorasi, pengujian, dan validasi lebih lanjut disesuaikan dengan kebutuhan perusahaan. Implementasi, evaluasi, dan pemantauan dapat memilih teknik yang efektif dan hasil akurat. Pemantauan diperlukan secara berkala agar alat dan teknik menjadi lebih utuh. Hasil penambangan data kepada para pengambil keputusan dapat divisualisasikan untuk menggambarkan keefektifan masalah yang diselesaikan (Gupta, 2014).

Data yang diolah pada penambangan data mempunyai perbedaan cara dan bergantung pada nilai sejumlah variabel yang disebut atribut. Terdapat atribut yang dikhususkan untuk memprediksi nilai atributnya pada kumpulan data yang belum pernah terlihat termasuk pada data berlabel dengan melakukan penambangan data *supervised learning*. Atribut yang bersifat kategoris akan menggunakan metode klasifikasi dan atribut bersifat numerik menggunakan metode regresi. Data yang tidak mempunyai atribut disebut data tidak berlabel dengan melaksanakan penambangan data *unsupervised learning*. Hubungan di antara nilai-nilai variabel dapat ditemukan dengan menggunakan metode aturan asosiasi dan pengelompokan nilai yang serupa dapat menggunakan metode *clustering* (Bramer, 2016).

Penambangan data memiliki tantangan tersendiri agar berjalan sesuai yang dibutuhkan. Banyaknya jenis data menghendaki sistem penambangan data untuk menangani data yang kompleks. Sistem penambangan data yang spesifik dapat dibangun untuk mendapatkan pengetahuan dalam basis data yang beragam. Waktu berjalan algoritma penambangan data mesti dapat diprediksi dan diterima dalam basis data besar.

Pengekspresian data ditentukan oleh pengguna akhir agar dapat memahami pengetahuan yang berasal dari suatu penambangan data. Penambangan data sebaiknya dibentuk secara interaktif dan fleksibel dari

berbagai sudut pandang. Informasi yang diperoleh dari berbagai sumber data diharapkan dapat menyanggupi ukuran basis data yang besar, distribusi data yang luas, dan kerumitan komputasi dari beberapa metode penambangan data. Beberapa data mesti dipastikan untuk memenuhi perlindungan privasi dan tindakan keamanan yang perlu dijalankan (Chen, 1996).

### 2.2.6 *Long Short-Term Memory*

Jaringan neural *Feed-Forward* konvensional yang mempunyai perluasan kemampuan mengelola masukan deret dengan panjang variabel dikenal dengan *Recurrent Neural Networks* (RNN). Model RNN mengadakan beberapa gerbang untuk menyimpan masukan sebelumnya dan memanfaatkan informasi berurutan dari masukan sebelumnya. Memori RNN khusus ini disebut *recurrent hidden states* (status tersembunyi berulang) dengan mengarahkan RNN untuk memprediksi masukan yang akan datang dalam urutan data masukan.

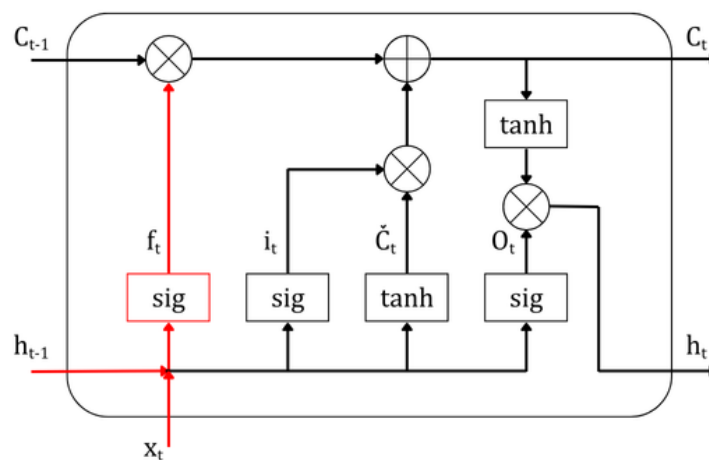
RNN mempunyai keterbatasan panjang informasi berurutan pada beberapa langkah mundur. Keterbatasan tersebut berhubungan pada informasi tentang masukan/gradien melalui banyak lapisan di mana gradien tersebut akan hilang saat mencapai lapisan akhir/awal yang dapat disebut *vanishing gradients*. Hal tersebut mengakibatkan RNN akan sulit untuk menangkap ketergantungan jangka panjang. *Vanishing gradient* memutuskan nilai yang lebih kecil kepada matriks yang digunakan dalam pelatihan RNN (*weight matrix*).

Adapun juga kekurangan lainnya terdapat pada informasi masukan/gradien melewati banyak lapisan yang terkumpul dan membentuk gradien yang sangat besar saat mencapai lapisan akhir/awal yang disebut *exploding gradients*. *Exploding gradient* pada algoritma pelatihan RNN memutuskan nilai yang lebih tinggi ke *weight matrix* (matriks bobot). Dua kekurangan ini mengakibatkan RNN rumit dilatih dan penyelesaian masalah dapat dilakukan dengan memotong gradien.



Dikarenakan kesulitan pada pembelajaran keterikatan jangka panjang, maka *Long Short-Term Memory* (LSTM) menjadi perluasan RNN yang cakap dalam mengatasi masalah gradien yang menghilang. Perluasan memori RNN di LSTM mempunyai kemampuan untuk mengingat informasi jangka waktu yang lebih lama sehingga dapat menunjang pembacaan, penulisan, dan penghapusan informasi dari memorinya. Pengambilan fitur penting dari masukan model LSTM dan menyimpan informasi dalam jangka waktu panjang dengan membuat keputusan untuk menghapus atau menyimpan informasinya berlandaskan penentuan nilai bobot pada informasi selama proses pelatihan. Maka dari itu, model LSTM mendalami informasi yang dapat disimpan atau dihapus.

Model LSTM terbentuk dari tiga gerbang (*gate*) yang dimulai dari *forget gate*. Keputusan dibangun pada *forget gate* untuk menyimpan/menghapus informasi yang ada dari memori LSTM yang berasal dari fungsi *sigmoid* (*sigmoid function*). Keputusan ini dibentuk berlandaskan nilai  $h_{t-1}$  (*hidden state* sebelumnya) dan  $x_t$  (masukan pada  $t$ ).

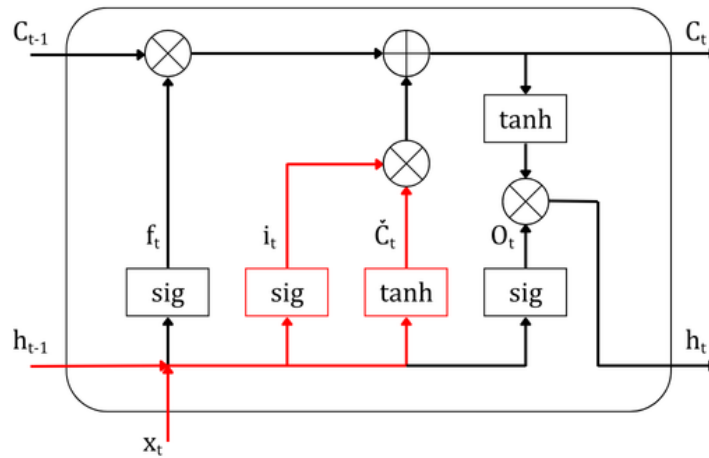


**Gambar 2.2** *Forget Gate*  
(Sumber: Singhal, 2020)

Keluaran gerbang ini ialah  $f_t$  (*forget gate* pada  $t$ ) dengan nilai antara 0 dan 1 di mana 0 menandakan penghapusan nilai yang dipelajari sepenuhnya dan 1 mengartikan penyimpanan seluruh nilai. Persamaan untuk *forget gate* terbentuk sebagai berikut.

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (1)$$

$\sigma$  menggambarkan fungsi *sigmoid*,  $W_f$  menggambarkan bobot untuk nilai masukan pada  $t$ , dan  $b_f$  menggambarkan bias *forget gate*. Pada gerbang kedua terdapat *input gate* dengan melaksanakan keputusan pemilihan informasi baru ditambahkan ke memori atau tidak. Gerbang ini mempunyai dua lapisan, yaitu lapisan *sigmoid* (*sigmoid layer*) untuk menentukan nilai yang perlu diperbarui dan lapisan *tanh* untuk membentuk vektor nilai kandidat baru yang akan ditambahkan ke memori LSTM.



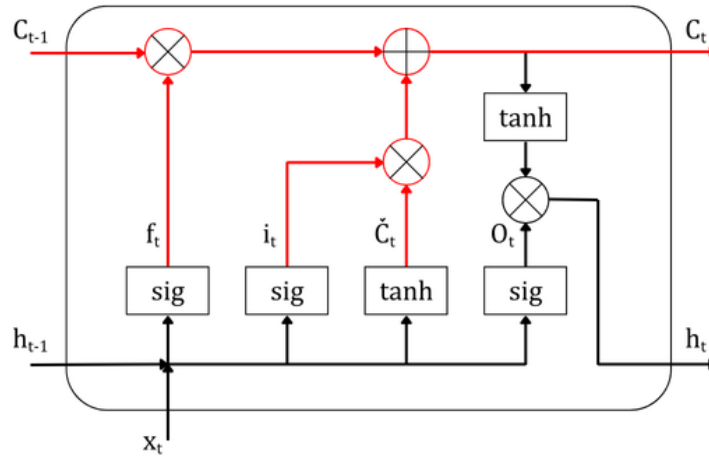
**Gambar 2.3** *Input Gate*  
(Sumber: Singhal, 2020)

Keluaran dari kedua lapisan (masing-masing persamaan untuk *input gate* dan kandidat *cell state* terbentuk) tersebut sebagai berikut.

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (2)$$

$$\check{c}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (3)$$

$i_t$  menunjukkan pemilihan nilai yang perlu diperbarui atau tidak dan  $\check{c}_t$  menunjukkan nilai kandidat baru yang akan ditambahkan ke memori LSTM.  $W_i$  menggambarkan bobot untuk nilai masukan pada  $t$ ,  $b_i$  menggambarkan bias pada *input gate*, *hyperbolic tangent* digambarkan dengan *tanh*,  $W_c$  menggambarkan bobot untuk masukan pada sel ke  $c$ , dan  $b_c$  menggambarkan bias pada sel ke  $c$ .

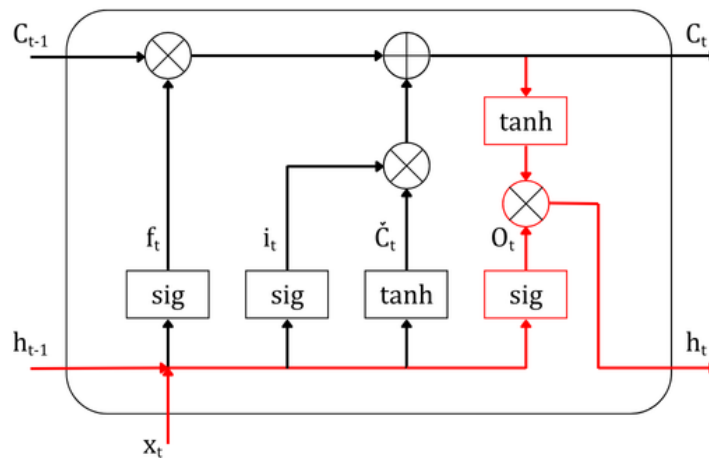


**Gambar 2.4 Cell State**  
(Sumber: Singhal, 2020)

Campuran kedua lapisan ini memberikan pembaruan untuk memori LSTM dengan nilai saat ini dilupakan menggunakan lapisan *forget gate* melalui perkalian nilai lama ( $C_{t-1}$ ) dibarengi dengan menambahkan kandidat baru  $i_t * \check{C}_t$ . Persamaan untuk *cell state* baru terbentuk sebagai berikut.

$$C_t = f_t * C_{t-1} + i_t * \check{C}_t \quad (4)$$

$C_t$  menggambarkan nilai memori *cell state* baru. Nilai *forget gate* dilambangkan dengan  $f_t$  yang memiliki nilai di antara 0 dan 1 di mana 0 memberitahukan penghapusan nilai sepenuhnya dan 1 mengartikan pemeliharaan nilai sepenuhnya. Gerbang terakhir bernama *output gate* yang pelaksanaan awalnya memakai lapisan *sigmoid* untuk membentuk keputusan terkait suatu bagian dari memori LSTM yang berperan pada keluaran.



**Gambar 2.5 Output Gate**  
(Sumber: Singhal, 2020)

*Output gate* mengerjakan fungsi *tanh* non-linear agar menempatkan nilai antara -1 dan 1 dan hasilnya akan dikalikan dengan keluaran lapisan *sigmoid*. Persamaan untuk *output gate* dan hasil akhir terbentuk sebagai berikut.

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (5)$$

$$h_t = o_t * \tanh(C_t) \quad (6)$$

$o_t$  menggambarkan nilai *output gate* (keluaran),  $W_o$  menggambarkan bobot untuk keluaran pada  $t$ ,  $b_o$  menggambarkan bias pada *output gate*, dan  $h_t$  menggambarkan nilai hasil akhir (Siarni-Namini dkk., 2019). Kumpulan data *time series* mempunyai urutan ketergantungan di antara variabel masukan. Sebagai jenis RNN, metode LSTM dapat menangani hal tersebut dengan menyimpan dan mempelajari urutan pengamatan yang panjang. Algoritma (*pseudocode*) yang dapat dijalankan dalam metode LSTM ialah sebagai berikut.

**Tabel 2.1** Algoritma LSTM

Algoritma LSTM
Masukan: Time series Keluaran: RMSE dari data yang di-forecasting # Membagi data dengan 70% training size $\leftarrow$ length(series) * 0.7 train $\leftarrow$ series[0...size] test $\leftarrow$ series[size...length(size)] # Mengatur random seed dengan nilai tetap set random.seed(7)  # Menyesuaikan model LSTM dengan data pelatihan Procedure fit_lstm(train, epoch, neurons) X $\leftarrow$ train y $\leftarrow$ train - X model = Sequential() model.add(LSTM(neurons), stateful=True))

```

model.compile(loss='mean_squared_error',
optimizer='adam')
for each I in range(epoch) do
    model.fit(X, y, epoch=1, shuffle=False)
    model.reset_states()
end for
return model

# Membuat forecast satu langkah
Procedure forecast_lstm(model, X)
yhat ← model.predict(X)
return yhat

epoch ← 1
neurons ← 4
predictions ← empty
# Menyesuaikan model LSTM
lstm_model = fit_lstm(train, epoch, neurons)
# Forecast kumpulan data training
lstm_model.predict(train)

# Memvalidasi data testing
for each i in range(length(test)) do
    # Membuat forecast satu langkah
    X ← test[i]
    yhat ← forecast_lstm(lstm_model, X)
    # Mencatat forecast
    predictions.append(yhat)
    expected ← test[i]
end for

```

```

MSE ← mean_squared_error(expected, predictions)

Return (RMSE ← sqrt(MSE))

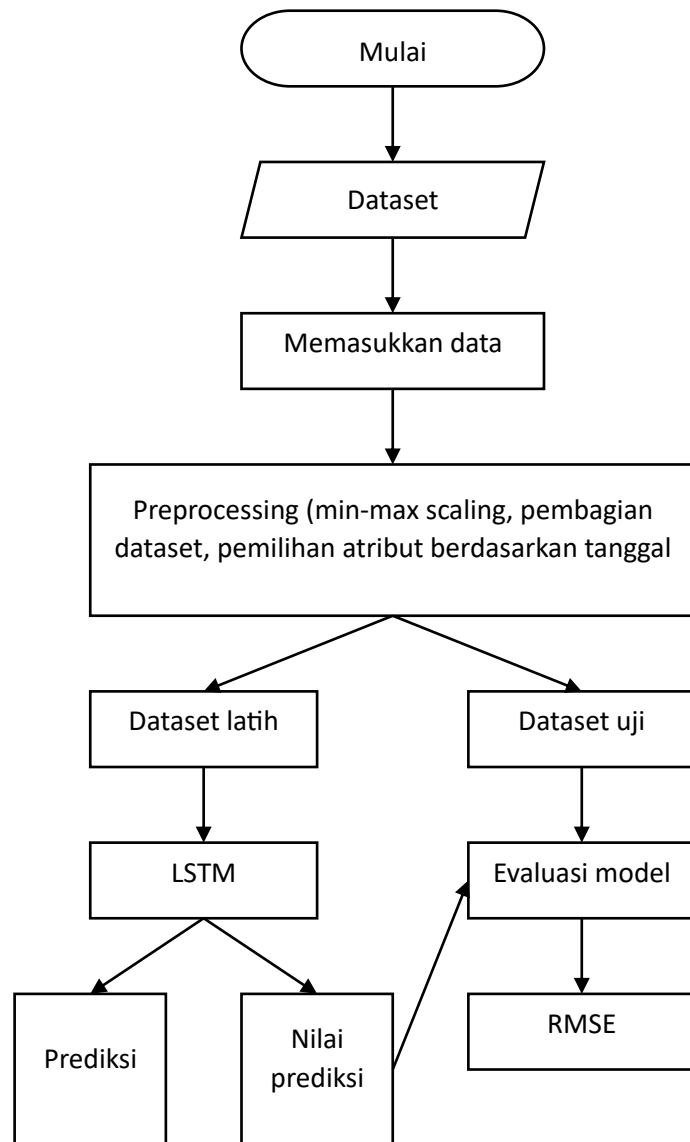
```

Sumber: Siami-Namini dkk, 2018

Algoritma membutuhkan pustaka Keras untuk memodelkan datanya. Algoritma dimulai dengan membagi dataset sebesar 70% untuk data pelatihan (*training*). Fungsi `fit_lstm` menggunakan data pelatihan, *epoch* (jumlah waktu dataset untuk melatih model), dan *neuron* (jumlah unit). Pengompilasian model menggunakan *loss function* 'mean\_squared\_error' dan Adam untuk optimasinya.

Penyesuaian model dengan data pelatihan dilakukan setelah kompilasi. Apabila lebih dari satu epoch, maka tahap pengaturan ulang *network* (jaringan) mesti diawasi pada model jaringan yang bersifat *stateful*. Parameter *shuffling false* diperlukan saat melatih model yang dioptimalkan menggunakan tahap sebelumnya untuk memaksimalkan mekanisme pembelajaran. *Epoch* akan menjadi iterasi berikutnya setelah status internal model pelatihan diatur ulang.

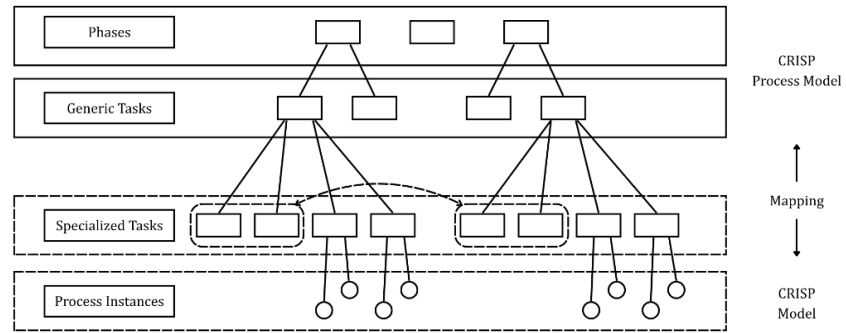
Fungsi `forecast_lstm` dapat memanggil model LSTM dan memprediksi langkah berikutnya dalam kumpulan datanya. Jumlah *epoch* dan *neuron* ditentukan, lalu model LSTM dibangun untuk memprediksi dan *forecasting* data pengujian (*testing*). Setelah itu, nilai metrik evaluasi RMSE ditampilkan. Terakhir, nilai RMSE ditampilkan untuk mengevaluasi model (Siami-Namini dkk., 2018). Visualisasi algoritma LSTM dengan *flowchart* adalah sebagai berikut (Albeladi dkk., 2023).



**Gambar 2.6** *Flowchart LSTM*  
(Sumber: Albeladi, 2023)

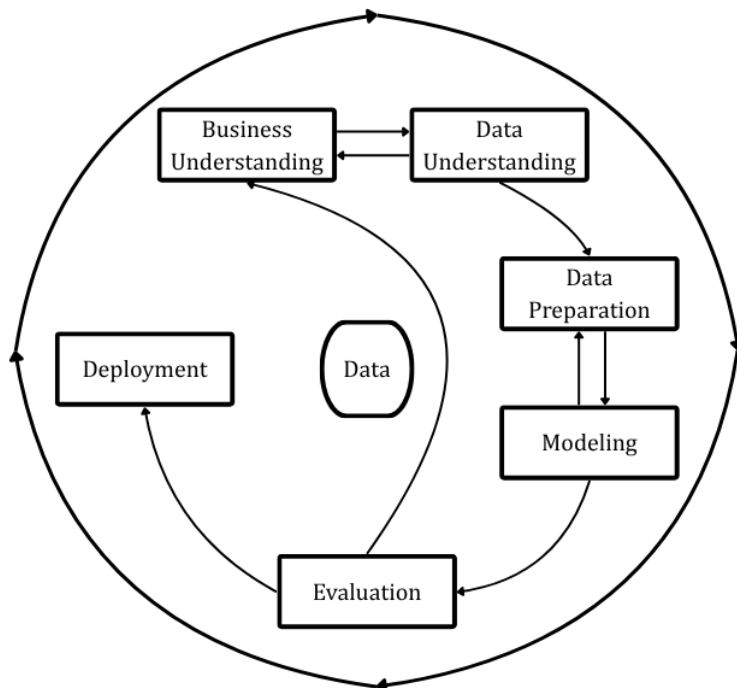
### 2.2.7 *Cross-Industry Standard Process for Data Mining*

CRISP-DM (*Cross-Industry Standard Process for Data Mining*) merupakan metodologi penambangan data dengan proses berjenjang yang meliputi penjabaran kumpulan tugas pada empat tingkat. Tingkat pertama dikoordinasikan ke dalam sejumlah *phases* (fase) di mana masing-masing fase terdiri dari beberapa tugas tingkat kedua. Tingkat kedua dinamakan *generic tasks* dengan mencakup proses penambangan data yang lengkap dan teknik model pada pemodelan baru yang stabil.



**Gambar 2.7** Tingkatan CRISP-DM  
(Sumber: Chapman dkk, 2000)

Tingkat ketiga dinamakan *specialized tasks* dengan mengatur tindakan tugas umum pada situasi tertentu secara spesifik dan memberikan perbedaan tugas dalam situasi yang berbeda. Tingkat keempat disebut *process instances* menyampaikan catatan, keputusan, dan hasil nyata pada penambangan data dalam keikutsertaan tertentu di tingkat yang lebih tinggi. Model CRISP-DM berisikan fase-fase proyek dan hubungan tugas antar tugas. Tugas-tugas penambangan data mempunyai hubungannya berlandaskan situasi, data, dan kebutuhan pengguna.



**Gambar 2.8** Model CRISP-DM  
(Sumber: Chapman dkk, 2000)



Siklus model CRISP-DM mempunyai enam fase. Urutan antar fase fleksibel dengan adanya perpindahan maju mundur antara fase yang berbeda. Pemilihan aksi fase ditentukan berlandaskan hasil setiap fase dan tugas tertentu dengan mengingat ketergantungan antar fase yang ditandai anak panah. Penambangan data terus dilanjutkan setelah solusi dilaksanakan yang ditandai pada lingkaran luar model. Pengetahuan yang dilatih dan solusi sebelumnya dapat mendorong topik bisnis baru yang lebih terarah.

Uraian setiap fase CRISP-DM dimulai dari *business understanding* di mana fase ini berkonsentrasi pada penjabaran masalah dan perancangan rencana awal yang berasal dari pemahaman tujuan dan persyaratan dari berbagai sudut pandang bisnis. Fase kedua bernama *data understanding* dengan melakukan pengumpulan data awal, mengenali masalah kualitas data, mendapatkan wawasan singkat data, dan mencari bagian yang menarik untuk membuat pendapat dari informasi yang tidak langsung terlihat.

*Data preparation* menjadi fase ketiga di mana perancangan dataset akhir dari data mentah, seperti pemilihan tabel, *instance*, dan atribut yang disertai perubahan dan pembersihan data untuk diarahkan kepada alat pemodelan. Teknik pemodelan ditunjuk dan dijalankan berlandaskan penyesuaian parameter dengan nilai optimal. Kadangkala pengerjaan balik kepada fase persiapan data dikarenakan persyaratan bentuk data yang khusus pada beberapa teknik pemodelan.

Pembangunan model menampilkan berbagai cara pandang dalam menganalisis data. Tujuan bisnis yang dicapai berasal dari langkah-langkah pemodelan dan evaluasi beberapa masalah yang terpecahkan pada model dan dapat diarahkan pada pengambilan keputusan dan eksekusi akhir model. *Deployment* menjadi fase terakhir dengan menyajikan pengetahuan tentang data dan penerapan model kepada pengguna. Aksi yang dilakukan pengguna perlu dipahami sebelumnya agar proses pengambilan keputusan dapat berlandaskan pada pemanfaatan model yang dikembangkan (Chapman dkk., 2000).

### 2.2.8 *Time Series Forecasting*

*Time series* (deret waktu) ialah sekumpulan titik data yang diurutkan berdasarkan waktu. Setiap titik data dipisahkan oleh selang waktu yang sama yang dapat tercatat dalam setiap jam atau menit dan diselaraskan setiap bulan atau tahun. Beberapa contoh yang menggambarkan deret waktu adalah konsumsi listrik rumah tangga, suhu di dalam ruangan, dan nilai pada suatu penutupan saham. Deret waktu dapat bersambung di masa mendatang dengan menganalisis tren tertentu (Peixeiro, 2022).

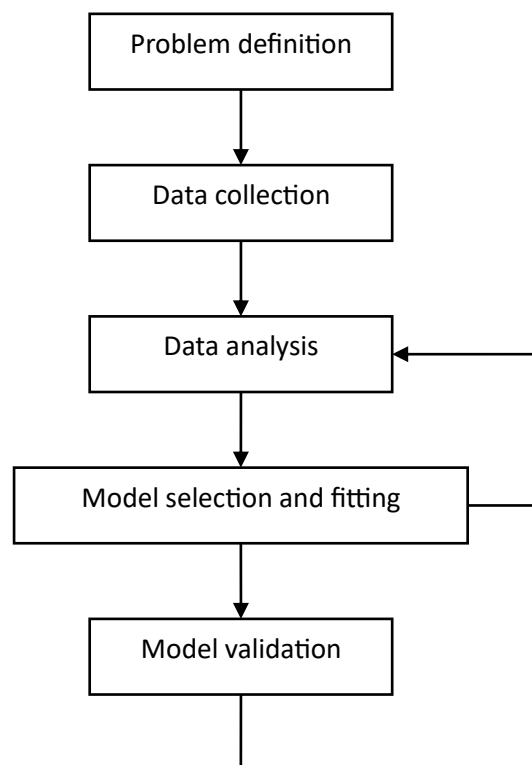
Komponen-komponen yang terjadi pada deret waktu terdiri dari *trend*, *seasonal*, *cyclical*, dan *irregular*. *Trend* (tren) merupakan perubahan jangka panjang dalam rata-rata suatu deret waktu yang berjalan stabil dalam arah tertentu. Apabila deret waktu naik, maka tren dapat dikatakan *upward or increasing trend*, dan sebaliknya. Salah satu contoh dari tren ialah pendapatan perusahaan selama bertahun-tahun. *Seasonal* (musimam) menggambarkan naik-turunnya deret waktu secara teratur dan berulang yang terkadang terjadi di setiap tahun. Beberapa contoh musimam adalah lonjakan penjualan eceran di hari raya keagamaan dan konsumsi pemanas ruangan di musim dingin.

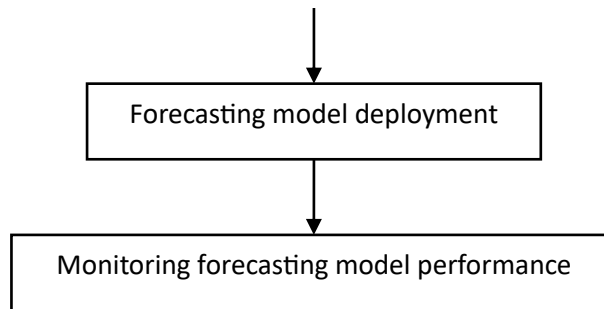
*Cyclical* (siklus) menandakan pola naik-turunnya deret waktu di setiap periode secara tidak teratur. Contohnya adalah resesi ekonomi yang terjadi selama siklus sepuluh tahun. *Irregular* (tidak teratur) merupakan deret waktu yang tidak sepenuhnya tidak dapat diprediksi dan dianggap tersisa. Berapapun banyak variabel yang ditambahkan kadangkala terdapat beberapa komponen sisa yang tidak teratur. Salah satu contohnya ialah suatu elemen yang tidak teratur pada kegiatan promosi dalam penjualan eceran (Joseph, 2022).

*Forecasting* (peramalan) dilakukan untuk memprediksi kelanjutan deret waktu dengan menggunakan data historis (masa lalu) dan pengetahuannya terkait kejadian di masa depan (Peixeiro, 2022). Metode peramalan dapat berasal dari prosedur baik bergantung maupun tidak bergantung pada suatu model deret waktu agar menemukan kemungkinan ramalan yang lebih baik (Chatfield, 2000).

Terdapat dua teknik peramalan yang umumnya terjadi, seperti kualitatif dan kuantitatif. Teknik peramalan kualitatif digunakan dalam kondisi di mana hanya sedikit data historis yang menjadi dasar peramalan dan membutuhkan penilai dari seseorang/kelompok ahli. Sedangkan, teknik peramalan kuantitatif merangkum pola data dan hubungan antar variabel dengan memakai data historis yang menjadikan suatu model deret waktu. Model tersebut dimanfaatkan untuk memperhitungkan pola data di masa depan dari perilaku masa lalu dan saat ini.

Proses peramalan merupakan susunan kegiatan yang terhubung di mana terdapat perubahan satu atau lebih masukan menjadi satu atau lebih keluaran. Aktivitas-aktivitas yang dilakukan dalam proses peramalan dimulai dengan *problem definition* (definisi masalah) yang mengaitkan penjabaran tujuan dan penggunaan sistem pengembangan oleh pengguna. Aktivitas ini juga dapat membantu dalam mengenali risiko pada penggunaan interval dan merumuskan komponen pada pemecahan masalah dengan tepat. Keberhasilan akhir model deret waktu sebagian besar dipastikan dari aktivitas definisi masalah ini.





**Gambar 2.9** Proses Peramalan  
(Sumber: Montgomery dkk, 2016)

Aktivitas kedua bernama *data collection* (pengumpulan data) dengan memperoleh data yang berguna termasuk informasi historis pada variabel target yang potensial. Pengumpulan data juga dapat direncanakan penangani dalam pengumpulan dan penyimpanan data di masa depan agar lebih fleksibel dan berkelanjutan. Analisis data penting dilakukan untuk menentukan variabel yang akan dipakai dengan cara memberikan ringkasan numerik dan eksplorasi variabel target potensial beserta hubungan antar variabel lainnya. Informasi pada analisis data dapat memberikan gambaran singkat terkait pemilihan metode pada pemodelan peramalan.

Pemilihan dan pemasangan model berisikan penggunaan satu atau lebih model peramalan dengan menimbang parameter yang tidak kelihatan. Hal tersebut dilakukan untuk menjelaskan pembeda dari berbagai model yang dibentuk. Model tersebut akan divalidasi dengan menggunakan evaluasi kinerja pada sistem tersebut. Galat pada pengujian dapat diupayakan mendekati pada titik data historis model.

*Forecasting model deployment* menyertakan penggunaan model dan peramalan yang dipahami oleh pengguna. Sumber data dan informasi lain dibutuhkan bagi pengguna dengan pemeliharaan model secara berkala. *Monitoring forecasting model performance* mesti dilaksanakan seiring perubahan waktu dan berdampak pada penurunan performa peramalan. Penyediaan catatan kesalahan dapat menjadi salah satu dalam dalam memantau kinerja model secara rutin (Montgomery dkk., 2016).

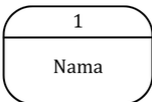
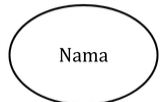
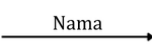
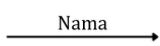
### 2.2.9 Data Flow Diagram

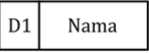
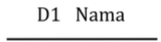
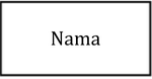
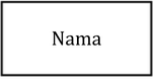
*Data Flow Diagram* (DFD) merupakan alat visual untuk menguraikan model logika dan mengutarakan transformasi data dalam suatu sistem. DFD membantu tahap analisis dan persyaratan desain sistem, mendukung teknik diagram dengan anotasi, dan memaparkan jaringan aktivitas/proses sistem target.

Terdapat empat elemen struktur DFD yang dimulai dari aktivitas/proses. Proses menerima aliran data sebagai masukan dan menghasilkan aliran data sebagai keluaran. Aktivitas dapat digambarkan lebih lanjut untuk membuat subprocess yang lebih menyeluruh. Label aktivitas/proses mesti bersifat kata kerja dan aktivitas dihubungkan kepada spesifikasi proses. Selanjutnya terdapat aliran data (*data flow*) yang merupakan elemen penghubung yang kedua ujungnya terhubung kepada aktivitas/proses, penyimpanan data, entitas eksternal, dan sejenisnya untuk menggambarkan aliran informasi. Anak panah pada *data flow* dapat dipisah/digabung dengan labelnya adalah kata benda.

Penyimpanan data (*data store*) menjadi tempat penyimpanan data tetap dan menampung basis data. Penyimpanan data dilayani oleh suatu proses, tidak ada aktivitas yang dapat melebihi kapasitas pengambilan data dasar, label *data store* bersifat kata benda. Entitas eksternal (*external entity*) menyiapkan koneksi menuju konteks sistem dan merupakan asal/tujuan aliran data eksternal. *External entity* hanya menerima/mengirim data dan labelnya bersifat kata benda. Adapun bentuk-bentuk dari elemen *data flow diagram* adalah sebagai berikut.

**Tabel 2.2** Elemen *Data Flow Diagram*

Elemen <i>Data Flow Diagram</i>	Simbol Gane dan Sarson	Simbol DeMarco dan Yourdon
<i>Process</i>		
<i>Data flow</i>		

<i>Data store</i>		
<i>External entity</i>		

Sumber: Dennis dkk, 2021

Aktivitas/proses tertentu dalam diagram induk pada DFD dapat diuraikan oleh diagram anak yang lebih mendetail. Satu paket DFD dapat meliputi diagram konteks, diagram level-0, dan diagram anak relatif (*relative child diagram*). Diagram konteks (*context diagram*) ialah DFD dari ruang lingkup sistem organisasi yang memperlihatkan batas-batas sistem, entitas eksternal yang berkomunikasi dengan sistem, dan aliran informasi utama antara entitas eksternal dan sistem.

Diagram level-0 (*level-0 diagram*) merupakan DFD yang menggambarkan proses utama sistem, aliran data, dan penyimpanan data pada tingkat detail yang tinggi. Sedangkan diagram anak DFD yang diuraikan (*decomposed DFD child diagrams*) ialah pengembangan dari turunan *level-0 diagram* yang meliputi perincian langkah demi langkah pembentukan kerangka kerja sistem (Q. Li & Chen, 2009).

#### 2.2.10 NoSQL Database

Basis data NoSQL (*Not Only SQL*) merupakan himpunan sistem manajemen data non-relasional yang beragam di mana tidak memerlukan SQL untuk manipulasi data. *NoSQL database* juga berpusat pada pemrosesan analitis pada kumpulan data berskala besar.

Basis data NoSQL mempunyai beberapa klasifikasi NoSQL yang dimulai dari *key-value stores* di mana basis data ini menyimpan hal sebagai kunci dan nilai yang terkait dalam tabel mandiri yang sederhana. Nilai dapat berupa string teks sederhana atau daftar yang lebih rumit. Pengambilan nilai pada *key-value stores* dapat dilakukan dengan cepat dan leluasa. Contoh dari *key-value stores* ialah Voldemort, Redis, BerkeleyDB, dan Dynamo.

Basis data dokumen (*document databases*) dibentuk untuk mengatur dan menyimpan dokumen yang dikodekan dalam format pertukaran data standar, seperti XML, JSON (Javascript Option Notation), atau BSON (Binary JSON). Kolom nilai dalam basis data dokumen berisi data semi terstruktur (tidak teratur) khususnya pasangan nilai atribut. Satu kolom dapat memuat ratusan atribut dan kunci maupun nilai dapat dicari secara utuh dalam basis data dokumen. Contoh dari basis data dokumen adalah CouchDB (JSON) dan MongoDB (BSON).

*Column family database* makai sturktur data berorientasi kolom terdistribusi khususnya data versi karena fungsi penandaan waktu yang melayani beberapa atribut per kunci. Klasifikasi NoSQL ini mempunyai pemrosesan data berorientasi *batch* berskala besar, seperti penyortiran, penguraian, dan konversi. Contoh dari *column family database* ialah Cassandra, SimpleDB (Amazon), dan Bigtable (Google), dan Hypertable.

Basis data grafik (*graph databases*) merubah tabel relasional dengan grafik relasional terstruktur dari pasangan kunci dan nilai yang saling berkaitan. *Graph databases* menjadi satu-satunya dari klasifikasi NoSQL lainnya yang berfokus pada relasi dan gambaran visual informasi yang ramah manusia. Contoh dari basis data grafik adalah Sones GraphDB, AllegroGraph, Infinite Graph, InfoGrid, dan Neo4j.

NoSQL dapat menangani data dalam jumlah besar, cukup cepat, dan seluruh *kluster* dari server komoditas berbagi sumber daya. Sebagian besar NoSQL adalah *open source* yang membantu perusahaan menyisihkan keuntungan harga daripada basis data komersial konvensional (Moniruzzaman & Hossain, 2013).

### 2.2.11 MongoDB

MongoDB merupakan basis data penyimpanan *document* (dokumen) yang menaruh data semi-terstruktur dan ditulis dalam bahasa JSON atau serupanya dengan format BSON (Binary JSON). MongoDB memberikan kestabilan yang tangguh dikarenakan operasi pembacaan akan

memberikan hasil yang diperbarui setelah penulisan berhasil dilakukan kepada memori (Srivastava dkk., 2015).

MongoDB bergerak berlandaskan konsep koleksi (*collection*) dan dokumen (*document*). Basis data (*database*) menjadi tempat untuk mengumpulkan *collection*. *Collection* ialah tempat pengumpulan *document* yang mirip dengan tabel di RDBMS (*Relational Database Management System*). Sedangkan *document* ialah senjumlah pasangan kunci dan nilai yang berisi kolom yang berbeda dan menampung berbagai jenis data.

Hubungan antara tabel-tabel yang terbentuk dalam basis data relasional tidak terjadi pada MongoDB. Kelebihan yang ditawarkan oleh MongoDB ialah struktur satu objek secara jelas, menggunakan memori internal untuk menyimpan set kerja (*windowed*) untuk pengaksesan data dengan lebih cepat, dan satu *collection* dapat mengakomodasi beragam *document* dengan jumlah bidang, konten, dan ukuran *document* yang berbeda antar masing-masing *document* (Chauhan, 2019).

Pemodelan data (*data modeling*) mengarah pada wadah data dalam basis data dan tautan antara entitas terkait. Data dalam MongoDB mempunyai dokumen dalam satu koleksi yang dapat memiliki kumpulan bidang dan tipe data yang berbeda-beda. Struktur dokumen dan cara sistem menggunakan data dari entitas terkait mesti dipertimbangkan saat membentuk model data. Penautan data terkait dapat dilakukan dengan *embedded data* dan *references*.

*Embedded data* menyimpan data terkait dalam satu struktur dokumen. Sebuah dokumen dapat berisi larik (*array*) dan sub-dokumen dengan data terkait. Adapun contoh pada *embedded data* seperti berikut.

**Tabel 2.3** *Embedded Data* pada MongoDB

<b><i>Embedded Data</i> pada MongoDB</b>
<pre>{   _id: &lt;ObjectId&gt;,   username: "123xyz",   contact: {     phone: "123-456-7890",     email: "xyz@example.com"</pre>



```

    },
    access: {
      level: 5,
      group: "dev"
    }
  }
}

```

Sumber: MongoDB, Inc., 2025

*References* menyimpan hubungan antara data dengan melibatkan tautan dari satu dokumen ke dokumen lainnya. Sistem dapat menuntaskan *references* ini untuk mengakses data terkait. contoh pada *references* adalah sebagai berikut (MongoDB Inc., 2025a).

**Tabel 2.4** *References* pada MongoDB

<b><i>References</i> pada MongoDB</b>
<p>user document</p> <pre> {   _id: &lt;ObjectId1&gt;,   username: "123xyz" } </pre> <p>contact document</p> <pre> {   _id: &lt;ObjectId2&gt;,   user_id: &lt;ObjectId1&gt;,   phone: "123-456-7890",   email: "xyz@example.com" } </pre> <p>access document</p> <pre> {   _id: &lt;ObjectId3&gt;,   user_id: &lt;ObjectId1&gt;, </pre>

```

    level: 5,
    group: "dev"
  }

```

Sumber: MongoDB, Inc., 2025

Model data MongoDB memberikan beberapa pilihan untuk memetakan hubungan antara berbagai entitas. Contoh hubungan data yang dapat diterapkan ialah model hubungan *One-to-One*, *One-to-Many*, atau *Many-to-Many* dengan *embedded data* dan *One-to-Many* dengan *references* (MongoDB Inc., 2025b).

### 2.2.12 Streamlit

Streamlit merupakan kerangka kerja (*framework*) publik yang menyediakan *deployment* sains data dan pembelajaran mesin berbasis web dengan penggunaan bahasa pemrograman Python. Streamlit dapat menampilkan informasi interaktif, memanfaatkan berbagai pilihan tema antarmuka, dan mendukung berbagai pustaka Python untuk pengolahan dan analisis data. *Deployment* Streamlit fleksibel terhadap beberapa platform *cloud*, seperti Amazon Web Services, Google Cloud, dan Streamlit Cloud. Streamlit Cloud memudahkan pengguna dalam menampilkan aplikasi daring tanpa pengelolaan lebih lanjut pada infrastruktur (*hosting*) (Surya & Aminuddin, 2024).

Pengambilan keputusan berlandaskan data dan pelaksanaan proses lebih efisien yang dilakukan perusahaan tidak luput dari peranan dan wawasan ilmuwan data. Kadangkala kehadiran aplikasi web melalui kerangka kerja Flask atau Django dengan platform *cloud* AWS atau lainnya mengalami keterlambatan masukan pengguna dan penyampaian informasi pada proses pengambilan keputusan terasa belum maksimal. Kehadiran Streamlit menjadikan salah satu solusi dalam mempresentasikan informasi yang berfokus pada data dengan cepat dan praktis (Richards, 2021). Beberapa perbedaan kerangka kerja berbasis web menggunakan bahasa pemrograman Python adalah sebagai berikut (Akkem dkk., 2023).

**Tabel 2.5** Perbedaan Beberapa *Framework* Web Python

<b>Fitur</b>	<b>Streamlit</b>	<b>Flask</b>	<b>Django</b>
Tujuan	Aplikasi web berfokus data dengan kode minimalis.	Aplikasi web ringan.	Kerangka kerja pengembangan web penuh.
Kemudahan penggunaan	Sangat mudah dan perancangan yang sederhana.	Sedang dan membutuhkan lebih banyak pengaturan manual.	Sedang sampai tinggi dengan meliputi banyak fitur bawaan.
Kecepatan pengembangan	Cepat.	Cepat.	Sedang sampai cepat.
Kasus penggunaan	Aplikasi berbasis data dan <i>dashboard</i> .	Aplikasi web ringan dan API.	Pengembangan web penuh dan aplikasi saling terkait.
Komponen bawaan	Himpunan komponen visualisasi data yang lengkap.	Minimalis dan dirancang untuk kemudahan perluasan.	Mencakup banyak komponen dan fitur bawaan.

Sumber: Akkem dkk, 2023

### 2.2.13 *System Usability Scale*

Skala Kegunaan Sistem (*System Usability Scale*) merupakan kuesioner yang digunakan untuk menilai kegunaan yang dirasakan. *System Usability Scale* (SUS) mempunyai 10 buah pernyataan yang disusun bergantian antara pernyataan positif dan negatif.

**Tabel 2.6** Pernyataan *System Usability Scale*

	1	2	3	4	5
1. I think that I would like to use this system frequently.					
2. I found the system unnecessarily complex.					
3. I thought the system was easy to use.					
4. I think that I would need the support of a technical person to be able to use this system.					
5. I found the various functions in this system were well integrated.					
6. I thought there was too much inconsistency in this system.					
7. I would imagine that most people would learn to use this system very quickly.					
8. I found the system very cumbersome to use.					
9. I felt very confident using the system					
10. I needed to learn a lot of things before I could get going with this system.					

Sumber: Brooke, 1996

Nilai yang dimasukkan pada SUS mempunyai langkah awal untuk mengubah skor mentah menjadi skor yang diselaraskan dalam kisaran dari 0 (peringkat terburuk) sampai 4 (peringkat terbaik) dengan penyesuaian yang berbeda untuk pernyataan bernomor ganjil dan genap. Semua pernyataan

membutuhkan peringkat dan jikalau responden memberikan pernyataan kosong maka mesti diberi skor mentah 3.

Pernyataan yang bernomor ganjil akan dikurangi 1 dari skor mentah dan bernomor genap akan dikurangi skor mentah dari 5. Jumlah skor yang disesuaikan dapat dikalikan dengan 2,5 untuk mendapatkan skor SUS. Persamaan skor SUS adalah sebagai berikut (J. R. Lewis, 2018).

$$SUM = 2,5(20 + SUM(SUS01, SUS03, SUS05, SUS07, SUS09) - SUM(SUS02, SUS04, SUS06, SUS08, SUS10)) \quad (7)$$

Sistem dapat dikatakan baik jikalau mempunyai skor SUS di atas 70 dengan sistem yang lebih baik mendapat skor di kisaran 70-an sampai 80-an. Sistem yang benar-benar unggul mendapat skor lebih dari 90. Sistem dengan skor kurang dari 50 mesti mendapatkan perhatian yang lebih dan dinilai tidak dapat diterima. Sedangkan sistem dengan skor kurang dari 70 mesti dilakukan pengawasan dan perbaikan berkelanjutan. Adapun penilaian kata sifat pada skor SUS adalah sebagai berikut (Bangor dkk., 2008).

**Tabel 2.7** Penilaian Kata Sifat pada Skor SUS

Penilaian Kata Sifat	Skor SUS
Best imaginable	100
Excellent	85,58
Good	72,75
OK	52,01
Poor	39,17
Awful	not applicable
Worst imaginable	25,00

Sumber: Bangor, 1996

## **BAB III**

### **METODE PENELITIAN**

#### **3.1 Pengembangan Sistem *Waterfall***

Bagian ini membahas tahap-tahap penyusunan yang menggambarkan pengembangan sistem analisis prediksi harga saham Hapag-Lloyd AG menggunakan LSTM. Fase-fase pengembangan sistem *waterfall* memberikan jabaran aktivitas mengenai konsep dan penerapan dalam pengembangan sistem tersebut.

##### **3.1.1 *Planning***

Setiap saham yang terdaftar di bursa memiliki karakteristik yang berbeda-beda. Adapun beberapa saham yang sangat dihargai tinggi, tetapi di sisi lain terdapat saham yang dihargai lebih rendah. Harga saham yang ditransaksikan mengalami perubahan yang signifikan seiring dengan aktifnya perdagangan saham tersebut. Investor mengalami kesulitan dalam memprediksi harga saham secara akurat.

Ketersediaan banyaknya data historis yang tersedia dan mudah diakses membuat investor berpikir tentang kegunaan dan kebermanfaatannya di masa mendatang. Sebagian pemegang modal tersebut juga mengabaikan pentingnya melihat data historis yang tercatat pada suatu saham. Masih sedikit sistem yang mudah digunakan untuk memprediksi harga saham menggunakan data historis yang cermat dan terkini.

Kehadiran sistem dalam menganalisis prediksi harga saham semakin dibutuhkan. Eksplorasi data historis harga saham mesti dilakukan oleh investor dengan pemanfaatan algoritma lebih lanjut. Pengembangan sistem ini mempunyai tujuan untuk membangun aplikasi dalam memprediksi harga saham secara langsung menggunakan metode LSTM dengan antarmuka Streamlit dan basis data MongoDB yang dapat dijangkau internet.

Prediksi harga saham menggunakan metode LSTM mementingkan banyak aspek, seperti akurasi prediksi, keterbatasan data historis, performa LSTM, dan kapasitas penyimpanan MongoDB. Sistem ini juga berketergantungan kepada *framework* Streamlit terhadap keterbatasan pada proyek kecil sampai menengah dan khususnya untuk aplikasi data.

Pengembangan menggunakan perangkat lunak yang umum, seperti bahasa pemrograman Python, pustaka Python yfinance, algoritma LSTM, kerangka kerja Streamlit, dan penyimpanan basis data MongoDB. Perangkat lunak tersebut memungkinkan pengeluaran biaya yang rendah dikarenakan sifatnya terbuka dan *deployment* Streamlit Cloud yang tidak berbayar.

Pembentukan aplikasi ini mengutamakan peningkatan keterjangkauan investor pemula dalam merasakan analisis prediksi harga saham yang sederhana dan lengkap. Dataset yang diambil dari Yahoo Finance bersifat legal dan dapat digunakan untuk riset/penelitian yang akan dilakukan penulis. Aplikasi akan mudah digunakan dengan antarmuka pengguna yang jelas dan ringan serta waktu pengembangannya diperkirakan terjadi selama satu bulan.

### 3.1.2 Analysis

Komunikasi kepada pengguna akhir sangat dibutuhkan agar tujuan dan manfaat sistem dapat terlaksana sesuai rencana pengembangan. Penggunaan bahasa mesti diperhatikan oleh pengembang dengan menerjemahkan konteks informatika supaya dimengerti oleh pengguna akhir.

Pembangunan aplikasi membutuhkan pengembang untuk menulis program menggunakan bahasa pemrograman Python dan merancang tampilan antarmuka yang akan diterapkan. Pembentukan aplikasi ini juga membutuhkan peranan dari investor sebagai pengguna akhir dalam mengevaluasi sistem yang telah dikembangkan.

Pengembang akan melakukan tinjauan terhadap dokumentasi pustaka Python yfinance yang menjadi modal dasar dalam pengumpulan dataset harga saham. Observasi alat-alat dalam melakukan prediksi dilakukan oleh pengembang dalam membentuk analisis yang diharapkan.

Performa pada sistem ini mempunyai dua kriteria yang salah satunya ialah nilai kinerja metrik evaluasi MAPE di bawah 10% dengan artian bahwa prediksi tersebut mempunyai akurasi yang tinggi dan tingkat galat yang rendah. Sistem juga memiliki kriteria lain dengan nilai *System Usability Scale* (SUS) di atas 70 yang mengartikan penilaian dengan hasil baik dan dapat diterima oleh pengguna akhir.

### 3.1.3 Design

Aplikasi analisis prediksi harga saham Hapag-Lloyd AG menggunakan metode LSTM dimulai dengan pengumpulan dataset harga saham yang berasal dari Yahoo Finance. Sistem akan memproses data dengan melakukan *data cleaning*, *normalization*, *data windowing*, dan *data splitting* supaya siap digunakan pada model LSTM.

Parameter model LSTM akan membentuk hasil prediksi harga saham dan menjadi modal pada *forecasting* harga saham mendatang. Antarmuka sistem dibangun dengan *framework* Streamlit yang dapat memberikan tampilan secara lugas dan jelas. Alur kerja sistem ini akan menggunakan *Data Flow Diagram* (DFD) dengan penyimpanan dan pembacaan hasil prediksi akan menggunakan MongoDB serta sistem akan di-*deploy* menggunakan Streamlit Cloud.

Pemodelan pada analisis prediksi dalam sistem ini mempunyai alternatif algoritma yang sudah umum dikenal, yaitu *Autoregressive Integrated Moving Average* (ARIMA). Akan tetapi, algoritma yang akan dipilih untuk pemodelan ialah LSTM dikarenakan beberapa penelitian terdahulu yang dilakukan (Lahboub & Benali, 2024) serta (Zhang, 2023) dengan menyatakan bahwa model LSTM dapat memberikan akurasi dan performa yang tinggi daripada ARIMA.

Setelah mengevaluasi konfigurasi yang tersedia maka penulis memutuskan untuk menggunakan pemodelan dengan algoritma LSTM, kerangka kerja Streamlit berbasis bahasa pemrograman Python, pengelolaan basis data MongoDB, dan deployment melalui Streamlit Cloud.



### 3.1.4 Implementation

Pengembangan sistem direncanakan akan memakan waktu selama satu bulan sesuai dengan perencanaan awal yang telah dibuat. Adapun rincian dari waktu pengembangan sistem tersebut, seperti satu minggu penyusunan tampilan pada sisi pengguna, satu minggu untuk pemodelan LSTM dan *server side*, satu minggu migrasi *local file-based database* kepada MongoDB, dan satu minggu untuk *deployment*, pengujian, dan pemeliharaan sistem.

Pengembang berkomunikasi kepada pengguna bahwa pembangunan sistem dilakukan dengan penyampaian analisis prediksi yang akan dimengerti oleh pengguna secara umum. Pengguna akan memberikan evaluasi sistem pada sistem yang nantinya telah di-*deploy* dan tersedia secara daring.

Lingkungan pembentukan sistem menggunakan bahasa pemrograman Python dengan pustaka-pustaka Python yang dapat disimpan dalam sebuah *virtual environment*. Jenis prosesor yang digunakan pada perangkat keras penulis adalah *Central Processing Unit* (CPU) dan sarana *deployment* yang akan dipakai bernama Streamlit Cloud.

Penggunaan algoritma LSTM berasal dari pustaka Python yang bernama Keras. Perancangan antarmuka aplikasi akan menggunakan alat yang bernama Figma dengan desain *low-fidelity wireframe*. *Low-fidelity wireframe* menggambarkan ide tampilan awal dengan tata letak dan elemen polos. Pengembang mendapatkan inspirasi tampilan aplikasi final dan mengadaptasi perubahan pada saat penggunaannya (Feng dkk., 2023).

Perwujudan antarmuka pengguna tersebut menggunakan kerangka kerja Streamlit yang fokus aplikasi berbasis data dan kode yang minimalis. Penyimpanan data pada sistem analisis prediksi harga saham ini berlandaskan *cloud storage* MongoDB yang bernama MongoDB Atlas.

Koleksi yang terdapat pada basis data terdiri atas *users* (daftar pengguna sistem), *model* (daftar model LSTM), *model\_metadata* (daftar metadata pada model LSTM), *registration\_access* (pengaturan akses registrasi), *password\_reset\_requests* (daftar permintaan reset kata sandi), dan *delete\_requests* (daftar permintaan penghapusan akun) Fasilitas fisik tidak diperlukan pada sistem ini dikarenakan semua komponen berbasis *cloud*.

Pengguna aplikasi akan diberikan informasi dalam setiap proses penambahan data yang dilakukan. Keberadaan validasi pada saat registrasi, masuk, dan melaksanakan analisis prediksi menjadi sangat penting untuk mengoperasikan sistem ini.

Sistem di-*deploy* pada Streamlit Cloud dan dilakukan *black box testing* supaya memastikan sistem berjalan dengan baik. Apabila pengujian memberikan hasil yang telah ditentukan maka sistem akan dipastikan untuk menjadi aktif dan terhubung dengan jaringan publik (internet).

*Black box testing* merupakan pengujian yang dapat dilakukan tanpa mempunyai pengetahuan terkait cara kerja internal aplikasi. *Black box testing* mempunyai nama lain, seperti *functional testing*, *close box testing*, dan *behavioural testing*. Pengujian ini hanya memeriksa aspek fundamental sistem dan tidak akan mempunyai akses ke kode sumber (Khan & Khan, 2012).

### 3.1.5 Use

Penggunaan aplikasi analisis prediksi harga saham Hapag-Lloyd AG menggunakan metode LSTM dilaksanakan secara daring di mana pengguna dapat mengakses aplikasi tersebut melalui web atau gawai. Hal tersebut dapat terjadi karena kerangka kerja Streamlit sudah mendukung tampilan aplikasi dengan *responsive*. Pengguna mesti memasukkan parameter pada pengaturan model, seperti *ticker*, tanggal awal dan akhir, serta frekuensi data agar aplikasi memberikan analisis, visualisasi, dan prediksi.

Evaluasi sistem akan diberikan kepada pengguna berlandaskan ruang lingkup mahasiswa/i Universitas Widyatama sebagai populasinya dan pemilihan sampel dipilih kepada mahasiswa/i program studi Informatika yang mempunyai minat atau ketertarikan pada dunia investasi. Kinerja sistem akan diperiksa menggunakan *System Usability Scale* (SUS) dan hasil evaluasi tersebut dapat digunakan untuk keputusan rekayasa ulang.

Pemeliharaan aplikasi dapat dilaksanakan dengan memperbarui model LSTM bilamana akurasi model tidak sesuai yang diharapkan dan model yang telah diperbarui dapat terlihat di halaman depan aplikasi tersebut. Perbaikan

diperlukan salah satunya apabila terdapat *bug* dalam antarmuka pengguna Streamlit tersebut.

### **3.2 *Cross-Industry Standard Process for Data Mining***

Bagian ini membahas tata cara penambangan data yang dilakukan pada topik skripsi ini. Rangkaian prosedur CRISP-DM akan dijelaskan guna mengetahui pekerjaan yang dilakukan di setiap langkahnya. Penjabaran prosedur ini memberikan suatu gambaran pada pengujian algoritma yang digunakan untuk menunjang analisis prediksi harga saham dengan menggunakan metode LSTM.

#### **3.2.1 *Business Understanding***

Dalam dunia investasi dan pasar modal, kemampuan untuk memprediksi harga saham secara akurat sangat penting bagi pengambilan keputusan yang tepat dan strategis. Oleh karena itu, dibutuhkan metode yang mampu menangkap pola waktu (*time series*) secara efektif. Long Short-Term Memory (LSTM) dirancang untuk mengenali pola dalam data berurutan dan memiliki keunggulan dalam mengatasi masalah *long-term dependencies*.

Kehadiran aplikasi analisis prediksi harga saham Hapag-Lloyd AG ditujukan untuk membantu investor dalam memprediksi harga saham menggunakan model *deep learning* berbasis LSTM. Analisis prediksi yang dilaksanakan ini mampu memberikan informasi pembelajaran mesin pada data historis harga saham dan memperoleh ramalan mendatang yang dapat dijadikan mendukung perencanaan dan manajemen risiko investasi.

#### **3.2.2 *Data Understanding***

Pengumpulan data dalam penelitian ini merupakan data harga saham Hapag-Lloyd AG yang dapat diakses pada *website* Yahoo Finance pada tautan <https://finance.yahoo.com/quote/HLAG.DE/history/>. Data yang diambil dari periode tanggal 6 November 2015 sampai dengan 31 Mei 2025 dengan frekuensi data menyesuaikan masukan pengguna yang terdiri dari harian, mingguan, atau bulanan. Penggunaan variabel dalam penelitian ini adalah variabel *Date* sebagai catatan tanggal perdagangan dan *Close* yang merupakan harga penutupan saham tersebut.

### 3.2.3 Data Preparation

Tahap awal yang dilakukan pada *data preparation* ialah transformasi data dengan mengubah atribut *multi-index* menjadi *single index* dan menambahkan *feature* Date yang berasal dari *index* dataset. Dataset akan difilter dengan mengambil *feature* Date dan *Close*. Selanjutnya akan dilakukan tahap *data cleaning* dengan memeriksa nilai yang hilang dan negatif pada *feature* *Close*.

*Exploratory Data Analysis* (EDA) dilaksanakan untuk mengetahui ringkasan statistik *feature* *Close*, mengetahui visualisasi distribusi data per tahun, dan grafik tren harga saham sepanjang waktu. Untuk mempersiapkan data yang mudah diolah saat pemodelan maka dijalankan normalisasi dengan membentuk data ke dalam rentang tertentu menggunakan *MinMaxScaler*. Persamaan untuk *MinMaxScaler* adalah sebagai berikut.

$$\text{minmax} = \frac{(X - X_{\min})}{(X_{\max} - X_{\min})} \quad (8)$$

Adapun  $X$  adalah data yang diuji dan  $X_{\min}$  serta  $X_{\max}$  adalah nilai atribut minimum dan maksimum dari data yang diuji (Pranolo dkk., 2024). Setelah dinormalisasikan maka akan dilanjutkan pada tahap data windowing dengan menyusun data menjadi *sequence* masukan ( $X$ ) dan target keluran ( $y$ ) berdasarkan *time step* (jumlah langkah waktu data poin ke belakang). Terakhir data akan dibagi sebesar 80% data latih (*training*) dan 20% data uji (*testing*).

### 3.3.4 Modeling

Model yang akan digunakan ialah LSTM dengan melatih data *training* berdasarkan parameter yang sudah ditentukan dan menampilkan visualisasi grafik *training loss* dan *validation loss* serta penggunaan *early stopping* untuk menghindari *overfitting*. Kemudian aplikasi menampilkan prediksi dari pemodelan dan menyimpan model dan metadatanya.

Pengguna dapat melakukan *model tuning* dengan kombinasi parameter *time step*, *epoch* (jumlah waktu dataset untuk melatih model), dan *batch size* (jumlah data yang diproses sekaligus saat pelatihan). Metode *hyperparameter tuning* yang dipilih adalah *grid search* dengan mempelajari semua

kemungkinan kombinasi parameter yang telah ditentukan agar dapat memberikan *training* data secara menyeluruh. *Pseudocode* yang dapat dijalankan dalam metode *grid search* ialah sebagai berikut (Pranolo dkk., 2024).

**Tabel 3.1** Metode *Grid Search*

<b><i>Hyperparameter tuning menggunakan grid search</i></b>
<pre> Function grid_search_LSTM(parameters_range):     best_params = None     best_score = +infinity     for each parameter combination in parameters_range:         model = create_LSTM_model(parameter_combination)         score = evaluate_model(model)         if score &lt; best_score:             best_score = score             best_params = parameter_combination     return best_params </pre>

Sumber: Pranolo dkk, 2024

Setelah melakukan *model tuning* maka akan ditampilkan visualisasi grafik *training loss* dan *validation loss* dari *model tuning* terbaik. Prediksi dari pemodelan *hyperparameter tuning* terbaik beserta penyimpanan model dan metadatanya akan dikemukakan selanjutnya.

### 3.3.5 Evaluation

Prediksi telah dilakukan oleh model LSTM dalam pengomparasian nilai prediksi dan nilai aktual dengan menggunakan data *testing*. Maka dari itu terdapat metrik evaluasi untuk menghitung dan menampilkan performa prediksi tersebut menggunakan RMSE (*Root Mean Squared Error*) dan MAPE (*Mean Absolute Percentage Error*). Persamaan untuk RMSE dan MAPE adalah sebagai berikut.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (F_i - A_i)^2} \quad (9)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|A_i - F_i|}{A_i} \quad (10)$$

RMSE dipandang bermakna dalam mendeteksi *outlier* yang menjelaskan kerentanan model terhadap titik data ekstrem dan MAPE memberikan wawasan mengenai kemampuan prediksinya yang menjadi indikator penting dari akurasi model. Adapun  $F_i$  adalah nilai prediksi,  $A_i$  merupakan nilai aktual, dan  $n$  ialah jumlah prediksi (Pranolo dkk., 2024). Metrik evaluasi MAPE mempunyai pemahaman dalam susunan rentang nilai sebagai berikut (C. D. Lewis, 1982).

**Tabel 3.2** Pemahaman Rentang Nilai MAPE

MAPE (%)	Rentang Nilai
< 10	Highly accurate forecasting
10 – 20	Good forecasting
20 – 50	Reasonable forecasting
> 50	Inaccurate forecasting

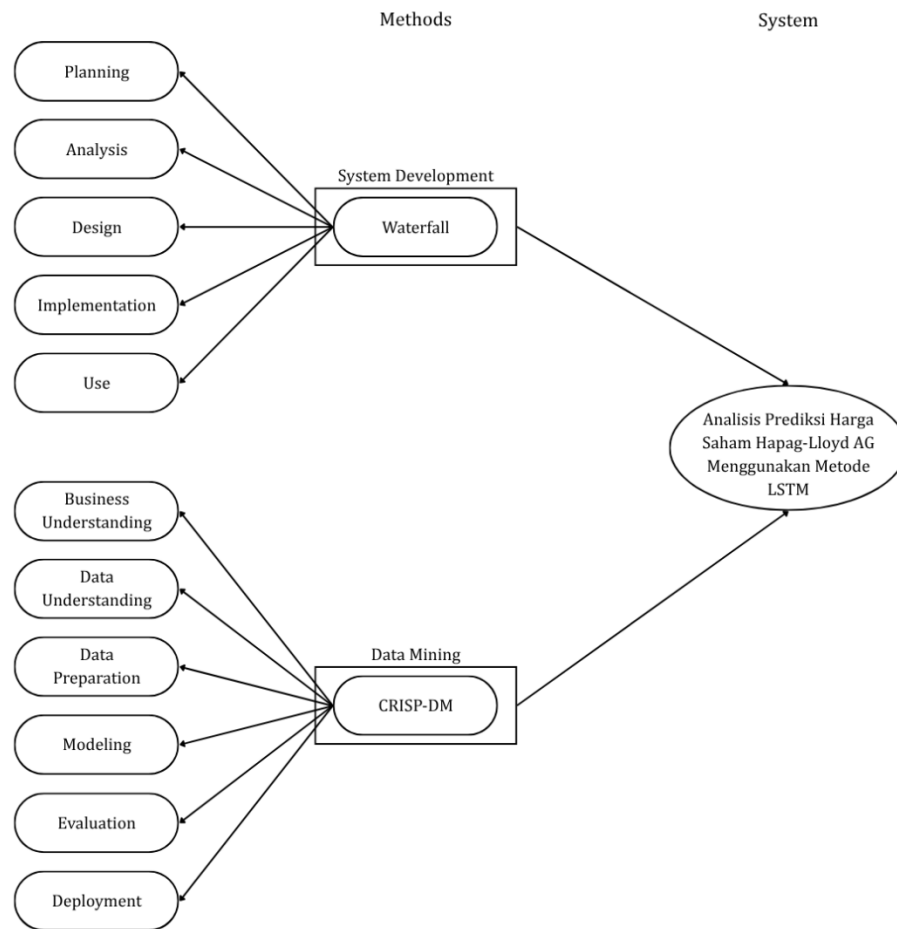
Sumber: Lewis, 1982

### 3.3.6 Deployment

*Deployment* analisis prediksi harga saham Hapag-Lloyd AG menggunakan metode LSTM akan menggunakan platform Streamlit Cloud dengan antarmuka aplikasi Streamlit. Pemodelan dapat dilakukan dengan memasukkan parameter, seperti *ticker*, tanggal awal dan akhir, serta frekuensi data supaya aplikasi memberikan analisis dan hasil prediksi. Aplikasi ini juga disertakan autentikasi pengguna yang dapat mengakses model dan metadata yang telah disimpan dengan lebih rinci.

### 3.3 Kerangka Penelitian

Metode pengembangan sistem *waterfall* dan metode penambangan data CRISP-DM telah dikemukakan sebelumnya. Maka dari itu, kerangka penelitian utuh diperlukan agar memudahkan pemahaman prosedur penelitian yang dilakukan oleh penulis. Kerangka penelitian analisis prediksi harga saham Hapag-Lloyd AG menggunakan metode LSTM adalah sebagai berikut.



**Gambar 3.1** Kerangka Penelitian  
(Sumber: Penulis, 2025)

## **BAB IV**

### **HASIL PENELITIAN DAN PEMBAHASAN**

#### **4.1 Analisis Kebutuhan**

##### **4.1.1 Kebutuhan Fungsional**

##### **4.1.2 Kebutuhan Sistem**

##### **4.1.3 Kebutuhan Non-Fungsional**

#### **4.2 Perancangan Sistem**

##### **4.2.1 Perancangan Perangkat Lunak**

##### **4.2.2 Perancangan Basis Data**

##### **4.2.3 Perancangan Antarmuka**

Desain antarmuka yang dirancang dalam pengembangan analisis prediksi harga saham Hapag-Lloyd AG menggunakan metode LSTM adalah sebagai berikut.



<p>Home Login Register Dashboard</p> <hr/> <p><b>Pengaturan Model</b></p> <p>Masukkan Ticker Saham <input type="text"/></p> <p>Tanggal Awal <input type="text"/></p> <p>Tanggal Akhir <input type="text"/></p> <p>Frekuensi Data <input type="text"/></p> <p>Metode Pemodelan  <input type="radio"/> Latih model baru  <input type="radio"/> Gunakan model dari database  <input type="checkbox"/> Aktifkan Model Tuning         </p> <p><input type="button" value="Mulai Prediksi"/></p>	<p>Analisis Prediksi Harga Saham Menggunakan Metode LSTM</p> <hr/> <p>Silakan isi parameter di sidebar, lalu tekan 'Mulai Prediksi'</p> <p>Cari Ticker Saham <input type="text"/></p> <div style="border: 1px solid black; padding: 5px; margin-top: 10px;"> <p>Daftar model dari database</p> <div style="height: 100px;"></div> </div>
--	--

(Sumber: Penulis, 2025)

Halaman *Home* berada di menu *Home* dengan menampilkan informasi daftar model dari basis data yang tersedia dan pesan menentukan pengaturan model untuk melakukan prediksi. Di *sidebar* terdapat empat menu, seperti *Home*, *Login*, *Register*, dan *Dashboard* di mana menu-menu tersebut akan selalu ada di antarmuka lainnya.

Di bawah keempat menu terdapat pengaturan model sebagai tempat pengguna memasukkan parameter yang dibutuhkan yang dimulai dari *ticker* saham, tanggal awal dan akhir, frekuensi data (harian, mingguan, atau bulanan), metode pemodelan, dan pilihan *tuning* model. Setelah memasukkan parameternya maka pengguna dapat menekan tombol “Mulai Prediksi” untuk menjalankan analisis prediksi harga sahamnya.

<p>Home Login Register Dashboard</p> <hr/> <p><b>Pengaturan Model</b></p> <p>Masukkan Ticker Saham <input type="text"/></p> <p>Tanggal Awal <input type="text"/></p> <p>Tanggal Akhir <input type="text"/></p> <p>Frekuensi Data <input type="text"/></p> <p>Metode Pemodelan  <input type="radio"/> Latih model baru  <input type="radio"/> Gunakan model dari database  <input type="checkbox"/> Aktifkan Model Tuning         </p> <p><input type="button" value="Mulai Prediksi"/></p>	<p>Analisis Prediksi Harga Saham Menggunakan Metode LSTM</p> <hr/> <p>1. Business Understanding</p> <p>_____</p> <p>_____</p> <p>_____</p> <p>2. Data Understanding</p> <div style="border: 1px solid black; padding: 5px; margin-top: 5px;"> <p style="text-align: center;">Dataset mentah</p> <div style="height: 150px; border: 1px solid black;"></div> </div>
--	--

(Sumber: Penulis, 2025)

Pada saat pengguna memasukkan parameter model dan menekan tombol “Mulai Prediksi” maka layar sistem akan menampilkan *business and data understanding*. *Business understanding* memberitahukan tujuan melakukan analisis prediksi menggunakan metode LSTM dan *data understanding* menampilkan dataset mentah dari *ticker* saham yang dimulai dari tanggal awal dan akhir yang dipilih serta frekuensi data sebagai interval setiap data poin.

<p>Home Login Register Dashboard</p> <hr/> <p><b>Pengaturan Model</b></p> <p>Masukkan Ticker Saham <input type="text"/></p> <p>Tanggal Awal <input type="text"/></p> <p>Tanggal Akhir <input type="text"/></p> <p>Frekuensi Data <input type="text"/></p> <p>Metode Pemodelan  <input type="radio"/> Latih model baru  <input type="radio"/> Gunakan model dari database  <input type="checkbox"/> Aktifkan Model Tuning         </p> <p><input type="button" value="Mulai Prediksi"/></p>	<p>Analisis Prediksi Harga Saham Menggunakan Metode LSTM</p> <hr/> <p>3. Data Preparation</p> <p>3.1 Data Transformation</p> <p>3.2 Data Filtering</p> <p>3.3 Data Cleaning</p> <p>3.4 Exploratory Data Analysis (EDA)</p> <p>3.5 Normalization</p> <p>3.6 Data Windowing</p> <p>3.7 Data Splitting</p>
--	---

(Sumber: Penulis, 2025)

*Data preparation* ditampilkan sebagai tahap eksplorasi dan analisis dataset yang dimulai dari transformasi data untuk mempersiapkan data yang difilter. *Data filtering* dilakukan untuk memilih *feature* data yang akan dianalisis dan *data cleaning* sebagai penanganan data apabila terdapat nilai yang hilang atau nilai negatif.

*Exploratory Data Analysis* (EDA) dikerjakan untuk mengetahui gambaran isi data dengan menampilkan visualisasi sederhana dan normalisasi untuk menyederhanakan nilai data agar dapat dimodelkan. *Data windowing* dilakukan untuk mempersiapkan data *time series* yang akan diberikan ke pemodelan LSTM dan *data splitting* untuk membagi data yang dijadikan *training* dan *testing*.

<ul style="list-style-type: none"> <li>Home</li> <li>Login</li> <li>Register</li> <li>Dashboard</li> </ul> <hr/> <p><b>Pengaturan Model</b></p> <p>Masukkan Ticker Saham</p> <input type="text"/> <p>Tanggal Awal</p> <input type="text"/> <p>Tanggal Akhir</p> <input type="text"/> <p>Frekuensi Data</p> <input type="text"/> <p>Metode Pemodelan</p> <p> <input type="radio"/> Latih model baru  <input type="radio"/> Gunakan model dari database  <input type="checkbox"/> Aktifkan Model Tuning         </p> <p>Mulai Prediksi</p>	<p>Analisis Prediksi Harga Saham Menggunakan Metode LSTM</p> <hr/> <p>4. Modeling</p> <ul style="list-style-type: none"> <li>4.1 Model Selection</li> <li>4.2 Model Training</li> <li>4.3 Model Inference</li> <li>4.4 Model Tuning             <div> <input type="text" value="Apabila mengaktifkan model tuning"/> </div> </li> <li>4.5 Best Model Selection</li> <li>4.6 Best Model Training</li> <li>4.7 Best Model Inference</li> </ul>
--	--

(Sumber: Penulis, 2025)

Tahap *modeling* menampilkan *model selection* untuk memberikan gambaran konfigurasi dan pemodelan LSTM. *Model training* menampilkan visualiasi *training epoch* dan *validation loss*. Lalu, *model inference* memberikan hasil prediksi dan distribusi galat. Apabila pengguna

mengaktifkan *model tuning* maka akan terjadi *hyperparameter tuning* kemudian menampilkan *model selection*, *model training*, dan *model inference* dari *model tuning* terbaik.

<p>Home Login Register Dashboard</p> <hr/> <p><b>Pengaturan Model</b></p> <p>Masukkan Ticker Saham <input type="text"/></p> <p>Tanggal Awal <input type="text"/></p> <p>Tanggal Akhir <input type="text"/></p> <p>Frekuensi Data <input type="text"/></p> <p>Metode Pemodelan  <input type="radio"/> Latih model baru  <input type="radio"/> Gunakan model dari database  <input type="checkbox"/> Aktifkan Model Tuning         </p> <p><input type="button" value="Mulai Prediksi"/></p>	<p>Analisis Prediksi Harga Saham Menggunakan Metode LSTM</p> <hr/> <p>5. Evaluation</p> <p>5.1 Final Model Evaluation</p> <table border="1" style="width: 100%;"> <tr> <td style="text-align: center;">Evaluasi model terbaik</td> </tr> <tr> <td style="height: 30px;"></td> </tr> </table> <p>5.2 Forecasting</p> <table border="1" style="width: 100%;"> <tr> <td style="text-align: center; height: 50px;">Prediksi harga saham ke depan</td> </tr> </table>	Evaluasi model terbaik		Prediksi harga saham ke depan
Evaluasi model terbaik				
Prediksi harga saham ke depan				

(Sumber: Penulis, 2025)

Selanjutnya terdapat *Evaluation* yang berisi *final model evaluation* untuk menampilkan metrik evaluasi pada hasil prediksi harga saham terbaik. Apabila tidak mengaktifkan *model tuning* maka hanya menampilkan *model evaluation* pada model awal saja. Sedangkan apabila mengaktifkannya maka akan menampilkan *model evaluation* dari model awal dan model yang di-*tuning*. *Forecasting* diberikan sebagai prediksi harga saham di masa mendatang.

<a href="#">Home</a> <a href="#">Login</a> <a href="#">Register</a> <a href="#">Dashboard</a>	<h3>Register</h3> <hr/> <p>Nama Akun</p> <input type="text"/> <p>Kata Sandi</p> <input type="password"/> <p>Konfirmasi Kata Sandi</p> <input type="password"/> <p>Kode Admin (Opsional)</p> <input type="text"/>
--	--

(Sumber: Penulis, 2025)

Halaman *register* memberikan tampilan kepada pengguna untuk membuat akun dengan memasukkan nama akun, kata sandi, dan konfirmasi kata sandi. Apabila pengguna mempunyai kode khusus admin maka peran pengguna tersebut akan menjadi *admin*, sedangkan tanpa kode khusus tersebut maka peran pengguna akan menjadi *user*.

<a href="#">Home</a> <a href="#">Login</a> <a href="#">Register</a> <a href="#">Dashboard</a>	<h3>Login</h3> <hr/> <p>Nama Akun</p> <input type="text"/> <p>Kata Sandi</p> <input type="password"/> <p>Reset Kata Sandi</p> <ol style="list-style-type: none"> <li>1. Reset Kata Sandi Langsung <input type="button" value="Reset dengan mengetahui kata sandi lama"/> </li> <li>2. Ajukan Permintaan Reset Kata Sandi <input type="button" value="Reset dengan tanpa kata sandi lama"/> </li> </ol> <p>Hapus Akun</p> <ol style="list-style-type: none"> <li>1. Hapus Akun Langsung <input type="button" value="Hapus dengan mengetahui kata sandi"/> </li> <li>2. Ajukan Permintaan Penghapusan Akun Langsung <input type="button" value="Hapus dengan tanpa kata sandi"/> </li> </ol>
--	--

(Sumber: Penulis, 2025)

Halaman *login* membantu pengguna untuk dapat mengakses halaman *dashboard* dengan sebelumnya pengguna memasukkan nama akun dan kata sandi. Selain itu, terdapat bantuan kepada pengguna seandainya ingin melakukan reset kata sandi atau menghapus akun.

Pengguna dapat melakukan reset kata sandi atau hapus akun langsung bilamana mengingat kata sandi yang digunakan. Jikalau pengguna telah melupakan kata sandi akunnya, maka pengajuan reset kata sandi atau hapus akun dapat dilakukan dengan memasukkan nama akunnya saja, tetapi langkah pasti untuk reset kata sandi atau hapus akun akan dilakukan oleh admin.

<div> <a href="#">Home</a>  <a href="#">Login</a>  <a href="#">Register</a>  <a href="#">Dashboard</a> </div> <hr/> <div> <p><b>Autentikasi</b></p> <div> <input type="text" value="Nama akun: {nama} {{peran}}"/> </div> <div> <input type="button" value="Keluar"/> </div> </div>	<div> <p><b>Dashboard</b></p> <hr/> <p><b>Manajemen Model</b></p> <div> <input type="button" value="Daftar model"/> </div> </div> <div> <p><b>Manajemen Metadata Model</b></p> <div> <input type="button" value="Daftar metadata model"/> </div> </div>
---	---

(Sumber: Penulis, 2025)

Halaman *dashboard* menunjukkan daftar model beserta metadatanya. Pengguna dapat menekan tombol “Keluar” pada *sidebar* ketika sudah selesai menggunakan aplikasi.

<p>Home Login Register Dashboard</p> <hr/> <p><b>Autentikasi</b></p> <p>Nama akun: {nama} ({peran})</p> <p>Keluar</p> <p><b>Kontrol Registrasi Admin</b></p> <p>Nonaktifkan Registrasi</p> <p><b>Tambah Akun Baru</b></p> <p>Nama Akun Baru</p> <p>Peran Akun</p> <p>Tambah</p>	<p><b>Dashboard</b></p> <hr/> <p><b>Manajemen Model</b></p> <p>Daftar model</p> <hr/> <p><b>Manajemen Metadata Model</b></p> <p>Daftar metadata model</p> <hr/> <p><b>Manajemen Pengguna</b></p> <p>1. Reset Kata Sandi dan Hapus Akun Langsung</p> <p>2. Permintaan Reset Kata Sandi dari Pengguna</p> <p>3. Permintaan Penghapusan Akun dari Pengguna</p>
---	---

(Sumber: Penulis, 2025)

Pada saat di halaman *dashboard* maka pengguna dengan peran sebagai admin dapat mengontrol registrasi pengguna dan menambah akun baru yang berada di *sidebar*. Admin juga dapat menghapus data model dan metadata model serta melakukan reset kata sandi dan hapus akun langsung di halaman *dashboard*. Pengguna yang telah mengajukan reset kata sandi dan penghapusan akun dapat disetujui admin pada bagian manajemen pengguna.

### 4.3 Implementasi Sistem

Implementasi sistem menjelaskan berbagai macam implementasi yang dilakukan dalam sistem yang telah dikembangkan dan diuji. Tujuannya adalah untuk memahami implementasi yang cukup dalam pengembangan dan pengujian sistem serta memberikan informasi mengenai spesifikasi perangkat yang harus dimiliki agar implementasi dapat berjalan dengan lancar.

#### 4.3.1 Batasan Implementasi

Adapun batasan implementasi pada aplikasi analisis prediksi harga saham Hapag-Lloyd AG menggunakan metode LSTM ini adalah sebagai berikut.

1. Aplikasi dikembangkan dengan bahasa pemrograman Python dan kerangka kerja Streamlit.
2. Basis data yang digunakan ialah MongoDB.
3. Pengujian dilakukan menggunakan perangkat keras penulis.

#### 4.3.2 Implementasi Perangkat Keras

Perangkat keras yang digunakan dalam pengembangan analisis prediksi harga saham Hapag-Lloyd AG menggunakan metode LSTM adalah sebagai berikut.

<i>Processor</i>	Intel Celeron N4020 1.1 GHz
<i>Memory</i>	4 GB
<i>Storage</i>	256 GB

#### 4.3.3 Implementasi Perangkat Lunak

Perangkat lunak yang digunakan dalam pengembangan analisis prediksi harga saham Hapag-Lloyd AG menggunakan metode LSTM adalah sebagai berikut.

Sistem Operasi	Windows 11 64-bit
Bahasa Pemrograman	Python
<i>Deploy</i>	Streamlit Cloud
<i>Browser</i>	Google Chrome

#### 4.3.4 Implementasi Antarmuka

Antarmuka yang dikembangkan dalam pengembangan analisis prediksi harga saham Hapag-Lloyd AG menggunakan metode LSTM adalah sebagai berikut.

### 4.4 Pengujian Sistem

#### 4.4.1 Pengujian Algoritma



#### 4.4.2 Black Box Testing

#### 4.4.3 System Usability Scale

Skor Hasil Hitung											
Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10		
4	3	4	4	4	3	4	2	4	4	36	90
4	2	4	2	4	2	2	2	4	2	28	70
4	3	3	4	3	3	4	3	2	4	33	82,5
5	2	5	2	4	1	4	2	5	5	35	87,5
3	1	4	1	4	1	4	1	4	2	25	62,5
3	2	4	2	4	2	4	2	4	2	29	72,5
1	1	5	1	4	3	5	1	5	3	29	72,5
5	1	5	1	5	1	5	1	5	2	31	77,5
5	2	4	1	5	1	5	1	5	3	32	80
4	2	4	3	4	3	4	2	3	4	33	82,5
Skor Rata-Rata (Hasil Akhir)										77,75	

## **BAB V**

### **PENUTUP**

#### **5.1 Kesimpulan**

#### **5.2 Saran**

## DAFTAR PUSTAKA

- Aggarwal, K. K., & Singh, Y. (2005). *Software Engineering*. New Age International Publishers.
- Akkem, Y., Kumar, B. S., & Varanasi, A. (2023). Streamlit Application for Advanced Ensemble Learning Methods in Crop Recommendation Systems – A Review and Implementation. *Indian Journal of Science and Technology*, 16(48), 4688–4702. <https://doi.org/10.17485/IJST/v16i48.2850>
- Albeladi, K., Zafar, B., & Mueen, A. (2023). Time Series Forecasting using LSTM and ARIMA. *International Journal of Advanced Computer Science and Applications*, 14(1), 313–320. <https://doi.org/10.14569/IJACSA.2023.0140133>
- Anandita, A., & Wahyuningsih, T. (2024). Prediksi Indeks Saham Syariah Menggunakan Model Long Short-Term Memory (LSTM). *Jurnal Ilmu Manajemen*, 9(1), 60–69. <https://doi.org/10.33474/jimmu.v9i1.21547>
- Aroral, H. K. (2021). Waterfall Process Operations in the Fast-paced World: Project Management Exploratory Analysis. *International Journal of Applied Business and Management Studies*, 6(1), 91–99. [https://www.ijabms.com/wp-content/uploads/2021/05/05\\_ARORAL\\_PB.pdf](https://www.ijabms.com/wp-content/uploads/2021/05/05_ARORAL_PB.pdf)
- Bangor, A., Kortum, P. T., & Miller, J. T. (2008). An Empirical Evaluation of the System Usability Scale. *International Journal of Human-Computer Interaction*, 24(6), 574–594. <https://doi.org/10.1080/10447310802205776>
- Bathla, G. (2020). Stock Price Prediction Using LSTM and SVR. *2020 Sixth International Conference on Parallel, Distributed and Grid Computing (PDGC)*, 211–214. <https://doi.org/10.1109/PDGC50313.2020.9315800>
- Bramer, M. (2016). *Principles of Data Mining*. Springer Nature.
- Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., & Wirth, R. (2000). *CRISP-DM 1.0: Step-by-Step Data Mining Guide*. SPSS.
- Chatfield, C. (2000). *Time-Series Forecasting*. CRC Press LLC.
- Chauhan, A. (2019). A Review on Various Aspects of MongoDB Databases. *International Journal of Engineering Research & Technology*, 8(5), 90–92. <https://doi.org/10.17577/IJERTV8IS050031>
- Chen, M.-S. (1996). Data mining: An Overview from a Database Perspective. *Institute of Electrical and Electronics Engineers*, 8(6), 866–883. <https://doi.org/10.1109/69.553155>
- Enriko, I. K. A., Gustiyana, F. N., & Putra, R. H. (2023). Komparasi Hasil Optimasi Pada Prediksi Harga Saham PT. Telkom Indonesia Menggunakan Algoritma Long Short Term Memory. *Jurnal Media Informatika Budidarma*, 7(2), 659–667. <https://doi.org/10.30865/mib.v7i2.5822>

- Feng, S., Yuan, M., Chen, J., Xing, Z., & Chen, C. (2023). Designing with Language: Wireframing UI Design Intent with Generative Large Language Models. *arXiv*, 1–21. <https://doi.org/10.48550/arXiv.2312.07755>
- Gupta, G. K. (2014). *Introduction to Data Mining with Case Studies*. PHI Learning Private Limited.
- Handharu Sworo, T., & Hermawan, A. (2024). Analysis and Prediction of Indonesia Stock Exchange (IDX) Stock Prices Using Long Short Term Memory (LSTM) Algorithm. *Journal of Computer Science and Technology Studies*, 6(2), 142–149. <https://doi.org/10.32996/jcsts.2024.6.2.16>
- Handika, B. D., & Sugianto. (2024). Perbandingan Prediksi Saham BTPN Dengan BBRI Menggunakan LSTM (Long Short Term-Memory). *Jurnal Ilmiah Teknologi Infomasi Dan Sains*, 4(2), 5–8. <https://doi.org/10.36815/submit.v4i2.3360>
- Hapag-Lloyd AG. (2025). *Our Company*. Hapag-Lloyd AG. <https://www.hapag-lloyd.com/en/company.html>
- Hin, L. T. (2008). *Panduan Berinvestasi Saham*. PT Elex Media Komputindo.
- Joseph, M. (2022). *Modern Time Series Forecasting with Python*. Packt Publishing.
- Khan, Mohd. E., & Khan, F. (2012). A Comparative Study of White Box, Black Box and Grey Box Testing Techniques. *International Journal of Advanced Computer Science and Applications*, 3(6), 12–15. <https://doi.org/10.14569/IJACSA.2012.030603>
- Kusuma, N. P. N. (2024). Prediksi Harga Saham Blue Chip Pada Indeks IDX30 Menggunakan Algoritma Recurrent Neural Network (RNN). *Jurnal Ekonomi & Bisnis*, 23(1), 90–97. <https://doi.org/10.32722/eb.v23i1.6609>
- Lahboub, K., & Benali, M. (2024). Assessing the Predictive Power of Transformers, ARIMA, and LSTM in Forecasting Stock Prices of Moroccan Credit Companies. *Journal of Risk and Financial Management*, 17(7), 293. <https://doi.org/10.3390/jrfm17070293>
- Lewis, C. D. (1982). *Industrial and Business Forecasting Methods: A Practical Guide to Exponential Smoothing and Curve Fitting*. Butterworth Scientific.
- Lewis, J. R. (2018). The System Usability Scale: Past, Present, and Future. *International Journal of Human-Computer Interaction*, 34(7), 577–590. <https://doi.org/10.1080/10447318.2018.1455307>
- Li, Q., & Chen, Y.-L. (2009). Data Flow Diagram. Dalam *Modeling and Analysis of Enterprise and Information Systems* (hlm. 85–97). Springer Berlin. [https://doi.org/10.1007/978-3-540-89556-5\\_4](https://doi.org/10.1007/978-3-540-89556-5_4)
- Li, Z., Yu, H., Xu, J., Liu, J., & Mo, Y. (2023). Stock Market Analysis and Prediction Using LSTM: A Case Study on Technology Stocks. *Innovations in Applied Engineering and Technology*, 2, 1–6. <https://doi.org/10.62836/iaet.v2i1.162>

- Mahbubi, F. A., Hermanto, T. I., & Lestari, C. D. (2025). Peramalan Penjualan Saham Nikel Menggunakan Algoritma Long Short Term Memory (LSTM). *Indonesia Journal Information System*, 8(1), 138–149. <https://doi.org/10.36080/idealis.v8i1.3254>
- McLeod, R. Jr., & Schell, G. (2007). *Management Information Systems*. Pearson Prentice Hall.
- MongoDB Inc. (2025a). *Data Modeling*. MongoDB, Inc. <https://www.mongodb.com/docs/manual/data-modeling/>
- MongoDB Inc. (2025b). *Document Relationships*. MongoDB, Inc. <https://www.mongodb.com/docs/manual/applications/data-models-relationships/>
- Moniruzzaman, A. B. M., & Hossain, S. A. (2013). NoSQL Database: New Era of Databases for Big data Analytics-Classification, Characteristics and Comparison. *International Journal of Database Theory and Application*, 6(4), 1–14. <https://doi.org/10.48550/arXiv.1307.0191>
- Montgomery, D. C., Jennings, C. L., & Kulahci, M. (2016). *Introduction to Time Series Analysis and Forecasting*. John Wiley & Sons, Inc.
- Nugraha, E. S., Alika, Z., & Amir Hamzah, D. (2024). Forecasting the Stock Price of PT Astra International Using the LSTM Method. *Rekayasa Sistem dan Teknologi Informasi*, 8(3), 431–437. <https://doi.org/10.29207/resti.v8i3.5699>
- Patriya, E., Latif, A., & Handayani. (2023). Peramalan Harga Saham Penutupan Indeks Harga Saham Gabungan (IHSG) Menggunakan Algoritma Long Short Term Memory (LSTM). *Jurnal Ilmiah Ekonomi Bisnis*, 28(2), 304–314. <https://doi.org/10.35760/eb.2023.v28i2.7964>
- Peixeiro, M. (2022). *Time Series Forecasting in Python*. Manning Publications Co.
- Pranolo, A., Zhou, X., Mao, Y., Pratolo, B. W., Wibawa, A. P., Utama, A. B. P., Ba, A. F., & Muhammad, A. U. (2024). Exploring LSTM-based Attention Mechanisms with PSO and Grid Search under Different Normalization Techniques for Energy demands Time Series Forecasting. *Knowledge Engineering and Data Science*, 7(1), 1–12. <https://doi.org/10.17977/um018v7i12024p1-12>
- Rahman, A., Istiyowati, L. S., Valentinus, Ivan, & Azis, Z. (2024). Implementasi Data Mining dalam Prediksi Harga Saham BBNi dengan Pemodelan Matematika Menggunakan Metode LSTM dengan Optimasi Adam. *Journal Education and Technology*, 5(2), 427–439. <https://doi.org/10.31932/jutech.v5i2.4137>
- Richards, T. (2021). *Getting Started with Streamlit for Data Science: Create and deploy Streamlit web applications from scratch in Python*. Packt Publishing.
- Ritu, N. A., Khandakar, S. N., Bhuiyan, Md. M., & Islam, Md. I. I. (2025). Optimizing Stock Market Prediction Using Long Short-Term Memory Networks. *Journal of Computer and Communications*, 13(2), 207–222. <https://doi.org/10.4236/jcc.2025.132013>
- Rosyd, A., Purnamasari, A. I., & Ali, I. (2024). Penerapan Metode Long Short Term Memory (LSTM) dalam Memprediksi Harga Saham PT Bank Central Asia. *Jurnal*

- Mahasiswa Teknik Informatika*, 8(1), 501–506.  
<https://doi.org/10.36040/jati.v8i1.8440>
- Saepulrohman, A., Chairunnas, A., Denih, A., & Yasibang, N. D. S. (2025). Optimization of Stock Price Prediction Using Long Short-Term Memory (LSTM) Algorithm and Cross-Industry Standard Process Approach for Data Mining (CRISP-DM). *International Journal of Electronics and Communications System*, 5(1), 19–30.  
<https://doi.org/10.24042/ijecs.v5i1.26727>
- Sahroni, M., Arif, M. F., & Misdram, M. (2024). Stock Price Prediction Using the Long Short-Term Memory Method. *Jurnal Teknik Informatika*, 5(6), 1769–1777.  
<https://doi.org/10.52436/1.jutif.2024.5.6.2615>
- Saputra, R., Gusti Alamsyah, A., Dandi, Tjoanda, M., Nick, K., Cornelius, A., & Fery Herdiatmoko, H. (2024). Analisis Prediksi Saham Tesla Menggunakan Algoritma Long Short Term Memory (LSTM). *Journal of Computer Science and Information Technology*, 2(1), 81–90. <https://doi.org/10.70248/jcsit.v2i1.1482>
- Siarni-Namini, S., Tavakoli, N., & Namin, A. S. (2019). The Performance of LSTM and BiLSTM in Forecasting Time Series. *2019 IEEE International Conference on Big Data (Big Data)*, 3285–3292. <https://doi.org/10.1109/BigData47090.2019.9005997>
- Siarni-Namini, S., Tavakoli, N., & Siarni Namin, A. (2018). A Comparison of ARIMA and LSTM in Forecasting Time Series. *Proceedings - 17th IEEE International Conference on Machine Learning and Applications, ICMLA 2018*, 1394–1401.  
<https://doi.org/10.1109/ICMLA.2018.00227>
- Situmorang, P., Mahardika, J., & Listiyarini, T. (2010). *Jurus-Jurus Berinvestasi Saham untuk Pemula*. TransMedia Pustaka.
- Sommerville, I. (2016). *Software Engineering*. Pearson Education Limited.
- Srivastava, P. P., Goyal, S., & Kumar, A. (2015). Analysis of Various NoSql Database. *2015 International Conference on Green Computing and Internet of Things (ICGCIoT)*, 539–544. <https://doi.org/10.1109/ICGCIoT.2015.7380523>
- Surya, J., & Aminuddin, F. H. (2024). *Pemrograman MYSQL Database With Streamlit Python*. PT Sonpedia Publishing.
- Syidi, J., Alam, S., & Jaelani, I. (2023). Stock Price Prediction Using Machine Learning With Long Short Therm Memory Method (LSTM). *KILAT*, 12(1), 64–78.  
<https://doi.org/10.33322/kilat.v12i1.1723>
- Tambunan, A. P. (2007). *Menilai Harga Wajar Saham (Stock Valuation)*. PT Elex Media Komputindo.
- Zhang, Z. (2023). Comparison of LSTM and ARIMA in Price Forecasting: Evidence from Five Indexes. *Proceedings of the 2023 2nd International Conference on Economics, Smart Finance and Contemporary Trade (ESFCT 2023)*, 40–46.  
[https://doi.org/10.2991/978-94-6463-268-2\\_6](https://doi.org/10.2991/978-94-6463-268-2_6)



## LAMPIRAN

### Lampiran 1. Riwayat Hidup

#### A. Identitas Diri

1	Nama Lengkap	:	Fransiscus Kristian Susanto
2	Tempat dan Tanggal Lahir	:	
3	Jenis Kelamin	:	Laki-laki
4	Program Studi	:	Informatika
5	NIM	:	40621100012
6	Alamat	:	
7	Kewarganegaraan	:	Indonesia
8	Agama	:	
9	E-Mail	:	
10	No HP	:	

#### B. Riwayat Pendidikan

	<b>SD</b>	<b>SMP</b>	<b>SMA</b>	<b>S1</b>
Nama Instansi				
Jurusan				
Tahun Masuk - Lulus				




## Lampiran 2. Kartu Bimbingan Tugas Akhir



UNIVERSITAS WIDYATAMA  
Program Studi Informatika - Program SI

### KARTU BIMBINGAN TUGAS AKHIR

Nama Mahasiswa	Fransiscus Kristian Susanto
N P M	40621100012
Alamat Mahasiswa	Komplek Permata Biru Blok H No. 93 RT 004 RW 015 Desa Cinunuk Kecamatan Cileunyi Kabupaten Bandung Telp. : +6281220894779 e-mail : kristian.fransiscus@widyatama.ac.id
Topik / Judul Tugas Akhir	Analisis Prediksi Harga Saham Hapag-Lloyd AG Menggunakan Metode Long Short-Term Memory (LSTM)
Konsentrasi	Applied Database
Dosen Pembimbing	Dr. Feri Sulianta, S.T., M.T.

BATAS WAKTU BIMBINGAN	PENGESAHAN PROGRAM STUDI
10 Februari 2025 s/d 21 Juni 2025	

PERPANJANGAN BIMBINGAN	
REKOMENDASI DOSEN PEMBIMBING	PERSETUJUAN PEMBIMBING
<p><i>Untuk perpanjangan kembali bimbingan, mahasiswa harus mengembalikan kartu ini ke jurusan sambil membawa foto kopi bukti <b>pembayaran registrasi dan Tugas Akhir</b>, disertai foto kopi FRS ( yg mencantumkan tugas akhir ). Kemudian kartu ini akan diganti dengan kartu perpanjangan bimbingan yang baru.</i></p>	

Versi/Revisi 1/1 -- 02/08/03

### CATATAN BIMBINGAN TUGAS AKHIR

NO.	T G L	POKOK BAHASAN	PARAF PEMBIMBING
01.	23/4/25	Identifikasi Struktur Skripsi	<i>[Signature]</i>
02.	30/4/25	Diskusi tinjauan Pustaka, Permodelan data, dan database	<i>[Signature]</i>
03.	14/4/25	Diskusi running dan komparasi Program	<i>[Signature]</i>
04.	21/4/25	Diskusi terkait Program yang berjalan dan Struktur Penulisan pada Skripsi	<i>[Signature]</i>
05.	13/6/25	Diskusi model database dan screening Singkat Bab 1-2	<i>[Signature]</i>
06.			
07.			
08.			
09.			
10.			
11.			
12.			
13.			
14.			
15.			

### PERSETUJUAN PENYELESAIAN BIMBINGAN TUGAS AKHIR

Bandung, \_\_\_\_\_

Menyetujui :  
Sek. Prodi Informatika,

Pembimbing,

#### CATATAN :

Jika proses bimbingan telah dinyatakan selesai dan mendapat persetujuan dosen pembimbing, maka mahasiswa segera mendaftarkan diri untuk sidang dengan menyerahkan kembali kartu ini ke jurusan dan menyelesaikan administrasinya di BAA.

Paling lambat 2 minggu setelah sidang, mahasiswa wajib mengumpulkan ke Jurusan :

- 1 buah laporan hasil revisi setelah sidang yang telah ditandatangani oleh pembimbing I & II

1 buah CD berlabel yg mencantumkan : nim, nama, judul TA, dan

nama pembimbing I & II (ditandatangani). CD berisi program dan laporan TA.

### Lampiran 3. Lembar Kuesioner *System Usability Scale*



## Kuesioner System Usability Scale

Halo, selamat datang 🌟

Perkenalkan, nama saya Fransiscus Kristian Susanto, mahasiswa Universitas Widyatama Bandung angkatan 2021.

Dalam rangka melengkapi data yang diperlukan untuk memenuhi Tugas Akhir, bersama ini peneliti menyampaikan kuesioner mengenai aplikasi Analisis Prediksi Harga Saham Hapag-Lloyd AG Menggunakan Metode Long Short-Term Memory (LSTM).

Sehubungan dengan hal tersebut, peneliti mengharapkan bantuan dan kerjasama dari teman-teman untuk mengisi kuesioner yang diberikan.

Waktu yang dibutuhkan untuk mengerjakan kuesioner ini sekitar 10-20 menit. Jawaban yang Anda isi dalam kuesioner ini tidak akan dinilai benar atau salah. Maka dari itu, diharapkan Anda mengisi sesuai dengan keadaan Anda yang sesungguhnya.

Seluruh data dan informasi yang Anda berikan dijamin kerahasiaannya serta hanya akan digunakan untuk kepentingan penelitian ini.

Atas kerjasama dan kesediaannya, peneliti mengucapkan terima kasih 🙏

Hormat peneliti,

Fransiscus Kristian Susanto

**\* Menunjukkan pertanyaan yang wajib diisi**

**Nama Lengkap \***

Jawaban Anda

**Institusi \***

☐ Universitas Widyatama

☐ Yang lain: \_\_\_\_\_

Apakah Anda memiliki minat atau ketertarikan dalam dunia investasi? \*

- ☐ Ya
- ☐ Tidak

Berikutnya

Kosongkan formulir

## Panduan Pengerjaan

Peneliti akan menampilkan aplikasi Analisis Prediksi Harga Saham Hapag-Lloyd AG Menggunakan Metode Long Short-Term Memory (LSTM) sebagai berikut.  
<https://lstm-stock-price-prediction-kristian-susanto.streamlit.app/>

Peneliti mempersilahkan Anda untuk mengakses tautan Streamlit tersebut. Agar tampilan aplikasi dapat terlihat seperti yang diharapkan, Anda dapat mengikuti langkah-langkah berikut.

1. Apabila ketika Anda membuka tautan tersebut dan menemukan kondisi berikut, Anda dapat menekan tombol yang bertuliskan Yes, get this app back up!. Kemudian, harap menunggu beberapa menit dan aplikasi akan tersedia dengan baik.



**Zzzzz**

This app has gone to sleep due to inactivity. Would you like to wake it back up?

Yes, get this app back up!

If you believe this is a bug, please [contact us](#) or [visit the Streamlit forums](#).

2. Apabila Anda menggunakan handphone, di bagian kiri atas ada tanda panah ">". Silahkan Anda klik untuk menampilkan menu dan pengaturan model. Sesudah itu, Anda dapat menyesuaikan masukan yang diinginkan dan kemudian dapat langsung tekan tombol **Mulai Prediksi** yang terletak di kiri bawah.



Apakah Anda sudah mengakses aplikasi Analisis Prediksi Harga Saham Hapag-Lloyd AG Menggunakan Metode Long Short-Term Memory (LSTM)? \*

☐ Ya

#### Pengisian System Usability Scale

System Usability Scale ini memiliki 10 pertanyaan dan 5 pilihan jawaban. Silahkan Anda dapat menentukan seberapa sesuai pernyataan tersebut dengan diri Anda.

1 : Sangat Tidak Setuju

2 : Tidak Setuju

3 : Ragu-Ragu

4 : Setuju

5 : Sangat Setuju

Contoh pengisian:

Pilih salah satu jawaban dari 1 - 5 yang sesuai dengan diri Anda.

\*Saya suka minum kopi (1 2 3 4 5)

Jika Anda memilih 4, maka jawaban Anda menunjukkan bahwa pernyataan "Saya suka minum kopi" sesuai dengan diri Anda.

Saya berpikir akan menggunakan sistem ini lagi

	1	2	3	4	5	
Sangat Tidak Setuju	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Sangat Setuju

Saya merasa sistem ini rumit untuk digunakan

	1	2	3	4	5	
Sangat Tidak Setuju	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Sangat Setuju

Saya merasa sistem ini mudah digunakan

	1	2	3	4	5	
Sangat Tidak Setuju	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Sangat Setuju

Saya membutuhkan bantuan dari orang lain atau teknisi dalam menggunakan sistem ini

	1	2	3	4	5	
Sangat Tidak Setuju	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Sangat Setuju

Saya merasa fitur-fitur sistem ini berjalan dengan semestinya

	1	2	3	4	5	
Sangat Tidak Setuju	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Sangat Setuju

Saya merasa ada banyak hal yang tidak konsisten (tidak serasi pada sistem ini)

	1	2	3	4	5	
Sangat Tidak Setuju	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Sangat Setuju

Saya merasa orang lain akan memahami cara menggunakan sistem ini dengan cepat

	1	2	3	4	5	
Sangat Tidak Setuju	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Sangat Setuju

Saya merasa sistem ini membingungkan

	1	2	3	4	5	
Sangat Tidak Setuju	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Sangat Setuju

Saya merasa tidak ada hambatan dalam menggunakan sistem ini

	1	2	3	4	5	
Sangat Tidak Setuju	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Sangat Setuju

Saya perlu membiasakan diri terlebih dahulu sebelum menggunakan sistem ini

	1	2	3	4	5	
Sangat Tidak Setuju	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Sangat Setuju

Apakah Anda sudah yakin bahwa data yang Anda masukkan ini sudah benar? \*

☐ Ya, benar

Kembali

Kirim

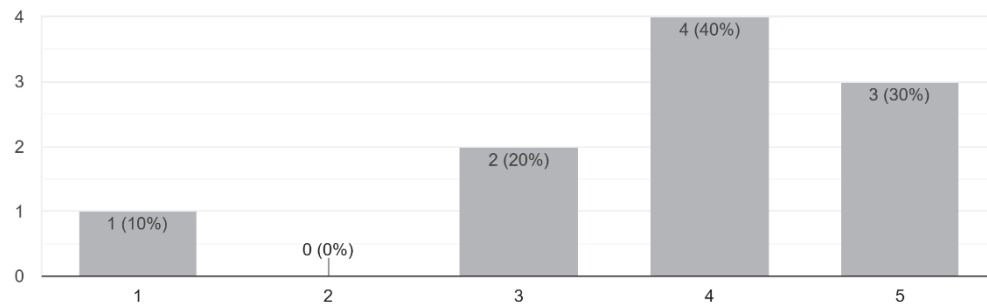
Kosongkan formulir



#### Lampiran 4. Rangkuman Hasil *System Usability Scale*

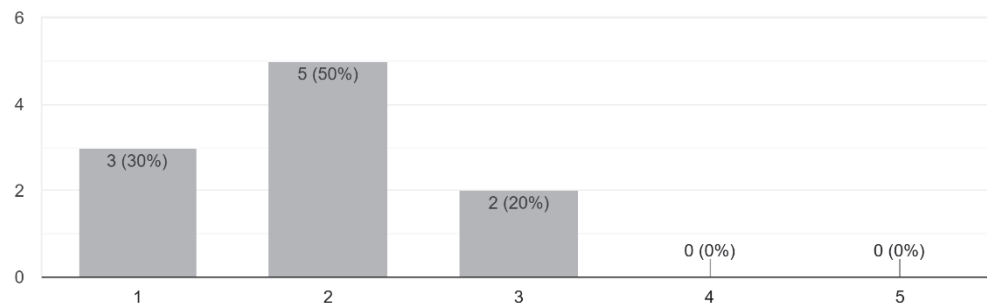
Saya berpikir akan menggunakan sistem ini lagi

10 jawaban



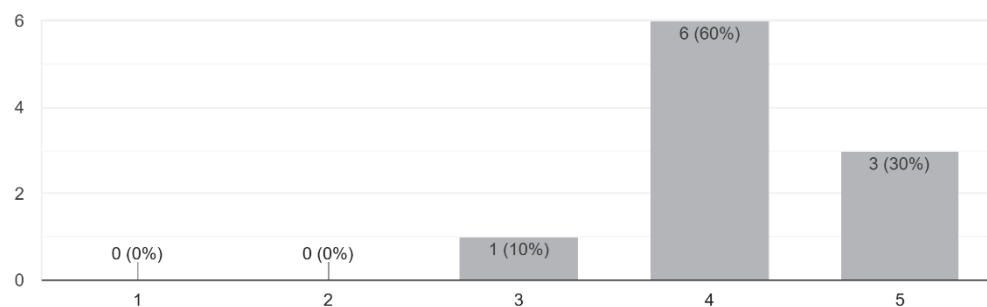
Saya merasa sistem ini rumit untuk digunakan

10 jawaban



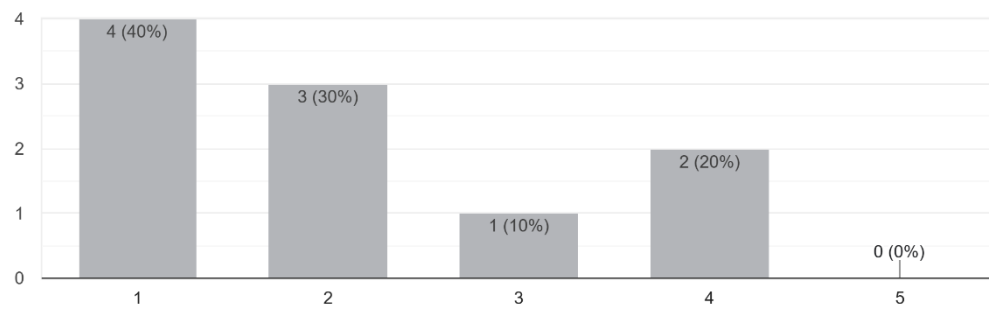
Saya merasa sistem ini mudah digunakan

10 jawaban



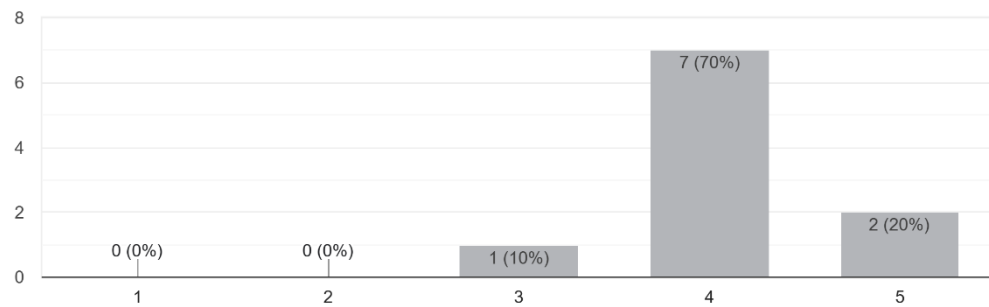
Saya membutuhkan bantuan dari orang lain atau teknisi dalam menggunakan sistem ini

10 jawaban



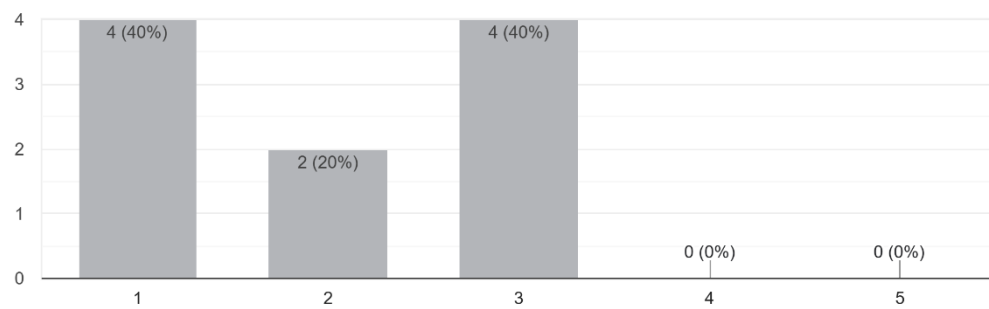
Saya merasa fitur-fitur sistem ini berjalan dengan semestinya

10 jawaban



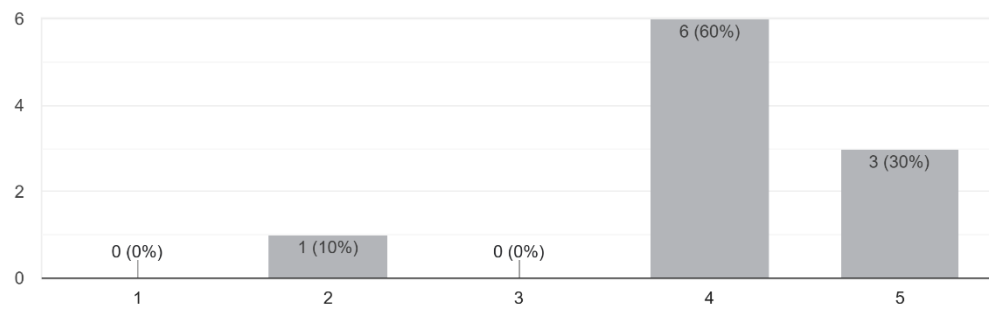
Saya merasa ada banyak hal yang tidak konsisten (tidak serasi pada sistem ini)

10 jawaban



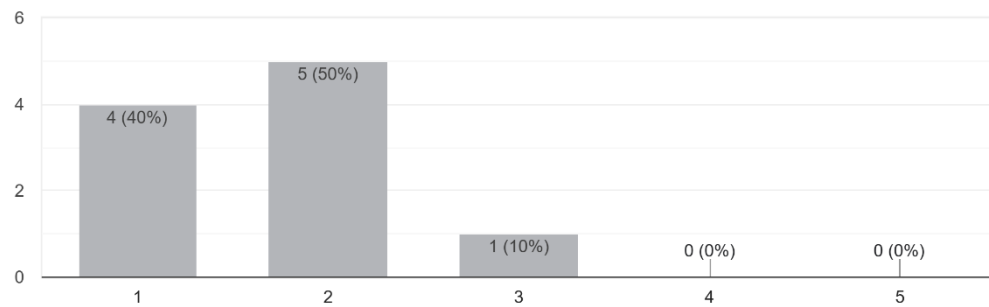
Saya merasa orang lain akan memahami cara menggunakan sistem ini dengan cepat

10 jawaban



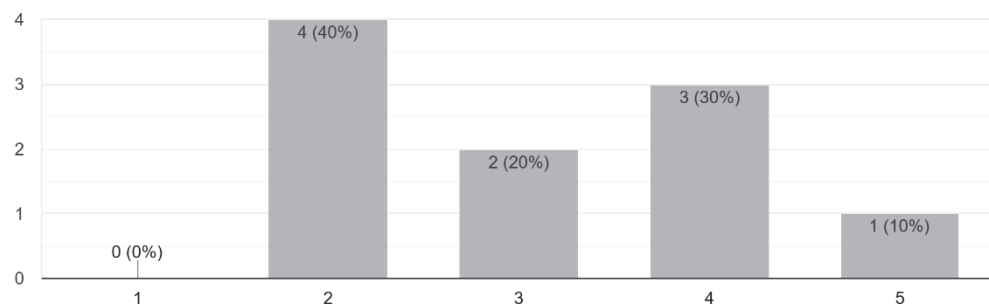
Saya merasa sistem ini membingungkan

10 jawaban



Saya perlu membiasakan diri terlebih dahulu sebelum menggunakan sistem ini

10 jawaban



Saya merasa tidak ada hambatan dalam menggunakan sistem ini

10 jawaban

