**ANALISIS PREDIKSI HARGA SAHAM HAPAG-LLOYD AG MENGGUNAKAN 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 MENGGUNAKAN 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,**  **Ari Purno Wahyu Wibowo, S.Kom., M.Kom.**  **NIDN. 0415078402** | **Dekan Fakultas Teknik,**  **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

[LEMBAR PENGESAHAN i](#_Toc200694491)

[SURAT PERNYATAAN ii](#_Toc200694492)

[ABSTRAK iii](#_Toc200694493)

[*ABSTRACT* iv](#_Toc200694494)

[KATA PENGANTAR v](#_Toc200694495)

[DAFTAR ISI vi](#_Toc200694496)

[DAFTAR GAMBAR viii](#_Toc200694497)

[DAFTAR TABEL ix](#_Toc200694498)

[DAFTAR LAMPIRAN x](#_Toc200694499)

[BAB I PENDAHULUAN 1](#_Toc200694500)

[1.1 Latar Belakang 1](#_Toc200694501)

[1.2 Rumusan Masalah 2](#_Toc200694502)

[1.3 Tujuan Penelitian 3](#_Toc200694503)

[1.4 Batasan Penelitian 3](#_Toc200694504)

[1.5 Manfaat Penelitian 4](#_Toc200694505)

[BAB II KAJIAN PUSTAKA 5](#_Toc200694506)

[2.1 Tinjauan Pustaka 5](#_Toc200694507)

[2.2 Landasan Teori 11](#_Toc200694508)

[2.2.1 Saham 11](#_Toc200694509)

[2.2.2 Hapag-Lloyd AG 13](#_Toc200694510)

[2.2.3 Pengembangan Sistem *Waterfall* 14](#_Toc200694511)

[2.2.4 *Data Mining* 16](#_Toc200694512)

[2.2.5 *Long Short-Term Memory* 17](#_Toc200694513)

[2.2.6 *Cross-Industry Standard Process for Data Mining* 25](#_Toc200694514)

[2.2.7 *Time Series Forecasting* 28](#_Toc200694515)

[2.2.8 Streamlit 31](#_Toc200694516)

[BAB III METODE PENELITIAN 33](#_Toc200694517)

[3.1 Pengembangan Sistem *Waterfall* 33](#_Toc200694518)

[3.2 *Cross-Industry Standard Process for Data Mining* 35](#_Toc200694519)

[3.3 *Flowchart* Penelitian 36](#_Toc200694520)

[BAB IV HASIL PENELITIAN DAN PEMBAHASAN 37](#_Toc200694521)

[4.1 Implementasi Sistem 37](#_Toc200694522)

[4.2 Implementasi Antarmuka 37](#_Toc200694523)

[4.3 Pengujian Sistem 37](#_Toc200694524)

[BAB V PENUTUP 38](#_Toc200694525)

[5.1 Kesimpulan 38](#_Toc200694526)

[5.2 Saran 38](#_Toc200694527)

[DAFTAR PUSTAKA 39](#_Toc200694528)

[LAMPIRAN 40](#_Toc200694529)

DAFTAR GAMBAR

DAFTAR TABEL

DAFTAR LAMPIRAN

[**Lampiran 1. Riwayat Hidup** 40](#_Toc197436132)

[**Lampiran 2. Kartu Bimbingan Tugas Akhir** 41](#_Toc197436133)

BAB I  
PENDAHULUAN

1.1 Latar Belakang

Investasi merupakan salah satu cara penting bagi individu maupun perusahaan untuk meningkatkan nilai aset mereka. 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 Farid dkk. (2023) mengimplementasikan CRISP-DM untuk menganalisis prediksi harga saham dengan menggunakan teknik penambangan data seperti klasifikasi. Penelitian tersebut menyimpulkan bahwa meskipun penambangan datadapat memberikan kontribusi dalam memprediksi harga saham, akurasi yang rendah menunjukkan perlunya model yang lebih menyeluruh untuk meningkatkan akurasi prediksi [1].

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. Berdasarkan penelitian yang dilakukan oleh Pawar dkk. (2019), hasil yang diperoleh menunjukkan bahwa model LSTM dapat memberikan prediksi harga saham yang lebih akurat dibandingkan dengan algoritma *machine learning* tradisional lainnya [2].

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 wawasan 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 [3].

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 [4].

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% [5].

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% [6].

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 [7].

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 (R2) sebesar 0,9928 dan MAPE sebesar 1,53% [8].

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% [9].

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 [10].

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 [11].

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 unit pada hidden layer. 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 [12].

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 preprocessingdata 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 [13].

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 lossMSE. 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 [14].

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 optimasi 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 [15].

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 [16].

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 [17].

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

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 badang 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 adlam menentukan kebijakan perusahaan. Apabila perusahaan dibubarkan, maka pemegang saham biasa akan membagi sisa aset perusahaan setelah dikurangi bagian pemegang saham preferen.

Saham preferen memeroleh hak istimewa dalam pembayaran dividen yang tetap setiap tahunnya dibandingkan saham biasa. Hak dividen pemegang saham preferen akan dikumpulkan apabila perusahaan pada tahun tesebut 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 [18].

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 [19].

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 memeroleh 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 [20].

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 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 [21].

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) [22].

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 [23].

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 [24].

2.2.4 *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 keefektifkan masalah yang diselesaikan [25].

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* [26].

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 mereka 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 [27].

2.2.5 *Long Short-Term Memory*

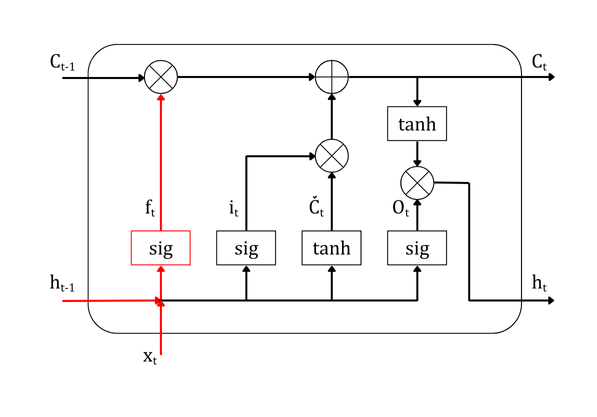
Jaringan neural *Feed-Forward* konvensional yang mempunyai perluasan kemampuan mengelola masukan deret dengan panjang variabel dikenal dengan *Recurrent Neural Networks* (RNN). Model RNN mengadakan beberapa gerbang untuk menyimpan masukan sebelumnya dan memanfaatkan informasi berurutan dari masukan sebelumnya. 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 (*hidden state* sebelumnya) dan (masukan pada ).



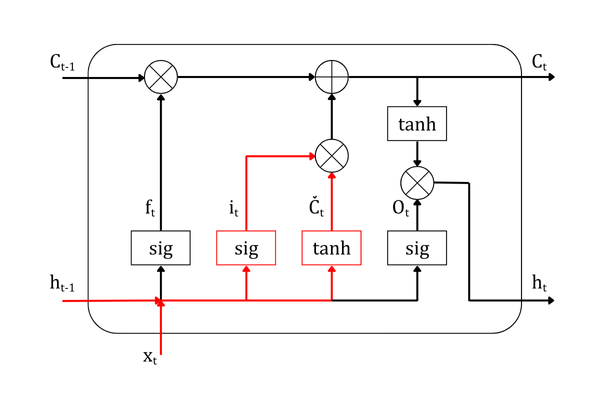
**Gambar 2.2** Forget Gate

(Sumber: Singhal, 2020)

Keluaran gerbang ini ialah (*forget gate* pada ) 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.

|  |  |
| --- | --- |
|  | (1) |

menggambarkan fungsi *sigmoid*, menggambarkan bobot untuk nilai masukan pada , dan 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 untuk membentuk vektor nilai kandidat baru yang akan ditambahkan ke memori LSTM.



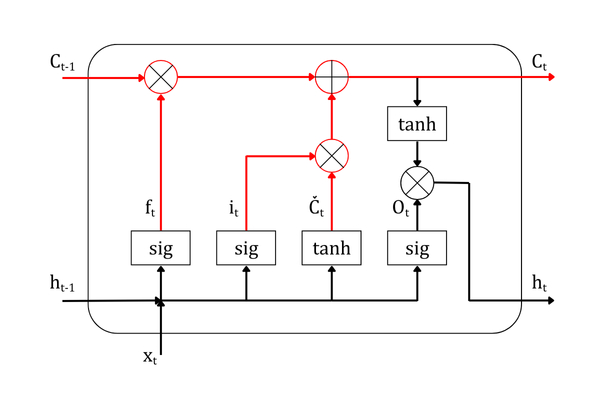
**Gambar 2.3** Input Gate

(Sumber: Singhal, 2020)

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

|  |  |
| --- | --- |
|  | (2) |
|  | (3) |

menunjukkan pemilihan nilai yang perlu diperbarui atau tidak dan menunjukkan nilai kandidat baru yang akan ditambahkan ke memori LSTM. menggambarkan bobot untuk nilai masukan pada , menggambarkan bias pada *input gate*, *hyperbolic tangent* digambarkan dengan , menggambarkan bobot untuk masukan pada sel ke , dan menggambarkan bias pada sel ke .



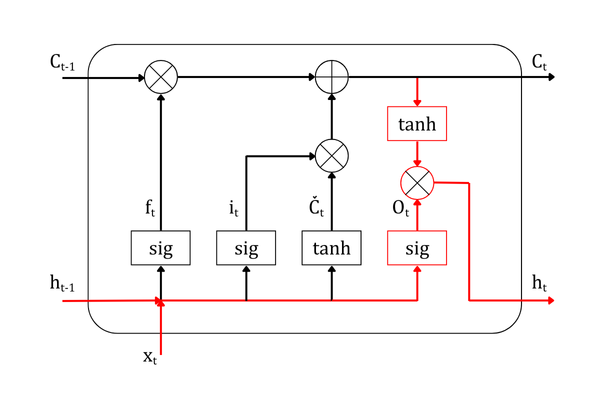
**Gambar 2.4** Cell State

(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 () dibarengi dengan menambahkan kandidat baru . Persamaan untuk *cell state* baru terbentuk sebagai berikut.

|  |  |
| --- | --- |
|  | (4) |

menggambarkan nilai memori *cell state* baru. Nilai *forget gate* dilambangkan dengan 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 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.

|  |  |
| --- | --- |
|  | (5) |
|  | (6) |

menggambarkan nilai *output gate* (keluaran), menggambarkan bobot untuk keluaran pada , menggambarkan bias pada *output gate*, dan menggambarkan nilai hasil akhir [28]. 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 ← length(series) \* 0.7  train ← series[0…size]  test ← 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 ← train  y ← 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 [29]. Visualisasi algoritma LSTM dengan *flowchart* adalah sebagai berikut [30].

Mulai

Dataset

Memasukkan data

Preprocessing (min-max scaling, pembagian dataset, pemilihan atribut berdasarkan tanggal

Dataset latih

Dataset uji

LSTM

Prediksi

Nilai prediksi

Evaluasi model

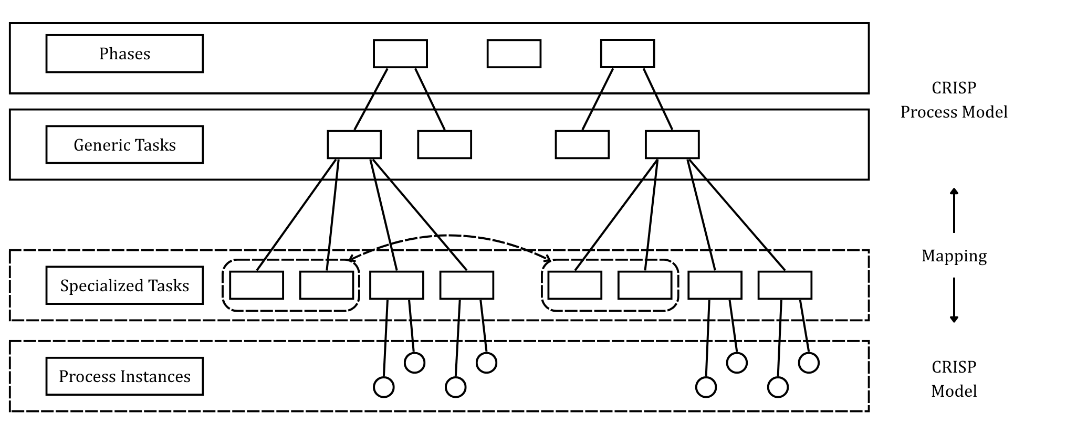
RMSE

**Gambar 2.6** Flowchart LSTM

(Sumber: Albeladi, 2023)

2.2.6 *Cross-Industry Standard Process for Data Mining*

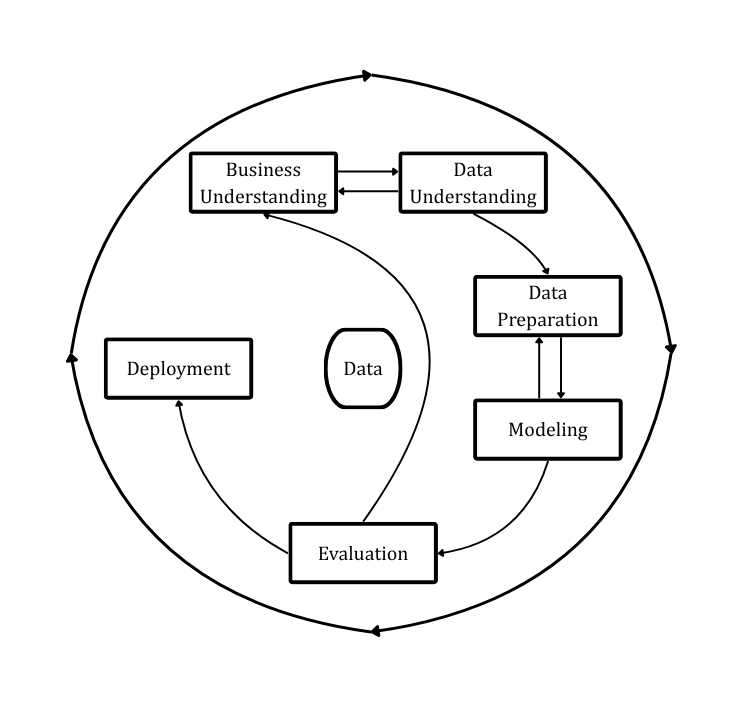
CRISP-DM (*Cross-Industry Standard Process for Data Mining*) merupakan metodologi penambangan data dengan proses berjenjang yang meliputi penjabaran kumpulan tugas pada empat tingkat. Tingkat pertama dikoordinasikan ke dalam sejumlah *phases* (fase) di mana masing-masing fase terdiri dari beberapa tugas 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 [31].

2.2.7 *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 [32].

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 [33].

*Forecasting* (peramalan) dilakukan untuk memprediksi kelanjutan deret waktu dengan menggunakan data historis (masa lalu) dan pengetahuannya terkait kejadian di masa depan [32]. Metode peramalan dapat berasal dari prosedur baik bergantung maupun tidak bergantung pada suatu model deret waktu agar menemukan kemungkinan ramalan yang lebih baik. [34].

Terdapatdua 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 pengubahan 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.

Problem definition

Data collection

Data analysis

Model selection and fitting

Model validation

Forecasting model deployment

Monitoring forecasting model performance

**Gambar 2.9** Proses Peramalan

(Sumber: Montgomery dkk, 2015)

Aktivitas kedua bernama *data collection* (pengumpulan data) dengan memeroleh 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 [35].

2.2.8 Streamlit

Streamlit merupakan kerangka kerja (*framework*) publik yang menyediakan *deployment* sains data dan pembelajaran mesin berbasis *website* dengan penggunaan bahasa pemrograman Python. Streamlit dapat menampilkan informasi interaktif, memanfaatkan berbagai pilihan tema pada antarmuka, dan mendukung berbagai pustaka Python untuk pengolahan dan analisis data. *Deployment* Streamlit fleksibel terhadap beberapa platform *cloud*, seperti AWS (Amazon Web Services), Google Cloud, dan Streamlit Cloud. Streamlit Cloud memudahkan pengguna dalam menampilkan aplikasi daring tanpa pengelolaan lebih lanjut pada infrastruktur (*hosting*) [36].

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 [37].

Beberapa perbedaan kerangka kerja berbasis web menggunakan bahasa pemrograman Python adalah sebagai berikut [38].

**Tabel 2.2** Perbedaan Beberapa Framework Web Python

|  |  |  |  |
| --- | --- | --- | --- |
| **Fitur** | **Streamlit** | **Flask** | **Django** |
| 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 *dashboard*. | 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

BAB III  
METODE PENELITIAN

3.1 Pengembangan Sistem *Waterfall*

3.1.1 *Planning*

Saham merupakan salah satu instrumen investasi yang dipercaya dapat memberikan profit yang besar bagi para investornya. Terdapat ribuan perusahaan terbuka di seluruh dunia telah menerbitkan saham perusahaannya kepada publik dengan maksud pengembangan modal bisnis, memperluas ekspansi pasar, dan mengupayakan pertumbuhan perusahaan tersebut.

Kadangkala investor merasa bimbang dalam menginvestasikan modalnya pada suatu perusahaan terbuka. Di samping hal ini juga perdagangan saham terus-menerus aktif dalam setiap hari perdagangan yang dapat memakan waktu dalam menganalisis sebagian besar perusahaan terbuka.

Setiap perusahaan yang terdaftar di bursa efek suatu negara memiliki berbagai macam data yang berhubungan dengan saham perusahaannya. 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.

Data historis saham dapat menggambarkan proses yang berhubungan dengan situasi yang terjadi dalam perusahaan. Perilaku internal, kondisi pasar, dan integritas perusahaan dapat diilustrasikan dengan perkembangan harga saham yang pernah berlangsung. Menangkap pola pikir harga saham dengan perkembangan ekonomi dengan perusahaan tersebut dapat menjadi cara yang masuk akal untuk ditelaah. Data historis harga saham diharapkan dapat menjadi pembelajaran yang menarik sembari mendalami kejadian masa lalu yang dapat berpengaruh pada waktu mendatang.

Pengumpulan data historis harga saham secara manual, seperti menulis susunan harga yang pernah terjadi maka akan membutuhkan waktu yang sangat panjang. Dalam pengumpulan harga historis saham tersebut juga sangat memungkinkan terjadi kesalahan dalam penulisan titik poin data tertentu. Investor dapat merasa kesulitan ketika mesti menyimpan susunan data harga saham dengan jangka waktu yang panjang.

Pengumpulan data harga saham yang paling efektif dapat dilakukan dengan mengakses suatu API (Application Programming Interface) aktif yang bernama Yahoo Finance. Penggunaan API Yahoo Finance telah banyak dilakukan penelitian terdahulu yang dapat dicermati dalam tinjauan pustaka pada penelitian ini.

Pengembangan sistem pada penelitian ini akan mengumpulkan data historis harga saham dengan menggunakan API Yahoo Finance. Sistem dikembangkan dengan kerangka kerja Streamlit yang memudahkan pengembang dan penerima dalam mengolah data historis harga saham pada suatu perusahaan terbuka khususnya Hapag-Lloyd AG yang menjadi fokus utama perusahaan terbuka dalam penelitian ini.

3.1.2 *Analysis*

Pengembangan aplikasi pada prediksi harga saham diselenggarakan untuk mendalami data historis harga suatu saham. Aplikasi ini nantinya akan diimplementasikan dan dijangkau melalui jaringan internet. Pengolahan data historis harga saham diproses menggunakan algoritma LSTM serta pemodalan data tersebut dapat disimpan ke dalam basis data dan diakses kembali oleh pengguna.

Pembangunan sistem membutuhkan pengguna yang mempunyai minat atau ketertarikan dalam dunia investasi. Evaluasi sistem diperlukan dengan tanggapan atau masukan dari pengguna terkait pengalaman penggunaan sistem tersebut. Pengembang menyusun dan membangun sistem dengan fase kerja secara berurutan dan pengolahan data menggunakan metodologi penambangan data CRISP-DM.

Pengembang menetapkan kebutuhan informasi dengan melakukan observasi terkait pemodelan data menggunakan algoritma LSTM. Penjajakan dokumentasi dikerjakan dengan meresapi data historis harga saham yang dihubungkan dengan penyimpanan basis data dan pengaplikasian kerangka kerja program.

Kinerja sistem hendak menampilkan akurasi prediksi model menggunakan metrik evaluasi yang menampilkan performa pengujian data harga saham. Hasil prediksi menghadirkan visualisasi dan ringkasan singkat terkait pemodelan data historis saham tersebut.

3.1.3 *Design*

3.1.4 *Implementation*

3.1.5 *Use*

3.2 *Cross-Industry Standard Process for Data Mining*

3.2.1 *Business Understanding*

3.2.2 *Data Understanding*

3.2.3 *Data Preparation*

3.3.4 *Modeling*

3.3.5 *Evaluation*

3.3.6 *Deployment*

3.3 *Flowchart* Penelitian

BAB IV  
HASIL PENELITIAN DAN PEMBAHASAN

4.1 Implementasi Sistem

4.2 Implementasi Antarmuka

4.3 Pengujian Sistem

BAB V  
PENUTUP

5.1 Kesimpulan

5.2 Saran

DAFTAR PUSTAKA

LAMPIRAN

**Lampiran 1. Riwayat Hidup**

1. Identitas Diri

|  |  |  |  |
| --- | --- | --- | --- |
| 1 | Nama Lengkap | : |  |
| 2 | Tempat dan Tanggal Lahir | : |  |
| 3 | Jenis Kelamin | : |  |
| 4 | Program Studi | : |  |
| 5 | NIM | : |  |
| 6 | Alamat | : |  |
| 7 | Kewarganegaraan | : |  |
| 8 | Agama | : |  |
| 9 | E-Mail | : |  |
| 10 | No HP | : |  |

1. Riwayat Pendidikan

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **SD** | **SMP** | **SMA** | **S1** |
| Nama Instansi |  |  |  |  |
| Jurusan |  |  |  |  |
| Tahun Masuk - Lulus |  |  |  |  |

**Lampiran 2. Kartu Bimbingan Tugas Akhir**

