

BAB II

LANDASAN TEORI

2.1 Penelitian Terkait

Dalam beberapa tahun terakhir, topik pengembangan model *Hybrid* semakin banyak dikaji dalam ranah prediksi *time series*, khususnya untuk meningkatkan akurasi dan efisiensi model. Tabel 2.1 merangkum sepuluh penelitian terdahulu yang menjadi dasar konseptual dalam penelitian ini, yang mayoritas menggabungkan algoritma LSTM dengan metode lain seperti CNN, SVM, dan GNN. Penelitian-penelitian tersebut memberikan bukti empiris bahwa pendekatan *Hybrid* memiliki potensi signifikan dalam menghasilkan prediksi yang lebih akurat.

Tabel 2. 1 Penelitian Terkait

No	Nama Jurnal	Judul Artikel	Penulis	Latar Belakang, Metode, dan Hasil
1	<i>Electronic s</i> 2022, 11, 3443	<i>Predicting Close Price in Emerging Saudi Stock Exchange Time Series Models [12]</i>	Al-Nefaie, A.H.; Aldhyani, T.H.H.	Memprediksi harga penutupan pasar saham Arab Saudi dengan menggunakan data dari tahun 2018-2022. Penelitian ini menggunakan algoritma <i>Multilayer Perceptron</i> (MLP) dan <i>Long Short-Term Memory</i> (LSTM), dengan hasil LSTM mengungguli MLP dengan nilai $R > 0.995$.
2	<i>Electronic s</i> 2023, 12(18), 3985	<i>Optimizing Long Short-Term Memory Network for Air Pollution Prediction Using a Novel Binary Chimp Optimization Algorithm [15]</i>	Baniasadi, S.; Salehi, R.; Soltani, S.; Martín, D.; Pourmand, P.; Ghafourian, E.	Penelitian ini memprediksi polusi udara pada atmosfer. Algoritma yang digunakan pada penelitian ini adalah <i>Long Short-Term Memory</i> (LSTM) untuk memprediksi data <i>time series</i> dan <i>Binary Chimp Optimization Algorithm</i> (BchOA) untuk mengoptimalkan arsitektur LSTM. Hasil menunjukkan <i>Hybrid</i> model ini menghasilkan akurasi 96.41%.
3	<i>Procedia Computer Science</i> 234 (2024) 333–340	<i>Development of a CNN-LSTM Approach with Images as Time Series Data Representation for Predicting Gold Prices [16]</i>	Salim, M.; Djunaidy, A.	Memprediksi harga saham dengan pendekatan model <i>Hybrid Convolutional Neural Network</i> (CNN) dan <i>Long Short-Term Memory</i> (LSTM) dengan bantuan <i>Gramian Angular Field</i> (GAF) untuk merepresentasikan data <i>time series</i> dalam bentuk gambar. Hasilnya menunjukkan model <i>Hybrid CNN-LSTM</i> memiliki akurasi yang lebih baik daripada LSTM dan CNN secara individu.

No	Nama Jurnal	Judul Artikel	Penulis	Latar Belakang, Metode, dan Hasil
4	<i>e-Prime-Advances in Electrical Engineering, Electronics and Energy</i> 9 (2024) 100636	<i>Hybrid machine learning model combining of CNN-LSTM-RF for time series forecasting of Solar Power Generation [17]</i>	Mobarak Abumohsen, Amani Yousef Owda, Majdi Owda, Ahmad Abumihsan	Memprediksi pembangkitan listrik tenaga matahari berdasarkan data <i>time series</i> . Penelitian ini dilakukan dengan <i>Hybrid</i> model CNN, LSTM, dan <i>Random Forest</i> . Hasil menunjukkan bahwa akurasi prediksi R^2 sebesar 92%, sehingga membuktikan bahwa model gabungan CNN-LSTM- <i>Random Forest</i> memiliki akurasi yang lebih baik dalam memprediksi data mengenai pembangkitan listrik tenaga matahari.
5	<i>EAI Endorsed Transactions on Scalable Information Systems</i> . Volume 10, Issue 1, e7	<i>Malware detection for Android application using Aquila optimizer and Hybrid LSTM-SVM classifier [14]</i>	M. Gracea, Dr. M. Sughasiny	Mendeteksi <i>malware</i> pada aplikasi android dengan menggunakan <i>Hybrid</i> model LSTM-SVM dan <i>Aquila Optimizer</i> sebagai pendukung, adapun algoritma lain yaitu LSTM, SVM, <i>Random Forest</i> , dan Naïve Bayes sebagai pembanding. Hasil menunjukkan model model <i>Hybrid</i> LSTM-SVM dengan bantuan <i>Aquila Optimizer</i> menghasilkan akurasi terbaik sekitar 97%.
6	<i>International Journal of Informatics and Computation (IJICOM)</i> Vol. 4, No.1, August 2022	<i>StockTM: Accurate Stock Price Prediction Model Using LSTM [13]</i>	Mohammad Diqi	Prediksi harga saham memiliki tantangan tersendiri karena sifat data <i>non linear</i> dan tergolong dinamis. Penelitian ini ingin menguji akurasi LSTM dalam melakukan prediksi terhadap data historis saham. Penelitian ini dilakukan dengan menguji model pada 2 buah <i>output</i> yaitu EMA10 dan EMA20. Hasil menunjukkan pada EMA10 evaluasi RMSE = 0.00714 dan MAPE = 0.07705, dan pada EMA20 menunjukkan RMSE = 0.00355 dan MAPE = 0.05273.
7	<i>May 2023 IJSDR Volume 8 Issue 5</i>	<i>Real-Time Stock Market Prediction Using ML and DL [18]</i>	Raghav Dholu, Prerna Mistri, Prerna Lohar, Sayali Bacchav Dr. Namita R. Kale	Penelitian ini dilakukan untuk membandingkan 2 model <i>Hybrid</i> yaitu CNN-LSTM dan SVM-LSTM. Pengujian dilakukan menggunakan data harga saham dari <i>yahoo finance</i> . Hasil menunjukkan bahwa CNN-LSTM memiliki akurasi yang lebih baik dalam memprediksi pergerakan harga saham.
8	<i>arXiv preprint (2024)</i>	<i>STOCK PRICE PREDICTION USING A HYBRID LSTM-GNN MODEL: INTEGRATING TIME-SERIES AND GRAPH-BASED ANALYSIS [19]</i>	Meet Satishbhai Sonani, Atta Badii, Armin Moin	Penelitian ini dilakukan untuk menggabungkan informasi temporal antar saham guna meningkatkan akurasi. Penelitian ini menggunakan <i>Hybrid</i> model LSTM-GNN dengan tambahan <i>expanding window validation</i> pada proses <i>training</i> . Hasil menunjukkan bahwa LSTM-GNN menghasilkan akurasi yang lebih tinggi

No	Nama Jurnal	Judul Artikel	Penulis	Latar Belakang, Metode, dan Hasil
				sebesar 10.6% dibandingkan algoritma LSTM.
9	<i>Wiley Journal of Probability and Statistics Volume 2025,</i>	<i>Time-Series Forecasting Using SVMD-LSTM: A Hybrid Approach for Stock Market Prediction [20]</i>	Sanskar Agarwal , Shivam Sharma , Kazi Newaj Faisal , Rishi Raj Sharma	Penelitian ini dilakukan untuk mengatasi pergerakan pasar saham yang non linear dengan menggabungkan algoritma <i>Successive Variational Mode Decomposition (SVMD)</i> dan LSTM. SVMD berperan sebagai dekomposisi dan LSTM berperan sebagai prediktor tiap komponen. Hasil menunjukan <i>Hybrid SVMD-LSTM</i> memiliki akurasi yang lebih tinggi daripada algoritma LSTM.
10	<i>Proceedings of the 5th Management Science Informatization and Economic Innovation Development Conference (MSIEID 2023)</i>	<i>Stock Prediction Based on LSTM [21]</i>	Min Li, Weize Liao, Jianrong Huang, Jiebo Jiang	Penelitian ini melakukan prediksi harga saham dengan menggunakan algoritma LSTM dengan 2 skenario yaitu dengan menggunakan 1 buah indikator yaitu harga penutupan dan dengan menggunakan beberapa indikator seperti <i>volume</i> penjualan dan <i>moving average</i> . Hasil menunjukan bahwa penggunaan lebih dari 1 variabel dapat menghasilkan model yang lebih akurat.

Berdasarkan sepuluh penelitian terdahulu yang tetera pada Tabel 2.1, mayoritas menggunakan pendekatan berbasis *deep learning* untuk menangani permasalahan prediksi data *time series*. Metode-metode tersebut diuji menggunakan metrik evaluasi standar seperti *Mean Squared Error (MSE)*, *Root Mean Squared Error (RMSE)*, *Mean Absolute Error (MAE)*, dan *Mean Absolute Percentage Error (MAPE)* sebagai tolok ukur kinerja model. Keberagaman domain penelitian, mulai dari prediksi harga saham, polusi udara , hingga deteksi *malware*, menunjukkan bahwa pendekatan ini memiliki fleksibilitas tinggi dalam pengolahan data sekuensial.

Sebagian besar penelitian tersebut menggunakan *Long Short-Term Memory* (LSTM) karena kemampuannya dalam mempelajari pola jangka panjang dalam data historis. Misalnya, pada penelitian [12] dan [13], LSTM terbukti mampu mengungguli metode lain seperti *Multilayer Perceptron* (MLP) dengan hasil $R > 0.995$ dan MAPE serendah 0.05273. Selain itu, penelitian [21] juga menunjukkan bahwa LSTM bekerja optimal ketika digunakan untuk memprediksi harga saham dengan berbagai indikator teknikal. Fakta ini menjadi dasar rasional dalam penelitian ini untuk menggunakan LSTM sebagai model utama dalam menangkap pola pergerakan harga saham Intel.

Meskipun LSTM memiliki performa yang baik, beberapa studi terdahulu mencoba meningkatkan akurasi dengan menerapkan model *Hybrid*. Penelitian [16], [17], dan [18] menunjukkan bahwa gabungan LSTM dengan CNN, *Random Forest* (RF), maupun SVM dapat meningkatkan kinerja model secara signifikan. Dalam penelitian [17], CNN-LSTM-RF menghasilkan nilai R^2 sebesar 92% dalam memprediksi pembangkit listrik tenaga surya. Sedangkan pada penelitian [16], penggunaan CNN-LSTM dengan input berbasis gambar melalui *Gramian Angular Field*/GAF mampu mengungguli model individual. Temuan-temuan ini membuktikan bahwa model *Hybrid* memiliki potensi besar dalam menghasilkan prediksi yang lebih akurat dan stabil.

Dalam konteks penelitian ini, model *Hybrid* yang digunakan adalah gabungan LSTM dan SVM. LSTM digunakan untuk mengekstrak fitur temporal dari data historis, sedangkan SVM digunakan untuk memproses fitur tersebut dalam melakukan prediksi akhir. Kombinasi ini bertujuan untuk mengatasi keterbatasan LSTM dalam klasifikasi dan memanfaatkan kekuatan SVM dalam memisahkan pola data secara optimal. Penelitian [14] secara khusus menunjukkan efektivitas *Hybrid* LSTM-SVM dalam mendeteksi *malware* pada aplikasi *Android*, dengan akurasi mencapai 97%. Struktur model seperti ini relevan untuk diterapkan dalam prediksi harga saham yang kompleks, karena mampu memadukan pemahaman pola temporal dengan klasifikasi hasil prediksi yang presisi.

Penguatan pendekatan *Hybrid* juga terlihat dari penelitian [19] yang menggabungkan LSTM dengan *Graph Neural Network* (GNN), menghasilkan peningkatan akurasi hingga 10.6% dibanding LSTM saja. Sementara itu, penelitian [20] menunjukkan bahwa kombinasi *Successive Variational Mode Decomposition* (SVMD) dan LSTM dapat mengurangi *error* secara signifikan dalam prediksi harga saham. Berdasarkan tinjauan terhadap sepuluh penelitian terdahulu, dapat disimpulkan bahwa meskipun model LSTM telah terbukti efektif dalam memproses data *time series*, dan model *Hybrid* seperti CNN-LSTM maupun LSTM-RF mampu meningkatkan akurasi, namun masih terdapat kekurangan dalam penerapan kombinasi model LSTM-SVM secara khusus dalam konteks prediksi harga saham berbasis regresi. Sebagian besar penelitian yang menggabungkan LSTM dan SVM lebih menekankan pada klasifikasi, bukan prediksi nilai numerik harga saham secara langsung. Selain itu, belum banyak studi yang secara langsung mengevaluasi performa kombinasi LSTM dan SVM berdasarkan rasio kontribusi masing-masing model terhadap hasil prediksi akhir. Maka dari itu, penelitian ini memberikan kontribusi dengan mengembangkan dan menguji model *Hybrid* LSTM-SVM dalam memprediksi harga penutupan saham Intel secara regresif, sekaligus menginvestigasi pengaruh perbedaan rasio bobot prediksi LSTM dan SVM terhadap akurasi model. Kontribusi ini diharapkan dapat memperluas pemahaman terhadap efektivitas model *Hybrid* dalam analisis data *time series* pasar saham, serta menjadi referensi dalam pengembangan sistem prediksi berbasis kecerdasan buatan yang lebih presisi dan adaptif.

2.2 Teori Penelitian

2.2.1 Saham

Saham merupakan salah satu bentuk instrumen keuangan yang mengindikasikan kepemilikan seseorang terhadap suatu perusahaan terbuka [22]. Dalam konteks pasar modal, saham diperdagangkan secara bebas dan nilainya ditentukan oleh mekanisme permintaan dan penawaran. Harga saham dapat berubah dalam waktu singkat dipengaruhi oleh berbagai faktor, seperti kondisi ekonomi makro, performa keuangan perusahaan, kebijakan pemerintah, hingga sentimen pasar [23], [24]. Oleh karena itu, prediksi harga saham menjadi

sebuah tantangan yang menarik dan kompleks karena karakteristiknya yang dinamis dan tidak linier.

Intel Corporation, sebagai salah satu perusahaan teknologi global terbesar, menjadi objek yang menarik dalam analisis pergerakan harga saham. Hal ini dikarenakan sektor teknologi memiliki volatilitas tinggi yang seringkali dipengaruhi oleh inovasi produk dan dinamika industri. Dengan data historis harga saham, analisis prediktif terhadap saham Intel menjadi relevan dalam upaya membantu investor mengambil keputusan berbasis data.

2.2.2 Machine Learning

Machine learning merupakan cabang dari kecerdasan buatan *Artificial Intelligence* yang berfokus pada pengembangan sistem yang mampu belajar secara otomatis dari data tanpa perlu diprogram secara langsung [25]. Dalam pendekatan ini, model dilatih menggunakan dataset tertentu untuk menemukan pola atau hubungan yang tersembunyi dan kemudian digunakan untuk membuat prediksi terhadap data baru. Beberapa algoritma populer dalam machine learning antara lain adalah *decision tree*, *random forest*, *naive bayes*, *k-nearest neighbors*, dan *support vector machine* (SVM).

Dalam konteks prediksi saham, *machine learning* memiliki potensi besar karena kemampuannya dalam menangani *dataset* berukuran besar dan menemukan hubungan kompleks antar variabel. Salah satu algoritma yang efektif dalam klasifikasi dan regresi adalah SVM, yang bekerja dengan mencari *hyperplane* optimal yang dapat memisahkan data ke dalam kelas-kelas tertentu [26]. Dengan bantuan *kernel trick*, SVM juga mampu mengatasi masalah non-linier yang sering muncul dalam data pasar saham.

2.2.3 Deep Learning

Deep learning adalah bagian dari *machine learning* yang menggunakan jaringan *neural network* untuk merekayasa cara kerja otak manusia dalam memproses informasi [27]. *Deep learning* unggul dalam menangani data yang kompleks dan tidak terstruktur, seperti gambar, suara, dan data sekuensial seperti teks dan *time series* [28]. Keunggulan ini menjadikan deep learning

sebagai metode yang banyak diterapkan dalam berbagai aplikasi berbasis kecerdasan buatan modern.

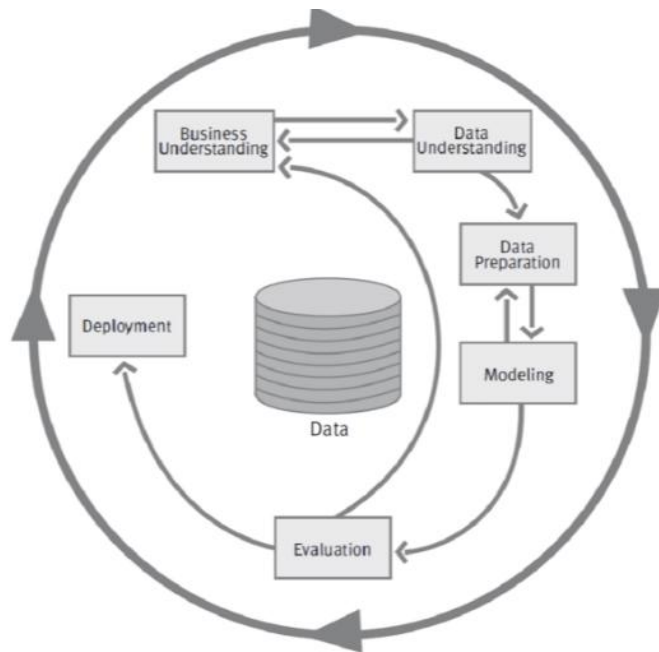
Salah satu model yang populer dalam deep learning adalah *Long Short-Term Memory* (LSTM), yang merupakan varian dari *Recurrent Neural Network* (RNN). LSTM dirancang untuk menangani masalah *long-term dependency* dalam data *time series* dengan mengandalkan mekanisme *gate* (*input gate*, *forget gate*, dan *output gate*) yang memungkinkan jaringan untuk menyimpan atau melupakan informasi yang tidak dibutuhkan [29]. Dalam konteks prediksi harga saham, LSTM sangat cocok digunakan karena mampu memahami urutan waktu dan dinamika data historis yang bersifat saling berkesinambungan [9].

Dalam pengembangan sistem prediksi yang lebih akurat, pendekatan *Hybrid* menjadi salah satu strategi yang banyak digunakan. Pada konteks penelitian ini, model *Hybrid* dibentuk dengan menggabungkan kekuatan LSTM dalam memahami hubungan sekuensial serta keunggulan SVM dalam klasifikasi atau regresi. Kombinasi ini diharapkan dapat meningkatkan kinerja prediksi harga saham secara lebih optimal dibandingkan penggunaan model tunggal.

2.3 Teori Framework dan Algoritma Penelitian

2.3.1 *Cross Industry Standard Process for Data Mining* (CRISP-DM)

Penelitian kali ini menggunakan *Cross Industry Standard Process for Data Mining* (CRISP-DM) yang alurnya dapat dilihat pada Gambar 2.1. CRISP-DM merupakan pendekatan yang lumrah digunakan dalam proses pengolahan data dalam berbagai bidang, karena mampu menguraikan langkah-langkah sistematis untuk memahami data, mempersiapkan, menganalisis, hingga menghasilkan informasi dari data tersebut [30], [31]. Pendekatan ini sangat cocok diterapkan dalam penelitian ini yang berfokus pada prediksi harga saham Intel, mengingat jumlah data yang digunakan cukup besar dan berbentuk *time series*. CRISP-DM terdiri dari enam tahap utama yang saling terhubung dan membentuk alur proses yang runtut.



Gambar 2. 1 CRISP DM

Sumber : [31]

1. *Business Understanding*

Tahapan awal ini berfokus pada melakukan pendalaman terhadap *scope* yang akan di teliti [31]. Dalam konteks ini, tujuan utamanya adalah membangun model prediksi harga saham Intel dengan pendekatan *Hybrid LSTM-SVM*. Permasalahan yang ingin diselesaikan adalah bagaimana menghasilkan prediksi yang akurat terhadap arah pergerakan harga saham berdasarkan data historis. Pemahaman terhadap latar belakang ini membantu dalam menyusun pertanyaan riset serta menentukan metode yang tepat untuk digunakan.

2. *Data Understanding*

Setelah memahami kebutuhan bisnis, tahap selanjutnya adalah mengumpulkan dan mengenali data yang akan digunakan [30]. Dalam penelitian ini, data yang digunakan adalah data historis saham Intel dari tahun 1980 hingga 2025, dengan jumlah lebih dari 11.000 baris data. Proses ini melibatkan eksplorasi awal terhadap data, seperti melihat distribusi harga, mengidentifikasi nilai ekstrem, dan mengecek adanya nilai yang

hilang. Dari sini, peneliti mendapatkan gambaran awal tentang struktur data yang akan dianalisis.

3. Data Preparation

Tahap ini sangat penting karena kualitas data akan sangat berpengaruh terhadap hasil akhir prediksi [31]. Data dipersiapkan dengan melakukan *cleaning data*, normalisasi nilai, serta pembentukan data urut sesuai kebutuhan model LSTM. Selain itu, dilakukan juga pembagian data menjadi data latih dan data uji. Semua tahapan ini dilakukan agar data siap digunakan dalam pelatihan model.

4. Modeling

Pada tahap ini, dilakukan penerapan dua algoritma utama, yaitu LSTM dan SVM. LSTM digunakan untuk mempelajari pola historis dari data harga saham dan mengekstrak fitur-fitur penting yang berkaitan dengan pergerakan harga. Hasil keluaran dari LSTM kemudian dijadikan input bagi SVM yang berperan sebagai algoritma klasifikasi. SVM akan mengolah hasil tersebut untuk menentukan arah pergerakan harga, apakah cenderung naik atau turun. Gabungan keduanya diharapkan menghasilkan model prediksi yang lebih akurat dan stabil.

5. Evaluation

Setelah model berhasil dibangun, tahap evaluasi dilakukan untuk mengetahui seberapa baik model bekerja dalam memprediksi harga saham. Evaluasi dilakukan dengan menggunakan metrik seperti RMSE, MAE, MAPE, dan *R-squared* (R^2). Hasil prediksi juga dibandingkan secara visual dengan data aktual melalui grafik. Tujuannya adalah untuk memastikan bahwa model tidak hanya baik secara angka, tetapi juga mampu mengikuti pola tren yang sebenarnya.

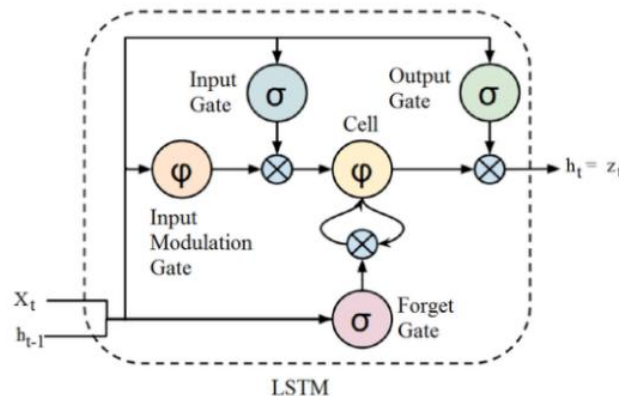
6. Deployment

Tahap akhir ini bertujuan untuk menempatkan model yang sudah dievaluasi ke dalam konteks penggunaan nyata [31]. Dalam penelitian ini, *deployment* dilakukan dalam bentuk analisis hasil prediksi yang kemudian digunakan untuk menjawab rumusan masalah penelitian. Meski tidak digunakan langsung dalam sistem investasi, model yang dikembangkan

dapat menjadi dasar dalam membangun sistem rekomendasi atau sistem prediksi harga saham berbasis kecerdasan buatan di masa depan.

2.3.2 Long Short Term Memory

Long Short-Term Memory (LSTM) adalah salah satu model turunan dari *Recurrent Neural Network* (RNN) yang dirancang untuk mengatasi permasalahan yang sering terjadi dalam RNN standar, yaitu *vanishing gradient* dan kesulitan dalam mempertahankan informasi *time series* jangka panjang (*long-term dependency*) [29], [32]. LSTM dipopulerkan oleh Hochreiter dan Schmidhuber pada tahun 1997 dan telah terbukti efektif dalam berbagai kasus data *time series*, terutama pada analisa harga saham.



Gambar 2. 2 Ilustrasi LSTM

Sumber : [33]

Pada Gambar 2.2. dapat dilihat struktur dasar dari LSTM yang memiliki arsitektur berupa *memory cell* yang bekerja dalam satu siklus waktu. Berbeda dengan RNN konvensional, LSTM memiliki komponen lainnya yang disebut *gate*, yang bertugas untuk mengontrol aliran data dalam jaringan [29]. Komponen utama pada LSTM terdiri dari tiga *gate* dan satu *state*, yaitu:

1. Forget Gate

Forget gate berfungsi untuk menentukan informasi mana dari *cell state* sebelumnya yang perlu dilupakan. Gerbang ini menggunakan fungsi aktivasi sigmoid yang menghasilkan nilai antara 0 hingga 1, di mana nilai 0

menyatakan “lupakan sepenuhnya” dan 1 menyatakan “pertahankan sepenuhnya.” Dengan cara ini, LSTM dapat membuang informasi yang tidak relevan dari memori jangka panjangnya dan fokus pada data penting. [9].

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

Sumber : [34]

Keterangan:

f_t = output dari *forget gate*

Wf = bobot *forget gate*

h_{t-1} = *hidden state* sebelumnya

x_t = input saat ini

bf = bias *forget gate*

σ = fungsi aktivasi sigmoid

2. *Input Gate*

Input gate bertugas mengatur informasi baru apa yang akan dimasukkan ke dalam *cell state*. Mekanisme ini terdiri dari dua bagian, yaitu fungsi sigmoid untuk menentukan elemen mana yang akan diperbarui, dan fungsi tanh untuk membuat vektor kandidat nilai baru. Kombinasi keduanya memungkinkan LSTM untuk menyimpan informasi baru yang dianggap penting dalam jangka waktu tertentu.[9].

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

Sumber : [34]

Keterangan:

i_t = output dari *input gate*

Wi = bobot *input gate*

b_i = bias *input gate*

h_{t-1} = *hidden state* sebelumnya

x_t = input saat ini

3. *Output Gate*

Output gate berperan dalam menentukan bagian mana dari *cell state* yang akan dijadikan *output* untuk langkah waktu berikutnya. Nilai ini dihitung berdasarkan kombinasi dari *cell state* yang diperbarui dan hasil dari fungsi sigmoid yang menentukan bagian mana yang akan diekspose. Output ini kemudian juga akan digunakan sebagai *hidden state* pada langkah berikutnya dalam jaringan LSTM. [9].

$$o_t = \sigma(W_o \cdot [h_t^{-1}, x_t] + b_o) \quad (3)$$

Sumber : [34]

Keterangan:

o_t = output dari *output gate*

h_t = *hidden state*

W_o = bobot *output gate*

b_o = bias *output gate*

4. *Cell State*

Cell state merupakan jalur utama yang membawa informasi memori jangka panjang sepanjang proses dalam jaringan LSTM. Nilainya diperbarui secara dinamis berdasarkan interaksi antara *forget gate* yang menentukan informasi mana yang dilupakan, dan *input gate* yang menentukan informasi baru yang akan ditambahkan. Dengan mekanisme ini, LSTM mampu mempertahankan atau memperbarui informasi secara selektif, sehingga memori tetap terjaga sepanjang urutan waktu.

$$C_t = f_t * C_t^{-1} + i_t * \hat{C}_t \quad (4)$$

Sumber : [34]

Keterangan:

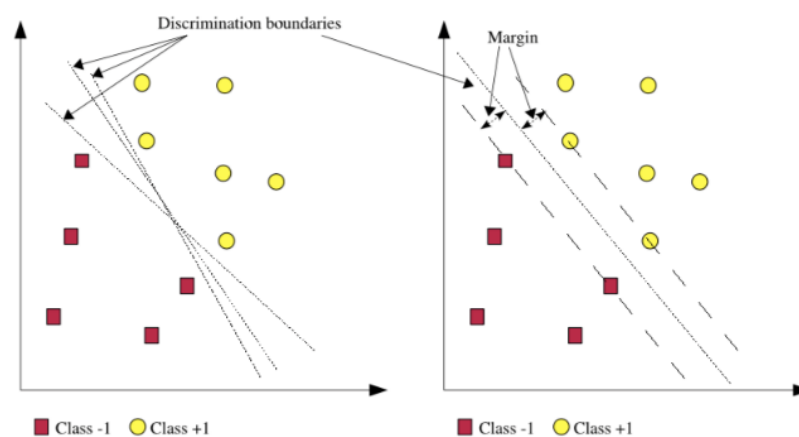
C_t = *cell state* saat ini

C_{t-1} = *cell state* sebelumnya

* = operasi perkalian per elemen

2.3.3 Support Vector Machine (SVM)

Support Vector Machine (SVM) merupakan salah satu algoritma *machine learning* yang digunakan untuk menyelesaikan permasalahan klasifikasi dan regresi. Mengacu kepada Gambar 2.3, SVM bekerja dengan mencari sebuah *hyperplane* atau garis pemisah terbaik yang mampu memisahkan data ke dalam dua kelas yang berbeda dengan margin maksimum [35]. Margin pada konteks ini adalah jarak antara *hyperplane* dengan titik data terdekat dari masing-masing kelas, yang dikenal sebagai *support vector*.



Gambar 2. 3 Ilustrasi SVM

Sumber : [35]

Semakin besar margin yang dihasilkan, maka semakin baik model dalam melakukan generalisasi terhadap data baru. Oleh karena itu, prinsip utama dari SVM adalah mencari *hyperplane* dengan margin maksimum, agar model tidak hanya mampu mengklasifikasikan data yang sudah diketahui, tetapi juga memiliki kemampuan prediksi yang baik terhadap data yang belum pernah dilihat sebelumnya.

2.4 Teori Tools dan Software Penelitian

2.4.1 Python

Python merupakan bahasa pemrograman yang dikenal terkenal karena fleksibilitas dan *object oriented program* [36]. Keunggulan utama dari Python terletak pada kode yang cukup singkat dan mudah dibaca. Python juga menyediakan *library* yang memungkinkan untuk membangun berbagai tipe aplikasi, mulai dari pengembangan *web*, kecerdasan buatan, hingga pengolahan data dan analisis statistik. Selain itu, *Python* dapat dijalankan secara lintas platform, baik pada sistem operasi *Windows*, *macOS*, maupun *Linux*, yang membuatnya sangat adaptif terhadap berbagai kebutuhan pengembangan perangkat lunak. Keunggulan lainnya adalah sifat *open source*, di mana kode sumber *Python* dapat diakses, dimodifikasi, dan dikembangkan secara lebih fleksibel oleh siapapun di seluruh dunia.

2.4.2 Jupyter Notebook

Jupyter Notebook adalah salah satu *tools* pemrograman berbasis web yang cukup lumrah digunakan dalam bidang *data science* dan pendidikan. Platform ini memungkinkan pengguna nya untuk menuliskan kode, menyisipkan dokumentasi, serta menampilkan visualisasi data secara terintegrasi dalam satu *interface* yang interaktif [37]. Fitur unggulan *Jupyter Notebook* antara lain adalah dukungan terhadap penulisan *markdown* untuk memberikan keterangan pada kode, serta kemampuannya dalam menyimpan *file* dalam berbagai *format* yang mendukung fleksibilitas bagi pengguna. Di samping itu, *Jupyter Notebook* juga mendukung proses transfer dan pelatihan ulang model antar *instance*, sehingga sangat memudahkan pengguna dalam implementasi machine learning dan penelitian yang lebih kompleks.