

BAB 1

PENDAHULUAN

1.1 Latar Belakang Masalah

Dalam beberapa tahun terakhir, perkembangan teknologi digital beserta kecerdasan buatan (*Artificial Intelligence/AI*) menunjukkan peningkatan yang sangat pesat. Implementasi AI pada berbagai sektor seperti industri manufaktur, layanan keuangan, transportasi, serta komputasi awan telah meningkatkan kebutuhan terhadap infrastruktur komputasi berperforma tinggi yang didukung oleh komponen semikonduktor. Menurut laporan industri, nilai pasar global semikonduktor diproyeksikan dapat mencapai lebih dari 1 triliun USD pada tahun 2030, meningkat hampir dua kali lipat dibandingkan nilai pasar sekitar 600 miliar USD pada awal dekade ini [1]. Semikonduktor merupakan elemen penting dalam perangkat elektronik modern seperti prosesor, unit pemrosesan grafis (GPU), serta memori komputer yang digunakan dalam sistem AI dan analisis data skala besar.

Peningkatan permintaan terhadap semikonduktor berdampak langsung pada kinerja perusahaan-perusahaan di sektor tersebut. Salah satu instrumen investasi yang merepresentasikan kinerja sektor semikonduktor adalah *Exchange Traded Fund* (ETF). *Exchange Traded Fund* (ETF) adalah produk investasi yang diperjualbelikan di pasar bursa dengan portofolio yang terdiri atas sejumlah saham untuk merepresentasikan suatu indeks maupun sektor tertentu. Dibandingkan reksa dana konvensional, ETF umumnya menawarkan likuiditas yang lebih baik serta biaya pengelolaan yang lebih efisien. [2]. *VanEck Semiconductor* ETF dengan kode *ticker* / symbol (SMH) merupakan salah satu ETF yang berfokus pada perusahaan-perusahaan semikonduktor global dan dikenal memiliki kinerja yang kuat, dengan total aset kelolaan mencapai lebih dari 20 miliar USD serta tingkat konsentrasi kepemilikan pada perusahaan teknologi besar yang mendominasi industri semikonduktor [3].

Pada pasar keuangan, proses memperkirakan harga aset bukanlah tugas yang sederhana karena perubahan nilainya memiliki pola yang tidak linear, terus berubah seiring waktu, serta dipengaruhi oleh banyak faktor, antara lain kondisi ekonomi global, sentimen investor, dan perkembangan teknologi. Data historis harga aset keuangan menunjukkan volatilitas yang tinggi serta pola fluktuasi jangka pendek dan jangka panjang yang sulit diprediksi menggunakan metode statistik

konvensional. Metode analisis tradisional seperti regresi linear sering kali memiliki keterbatasan dalam menangkap hubungan temporal yang panjang pada data runtun waktu (*time series*). Oleh karena itu, pendekatan berbasis *machine learning* dan *deep learning* mulai banyak digunakan dalam prediksi harga saham dan ETF.

Ketidakmampuan dalam memprediksi pergerakan harga SMH secara memadai dapat menimbulkan dampak terhadap pengambilan keputusan investasi. Kesalahan prediksi dapat menyebabkan investor terlambat mengambil posisi beli, menjual aset pada waktu yang kurang tepat, atau salah memperkirakan risiko yang sedang berlangsung. Pada kondisi pasar yang mengalami volatilitas tinggi, kesalahan tersebut dapat memperbesar potensi kerugian dan mengurangi peluang memperoleh keuntungan.

Selain itu, tanpa model prediksi yang mampu memanfaatkan pola historis secara sistematis, analisis pergerakan SMH hanya bergantung pada interpretasi subjektif atau indikator yang tidak selalu mampu menangkap hubungan temporal jangka panjang. Oleh karena itu, penelitian ini penting dilakukan untuk menyediakan pendekatan berbasis data yang dapat membantu mengukur pola *return* dan harga SMH secara lebih objektif. Model yang dihasilkan tidak dimaksudkan sebagai jaminan keuntungan, tetapi sebagai alat bantu analisis dalam memahami kemungkinan pergerakan harga pada beberapa *horizon* prediksi.

Long Short-Term Memory (LSTM) merupakan salah satu model *deep learning* yang dirancang untuk memproses data berurutan dengan mempertahankan informasi penting dari periode sebelumnya. Model ini merupakan pengembangan dari *Recurrent Neural Network* (RNN) yang mampu mengatasi permasalahan *vanishing gradient* sehingga lebih efektif dalam mempelajari ketergantungan jangka panjang pada data [4]. Penelitian empiris menunjukkan bahwa model berbasis *deep learning* dapat meningkatkan akurasi prediksi harga aset keuangan dibandingkan metode tradisional. Chen (2025) membandingkan beberapa model *machine learning* untuk memprediksi arah perubahan harga VanEck Semiconductor ETF (SMH) pada *horizon* H+5. Hasil penelitian menunjukkan bahwa model dengan performa terbaik memperoleh *test accuracy* sekitar 56,64% dan AUC sebesar 58,76% [5]. Hasil tersebut menunjukkan bahwa prediksi arah pergerakan SMH tetap merupakan permasalahan yang sulit, sehingga diperlukan penelitian tambahan dengan pendekatan regresi nilai kontinu. Selain itu, penelitian lain menunjukkan bahwa model LSTM yang dikombinasikan dengan variabel sentimen pasar dapat meningkatkan kemampuan prediksi tren industri semikonduktor secara signifikan dibandingkan model statistik dasar [6].

Meskipun berbagai penelitian telah mengkaji penggunaan *machine learning* dalam prediksi harga saham, penelitian yang secara khusus berfokus pada prediksi harga ETF sektor semikonduktor menggunakan model *deep learning* masih relatif terbatas. Oleh karena itu, penelitian ini bertujuan untuk mengimplementasikan model *Long Short-Term Memory* (LSTM) dalam memprediksi *return* harga VanEck *Semiconductor* ETF (SMH) berdasarkan data historis, yang kemudian dikonversi menjadi harga prediksi.

Selain membandingkan model yang digunakan, penelitian-penelitian terdahulu juga mengevaluasi performa model menggunakan metrik evaluasi seperti *Mean Absolute Error* (MAE), *Root Mean Squared Error* (RMSE), dan *Mean Absolute Percentage Error* (MAPE) [7]. Penggunaan ketiga metrik tersebut memungkinkan peneliti melakukan perbandingan performa model secara objektif karena masing-masing metrik memiliki karakteristik yang berbeda dalam mengukur tingkat kesalahan prediksi. Oleh karena itu, penelitian ini menggunakan ketiga metrik tersebut agar hasil evaluasi dapat dibandingkan dengan penelitian-penelitian sebelumnya secara lebih komprehensif.

MAE digunakan untuk menunjukkan rata-rata besar kesalahan prediksi dalam satuan harga sehingga mudah diinterpretasikan. RMSE digunakan untuk mengetahui apakah terdapat kesalahan prediksi yang besar karena metrik ini memberikan penalti lebih tinggi terhadap *error* yang besar. Sementara itu, MAPE digunakan untuk menunjukkan besar kesalahan dalam bentuk persentase sehingga memudahkan interpretasi tingkat akurasi model. Dengan menggunakan ketiga metrik tersebut secara bersamaan, penelitian ini tidak hanya mengetahui rata-rata kesalahan model, tetapi juga dapat memahami distribusi kesalahan prediksi secara lebih menyeluruh, termasuk apakah model menghasilkan *error* yang relatif stabil atau masih terdapat beberapa prediksi yang menyimpang jauh dari nilai aktual.

Dengan demikian, distribusi *error* dapat dianalisis secara lebih menyeluruh. Apabila nilai RMSE jauh lebih besar daripada MAE, terdapat beberapa prediksi dengan kesalahan yang relatif ekstrem. Apabila MAE dan RMSE memiliki nilai yang berdekatan, distribusi *error* cenderung lebih stabil. MAPE kemudian memberikan gambaran mengenai proporsi rata-rata kesalahan terhadap harga aktual.

Sebagian besar penelitian *deep learning* pada prediksi harga saham menggunakan *dataset* dengan periode yang relatif panjang [5]. Namun, dalam praktiknya, tidak semua instrumen investasi memiliki data historis yang panjang atau relevan dengan kondisi pasar terkini. Oleh karena itu, penelitian ini

mengevaluasi kemampuan model LSTM menggunakan data historis selama lima tahun pada VanEck *Semiconductor* ETF (SMH), sekaligus menganalisis apakah model masih mampu menghasilkan prediksi dengan tingkat kesalahan yang rendah.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, maka rumusan masalah dalam penelitian ini adalah sebagai berikut:

1. Bagaimana memprediksi harga VanEck *Semiconductor* ETF (SMH) menggunakan *Long Short-Term Memory* (LSTM)?
2. Berapa tingkat kesalahan (error) *Long Short-Term Memory* (LSTM) dalam memprediksi *return* harga VanEck *Semiconductor* ETF (SMH) diukur dengan metrik *Mean Absolute Error* (MAE), *Root Mean Squared Error* (RMSE), dan *Mean Absolute Percentage Error* (MAPE)?

1.3 Batasan Permasalahan

Agar penelitian yang dilakukan lebih terarah dan memiliki ruang lingkup yang jelas, maka penelitian ini memiliki beberapa batasan permasalahan sebagai berikut:

1. Penelitian ini memprediksi *return* harga VanEck *Semiconductor* ETF (SMH) sebagai variabel *target* dalam proses pemodelan, yang selanjutnya dikonversi menjadi nilai harga prediksi.
2. Data yang digunakan dalam penelitian ini merupakan data runtun waktu (*time series*) historis yang diperoleh dari sumber data pasar keuangan publik dengan rentang waktu dari 5 Juni 2021 hingga 18 April 2026.
3. Penelitian ini tidak mempertimbangkan faktor eksternal seperti sentimen pasar, berita ekonomi, maupun indikator makroekonomi lainnya yang dapat mempengaruhi pergerakan harga ETF.

1.4 Tujuan Penelitian

Tujuan penelitian ini adalah sebagai berikut:

1. Memprediksi *return* harga VanEck *Semiconductor* ETF (SMH) menggunakan *Long Short-Term Memory* (LSTM).
2. Mengevaluasi tingkat kesalahan (*error*) *Long Short-Term Memory* (LSTM) dalam memprediksi *return* harga VanEck *Semiconductor* ETF (SMH) diukur dengan metrik *Mean Absolute Error* (MAE), *Root Mean Squared Error* (RMSE), dan *Mean Absolute Percentage Error* (MAPE).

1.5 Manfaat Penelitian

Manfaat utama yang ingin didapatkan dari penelitian ini adalah sebagai berikut:

1. Memberikan wawasan mengenai penerapan *Long Short-Term Memory* (LSTM) dalam membangun model prediksi *return* harga *Exchange Traded Fund* (ETF) sektor semikonduktor berdasarkan data historis, yang kemudian dikonversi menjadi harga prediksi.
2. Memberikan gambaran mengenai pola pergerakan harga VanEck *Semiconductor* ETF (SMH) yang dapat dipelajari melalui pemodelan data runtun waktu menggunakan pendekatan *deep learning*.
3. Menjadi referensi awal untuk penelitian selanjutnya yang berkaitan dengan prediksi harga aset keuangan, khususnya ETF sektor teknologi, menggunakan metode *deep learning*.

1.6 Sistematika Penulisan

Sistematika penulisan laporan adalah sebagai berikut:

- Bab 1 PENDAHULUAN
Bab ini menjelaskan gambaran umum penelitian yang meliputi latar belakang, rumusan masalah, batasan penelitian, tujuan penelitian, serta manfaat yang diharapkan. Selain itu, bab ini menguraikan alasan pentingnya penelitian mengenai prediksi harga aset keuangan, khususnya *Exchange Traded Fund* (ETF) pada sektor semikonduktor, beserta pendekatan *deep learning* yang digunakan sebagai solusi dalam penelitian ini.

- Bab 2 LANDASAN TEORI

Bab ini menyajikan landasan teori dan kajian pustaka yang menjadi dasar pelaksanaan penelitian. Pembahasan mencakup konsep *time series*, *Exchange Traded Fund* (ETF), *deep learning*, *Recurrent Neural Network* (RNN), *Long Short-Term Memory* (LSTM), metrik evaluasi performa prediksi, serta berbagai penelitian terdahulu yang memiliki keterkaitan dengan prediksi harga aset pada pasar keuangan.

- Bab 3 METODOLOGI PENELITIAN

Bab ini menguraikan metodologi penelitian yang diterapkan dalam pengembangan model prediksi. Pembahasan meliputi tahapan penelitian, proses pengumpulan data historis harga VanEck *Semiconductor* ETF (SMH), tahap prapemrosesan data, pembentukan *sliding window*, perancangan arsitektur model LSTM, proses pelatihan model, hingga metode yang digunakan untuk mengevaluasi performa model prediksi.

- Bab 4 HASIL DAN DISKUSI

Bab ini membahas implementasi model yang telah dikembangkan beserta analisis terhadap hasil prediksi yang diperoleh. Uraian di dalamnya meliputi proses prapemrosesan data, eksplorasi dan visualisasi data, pembangunan serta pelatihan model LSTM, hasil prediksi harga ETF, dan evaluasi kinerja model berdasarkan metrik kesalahan prediksi yang digunakan dalam penelitian.

- Bab 5 KESIMPULAN DAN SARAN

Bab terakhir menyampaikan kesimpulan yang diperoleh berdasarkan hasil penelitian serta memberikan rekomendasi untuk penelitian berikutnya. Selain merangkum temuan utama mengenai kemampuan model LSTM dalam memprediksi harga ETF sektor semikonduktor, bab ini juga membahas peluang pengembangan metode prediksi agar dapat menghasilkan performa yang lebih baik pada penelitian selanjutnya.