

BAB I

PENDAHULUAN

1.1 Latar Belakang

Saham menjadi instrumen utama yang banyak diminati di pasar keuangan serta memberikan kontribusi penting dalam pergerakan aktivitas ekonomi [1]. Meskipun investasi saham berpotensi memberikan keuntungan yang besar, instrumen ini juga memiliki risiko yang tinggi akibat fluktuasi nilai yang dapat berubah secara tajam [2]. Penawaran dan permintaan menjadi faktor utama yang menentukan fluktuasi harga saham, di mana meningkatnya minat beli akan mendorong harga naik, sedangkan tingginya keinginan menjual akan menekan harga turun [3]. Kondisi ini membuat lembaga sekuritas maupun investor memiliki ketertarikan yang tinggi dalam mengidentifikasi dan menganalisis tren pergerakan harga saham [4]. Namun, perubahan harga saham umumnya menunjukkan pola yang bersifat non-linier, sehingga proses peramalan menjadi lebih kompleks dan sulit dilakukan secara akurat [5]. Oleh karena itu, diperlukan metode prediksi harga saham yang tepat agar dapat mendukung pengambilan keputusan investasi secara lebih optimal.

Berbagai penelitian telah dilakukan untuk mengembangkan metode prediksi harga saham. Salah satu peneliti membandingkan algoritma *Random Forest* dan *Support Vector Machine* (SVM). Hasil penelitian menunjukkan bahwa *Random Forest* memberikan akurasi yang lebih tinggi dibandingkan SVM, yaitu sekitar 80.8% untuk *Random Forest* dan 78.7% untuk SVM pada data pengujian [6]. Namun, karena *Random Forest* merupakan kombinasi banyak pohon keputusan, model ini cenderung kurang mudah dipahami [7]. Penelitian lainnya menggunakan *Long Short Term Memory* (LSTM) untuk memprediksi 10 saham acak. Hasil penelitian menunjukkan bahwa rata-rata akurasi prediksi mencapai sekitar 83% pada seluruh saham yang diuji, dengan akurasi tertinggi sebesar 83,33% diperoleh pada saham SBI [8]. Model ini menunjukkan kinerja yang baik dalam menangani beragam permasalahan yang berkaitan dengan prediksi deret waktu pada berbagai bidang [9]. Meskipun demikian, LSTM memiliki keterbatasan dalam memahami konteks secara menyeluruh karena hanya memproses data ke arah maju (*forward*) [10]. Akibatnya, informasi dari masa mendatang tidak dapat dimanfaatkan secara optimal sehingga berpotensi menurunkan akurasi prediksi pada data deret waktu yang kompleks [11].

Metode *Bidirectional Long Short-Term Memory* (BiLSTM) diterapkan untuk mengatasi keterbatasan LSTM, karena metode ini memproses data ke arah maju (*forward*) dan ke arah mundur (*backward*), sehingga mampu menangkap konteks lebih menyeluruh

dan meningkatkan akurasi prediksi [12]. Sebuah penelitian menggunakan BiLSTM untuk memprediksi harga penutupan indeks keuangan CSI 300 di China. Hasilnya BiLSTM mencapai akurasi prediksi tertinggi dibanding LSTM, SVR dan Arima dengan RMSE sebesar 0.6056 dan MAPE sebesar 0.4482 [13]. Penelitian lainnya menunjukkan bahwa model BiLSTM menghasilkan kinerja prediksi harga saham syariah yang lebih baik dibandingkan LSTM, dengan nilai MAPE masing-masing sebesar 2,59% pada saham ANTM.JK, 1,77% pada UNVR.JK, dan 1,05% pada INDF.JK, yang seluruhnya termasuk dalam kategori sangat akurat [14]. Berdasarkan penelitian tersebut, model BiLSTM menunjukkan kinerja yang lebih unggul dalam melakukan prediksi. Namun, model ini juga rentan terhadap kinerja yang buruk jika konfigurasi *hyperparameter* tidak tepat sehingga dapat berdampak pada hasil prediksi [15].

Untuk mengatasi permasalahan tersebut, teknik optimasi diperlukan sebagai upaya mengefisienkan pencarian *hyperparameter* terbaik [16]. Salah satu pendekatan yang digunakan ialah *Improved Firefly Algorithm* (IFA) yang dimanfaatkan untuk menyempurnakan struktur jaringan saraf tiruan sehingga dapat meningkatkan akurasi dan performa model [17]. Algoritma IFA telah terbukti memiliki performa yang lebih unggul dibandingkan berbagai algoritma optimasi lainnya, seperti *simulated annealing* (SA), *niche genetic algorithm* (NGA), *improved discrete cuckoo search* (IDCS), *mutant firefly algorithm* (MFA), *cuckoo–firefly search* (CSFA), serta *best-neighborhood firefly algorithm* (BNFA) [18]. Pada sebuah penelitian, metode IFA diterapkan untuk mengoptimalkan *hyperparameter* model BiLSTM dalam melakukan prediksi harga Bitcoin. Hasil penelitian menunjukkan bahwa model IFA-BiLSTM mencapai RMSE sebesar 2051,55 dan MAPE sebesar 3,48% [19].

Berdasarkan uraian latar belakang tersebut, maka akan dilakukan penelitian yang berfokus pada **“Implementasi BiLSTM dengan Improved FA untuk Prediksi Harga Saham”**.

1.2 Rumusan Masalah

Berdasarkan uraian pada latar belakang, maka rumusan masalah dalam penelitian ini dapat dinyatakan sebagai berikut:

1. Bagaimana memperoleh konfigurasi *hyperparameter* yang optimal agar model BiLSTM menghasilkan performa prediksi harga saham yang lebih baik?
2. Apakah optimasi *hyperparameter* menggunakan *Improved Firefly Algorithm* (IFA) dapat meningkatkan performa model BiLSTM dibandingkan tanpa optimasi?

1.3 Tujuan

Penelitian ini bertujuan memperoleh konfigurasi *hyperparameter* yang optimal pada model BiLSTM untuk prediksi harga saham menggunakan *Improved Firefly Algorithm* (IFA) serta menganalisis peningkatan performa model BiLSTM dibandingkan model tanpa optimasi.

1.4 Manfaat

Adapun manfaat yang diharapkan dari penelitian ini adalah sebagai berikut:

1. Menghasilkan model BiLSTM dengan konfigurasi *hyperparameter* yang optimal melalui penerapan *Improved Firefly Algorithm* (IFA) sehingga diperoleh performa prediksi yang lebih baik.
2. Menjadi acuan bagi peneliti selanjutnya dalam pengembangan metode prediksi deret waktu, khususnya melalui penerapan BiLSTM yang dioptimalkan menggunakan *Improved Firefly Algorithm* (IFA).

1.5 Ruang Lingkup

Penelitian ini memiliki ruang lingkup sebagai berikut:

1. Dataset historis harga saham diperoleh dari situs Yahoo Finance untuk saham PT Aneka Tambang Tbk (ANTM.JK), PT Indofood Sukses Makmur Tbk (INDF.JK), dan PT Unilever Indonesia Tbk (UNVR.JK) dengan periode 2 Januari 2020 sampai dengan 30 Desember 2025.
2. Dataset yang digunakan merupakan data historis saham ANTM.JK, INDF.JK, dan UNVR.JK yang diperdagangkan dalam mata uang Indonesian Rupiah (IDR).
3. Setiap dataset terdiri dari sekitar 1447 baris dengan enam atribut utama, yaitu: *date*, *open*, *high*, *low*, *close*, dan *volume*.
4. Dataset dibagi menjadi tiga bagian, yaitu 70% sebagai data pelatihan (*training*), 10% sebagai data validasi (*validation*), dan 20% sebagai data pengujian (*testing*), dengan proses pembagian dilakukan secara berurutan sesuai urutan waktu data.
5. *Hyperparameter* yang disesuaikan menggunakan IFA meliputi jumlah *neuron* pada *hidden layer*, *batch size*, serta persentase *dropout* yang digunakan.
6. Pengujian terhadap model dilakukan menggunakan metrik MAE, RMSE, dan MAPE untuk mengevaluasi kinerja dari model yang diusulkan.