

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Pasar modal memiliki peran strategis dalam perekonomian Indonesia dengan Bursa Efek Indonesia (BEI) sebagai wadah utamanya. Di antara berbagai instrumen yang tersedia, saham BBCA konsisten menarik perhatian investor karena berstatus sebagai *blue-chip* dengan kapitalisasi pasar terbesar di BEI [1]. Saham BBCA menampilkan karakteristik yang unik sekaligus menantang, dimana tren pertumbuhannya linear jangka panjang yang stabil berkat fundamentalnya yang kokoh, namun secara simultan diwarnai oleh volatilitas non-linear jangka pendek akibat sensitivitasnya terhadap makro ekonomi dan arus modal asing. Kehadiran pola linear dan non-linear yang terjadi bersamaan ini menjadikan peramalan harga saham BBCA sebagai objek yang ideal untuk menguji keandalan model kombinasi.

Secara umum, data harga saham merupakan *time series* kompleks yang bersifat non-linear, tidak stasioner, dan bervolatilitas tinggi, sehingga pemodelan menggunakan metode linear tunggal sering kali menghasilkan akurasi yang sub-optimal [2]. Akibatnya, pemodelan dengan satu pendekatan tunggal sering kali tidak memadai. Model ARIMA sangat andal dalam menangkap pola linear stokastik dan autokorelasi, tetapi gagal memetakan dinamika non-linear [2]. Metode ES efektif dalam menangkap tren dan musiman jangka panjang, namun memiliki keterbatasan linearitas yang sama dengan ARIMA [3]. Sementara itu, arsitektur *deep learning* seperti LSTM terbukti unggul dalam mempelajari dependensi jangka panjang dan pola non-linear yang rumit [4]. Meski demikian, pada fase data yang didominasi tren linear kuat, LSTM terkadang kurang efektif dibandingkan model statistik konvensional.

Keterbatasan model tunggal ini mendorong berkembangnya pendekatan *hybrid parallel* untuk menggabungkan kekuatan masing-masing metode [2]. Namun, tantangan utamanya terletak pada penentuan bobot kontribusi setiap model secara dinamis. Penggunaan bobot yang sama (*equal weighting*) atau penentuan manual sering kali menghasilkan prediksi sub-optimal karena gagal beradaptasi dengan perubahan karakteristik data [5]. Untuk mengatasi masalah tersebut, GA digunakan secara spesifik dalam penelitian ini untuk mengoptimasi bobot

kombinasi akhir ( $w_1, w_2, w_3$ ) dengan kendala matematis  $w_1 + w_2 + w_3 = 1$ , alih-alih digunakan untuk menyetel *hyperparameter* internal model. Melalui proses seleksi, *crossover*, dan mutasi secara iteratif, GA bertindak sebagai pencari solusi global yang meminimalkan fungsi kesalahan prediksi [6].

Meskipun kajian model kombinasi pada data keuangan telah banyak dilakukan, terdapat beberapa kesenjangan penelitian (*research gap*) yang mendasar dalam literatur terdahulu. Zhang dan Peter (2003) mengembangkan model *hybrid* berbasis ARIMA dan *Artificial Neural Network* (ANN) sederhana untuk menangkap pola linear dan non-linear, namun model ini memiliki keterbatasan krusial dalam menangkap ketergantungan jangka panjang (*long-term dependencies*) pada data [2]. Untuk menangkap tren makro, Smyl dan Slawek (2020) mengintegrasikan model ES dan *Recurrent Neural Networks* (RNN), tetapi arsitektur tersebut tidak secara eksplisit memasukkan model autokorelasi linear ARIMA yang sangat penting untuk karakteristik data stokastik [7].

Di sisi lain, peneliti yang mulai menyadari pentingnya mengintegrasikan ketiga model heterogen ini umumnya masih terjebak pada metode kombinasi konvensional. Sebagai contoh, Safi et al. (2021) mengombinasikan ARIMA, ES, dan ANN (*Neural Network Autoregressive* (NNAR)), namun penentuan bobot akhir prediksinya masih menggunakan pendekatan *equal weight* yang kaku [5]. Pendekatan tersebut terbukti memberikan hasil yang kurang optimal dibandingkan model tunggal karena gagal beradaptasi dengan dinamika data yang berubah. Sementara itu, potensi penggunaan algoritma evolusioner untuk optimasi pembobotan telah ditunjukkan oleh Ahn et al. (2006) yang sukses menerapkan metode GA untuk pencarian bobot otomatis dan pemilihan fitur secara simultan, di mana pendekatan tersebut terbukti mampu meningkatkan akurasi estimasi secara signifikan pada model *hybrid* konvensional [8]. Sepanjang penelusuran literatur yang dilakukan, belum ada penelitian yang mengadaptasi kekuatan optimasi bobot berbasis GA tersebut ke dalam arsitektur *parallel* ARIMA-ES-LSTM, khususnya yang diimplementasikan untuk data saham BBKA.

Celah empiris inilah yang menjadi landasan utama penelitian ini. Kebaruan (*novelty*) yang ditawarkan adalah pengembangan arsitektur *Hybrid Parallel Forecasting* yang memproses data aktual saham BBKA secara bersamaan pada ketiga model untuk menghasilkan prediksi independen, yang kemudian dikombinasikan berdasarkan pembobotan optimal secara otomatis untuk meminimalkan kesalahan akumulatif [9]. Pendekatan *parallel* ini berbeda dengan arsitektur *hybrid serial* konvensional yang umumnya hanya menggunakan model

non-linear untuk mengoreksi sisaan (*error residual*). Selain itu, pengujian performa dilakukan secara komprehensif melintasi rentang waktu multi periode, meliputi jangka pendek (7 hari), jangka menengah (30 hari), dan jangka panjang (90 hari) untuk menguji resiliensi model terhadap perubahan periode prediksi.

Berdasarkan uraian tersebut, penelitian ini bertujuan melakukan komparasi sistematis antara model tunggal (ARIMA, ES, LSTM) dengan model *Hybrid Parallel* ARIMA-ES-LSTM berbasis optimasi GA. Evaluasi performa model diuji secara rigor menggunakan metrik akurasi standar, yaitu *Root Mean Square Error* (RMSE), *Mean Square Error* (MAE) dan *Mean Absolute Percentage Error* (MAPE) guna membuktikan keunggulan strategi kombinasi ini dalam menghadapi kompleksitas data saham BBKA.

## 1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, maka permasalahan yang akan dikaji dalam penelitian ini dapat dirumuskan sebagai berikut:

1. Bagaimana membangun arsitektur model *hybrid* yang mengintegrasikan model ARIMA, ES, dan LSTM secara simultan untuk menangkap karakteristik linear dan non-linear pada data *time series*?
2. Sejauh mana efektivitas dan peningkatan akurasi model *hybrid* yang dioptimasi dengan GA dibandingkan dengan penggunaan model tunggal (ARIMA, ES, atau LSTM) saja?

## 1.3 Batasan Masalah

Dari rumusan masalah yang ada, terdapat beberapa batasan masalah yang menjadi fokus pembahasan dalam penelitian ini, yakni sebagai berikut.

1. Data yang digunakan adalah data sekunder berupa data *time series* dari harga harian penutupan saham BBKA dari tanggal 28 Mei 2018 sampai dengan 21 Mei 2026.
2. Prediksi dan analisis hasil pemodelan dibatasi untuk tiga jangka waktu masa depan (periode), yaitu 7 hari (jangka pendek), 30 hari (jangka menengah), dan 90 hari (jangka panjang).
3. Model ARIMA dibatasi pada penentuan parameter  $(p, d, q)$  menggunakan pendekatan analisis *Autocorrelation Function* (ACF) dan *Partial Autocorrelation Function* (PACF).

4. Model ES yang digunakan adalah model yang relevan dengan pola data yang dipakai.
5. Model LSTM menggunakan *single layer* LSTM dengan menggunakan mekanisme *early stopping* dalam penentuan epoch optimal.
6. *Genetic Algorithm* hanya digunakan untuk mencari bobot kontribusi akhir dari ketiga model, bukan untuk mengoptimasi arsitektur internal LSTM.
7. Uji Pengukuran akurasi yang digunakan adalah RMSE, MAE dan MAPE.
8. Data yang digunakan dibagi menjadi dua dan tiga bagian utama, tergantung model yang dipakai, yaitu data latih, data validasi dan data uji/*testing*.
9. Simulasi menggunakan bahasa pemrograman Python.

#### 1.4 Tujuan Penelitian

Berdasarkan rumusan masalah, maka tujuan penelitian ini adalah sebagai berikut.

1. Mengembangkan kerangka kerja peramalan *hybrid* yang menggabungkan kekuatan statistik klasik dari ARIMA dan ES dengan kekuatan komputasi non-linear dari LSTM.
2. Mengevaluasi dan membandingkan kinerja model yang diusulkan untuk membuktikan bahwa integrasi ketiga model dengan optimasi GA memberikan hasil peramalan yang lebih akurat.

Hasil penelitian ini diharapkan dapat memberikan manfaat seperti memperkaya kajian dalam bidang peramalan, khususnya pada analisis model *hybrid* kombinasi ARIMA, ES dan LSTM dengan optimasi GA. Lalu, dapat menjadi referensi dan dasar pengembangan penelitian selanjutnya yang berkaitan dengan model *hybrid* yang lebih optimal.

#### 1.5 Metodologi Penelitian

Metodologi penelitian yang digunakan pada skripsi ini di antaranya:

##### 1. Studi Literatur

Penelitian ini diawali dengan studi literatur terhadap teori dan penelitian terdahulu yang berkaitan dengan peramalan *time series*, khususnya metode ARIMA, ES dan ANN (NNAR) serta pendekatan *hybrid* ARIMA–ES–ANN

(NNAR). Studi literatur ini menggabungkan beberapa referensi yang relevan dengan penelitian skripsi ini. Referensi yang digunakan berasal dari buku, jurnal, artikel ilmiah, dsb.

## 2. Analisis

Pada Studi Literatur ditemukan beberapa evaluasi terkait model NNAR yang memiliki kelemahan dalam menangkap ketergantungan jangka panjang (*long-term dependencies*) karena struktur *feedback* yang terbatas serta pendekatan *hybrid* menggunakan *equal weight* yang menganggap setiap model memiliki performa yang selalu sama, tanpa memperdulikan pola data yang berubah-ubah. Maka dari itu, pada penelitian skripsi ini akan mengganti salah satu model dan juga pendekatan *hybrid* yang digunakan, setelah itu akan dijabarkan analisis konsep dan mekanisme kerja model ARIMA, ES dan LSTM, termasuk bagaimana ke tiga model itu dikombinasikan dengan pendekatan GA untuk menangkap pola linear dan nonlinear dalam data *time series*. Analisis juga mencakup proses pembentukan model *hybrid* ARIMA-ES-LSTM yang menjadi dasar dalam tahap simulasi berikutnya.

## 3. Simulasi Numerik

Simulasi numerik akan dilakukan untuk melakukan peramalan serta membandingkan kinerja model tunggal dan model *hybrid* pada data *time series* yang memiliki pola linear dan nonlinear . Simulasi dilakukan dengan menggunakan bantuan *software* seperti python.

## 4. Kesimpulan dan Saran

Penarikan kesimpulan berdasarkan analisis yang telah dilakukan, serta memberikan rekomendasi untuk penelitian lanjutan nantinya.

### 1.6 Sistematika Penulisan

Sistematika penulisan dalam penelitian ini adalah:

#### **BAB I PENDAHULUAN**

Bab ini berisi mengenai latar belakang masalah, rumusan masalah, batasan masalah, tujuan penelitian, metode penelitian, dan sistematika penulisan.

## **BAB II LANDASAN TEORI**

Bab ini berisi tentang teori-teori yang melandasi pembahasan dalam studi literatur ini seperti Peramalan, *Time Series*, ANN, Uji Normalitas dan Analisis Residual, *Maximum Likelihood Estimation* (MLE) dan *Hybrid Model*. Teori inilah yang nantinya digunakan dalam pembahasan bab berikutnya.

## **BAB III METODE PENELITIAN**

Bab ini berisi penjelasan tentang ARIMA, ES, LSTM, GA dan model *Hybrid*.

## **BAB IV IMPLEMENTASI DAN SIMULASI**

Bab ini berisi penerapan studi kasus, hasil perhitungan serta analisis dari metode yang digunakan.

## **BAB V KESIMPULAN DAN SARAN**

Berisi kesimpulan dari pembahasan yang telah dianalisis serta saran untuk pengembangan penelitian selanjutnya.

