

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Pasar modal berperan sebagai mekanisme intermediasi keuangan yang menyalurkan dana dari pihak yang memiliki kelebihan dana kepada pihak yang membutuhkan pembiayaan untuk kegiatan investasi produktif. Mekanisme tersebut memungkinkan alokasi sumber daya ke sektor ekonomi yang memiliki peluang pertumbuhan sehingga meningkatkan efisiensi dalam sistem keuangan. Harga saham di pasar modal mencerminkan ekspektasi pelaku pasar terhadap kondisi ekonomi dan prospek kinerja perusahaan pada masa mendatang. Perubahan informasi ekonomi, kebijakan moneter, serta kondisi ekonomi global dapat mempengaruhi persepsi risiko investor sehingga memicu fluktuasi harga saham di pasar keuangan (Cheng et al., 2024).

Pergerakan harga saham dipengaruhi oleh interaksi antara faktor fundamental, faktor teknikal, dan sentimen pasar. Faktor fundamental merepresentasikan kondisi ekonomi makro dan kinerja keuangan perusahaan yang menjadi dasar dalam penilaian nilai intrinsik suatu saham. Faktor teknikal tercermin melalui pola historis harga dan volume perdagangan yang digunakan untuk mengidentifikasi tren dan perilaku pasar dari waktu ke waktu. Volatilitas harga saham menunjukkan karakteristik persistensi, di mana fluktuasi pada suatu periode cenderung berlanjut pada periode berikutnya. Penelitian berbasis model *Generalized Autoregressive Conditional Heteroskedasticity* (GARCH) dan *Markov Switching* menunjukkan bahwa volatilitas harga saham bersifat berkelanjutan dalam jangka pendek sebagai

respons pasar terhadap informasi ekonomi yang berkembang secara dinamis (Enow, 2023).

Prediksi harga saham termasuk dalam permasalahan *non-linear time series* karena data pasar saham bersifat dinamis, mengandung tingkat *noise* yang tinggi, serta mengalami perubahan rezim pasar (*regime shift*). Model statistik konvensional seperti regresi linier dan *Autoregressive Integrated Moving Average* (ARIMA) memiliki keterbatasan dalam menangkap hubungan non-linear antarvariabel pada data harga saham. Pendekatan *machine learning* menawarkan fleksibilitas yang lebih tinggi dalam memodelkan kompleksitas tersebut melalui pembelajaran pola dari data historis secara adaptif (Osman et al., 2025).

Data *Open, High, Low, Close, dan Volume* (OHLCV) merepresentasikan informasi utama mengenai pergerakan harga dan aktivitas perdagangan saham. Data tersebut mencerminkan dinamika harga serta tekanan permintaan dan penawaran yang terjadi di pasar. Sejumlah penelitian menunjukkan bahwa pemanfaatan data OHLCV pada model *machine learning* mampu meningkatkan akurasi prediksi harga saham dibandingkan metode statistik tradisional. Penelitian empiris membuktikan bahwa algoritma *Extreme Gradient Boosting* (XGBoost) efektif dalam memodelkan hubungan non-linear dan volatilitas harga saham berbasis data OHLCV (William et al., 2025).

Penggunaan data OHLCV menghadapi masalah multikolinearitas tinggi, terutama antarvariabel harga seperti *Open, High, dan Low* yang memiliki korelasi kuat. Kondisi ini berpotensi menurunkan stabilitas model dan meningkatkan risiko *overfitting*. *Principal Component Analysis* (PCA) digunakan sebagai metode reduksi dimensi untuk mentransformasikan variabel yang saling berkorelasi menjadi

sejumlah komponen utama yang tidak berkorelasi. Penerapan PCA terbukti mengurangi redundansi informasi dan meningkatkan efisiensi pembelajaran model prediksi (Mostafavi & Hooman, 2025).

Integrasi PCA dengan algoritma berbasis *tree* seperti *Extreme Gradient Boosting* (XGBoost) menunjukkan peningkatan performa prediksi pada data teknikal berdimensi tinggi. Transformasi fitur menggunakan PCA sebelum proses pelatihan memungkinkan model XGBoost bekerja pada ruang fitur yang lebih stabil dan informatif, sehingga meningkatkan kemampuan generalisasi model dalam memprediksi pergerakan harga saham (Nobre & Neves, 2019).

*Extreme Gradient Boosting* (XGBoost) merupakan algoritma *ensemble* berbasis *gradient boosting* yang banyak digunakan dalam prediksi harga saham karena efisiensi komputasi dan kemampuan generalisasi yang baik. Keunggulan XGBoost terletak pada penerapan *regularization*, kemampuan menangani data *sparse*, serta mekanisme penilaian kepentingan fitur (*feature importance*) yang memungkinkan model menangkap hubungan non-linear secara efektif (Chen & Guestrin, 2016).

Kinerja XGBoost sangat dipengaruhi oleh konfigurasi *hyperparameter* seperti *learning rate*, *max depth*, *n\_estimators*, dan *min\_child\_weight*. Penetapan nilai *hyperparameter* yang tidak optimal dapat menyebabkan model gagal menangkap pola non-linear secara akurat atau mengalami *overfitting*. Penyesuaian *hyperparameter* secara manual melalui metode *grid search* dan *random search* memerlukan waktu komputasi besar dan tidak selalu menghasilkan kombinasi parameter terbaik, terutama pada data harga saham harian yang berdimensi tinggi (Verma, 2024).

Keterbatasan tersebut mendorong penggunaan metode optimasi otomatis yang lebih efisien. Optuna merupakan *framework* optimasi *hyperparameter* berbasis *Sequential Model-Based Optimization* (SMBO) yang mampu menyesuaikan ruang pencarian secara adaptif berdasarkan hasil percobaan sebelumnya. Pendekatan ini memungkinkan proses pencarian *hyperparameter* lebih cepat dan efektif dibandingkan metode konvensional pada data keuangan yang volatil dan non-linear (Akiba et al., 2019)

Pemilihan PT Indofood Sukses Makmur Tbk sebagai objek penelitian didasarkan pada karakteristik fundamental perusahaan yang kuat dan stabil. Laporan Tahunan PT Indofood Sukses Makmur Tbk tahun 2023 mencatat penjualan neto sebesar Rp111,70 triliun dan peningkatan laba tahun berjalan dari Rp5,1 triliun pada tahun 2022 menjadi Rp6,3 triliun pada tahun 2023. Kinerja ini menunjukkan ketahanan fundamental perusahaan di tengah tekanan inflasi dan volatilitas ekonomi global, sehingga pergerakan harga saham Indofood relevan untuk dianalisis secara kuantitatif (PT Indofood Sukses Makmur Tbk, 2023)

Penelitian ini mengaplikasikan model XGBoost dengan pra-pemrosesan PCA dan optimasi *hyperparameter* menggunakan Optuna untuk memprediksi harga penutupan saham PT Indofood Sukses Makmur Tbk (INDF) berdasarkan data *OHLCV*. Evaluasi kinerja model dilakukan menggunakan metrik *Root Mean Squared Error* (RMSE), *Mean Absolute Error* (MAE), dan *Mean Absolute Percentage Error* (MAPE). Penelitian ini diharapkan memberikan kontribusi empiris dalam pengembangan model prediksi harga saham berbasis *machine learning* serta mendukung pengambilan keputusan investasi berbasis data di pasar modal Indonesia.

## 1.2 Rumusan Masalah

Rumusan masalah berdasarkan latar belakang tersebut adalah sebagai berikut:

1. Bagaimana performa model XGBoost dengan pengaturan *hyperparameter* default dalam memprediksi harga saham PT Indofood Sukses Makmur Tbk (INDF)?
2. Seberapa besar peningkatan akurasi yang diperoleh dari penerapan metode optimasi otomatis menggunakan Optuna dalam penentuan *hyperparameter* XGBoost?
3. Seberapa efisien metode Optuna berbasis *Bayesian Optimization* dalam menemukan kombinasi *hyperparameter* optimal untuk XGBoost dibandingkan dengan *Grid Search Optimization*?

## 1.3 Batasan Masalah

Batasan masalah dalam penelitian ini adalah:

1. Penelitian ini hanya menggunakan data historis harga saham harian PT Indofood Sukses Makmur Tbk periode 2 Januari 2019 - 29 November 2024 yang diperoleh dari Yahoo Finance, tanpa memasukkan variabel makroekonomi, sentimen pasar, maupun faktor fundamental perusahaan.
2. Model yang digunakan terbatas pada XGBoost dengan optimasi hyperparameter menggunakan Optuna serta preprocessing PCA, tanpa membandingkan dengan model deep learning atau metode ensemble lainnya.
3. Penelitian ini hanya memprediksi harga penutupan saham untuk satu periode ke depan ( $t+1$ ) sehingga tidak membahas prediksi jangka menengah atau jangka panjang maupun strategi trading aktual.

#### 1.4 Tujuan Penelitian

Tujuan yang akan dicapai dalam penelitian ini yaitu:

1. Menganalisis hasil prediksi harga saham PT Indofood Sukses Makmur Tbk (INDF) dengan membangun model *Extreme Gradient Boosting* (XGBoost) tanpa penerapan optimasi otomatis pada *hyperparameter*.
2. Menganalisis hasil prediksi harga saham PT Indofood Sukses Makmur Tbk (INDF) menggunakan model XGBoost yang dioptimalkan dengan algoritma Optuna sebagai metode *automated hyperparameter tuning*.
3. Membandingkan performa dan tingkat akurasi prediksi antara model XGBoost tanpa optimasi dan model XGBoost yang dioptimalkan menggunakan Optuna untuk menilai efektivitas pendekatan optimasi otomatis dalam memprediksi harga saham.