

BAB I

PENDAHULUAN

1.1 Latar Belakang

Kemajuan dalam teknologi informasi dan peningkatan kapasitas komputasi, penggunaan teknologi berbasis data semakin luas diterapkan di berbagai sektor, salah satunya adalah pasar modal (Rohyati dkk., 2024). Aktivitas perdagangan saham menghasilkan data historis dalam jumlah besar dengan karakteristik yang dinamis dan kompleks (Ramadhani, 2024). Volume data yang terus meningkat tersebut menyebabkan proses pengolahan data menjadi semakin menantang apabila dilakukan secara manual (Arif dkk., 2025).

Data historis saham memiliki peran penting dalam proses analisis pergerakan harga saham karena mencerminkan pola dan perilaku pasar di masa lalu (Zapar dkk., 2024). Informasi yang dihasilkan dari data ini memainkan peran penting dalam mendukung keputusan investasi yang diambil oleh investor (Edi dkk., 2023). Namun, tingginya fluktuasi harga saham serta kompleksitas hubungan antarvariabel harga menyebabkan proses analisis harga saham menjadi sulit apabila dilakukan secara manual.

Saham BRI merupakan salah satu contoh saham yang memiliki karakteristik data sebagaimana dibahas dalam penelitian ini. Perusahaan ini tergolong sebagai Badan Usaha Milik Negara (BUMN) perbankan terbesar di Indonesia dan memiliki kontribusi yang signifikan dalam mendukung stabilitas ekonomi nasional dengan fokus pada pelayanan kepada sektor usaha mikro, kecil, dan menengah (UMKM) (Jaya dkk., 2023). Dengan kode emiten BBRI, saham tersebut termasuk dalam kategori saham lapis satu (*blue chip*) dan secara konsisten menempati jajaran saham dengan nilai transaksi tertinggi di Bursa Efek Indonesia (Romdhon dkk., 2025). Per April 2026, kapitalisasi pasar BRI tercatat sebesar Rp518,33 triliun dengan pertumbuhan sejak 2004 mencapai +2.779% dan rata-rata volume perdagangan harian mencapai 237 juta lembar saham. Tingginya nilai kapitalisasi dan likuiditas tersebut menunjukkan kuatnya minat investor terhadap saham BRI (TradingView, 2026), sekaligus menegaskan relevansi BRI sebagai objek penelitian dalam analisis prediksi harga saham berbasis data historis.

Secara umum, analisis pergerakan harga saham dilakukan dengan mengandalkan dua pendekatan utama, yakni analisis fundamental dan analisis teknikal (Huang dkk., 2021). Pendekatan fundamental berfokus pada penilaian aspek internal perusahaan, seperti performa keuangan dan potensi pertumbuhan bisnis, sedangkan pendekatan teknikal menekankan pengamatan pola harga saham yang terbentuk dari data historis (Zou dkk., 2023). Namun demikian, kedua metode tersebut memiliki keterbatasan dalam menangani data berukuran besar dan bersifat kompleks. Kondisi ini mendorong perlunya penggunaan pendekatan yang lebih fleksibel dan mampu menyesuaikan diri dengan karakteristik data yang ada.

Seiring dengan perkembangan teknologi di bidang *data science*, pendekatan *machine learning* mulai banyak diterapkan dalam prediksi harga saham (Li dkk., 2024), karena kemampuannya mempelajari pola data serta menangkap hubungan linier maupun nonlinier yang sulit diidentifikasi melalui metode analisis konvensional (Ouf dkk., 2024). Regresi Linear merupakan salah satu algoritma *machine learning* yang umum diterapkan dalam pemodelan dan prediksi data numerik, karena memiliki struktur model yang sederhana, bersifat stabil, serta mudah untuk diinterpretasikan.

Hasil penelitian Priyatno dkk. (2023) menunjukkan bahwa algoritma Regresi Linear memiliki tingkat performa prediksi yang baik, berdasarkan nilai metrik evaluasi pada saham BBKA. Sementara itu, penelitian oleh Lumbantoruan dkk. (2026) juga menegaskan bahwa Regresi Linear mampu menghasilkan prediksi yang konsisten, sehingga layak dipertimbangkan sebagai salah satu pendekatan yang sesuai dalam prediksi harga saham menggunakan data historis.

Berbagai studi sebelumnya telah membuktikan potensi *machine learning* dalam melakukan prediksi harga saham. Namun demikian, efektivitas model yang dihasilkan sangat bergantung pada sifat data dan karakteristik saham yang menjadi objek analisis. Kondisi ini menunjukkan bahwa tidak semua metode prediksi dapat memberikan hasil yang optimal untuk setiap *Dataset*. Dengan demikian, diperlukan penelitian lanjutan yang berfokus pada penerapan satu metode prediksi yang selaras dengan karakteristik data historis saham. Atas dasar permasalahan tersebut, penelitian ini berjudul **“Prediksi Harga Saham Bank Rakyat Indonesia (BRI) Menggunakan Algoritma Regresi Linear”**. Diharapkan penelitian ini mampu

memberikan pemahaman yang lebih komprehensif mengenai kinerja algoritma Regresi Linear dalam memprediksi harga saham BRI dengan memanfaatkan data historis.

1.2 Rumusan Masalah

Berdasarkan uraian latar belakang yang telah disampaikan, penelitian ini merumuskan beberapa permasalahan sebagai berikut:

1. Bagaimana implementasi algoritma Regresi Linear untuk memprediksi harga penutupan (*Close*) saham BRI pada hari perdagangan berikutnya berdasarkan data historis?
2. Sejauh mana kinerja algoritma Regresi Linear dalam menghasilkan prediksi harga saham BRI menggunakan metrik evaluasi MAE, RMSE, MAPE, dan Koefisien Determinasi (R^2)?

1.3 Tujuan Penelitian

Mengacu pada permasalahan penelitian yang telah ditetapkan, tujuan yang ingin dicapai dalam penelitian ini adalah sebagai berikut:

1. Mengimplementasikan algoritma Regresi Linear untuk membangun model prediksi harga penutupan (*Close*) saham BRI pada hari perdagangan berikutnya berdasarkan data historis saham.
2. Mengevaluasi kinerja model Regresi Linear dalam memprediksi harga penutupan saham BRI menggunakan metrik evaluasi MAE, RMSE, MAPE, dan Koefisien Determinasi (R^2).

1.4 Manfaat Penelitian

Manfaat yang diharapkan dari penelitian ini adalah sebagai berikut:

1. Memberikan informasi bagi investor atau pihak terkait mengenai prediksi harga saham yang dapat digunakan sebagai bahan pertimbangan tambahan dalam melakukan analisis pergerakan harga saham.

2. Menjadi bahan evaluasi penggunaan metode Regresi Linear serta referensi bagi penelitian selanjutnya dalam prediksi harga saham berbasis data historis.

1.5 Batasan Penelitian

Untuk menjaga agar penelitian tetap terarah dan selaras dengan tujuan yang telah ditetapkan, maka batasan permasalahan dalam penelitian ini ditentukan sebagai berikut:

1. *Dataset* yang digunakan berupa data historis saham Bank Rakyat Indonesia (BRI) yang diperoleh dari *website investing.com* dengan time frame harian (1D) pada periode 2020–2025 dan berjumlah sekitar 1.447 data perdagangan. Setiap baris data merepresentasikan satu hari perdagangan saham.
2. Variabel independen yang digunakan terdiri atas data historis *Open*, *High*, *Low*, *Close*, dan *Volume* dari dua hari perdagangan sebelumnya, sedangkan variabel dependen adalah harga penutupan (*Close*) pada hari perdagangan berikutnya.
3. Metode yang digunakan dalam penelitian ini adalah Regresi Linear Berganda yang diimplementasikan menggunakan bahasa pemrograman *Python* dengan bantuan *library Pandas*, *NumPy*, *Scikit-learn*, dan *Matplotlib*.
4. Model digunakan untuk memprediksi harga penutupan (*Close*) pada hari perdagangan berikutnya berdasarkan data historis dari dua hari perdagangan sebelumnya.
5. Model ini secara spesifik dilatih menggunakan karakteristik data historis saham Bank Rakyat Indonesia (BRI), sehingga koefisien regresi yang dihasilkan tidak direkomendasikan untuk digunakan dalam memprediksi harga saham lain.
6. Penelitian tidak membahas faktor fundamental perusahaan, sentimen berita, kebijakan pemerintah, maupun kondisi makroekonomi yang dapat memengaruhi harga saham.

7. Evaluasi performa model dilakukan menggunakan *Mean Absolute Error* (MAE), *Root Mean Square Error* (RMSE), *Mean Absolute Percentage Error* (MAPE), dan Koefisien Determinasi (R^2).
8. Penelitian ini tidak melakukan perbandingan performa dengan algoritma prediksi lainnya dan tidak menggunakan data *real-time* sebagai data evaluasi utama.

