

BAB I

PENDAHULUAN

1.1 Latar Belakang

Hingga saat ini, pergerakan harga saham masih sangat sulit diprediksi [1]. Bahkan dengan perkembangan teknologi AI, harga saham tetap sulit diprediksi dengan akurat. Hal ini disebabkan adanya variabel di luar data pasar, seperti kondisi politik, ekonomi, dan kestabilan masyarakat di suatu negara [2].

Prediksi harga saham dapat membantu investor atau perusahaan dalam pengambilan keputusan berinvestasi pada sebuah perusahaan. Selain itu, prediksi saham juga dapat digunakan untuk meninjau gejala pasar guna mengatasi risiko *market bubble* atau gelembung pasar yang bisa memicu kontraksi tajam [3]. Dalam skala yang lebih luas, kondisi saham juga dapat menjadi parameter dalam menilai kondisi ekonomi sebuah negara [4].

Oleh karena itu, penelitian ini dilakukan untuk menguji penggunaan data sentimen publik masyarakat Indonesia di media sosial X. Di dalamnya terdapat opini masyarakat terhadap sebuah perusahaan, khususnya dalam konteks gejala politik dan kondisi masyarakat Indonesia pada tahun 2024–2025.

Penelitian ini berfokus pada perusahaan Alfamart. Fokus utama penelitian ini adalah pada sentimen publik dan teknik prediksi menggunakan algoritma *Deep Learning*. Penelitian ini menganalisis pengaruh sentimen publik yang diperoleh dari media sosial X terhadap pergerakan harga saham Alfamart.

Sentimen adalah proses analitis dalam mengklasifikasikan opini-opini dalam bentuk potongan teks. Sentimen tersebut dapat dikategorikan menjadi negatif, netral, dan positif. Dalam konteks harga saham, sentimen yang disampaikan di media sosial dapat menjadi gambaran pandangan publik terhadap suatu perusahaan. Pandangan tersebut berpotensi memengaruhi harga saham [5].

Penggunaan algoritma *deep learning* dalam penelitian ini didasarkan pada kemampuannya yang unggul dalam mengenali pola non-linier dan menangani data sekuensial yang kompleks seperti deret waktu harga saham. Berbeda dengan algoritma *Machine Learning* konvensional yang memerlukan rekayasa fitur manual, *deep learning* memungkinkan model mengekstraksi fitur secara otomatis dari data besar dan tidak terstruktur.

Dalam penelitian ini digunakan tiga algoritma *deep learning* yang telah terbukti efektif dalam prediksi harga saham, yaitu *Recurrent Neural Network* (RNN), *Long Short-Term Memory* (LSTM), dan *Extreme Learning Machine* (ELM). RNN sangat efektif untuk permasalahan yang melibatkan data berurutan dan memiliki hubungan temporal, seperti prediksi deret waktu dan analisis teks. Dalam sebuah penelitian, algoritma RNN mampu mencapai akurasi sebesar 94,16% pada data latih dan 55,26% pada data uji dalam memprediksi harga saham [6].

LSTM, yang merupakan pengembangan dari RNN, dirancang untuk mengatasi masalah *vanishing gradient* dan *exploding gradient*. Hal ini membuatnya mampu menangani urutan data yang panjang dengan lebih stabil. Dalam penelitian terdahulu, LSTM menghasilkan nilai akurasi sebesar 95,78% saat memprediksi harga saham, menunjukkan tingkat kesalahan yang sangat rendah [7].

Sementara itu, ELM dikenal sebagai algoritma yang sangat cepat dalam proses pelatihan dan memiliki kemampuan generalisasi yang baik. Pada penelitian sebelumnya terhadap prediksi harga saham, ELM menghasilkan akurasi sebesar 98,7% terhadap data latih dan 95,5% terhadap data uji [8]. Dengan mempertimbangkan keunggulan masing-masing algoritma, pemilihan RNN, LSTM, dan ELM dalam penelitian ini dimaksudkan untuk mengeksplorasi model prediksi harga saham yang tidak hanya akurat, tetapi juga efisien dan mampu mengadaptasi kompleksitas data pasar ritel Indonesia.

Dataset yang digunakan dalam penelitian ini adalah *dataset* bernama AMRT.csv yang berasal dari GitHub. *Dataset* tersebut dipilih karena bersumber dari BEI (Bursa Efek Indonesia), sehingga data yang diperoleh bersifat aktual. *Dataset* ini memiliki data-data yang dibutuhkan untuk melatih model *deep learning* yang telah disebutkan. Data tersebut meliputi *high price*, *low price*, *open price*, *close price*, *volume*, dan perubahan harga setiap harinya selama lima tahun terakhir yang diperbarui setiap hari kerja.

Pada penelitian ini, ketiga algoritma prediksi (RNN, LSTM, dan ELM) akan diuji menggunakan data teknis dari *dataset* AMRT.csv. Kemudian, dari setiap algoritma akan dipilih model dengan akurasi tertinggi. Ketiga model dengan akurasi tertinggi tersebut akan diuji kembali dengan *dataset* yang telah ditambahkan data sentimen publik dari media sosial X yang membahas Alfamart.

Data sentimen yang telah diperoleh dari X akan diolah menggunakan tiga algoritma, yaitu Neural Network (NN), Lexicon, dan BERT. Ketiganya digunakan untuk mengklasifikasikan opini menjadi opini positif, netral, dan negatif. Setelah itu, setiap opini akan diubah menjadi data numerik, di mana angka 1 mewakili opini positif, 0 untuk netral, dan -1 untuk negatif.

Data numerik dalam satu hari kemudian dijumlahkan untuk menghasilkan nilai sentimen harian. Nilai ini dianggap mewakili sentimen masyarakat terhadap Alfamart pada hari tersebut. Setelah proses pengolahan selesai, data sentimen tersebut akan digabungkan dengan *dataset* AMRT.csv untuk digunakan sebagai fitur tambahan dalam pelatihan model prediksi harga saham.

Setelah model-model dengan akurasi terbaik diperoleh, dilakukan analisis untuk menentukan model mana yang paling unggul. Analisis juga dilakukan untuk menilai apakah integrasi data sentimen publik dari media sosial X dapat meningkatkan akurasi prediksi saham Alfamart.

Keaslian dari penelitian ini dibandingkan penelitian sebelumnya adalah fokusnya yang secara khusus ditujukan pada prediksi harga saham Alfamart. Penelitian ini menggunakan data teknis dari AMRT.csv yang bersumber dari GitHub dan data sentimen publik dari media sosial X. Media sosial X dipilih karena merupakan platform terbuka dengan lalu lintas opini yang sangat tinggi [9]. Oleh karena itu, media ini menyediakan banyak opini terhadap perusahaan yang dapat digunakan sebagai variabel dalam prediksi harga saham menggunakan algoritma *deep learning* untuk mengkaji pengaruh sentimen publik terhadap pergerakan harga saham..

1.2 Rumusan Masalah

Berdasarkan latar belakang di atas, maka didapat rumusan masalah seperti berikut :

1. Bagaimana mengembangkan dan mengimplementasikan model prediksi harga saham Alfamart menggunakan algoritma *deep learning* terhadap data historis harga saham AMRT?
2. Bagaimana proses pengumpulan dan analisis sentimen masyarakat Indonesia terkait Alfamart dari media sosial X?
3. Bagaimana cara mengintegrasikan data hasil analisis sentimen dari media sosial X ke dalam model *deep learning* untuk memprediksi harga saham Alfamart?
4. Bagaimana pengaruh penggunaan data sentimen publik dari media sosial X terhadap akurasi model prediksi harga saham Alfamart berbasis *deep learning*?

1.3 Tujuan Penelitian

Tujuan dari penelitian ini adalah :

1. Mengembangkan dan mengimplementasikan model prediksi harga saham Alfamart menggunakan algoritma *deep learning* berdasarkan data historis harga saham AMRT.
2. Melakukan proses pengumpulan, pengolahan, dan analisis sentimen masyarakat Indonesia terhadap Alfamart dari media sosial X.

3. Mengintegrasikan data hasil analisis sentimen masyarakat dari media sosial X ke dalam model *deep learning* untuk prediksi harga saham Alfamart.
4. Mengevaluasi pengaruh integrasi data sentimen publik terhadap peningkatan akurasi model prediksi harga saham Alfamart.

1.4 Batasan Penelitian

1. Data yang digunakan dalam prediksi harga saham adalah *dataset* AMRT.csv yang berisi data saham Alfamart dari Bursa Efek Indonesia selama lima tahun ke belakang yaitu dari 2020 sampai 2025, tetapi data yang digunakan adalah data selama satu tahun ke belakang yaitu tahun 2024 sampai 2025.
2. Data sentimen publik yang diambil dari media sosial X adalah data *tweet* berbahasa Indonesia yang memiliki kata Alfamart dari tahun 2024 sampai 2025.
3. Penelitian ini dilakukan dengan perangkat lunak Visual Studio Code yang dijalankan pada laptop dengan *processor* Ryzen 5 8645HS, GPU RTX 4050 dengan VRAM 6GB dan RAM 16GB.
4. Proses pengolahan data dan *crawling* data dilakukan menggunakan bahasa pemrograman Python.
5. Algoritma *deep learning* yang akan digunakan adalah algoritma RNN, LSTM, dan ELM.
6. Algoritma yang akan digunakan untuk mengklasifikasikan data sentimen publik adalah algoritma NN (Neural Network), pendekatan Lexicon, dan BERT (Bidirectional Encoder Representations from Transformers).
7. Proses *crawling* data dilakukan di Google Colab dan bahasa pemrograman Python sedangkan proses *training* dan prediksi data akan dilakukan menggunakan Visual Studio Code.
8. Metrik evaluasi yang digunakan dalam evaluasi model prediksi *deep learning* adalah MAPE, RMSE, dan R2.

1.5 Manfaat Penelitian

Penelitian ini diharapkan memberikan manfaat bagi berbagai pihak yang berkepentingan. Berikut adalah manfaat penelitian ini :

1. Sebagai pertimbangan penyusunan strategi pembelian dan penjualan saham bagi investor.
2. Sebagai pertimbangan untuk pihak *social media specialist* dan tim *marketing* Alfamart dalam melakukan kampanye promosi di media sosial.
3. Sebagai pertimbangan untuk peneliti dan *AI engineer* untuk menggunakan data sentimen media sosial X untuk pengembangan AI untuk memprediksi harga saham pada perusahaan ritel seperti Alfamart.

1.6 Struktur Penulisan

Sistematika penulisan skripsi ini adalah sebagai berikut:

BAB I PENDAHULUAN

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

BAB II KAJIAN PUSTAKA

Bab ini berisi tentang teori dan konsep terkait penelitian.

BAB III METODE PENELITIAN

Bab ini berisi langkah-langkah yang akan dilakukan dalam penelitian.

BAB IV HASIL PENELITIAN DAN PEMBAHASAN

Bab ini berisi uraian dan pembahasan terhadap penelitian yang dilakukan.

BAB V KESIMPULAN DAN SARAN

Bab ini berisi kesimpulan terkait penelitian dan saran dari penulis untuk kegiatan penelitian selanjutnya terkait dengan topik yang sedang dibahas