

BAB I

PENDAHULUAN

1.1 Latar Belakang

Kemajuan teknologi dalam bidang analisis sentimen telah menghadirkan berbagai model berbasis machine learning dan deep learning yang dirancang untuk memahami sentimen dalam teks keuangan secara akurat. Salah satu model yang menjadi terobosan dalam analisis sentimen keuangan adalah *RoBERTa* yang menggunakan mekanisme attention untuk memahami hubungan antara kata-kata dalam teks (Semary et al., 2023), menghasilkan evaluasi sentimen yang lebih akurat dibandingkan model analisis sentimen lain.

Penelitian (Karanikola et al., 2023) menunjukkan bahwa performa *RoBERTa* lebih unggul dibandingkan dengan model *BERT* dan *FinancialBERT*. Berdasarkan hasil evaluasi, *RoBERTa* mencapai rata-rata skor 0.775877 pada berbagai matrik, lebih tinggi dibandingkan *BERT* dengan rata-rata skor 0.74166, dan *FinancialBERT* dengan rata-rata skor 0.704876, sehingga sangat relevan untuk analisis sentimen berita keuangan. Meskipun demikian, analisis sentimen sendiri memiliki keterbatasan dalam memahami dinamika pasar secara menyeluruh, karena tidak mempertimbangkan faktor teknikal seperti tren harga saham atau volume perdagangan. Tanpa data historis, model prediksi saham berbasis sentimen tidak dapat sepenuhnya memahami dinamika pasar, sehingga akurasi dalam memprediksi pergerakan harga saham tidak maksimal (Peivandizadeh et al., 2024).

Data historis saham memainkan peran kunci dalam memprediksi pola pergerakan harga. Sebagai indikator utama dinamika pasar, data historis

menawarkan wawasan tentang pola fluktuasi harga berdasarkan *volume* perdagangan, harga pembukaan dan penutupan, serta volatilitas (Rashid et al., 2023). Penelitian sebelumnya menunjukkan bahwa data historis mampu memberikan wawasan signifikan tentang perilaku harga di masa depan, terutama jika dikombinasikan dengan model prediksi canggih berbasis machine learning atau deep learning, (Haryono, Sarno, & Sungkono, 2024) menunjukkan bahwa Model *Gated Recurrent Unit*, merupakan algoritma yang unggul dalam peramalan saham di Bursa Efek Indonesia (IDX), dengan performa unggul berdasarkan berbagai matrik evaluasi, yaitu $MAE = 127,978,064.19$, $RMSE = 553.327$, dan $MAPE = 0.858$. Namun, pada studi ini hanya berfokus pada data historis tanpa mempertimbangkan faktor eksternal yang dapat memengaruhi harga saham, seperti sentimen pasar yang dipengaruhi oleh berita atau laporan keuangan.

Forecasting harga saham sendiri memiliki nilai strategis dalam dunia keuangan karena dampaknya yang signifikan terhadap pengambilan keputusan investasi (Khattak et al., 2023). Dengan memanfaatkan hasil prediksi yang akurat, investor dapat mengoptimalkan peluang keuntungan sekaligus meminimalkan risiko kerugian (Tang et al., 2023). Hasil penelitian (Zhang et al., 2024) menunjukkan bahwa *forecasting* berbasis data historis dapat memberikan kontribusi besar terhadap penguatan strategi investasi berbasis data. Namun demikian, integrasi antara analisis sentimen dengan data historis masih belum banyak diterapkan secara menyeluruh untuk meningkatkan akurasi prediksi, khususnya untuk saham unggulan di pasar Indonesia.

Situs berita mengenai bisnis di Indonesia yang memiliki kredibilitas tinggi, seperti Bisnis.com dan CNBC Indonesia, dapat dijadikan sebagai rujukan untuk mendapatkan informasi terkini tentang berita keuangan (Afifah, 2023; Ahsyam et al., 2024). Kedua situs tersebut merupakan sumber informasi bisnis dan keuangan yang sering digunakan, dengan laporan yang mencakup analisis berbagai aspek ekonomi, termasuk perkembangan pasar saham. Dalam konteks penelitian ini, data yang diambil dari kedua situs tersebut melalui proses scraping menggunakan kata kunci relevan, seperti kode saham BBRI (PT Bank Rakyat Indonesia Tbk), BMRI (PT Bank Mandiri Tbk), dan BBCA (PT Bank Central Asia Tbk), dipilih karena telah terbukti dalam penelitian sebelumnya memiliki kredibilitas tinggi dalam mendukung analisis sentimen yang berpotensi memengaruhi pergerakan harga saham (Haryono, Sarno, Anggraini, et al., 2024). Pemilihan ketiga bank tersebut didasarkan pada dominasinya dalam sektor perbankan dengan kapitalisasi pasar terbesar di Indonesia, menjadikannya saham yang paling banyak diperhatikan investor dan sering dibahas dalam pemberitaan keuangan.

Integrasi antara analisis sentimen berita dan data historis saham merupakan pendekatan inovatif yang menggabungkan data kuantitatif dari histori harga dengan data kualitatif dari sentimen pasar (Sai et al., 2023). Pada penelitian (John & Latha, 2023) membuktikan bahwa sentimen berita berpengaruh terhadap pergerakan harga saham, sementara pola historis menjadi dasar analisis tren. Studi ini menunjukkan bahwa model HyRNN yang hanya menggunakan data historis menghasilkan RMSE 14.62, RMSE 19.12, dan R^2 0.984. Setelah integrasi dengan sentimen berita dilakukan, performa model meningkat menjadi RMSE 13.102, RMSE 18.142, dan

R^2 0.987, membuktikan bahwa kombinasi data historis dan sentimen dapat meningkatkan akurasi prediksi saham. Namun, penelitian ini masih menggunakan VADER sebagai metode analisis sentimen, yang berbasis leksikon dan memiliki keterbatasan dalam memahami konteks serta nuansa bahasa dalam berita keuangan.

Penelitian ini bertujuan untuk mengatasi celah dengan menggabungkan *Financial RoBERTa* (*Robustly Optimized BERT Pretraining Approach*), algoritma yang unggul dalam analisis sentimen di sektor keuangan (Karanikola et al., 2023), dengan *GRU* (*Gated Recurrent Unit*), yang telah terbukti efektif dalam peramalan saham berdasarkan studi (Haryono, Sarno, & Sungkono, 2024). Dengan memadukan kekuatan *Financial RoBERTa* dalam analisis sentimen dan keunggulan *GRU* dalam peramalan, diharapkan pendekatan ini dapat meningkatkan akurasi prediksi saham sekaligus berkontribusi secara signifikan pada pengembangan model prediksi di pasar keuangan Indonesia.

1.2 Rumusan Masalah

- a. Bagaimana membangun pemodelan yang mengintegrasikan *Financial RoBERTa* dan *forecasting* saham untuk memprediksi harga saham perbankan?
- b. Bagaimana pengaruh dari integrasi analisis sentimen dari media berita finansial dengan data historis saham terhadap akurasi prediksi pergerakan harga saham di sektor perbankan Indonesia?

1.3 Tujuan Penelitian

- a. Membangun pemodelan yg mengintegrasikan analisis sentimen dari berita finansial, dan data historis saham untuk memprediksi saham perbankan.

- b. Menganalisis pengaruh analisis sentimen dari media berita finansial terhadap akurasi prediksi pergerakan harga saham, serta mengidentifikasi hubungan antara sentimen pasar dan fluktuasi harga saham.

1.4 Manfaat Penelitian

- a. Bagi akademisi, penelitian ini dapat dijadikan alternatif dalam pengembangan model prediksi saham berbasis *deep learning* dengan menggabungkan analisis dan data historis, serta memberikan pemahaman tentang dampak sentimen pasar terhadap saham.
- b. Bagi investor, penelitian ini dapat digunakan sebagai acuan dalam membuat keputusan investasi berbasis data yang mencakup faktor eksternal seperti sentimen pada berita finansial.
- c. Bagi perusahaan perbankan, hasil penelitian ini dapat digunakan untuk mengukur dan memahami bagaimana sentimen publik dan berita keuangan mempengaruhi harga saham mereka, sehingga bisa dijadikan dasar strategi dalam pengelolaan saham dan bisnis.

1.5 Batasan Penelitian

- a. Penelitian ini terbatas pada data saham perbankan dari tiga bank besar di Indonesia, yaitu BMRI, BBKA, dan BBRI. Data saham harian diambil dari situs Investing.com dengan periode yang dianalisis mulai dari 19 September 2022 hingga 19 September 2024.
- b. Analisis sentimen dalam penelitian ini akan menggunakan data yang diambil dari dua media berita keuangan, yaitu Bisnis.com dan CNBCIndonesia. Fokus analisis akan diarahkan pada sentimen yang

terdapat dalam berita dan postingan yang berbahasa Indonesia. Digunakan juga dataset dari Hugging Face bernama *intanm/indonesian-Financial phrasebank*, sebagai data pelatihan tambahan untuk kebutuhan fine-tuning model.

- c. Model yang digunakan dalam penelitian ini mencakup *Financial RoBERTa* untuk mengekstraksi dan menganalisis sentimen dari teks berita keuangan. Untuk prediksi atau *forecasting* harga saham, model *GRU* akan digunakan dengan memanfaatkan data sentimen dan data historis sebagai input utama.
- d. Ruang lingkup penelitian ini terbatas pada prediksi harga saham di sektor perbankan Indonesia, khususnya pada saham BMRI, BBKA, dan BBRI. Penelitian ini tidak mencakup analisis terhadap sektor saham lainnya atau pengaruh kebijakan ekonomi global secara langsung.
- e. Evaluasi performa model untuk analisis sentimen akan menggunakan F1 Score, Akurasi, Presisi dan Recall, sedangkan dalam prediksi harga saham akan menggunakan *Mean Absolute Error (RMSE)*, *Mean Absolute Percentage Error (MAPE)*, dan *Mean Squared Error (MAE)*.