

Analisis Sentimen Berita Keuangan terhadap Pergerakan Harga Saham BBKA Menggunakan Model Zero-Shot mDeBERTa-v3

Yoga Effendi Saputra*¹, Abdussalam²

^{1,2}Universitas Dian Nuswantoro, Indonesia

Email: ¹111202214129@mhs.dinus.ac.id, ²abdussalam@dsn.dinus.ac.id

Abstrak

Penelitian ini mengkaji hubungan antara sentimen berita keuangan dan pergerakan harga saham PT Bank Central Asia Tbk (BBCA), salah satu emiten dengan kapitalisasi pasar terbesar di Indonesia, pada periode Mei 2025 hingga Mei 2026. Sebanyak 3.113 artikel berita diklasifikasikan menggunakan model zero-shot berbasis mDeBERTa-v3 ke dalam tiga kategori sentimen, yaitu positif (66,1%), negatif (31,3%), dan netral (2,6%). Penggunaan mDeBERTa-v3 tanpa proses fine-tuning dan tanpa kebutuhan data berlabel menjadi kontribusi penelitian ini dalam penerapan metode analisis sentimen yang lebih efisien pada pasar modal Indonesia. Compound score harian dihitung sebagai agregat skor sentimen dan dianalisis korelasinya terhadap harga penutupan dan return harian BBCA. Hasil analisis Pearson menunjukkan korelasi positif signifikan antara compound score harian dan return harian ($r = 0,2446$; $p < 0,001$), sedangkan korelasi sentimen lag D-1 terhadap return harian tidak signifikan ($r = -0,0717$; $p = 0,2715$), yang mengindikasikan bahwa pasar bereaksi pada hari yang sama dengan munculnya berita. Rolling correlation 30 hari menunjukkan hubungan yang bervariasi, berkisar antara -0,15 hingga 0,53 selama periode pengamatan. Evaluasi arah prediksi menggunakan sentimen lag D-1 menghasilkan akurasi sebesar 45,1%, sehingga sentimen berita tidak dapat diandalkan sebagai prediktor tunggal arah harga saham. Temuan ini menunjukkan bahwa sentimen berita memiliki hubungan yang signifikan dengan kondisi pasar pada hari yang sama, namun lebih tepat dimanfaatkan sebagai fitur pelengkap dalam model prediksi harga saham multivariat yang menggabungkan faktor teknikal, fundamental, dan variabel ekonomi lainnya. Dengan demikian, penelitian ini memberikan alternatif pendekatan analisis sentimen berbasis zero-shot yang lebih efisien untuk mendukung pengambilan keputusan di pasar modal Indonesia.

Kata kunci: analisis sentimen, berita keuangan, harga saham, mDeBERTa-v3, zero-shot classification, BBKA

Abstract

This study examines the relationship between financial news sentiment and stock price movements of PT Bank Central Asia Tbk (BBCA), one of the issuers with the largest market capitalization in Indonesia, from May 2025 to May 2026. A total of 3,113 news articles were classified using a zero-shot model based on mDeBERTa-v3 into three sentiment categories: positive (66.1%), negative (31.3%), and neutral (2.6%). The use of mDeBERTa-v3 without fine-tuning and without the need for labeled data contributes to the application of a more efficient sentiment analysis method in the Indonesian capital market. The daily compound score was calculated as an aggregate of sentiment scores and analyzed for its correlation with BBCA's closing price and daily returns. The Pearson analysis results showed a significant positive correlation between the daily compound score and daily returns ($r = 0.2446$; $p < 0.001$), while the correlation of D-1 lag sentiment to daily returns was insignificant ($r = -0.0717$; $p = 0.2715$), indicating that the market reacted on the same day as the news emerged. The 30-day rolling correlation shows a variable relationship, ranging from -0.15 to 0.53 over the observation period. Evaluation of the predicted direction using D-1 lag sentiment yielded an accuracy of 45.1%, indicating that news sentiment cannot be relied upon as a sole predictor of stock price direction. These findings suggest that news sentiment has a significant relationship with market conditions on the same day, but it is more appropriately utilized as a complementary feature in a multivariate stock price prediction model that incorporates technical, fundamental, and other economic variables. Thus, this study provides a more efficient alternative zero-shot sentiment analysis approach to support decision-making in the Indonesian capital market.

Keywords: sentiment analysis, financial news, stock price, mDeBERTa-v3, zero-shot classification, BBKA

1. PENDAHULUAN

Pasar modal merupakan salah satu instrumen investasi yang sangat sensitif terhadap berbagai informasi, baik yang bersifat fundamental maupun sentimen pasar. Dalam era digital, informasi tentang suatu perusahaan dapat menyebar dengan sangat cepat melalui berbagai kanal berita daring, media sosial, dan platform finansial. PT Bank Central Asia Tbk (BBCA) merupakan salah satu emiten unggulan di Bursa Efek Indonesia (BEI) dan menjadi salah satu saham yang paling aktif diperdagangkan serta dibahas dalam ekosistem berita keuangan Indonesia.

Analisis sentimen (sentiment analysis) atau yang dikenal juga sebagai opinion mining adalah cabang pemrosesan bahasa alami (Natural Language Processing/NLP) yang berfokus pada identifikasi dan ekstraksi opini subjektif dari teks [1]. Dalam konteks keuangan, analisis sentimen berita telah terbukti dapat memberikan sinyal yang relevan terhadap perilaku pasar saham. Penelitian terbaru menunjukkan bahwa integrasi model bahasa modern seperti FinBERT dan GPT-4 mampu meningkatkan kemampuan prediksi harga saham berbasis informasi tekstual [2]. Hipotesis efisiensi pasar bentuk semi-kuat (semi-strong form) menyatakan bahwa harga saham mencerminkan semua informasi yang tersedia untuk publik, termasuk informasi dari berita [3]. Dengan demikian, ketika sentimen berita berubah, harga saham pun diharapkan turut merespons perubahan tersebut.

Perkembangan model bahasa berskala besar (Large Language Model/LLM) berbasis Transformer, khususnya model BERT dan variannya, telah merevolusi pendekatan analisis sentimen. Model-model ini mampu memahami konteks teks secara lebih mendalam dibandingkan pendekatan berbasis leksikon atau machine learning tradisional [4]. Salah satu kemajuan terkini adalah teknik zero-shot classification, di mana model dapat mengklasifikasikan teks ke dalam kategori yang belum pernah dilihat saat pelatihan hanya dengan deskripsi label [5]. Teknik ini sangat relevan untuk analisis sentimen berita keuangan multibahasa karena tidak memerlukan data berlabel yang besar.

Kajian terbaru juga menunjukkan bahwa Large Language Model tidak hanya digunakan untuk analisis sentimen, tetapi juga telah diterapkan pada berbagai tugas di pasar modal seperti prediksi return, manajemen portofolio, ekstraksi informasi keuangan, dan pengambilan keputusan investasi berbasis data tekstual [6]. Pendekatan Transformer modern didasarkan pada mekanisme attention yang diperkenalkan oleh Vaswani et al. dan telah diadopsi secara luas dalam model-model bahasa mutakhir [7]. Arsitektur ini memungkinkan model memahami hubungan kontekstual antar kata secara lebih efektif dibandingkan pendekatan sequence tradisional. Dalam domain keuangan, pendekatan berbasis Transformer telah banyak digunakan untuk analisis sentimen dan prediksi pasar saham karena kemampuannya memahami konteks linguistik yang kompleks [8]. Selain itu, model berbasis FinBERT juga menunjukkan performa yang baik dalam menghubungkan sentimen berita dengan pergerakan harga saham melalui representasi bahasa yang lebih spesifik pada domain keuangan [9].

Penelitian sebelumnya menunjukkan bahwa sentimen media dan opini publik memiliki hubungan dengan dinamika pasar saham. Zhang et al. [10] menunjukkan bahwa model berbasis sentimen FinBERT yang diintegrasikan dengan faktor-faktor aset pricing mampu menjelaskan variasi return saham harian secara signifikan, sementara Linhares et al. [11] menemukan bahwa model LLM seperti OPT mampu memprediksi return pasar saham dengan akurasi tinggi berdasarkan sentimen berita keuangan. Selain itu, pendekatan berbasis NLP pada berita finansial juga telah digunakan secara luas dalam sistem prediksi pasar saham modern [12].

Meskipun penelitian sebelumnya menunjukkan bahwa sentimen media dan teknik pemrosesan bahasa alami (Natural Language Processing/NLP) dapat menjelaskan dan memprediksi pergerakan pasar saham, sebagian besar masih berfokus pada pasar global serta menggunakan model yang memerlukan fine-tuning dan data berlabel dalam jumlah besar. Sementara itu, pemanfaatan model zero-shot multilingual seperti mDeBERTa-v3 untuk menganalisis hubungan sentimen berita dengan pergerakan harga saham emiten tertentu masih terbatas. Model ini dipilih karena memiliki kemampuan disentangled attention yang menghasilkan representasi semantik lebih baik dibandingkan BERT konvensional serta memungkinkan klasifikasi tanpa fine-tuning. Oleh karena itu, penelitian ini berfokus pada analisis hubungan sentimen berita terhadap pergerakan harga saham BBCA melalui pendekatan zero-shot classification yang lebih efisien dan tidak bergantung pada data berlabel dalam jumlah besar. Mengingat BBCA merupakan salah satu emiten dengan kapitalisasi pasar terbesar dan berpengaruh terhadap IHSG, penelitian ini diharapkan dapat menjadi alternatif pendekatan dalam memahami pengaruh sentimen berita terhadap dinamika pasar modal Indonesia.

Dalam beberapa tahun terakhir, model bahasa khusus domain finansial seperti FinBERT juga dikembangkan untuk meningkatkan performa analisis sentimen keuangan [13][14]. Penelitian terbaru menunjukkan bahwa Large Language Model mampu meningkatkan kualitas ekstraksi sentimen keuangan dan menghasilkan strategi perdagangan yang lebih adaptif dibandingkan pendekatan berbasis leksikon tradisional [15]. Kehadiran model-model tersebut menunjukkan bahwa pemrosesan bahasa alami berbasis Transformer memiliki potensi besar dalam mendukung analisis pasar modal dan pengambilan keputusan investasi.

Model mDeBERTa-v3 (multilingual DeBERTa version 3) dipilih karena memiliki beberapa keunggulan dibandingkan model NLP lainnya. DeBERTa menerapkan disentangled attention yang memisahkan representasi

konten dan posisi kata sehingga mampu menangkap hubungan semantik secara lebih baik. Selain itu, versi multilingual mDeBERTa-v3 telah dilatih pada banyak bahasa termasuk Bahasa Indonesia sehingga dapat digunakan dalam skenario zero-shot classification tanpa memerlukan proses fine-tuning [4][5]. Dibandingkan model monolingual seperti FinBERT yang berfokus pada bahasa Inggris, mDeBERTa-v3 lebih sesuai untuk menganalisis berita keuangan Indonesia. Kemampuan ini menjadi penting karena ketersediaan korpus sentimen keuangan berbahasa Indonesia masih terbatas, sehingga pendekatan zero-shot menawarkan solusi yang lebih efisien dari sisi data dan komputasi.

Penelitian ini bertujuan untuk: (1) mengklasifikasikan sentimen berita keuangan BBCA selama periode Mei 2025 hingga Mei 2026 menggunakan model zero-shot mDeBERTa-v3; (2) menganalisis korelasi antara sentimen berita dan pergerakan harga saham BBCA; (3) mengevaluasi kemampuan prediktif sentimen berita terhadap arah pergerakan harga saham pada hari berikutnya; dan (4) mengkaji dinamika temporal hubungan sentimen-harga melalui analisis rolling correlation.

2. METODE PENELITIAN

2.1. Pengumpulan Data

Penelitian ini menggunakan dua jenis data utama: (1) data historis harga saham BBCA dan (2) data berita keuangan terkait BBCA. Data harga saham diperoleh dari Yahoo Finance melalui library yfinance untuk periode Mei 2025 hingga Mei 2026, menghasilkan 238 titik data perdagangan dengan variabel harga pembukaan, penutupan, tertinggi, terendah, dan volume.

Data berita dikumpulkan dari Google News RSS Feed menggunakan kata kunci 'BBCA' dan 'Bank Central Asia' untuk periode yang sama, menghasilkan total 3.113 artikel berita. Berita diproses untuk menghapus duplikat dan berita yang tidak relevan, kemudian judul berita digunakan sebagai unit analisis sentimen mengingat representasinya yang ringkas namun informatif. Data yang diperoleh selanjutnya melalui tahap pembersihan data (data clearing). Tahap ini meliputi penghapusan artikel duplikat berdasarkan kesamaan judul dan tanggal publikasi, penghapusan berita yang tidak berkaitan langsung dengan PT Bank Central Asia Tbk (BBCA), serta penghapusan berita dengan judul yang tidak memuat informasi ekonomi, keuangan, atau aktivitas perusahaan. Artikel yang hanya berisi pengumuman singkat, iklan, atau berita dengan konteks berbeda meskipun memuat kata kunci "BBCA" juga dikeluarkan dari dataset. Setelah proses tersebut, diperoleh 3.113 artikel yang digunakan dalam analisis.

Setelah proses pembersihan data, judul berita digunakan sebagai unit analisis sentimen. Pemilihan judul didasarkan pada pertimbangan bahwa judul merupakan ringkasan singkat yang merepresentasikan informasi utama dan menjadi bagian pertama yang diterima oleh investor. Penggunaan judul juga memungkinkan pemrosesan data yang lebih efisien serta mengurangi kompleksitas akibat variasi panjang dan struktur isi artikel.

2.2. Model Klasifikasi Sentimen Zero-Shot

Klasifikasi sentimen dilakukan menggunakan model facebook/bart-large-mnli dan MoritzLaurer/mDeBERTa-v3-base-mnli-xnli yang diakses melalui library Hugging Face Transformers. Dalam penelitian ini, model mDeBERTa-v3 dipilih sebagai model utama karena kemampuan multibahasanya yang superior, terutama untuk Bahasa Indonesia. Penelitian ini menggunakan pendekatan zero-shot classification, sehingga tidak dilakukan proses fine-tuning maupun validasi menggunakan data berlabel khusus. Oleh karena itu, hasil klasifikasi sentimen dalam penelitian ini bersifat eksploratif dan dimaksudkan untuk mengidentifikasi pola hubungan sentimen dengan pergerakan harga saham.

Pipeline zero-shot classification dikonfigurasi dengan tiga label kandidat: 'positif', 'negatif', dan 'netral'. Untuk setiap berita, model menghasilkan skor probabilitas untuk masing-masing label. Sentimen yang ditetapkan adalah label dengan probabilitas tertinggi. Compound score dihitung menggunakan formula:

$$\text{Compound Score} = \text{Score}_{\text{Positif}} - \text{Score}_{\text{Negatif}} \quad (1)$$

di mana $\text{Score}_{\text{Positif}}$ dan $\text{Score}_{\text{Negatif}}$ adalah skor probabilitas yang dihasilkan model untuk label positif dan negatif. Nilai compound score berkisar antara -1 (sangat negatif) hingga +1 (sangat positif).

Compound score harian (D_0) dihitung sebagai rata-rata compound score seluruh berita pada tanggal yang sama dengan hari perdagangan:

$$\text{Compound_Mean}_t = (1/n) \times \sum \text{Compound Score}_i \quad (2)$$

di mana n adalah jumlah berita pada hari t dan Compound Score_i adalah compound score berita ke- i .

Penelitian ini tidak melakukan pengujian akurasi klasifikasi sentimen menggunakan data berlabel yang dibuat secara manual. Oleh karena itu, hasil klasifikasi sentimen yang diperoleh dari model mDeBERTa-v3 diperlakukan sebagai representasi sentimen otomatis yang digunakan untuk menganalisis hubungan dengan pergerakan harga saham, bukan sebagai ukuran akurasi klasifikasi sentimen yang absolut.

2.3. Analisis Korelasi

Hubungan antara sentimen berita dan pergerakan harga saham dianalisis menggunakan tiga metode korelasi: (1) Pearson correlation untuk mengukur hubungan linear, (2) Spearman rank correlation sebagai ukuran non-parametrik, dan (3) Kendall tau correlation untuk validasi tambahan. Return harian BBKA dihitung menggunakan formula:

$$Return_t = (Close_t - Close_{t-1}) / Close_{t-1} \times 100\% \quad (3)$$

Untuk mengevaluasi efek prediktif, analisis dilakukan dengan lag D-1, yaitu menggunakan compound score hari sebelumnya sebagai prediktor return hari ini:

$$Lag\ D-1\ Compound = Compound_Mean_{t-1} \quad (4)$$

2.4. Rolling Correlation Analysis

Untuk memahami dinamika temporal hubungan sentimen-harga, dilakukan analisis rolling correlation dengan window 30 hari perdagangan. Rolling correlation menghitung korelasi Pearson pada setiap window waktu yang bergeser satu hari, menghasilkan time series korelasi yang menggambarkan perubahan kekuatan hubungan sepanjang periode pengamatan.

2.5. Evaluasi Prediksi Arah Harga

Evaluasi kemampuan prediktif sentimen dilakukan dengan menggunakan compound score lag D-1 untuk memprediksi arah pergerakan harga pada hari berikutnya. Aturan prediksi adalah: jika compound score lag D-1 positif, prediksi harga naik; jika negatif, prediksi harga turun. Akurasi prediksi dievaluasi menggunakan confusion matrix.

3. HASIL DAN PEMBAHASAN

3.1. Distribusi Sentimen Berita BBKA

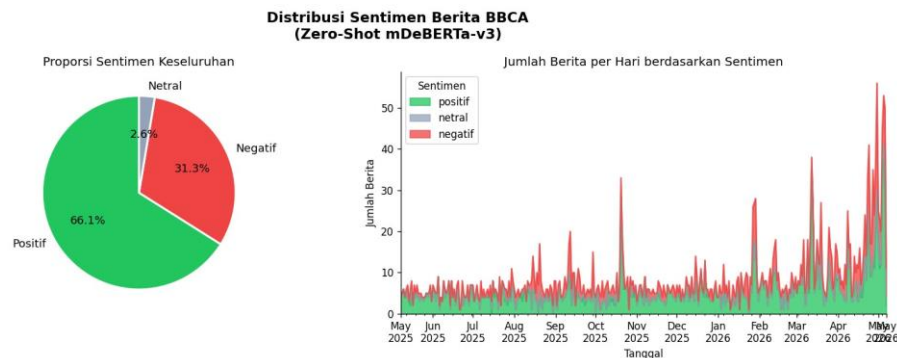
Hasil klasifikasi sentimen terhadap 3.113 berita BBKA menggunakan model zero-shot mDeBERTa-v3 ditampilkan pada Tabel 1. Analisis menunjukkan dominasi berita bersentimen positif, yang mencerminkan citra fundamental BBKA sebagai bank terbesar dan paling stabil di Indonesia.

Tabel 1. Distribusi Sentimen Berita BBKA (Mei 2025 – Mei 2026)

Kategori Sentimen	Jumlah Berita	Persentase (%)	Deskripsi
Positif	2.057	66,1%	Berita dengan nuansa optimistis terhadap BBKA
Negatif	975	31,3%	Berita dengan nuansa pesimistis atau risiko
Netral	81	2,6%	Berita informatif tanpa polaritas dominan
Total	3.113	100%	Keseluruhan berita periode Mei 2025 – Mei 2026

Berdasarkan Tabel 1, dari total 3.113 berita yang dianalisis, sebesar 66,1% diklasifikasikan bersentimen positif (2.057 berita), 31,3% negatif (975 berita), dan 2,6% netral (81 berita). Dominasi sentimen positif ini konsisten dengan posisi fundamental BBKA sebagai bank dengan kapitalisasi pasar terbesar di Indonesia dan kinerja keuangan yang konsisten stabil.

Pada panel time series (Gambar 1, kanan), terlihat adanya peningkatan jumlah berita yang signifikan pada periode akhir pengamatan (April–Mei 2026), dengan lonjakan mencapai lebih dari 50 berita per hari. Hal ini bertepatan dengan kondisi pasar yang lebih volatile, kemungkinan dipicu oleh faktor makroekonomi global seperti ketidakpastian kebijakan moneter Federal Reserve dan volatilitas nilai tukar Rupiah.



Gambar 1. Distribusi Sentimen Berita BBKA – Proporsi Keseluruhan dan Jumlah Berita Harian

Dominasi sentimen positif pada pemberitaan BBKA tidak hanya mencerminkan persepsi publik terhadap kinerja fundamental perusahaan, tetapi juga dapat dipengaruhi oleh kecenderungan media untuk memberikan eksposur lebih besar terhadap perusahaan dengan reputasi kuat dan stabilitas tinggi. Fenomena ini dikenal sebagai media positivity bias, di mana perusahaan dengan fundamental baik cenderung lebih sering diberitakan secara optimistik dibandingkan perusahaan dengan risiko tinggi. Namun demikian, dominasi sentimen positif ini tidak serta merta berbanding lurus dengan kinerja harga saham, karena sentimen berita lebih merepresentasikan persepsi informasi, sedangkan harga saham dipengaruhi oleh kombinasi faktor mikro dan makroekonomi.

3.2. Dinamika Harga Saham dan Sentimen Harian

Gambar 2 menampilkan overlay antara harga penutupan BBKA dan compound score sentimen harian dalam periode pengamatan. Terlihat tren penurunan harga saham BBKA dari sekitar Rp 8.500 per saham pada Mei 2025 menjadi sekitar Rp 6.000–6.200 per saham pada Mei 2026, merepresentasikan koreksi signifikan sekitar 27–30% sepanjang tahun.



Gambar 2. Harga Saham BBKA dan Compound Score Sentimen Harian (Mei 2025 – Mei 2026)

Meskipun terjadi tren penurunan harga secara keseluruhan, compound score sentimen sebagian besar tetap berada di zona positif (di atas 0). Hal ini menunjukkan bahwa berita mengenai BBKA cenderung bernuansa positif meski harga saham mengalami tekanan. Kondisi ini dapat dijelaskan oleh fakta bahwa penurunan harga BBKA lebih banyak disebabkan oleh faktor eksternal (sentimen pasar global, rotasi sektor) daripada penurunan fundamental perusahaan.

Beberapa episode compound score negatif yang dalam (mencapai -0,5 hingga -0,7) teridentifikasi pada periode Agustus–September 2025 dan Januari–Februari 2026, yang berkorelasi dengan periode volatilitas harga yang lebih tinggi.

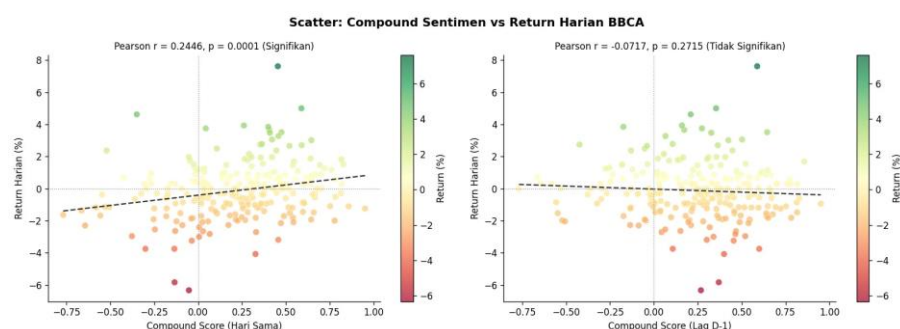
Perbedaan arah antara tren harga saham yang menurun dan dominasi sentimen positif menunjukkan adanya potensi decoupling antara informasi berbasis berita dan pergerakan harga saham. Kondisi ini dapat terjadi karena pergerakan harga saham tidak hanya dipengaruhi oleh informasi spesifik perusahaan, tetapi juga oleh faktor eksternal seperti kebijakan suku bunga, arus modal asing, serta kondisi pasar global yang dapat mendominasi sentimen investor. Dengan demikian, meskipun berita mengenai BBCA cenderung positif, tekanan dari faktor eksternal dapat menyebabkan harga saham tetap mengalami koreksi dalam periode tertentu.

3.3. Analisis Korelasi Sentimen dan Return Saham

Tabel 2 merangkum hasil analisis korelasi antara berbagai variabel sentimen dan variabel harga saham BBCA. Scatter plot hubungan antara compound score dan return harian ditampilkan pada Gambar 3.

Tabel 2. Hasil Analisis Korelasi Pearson: Variabel Sentimen vs Harga Saham BBCA

Variabel Sentimen	Variabel Target	Pearson r	p-value	Signifikan
Compound Score Harian (D0)	Return Harian (%)	0,2446	0,0001	Ya (✓)
Compound Score Harian (D0)	Harga Penutupan	0,2548	0,0001	Ya (✓)
Compound Score Lag D-1	Return Harian (%)	-0,0717	0,2715	Tidak (X)
Compound Score Lag D-1	Harga Penutupan	0,2322	0,0003	Ya (✓)
Rasio Berita Positif (D0)	Return Harian (%)	0,1403	0,0309	Ya (✓)
Rasio Berita Positif (D0)	Harga Penutupan	0,2636	<0,001	Ya (✓)
Rasio Berita Negatif (D0)	Return Harian (%)	-0,1692	0,0091	Ya (✓)
Rasio Berita Negatif (D0)	Harga Penutupan	-0,2566	0,0001	Ya (✓)
Rasio Positif Lag D-1	Return Harian (%)	-0,0274	0,6753	Tidak (X)
Rasio Negatif Lag D-1	Return Harian (%)	0,0308	0,6364	Tidak (X)



Gambar 3. Scatter Plot: Compound Score Sentimen vs Return Harian BBCA (Hari Sama dan Lag D-1)

Hasil analisis korelasi pada Tabel 2 menunjukkan beberapa temuan penting. Pertama, compound score harian menunjukkan korelasi positif yang signifikan secara statistik terhadap return harian ($r = 0,2446$; $p < 0,001$) maupun harga penutupan ($r = 0,2548$; $p < 0,001$). Ini mengindikasikan bahwa pada hari-hari dengan sentimen berita lebih positif, return BBCA cenderung lebih tinggi.

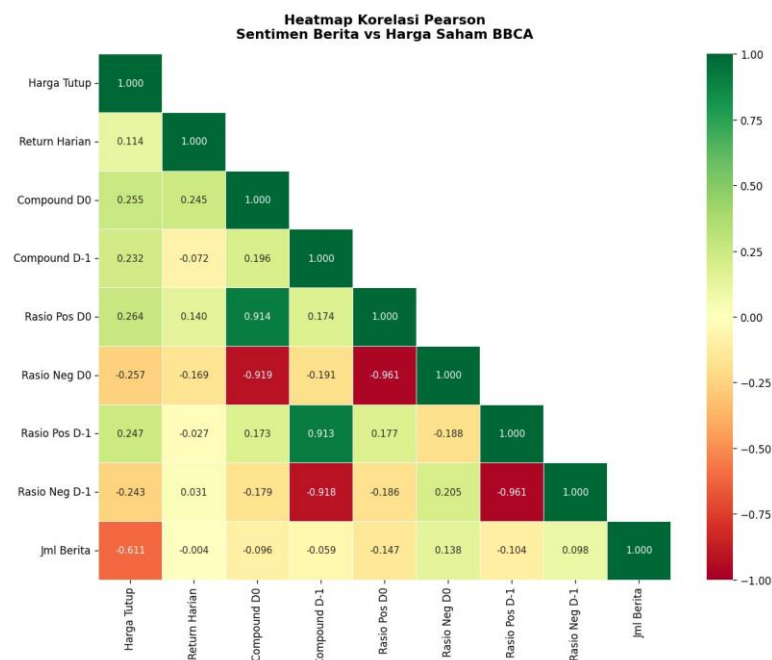
Kedua, compound score lag D-1 tidak menunjukkan korelasi signifikan terhadap return harian ($r = -0,0717$; $p = 0,2715$), meskipun berkorelasi signifikan dengan harga penutupan ($r = 0,2322$; $p < 0,001$). Temuan ini mengindikasikan bahwa pasar merespons sentimen berita pada hari yang sama (same-day effect) dan bukan pada hari berikutnya, konsisten dengan hipotesis pasar efisien semi-kuat.

Meskipun hasil menunjukkan korelasi positif yang signifikan antara compound score dan return harian, besarnya korelasi yang relatif rendah ($r = 0,2446$) mengindikasikan bahwa sentimen berita hanya menjelaskan

sebagian kecil variasi pergerakan return saham BBCA. Secara kuantitatif, nilai tersebut setara dengan R^2 sekitar 6%, yang berarti lebih dari 90% variasi return dipengaruhi oleh faktor lain di luar sentimen berita. Temuan ini menunjukkan bahwa meskipun informasi publik memiliki pengaruh terhadap pergerakan harga saham, sentimen berita bukan merupakan faktor dominan dalam pembentukan return, melainkan hanya salah satu komponen dalam mekanisme pembentukan harga yang lebih kompleks.

Selain itu, signifikansi statistik yang diperoleh tidak selalu mencerminkan signifikansi ekonomi, sehingga hubungan yang ditemukan lebih tepat diinterpretasikan sebagai hubungan lemah hingga moderat dalam konteks praktis investasi. Temuan penelitian ini sejalan dengan Zhang et al. [8] yang menunjukkan bahwa sentimen berbasis model Transformer memiliki hubungan yang signifikan terhadap pergerakan return saham. Namun demikian, nilai korelasi yang diperoleh dalam penelitian ini ($r = 0,2446$) relatif lebih rendah dibandingkan beberapa penelitian pada pasar saham negara maju yang melaporkan hubungan yang lebih kuat antara sentimen dan return saham. Perbedaan tersebut dapat disebabkan oleh karakteristik pasar modal Indonesia yang berbeda, penggunaan pendekatan zero-shot tanpa fine-tuning, serta pengaruh faktor makroekonomi yang lebih dominan selama periode observasi.

3.4. Heatmap Korelasi Komprehensif



Gambar 4. Heatmap Korelasi Pearson: Variabel Sentimen dan Variabel Harga Saham BBCA

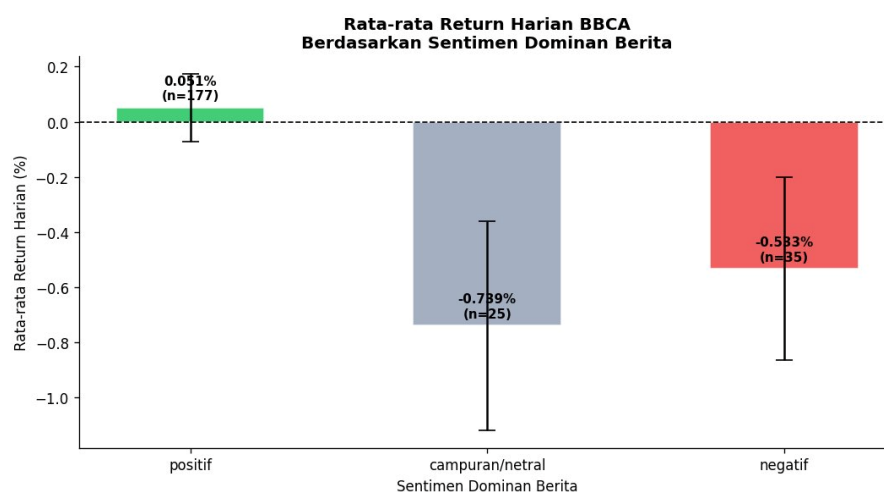
Gambar 4 menampilkan heatmap korelasi Pearson antara seluruh variabel sentimen dan variabel harga saham. Beberapa pola penting teridentifikasi: (1) korelasi negatif yang kuat antara harga penutupan dan jumlah berita (-0,611), mengindikasikan bahwa semakin banyak berita yang muncul cenderung bersamaan dengan periode harga yang lebih rendah; (2) korelasi tinggi antar variabel sentimen internal (Compound D0 dengan Rasio Positif D0: 0,914), menunjukkan konsistensi representasi sentimen; dan (3) korelasi antara harga penutupan dengan compound score (0,255) lebih tinggi dibandingkan korelasi dengan return harian (0,114).

Tingginya korelasi antar variabel sentimen, khususnya antara compound score dan rasio sentimen positif, menunjukkan adanya redundansi informasi dalam fitur berbasis sentimen. Hal ini mengindikasikan bahwa variabel-variabel tersebut memiliki sumber informasi yang serupa sehingga tidak sepenuhnya independen satu sama lain. Dalam konteks analisis data keuangan, kondisi ini menunjukkan bahwa penggunaan seluruh variabel sentimen secara simultan dalam model prediktif dapat menimbulkan masalah multikolinearitas yang berpotensi mengurangi stabilitas model.

3.5. Return Harian Berdasarkan Sentimen Dominan

Tabel 3. Rata-rata Return Harian BBKA Berdasarkan Sentimen Dominan Berita

Sentimen Dominan	Jumlah Hari (n)	Rata-rata Return (%)	Keterangan
Positif	177	+0,051%	Return rata-rata positif
Campuran/Netral	25	-0,739%	Return rata-rata negatif terbesar
Negatif	35	-0,533%	Return rata-rata negatif



Gambar 5. Rata-rata Return Harian BBKA Berdasarkan Sentimen Dominan Berita

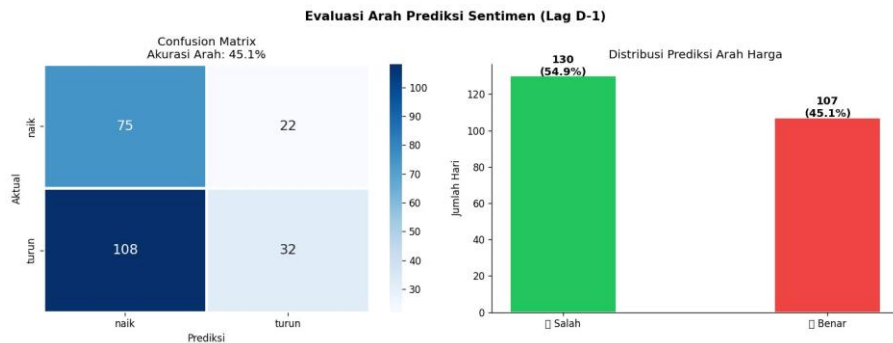
Analisis return harian berdasarkan sentimen dominan berita (Tabel 3) mengungkap pola yang konsisten. Hari-hari dengan sentimen dominan positif ($n = 177$) menghasilkan rata-rata return sebesar +0,051%, sementara hari dengan sentimen dominan negatif ($n = 35$) menghasilkan rata-rata return sebesar -0,533%. Hari dengan sentimen campuran/netral ($n = 25$) justru menunjukkan rata-rata return paling negatif (-0,739%), kemungkinan mencerminkan ketidakpastian investor ketika sinyal pasar tidak jelas.

Perbedaan rata-rata return harian berdasarkan sentimen dominan menunjukkan bahwa arah sentimen memiliki pengaruh terhadap kecenderungan return, meskipun besarnya efek relatif kecil. Kondisi ini mengindikasikan bahwa pasar merespons agregasi informasi sentimen secara kolektif, bukan pada tingkat berita individual. Namun demikian, besarnya perbedaan return antar kategori sentimen masih terbatas, sehingga dapat disimpulkan bahwa sentimen hanya memberikan pengaruh marginal terhadap pembentukan return harian.

3.6. Evaluasi Kemampuan Prediksi Arah Harga

Tabel 4. Evaluasi Prediksi Arah Harga BBKA Menggunakan Sentimen Lag D-1

Metrik Evaluasi	Nilai	Keterangan
Akurasi Prediksi Arah (Lag D-1)	45,1%	Di bawah baseline acak (50%)
Prediksi Benar	107 hari	Dari total 237 hari perdagangan
Prediksi Salah	130 hari	Mayoritas salah prediksi arah
True Positive (naik benar)	32	Prediksi naik, aktual naik
True Negative (turun benar)	75	Prediksi turun, aktual turun
False Positive	22	Prediksi naik, aktual turun
False Negative	108	Prediksi turun, aktual naik



Gambar 6. Confusion Matrix dan Distribusi Prediksi Arah Harga BBCA (Lag D-1)

Tabel 4 dan Gambar 6 menunjukkan hasil evaluasi kemampuan prediksi arah harga menggunakan sentimen lag D-1. Akurasi prediksi hanya mencapai 45,1%, di bawah baseline random 50%. Dari 237 hari perdagangan yang dievaluasi, hanya 107 prediksi yang benar, sementara 130 prediksi salah.

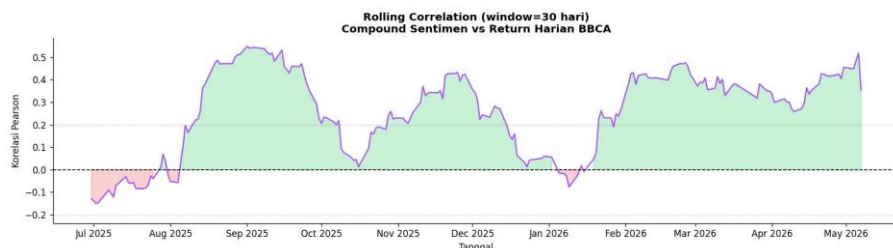
Confusion matrix menunjukkan bahwa model berbasis sentimen lag D-1 cenderung memprediksi harga akan naik (183 dari 237 hari), namun 108 dari prediksi naik tersebut ternyata salah. Temuan ini mengkonfirmasi bahwa sentimen berita tidak dapat digunakan sebagai prediktor tunggal yang andal untuk arah pergerakan harga saham BBCA.

Hasil evaluasi menunjukkan bahwa penggunaan sentimen lag D-1 sebagai prediktor arah harga tidak memberikan performa yang lebih baik dibandingkan model acak. Hal ini mengindikasikan bahwa informasi yang terkandung dalam sentimen berita telah sepenuhnya terefleksi pada harga pada hari yang sama, sehingga tidak memberikan nilai tambah dalam memprediksi pergerakan harga pada periode berikutnya. Temuan ini konsisten dengan hipotesis pasar efisien bentuk semi-kuat yang menyatakan bahwa informasi publik akan segera tercermin dalam harga aset.

3.7. Analisis Rolling Correlation

Tabel 5. Rangkuman Rolling Correlation (Window 30 Hari): Compound Score vs Return Harian BBCA

Periode	Rata-rata Korelasi	Interpretasi
Mei – Jul 2025	-0,05 s.d. 0,05	Korelasi sangat lemah / tidak ada
Agustus – Oktober 2025	0,20 – 0,53	Korelasi positif moderat – kuat
November – Desember 2025	0,05 – 0,42	Korelasi positif lemah – moderat
Januari 2026	-0,12 s.d. 0,05	Korelasi melemah / negatif sementara
Februari – Mei 2026	0,20 – 0,52	Korelasi positif moderat – kuat kembali



Gambar 7. Rolling Correlation 30 Hari: Compound Score Sentimen vs Return Harian BBCA

Analisis rolling correlation mengungkap dinamika temporal yang penting dari hubungan sentimen-return. Terdapat beberapa fase yang dapat diidentifikasi: (1) Fase awal (Mei–Juli 2025): korelasi sangat lemah bahkan negatif; (2) Fase kuat (Agustus–Oktober 2025): korelasi melonjak hingga 0,53; (3) Fase melemah (November

2025–Januari 2026): korelasi berfluktuasi antara -0,12 hingga 0,42; dan (4) Fase pemulihan (Februari–Mei 2026): korelasi kembali menguat ke kisaran 0,20–0,52.

Fluktuasi rolling correlation ini mengimplikasikan bahwa hubungan antara sentimen berita dan return saham tidak bersifat stabil atau konstan. Model prediksi yang adaptif (adaptive/time-varying model) akan lebih efektif dibandingkan model statis.

Fluktuasi rolling correlation menunjukkan bahwa hubungan antara sentimen berita dan return saham bersifat dinamis dan tidak stabil sepanjang waktu. Variasi ini dapat dijelaskan oleh perubahan kondisi pasar yang bersifat regime-dependent, di mana pada periode tertentu ketika volatilitas pasar meningkat, sensitivitas investor terhadap informasi berita menjadi lebih tinggi sehingga korelasi meningkat. Sebaliknya, pada periode pasar yang relatif stabil, pengaruh sentimen terhadap return menjadi lebih lemah. Hal ini menunjukkan bahwa hubungan antara sentimen dan return tidak bersifat konstan, sehingga model linier statis kurang mampu menangkap dinamika tersebut secara optimal.

3.8. Diskusi dan Implikasi

Hasil penelitian ini memberikan beberapa kontribusi penting. Pertama, model zero-shot mDeBERTa-v3 terbukti dapat mengklasifikasikan sentimen berita keuangan berbahasa Indonesia dengan hasil yang koheren dan masuk akal, tanpa memerlukan data berlabel khusus.

Kedua, temuan korelasi signifikan hari yang sama (same-day correlation) antara sentimen dan return ($r = 0,2446$) mengindikasikan adanya kandungan informasi dalam berita keuangan yang tercermin dalam harga. Tidak adanya korelasi prediktif (lag D-1) mengkonfirmasi bahwa pasar Indonesia tergolong semi-efisien dalam hal informasi berita.

Ketiga, akurasi prediksi arah yang hanya 45,1% menunjukkan keterbatasan sentimen sebagai prediktor tunggal. Berbagai faktor fundamental, teknikal, dan makroekonomi memiliki pengaruh yang jauh lebih besar terhadap pergerakan harga saham.

Keempat, pola rolling correlation yang time-varying mengimplikasikan perlunya pendekatan model yang adaptif. Selama periode sentimen relevan tinggi (korelasi $> 0,4$), bobot sentimen dalam model prediksi dapat ditingkatkan, dan sebaliknya.

Penelitian ini memberikan kontribusi dalam tiga aspek utama. Pertama, secara metodologis, penelitian ini menunjukkan bahwa pendekatan zero-shot berbasis mDeBERTa-v3 dapat digunakan untuk analisis sentimen pasar modal tanpa memerlukan proses fine-tuning maupun data berlabel. Kedua, secara empiris, hasil penelitian menunjukkan bahwa pengaruh sentimen berita terhadap return saham BBKA bersifat contemporaneous (same-day effect) dan tidak memiliki daya prediktif untuk periode berikutnya. Ketiga, secara teoritis, temuan ini mendukung hipotesis pasar efisien bentuk semi-kuat, di mana informasi publik tercermin secara cepat dalam harga saham. Dengan demikian, sentimen berita lebih tepat digunakan sebagai variabel pendukung dalam model prediksi multivariat dibandingkan sebagai indikator tunggal dalam pengambilan keputusan investasi.

Dari perspektif praktis, hasil penelitian ini memberikan implikasi bagi investor dan analis pasar bahwa sentimen berita dapat digunakan sebagai indikator tambahan untuk memahami kondisi pasar jangka pendek, terutama pada hari yang sama ketika informasi dipublikasikan. Namun demikian, karena kemampuan prediktifnya terhadap return hari berikutnya relatif rendah, penggunaan sentimen berita sebaiknya dikombinasikan dengan indikator teknikal, fundamental, dan faktor makroekonomi lainnya agar menghasilkan keputusan investasi yang lebih komprehensif.

4. KESIMPULAN

Penelitian ini menganalisis hubungan antara sentimen berita keuangan dan pergerakan harga saham BBKA menggunakan model zero-shot mDeBERTa-v3 pada periode Mei 2025 hingga Mei 2026. Hasil penelitian menunjukkan bahwa model zero-shot mampu mengklasifikasikan 3.113 berita keuangan berbahasa Indonesia ke dalam kategori sentimen positif, negatif, dan netral tanpa memerlukan proses fine-tuning maupun data berlabel khusus. Distribusi sentimen yang diperoleh menunjukkan dominasi sentimen positif sebesar 66,1%, diikuti sentimen negatif sebesar 31,3% dan sentimen netral sebesar 2,6%.

Analisis statistik menunjukkan adanya korelasi positif yang signifikan antara compound score sentimen harian dan return harian BBKA ($r = 0,2446$; $p < 0,001$). Temuan ini mengindikasikan bahwa informasi yang terkandung dalam berita keuangan memiliki hubungan dengan kondisi pasar pada hari yang sama. Namun demikian, korelasi sentimen lag D-1 terhadap return harian tidak signifikan dan akurasi prediksi arah harga hanya mencapai 45,1%, yang menunjukkan bahwa sentimen berita tidak dapat diandalkan sebagai prediktor tunggal pergerakan harga saham pada periode berikutnya. Hasil tersebut mendukung hipotesis pasar efisien bentuk semi-

kuat, di mana informasi publik cenderung tercermin secara cepat ke dalam harga saham. Korelasi sentimen lag D-1 terhadap return harian tidak signifikan ($r = -0,0717$; $p = 0,2715$), mengkonfirmasi bahwa pasar merespons informasi berita pada hari yang sama dan sentimen tidak memiliki daya prediksi ke depan secara mandiri. Akurasi prediksi arah harga menggunakan sentimen lag D-1 hanya 45,1%, di bawah baseline random 50%, sehingga sentimen berita tidak dapat diandalkan sebagai prediktor tunggal arah harga saham.

Analisis rolling correlation juga menunjukkan bahwa hubungan antara sentimen berita dan return saham bersifat dinamis dan berubah sepanjang waktu, dengan nilai korelasi berkisar antara -0,15 hingga 0,53. Temuan ini menunjukkan bahwa relevansi sentimen terhadap pasar bersifat time-varying sehingga pendekatan model prediksi yang adaptif berpotensi memberikan hasil yang lebih baik dibandingkan model statis. Secara praktis, sentimen berita dapat dimanfaatkan sebagai variabel pendukung dalam sistem analisis maupun prediksi harga saham, namun penggunaannya perlu dikombinasikan dengan faktor teknikal, fundamental, dan makroekonomi lainnya. Penelitian selanjutnya disarankan untuk melakukan fine-tuning model pada domain keuangan Indonesia, melakukan validasi manual terhadap hasil klasifikasi sentimen, serta memperluas objek penelitian pada sektor dan emiten lain untuk memperoleh pemahaman yang lebih komprehensif mengenai pengaruh sentimen berita terhadap pasar modal Indonesia.

DAFTAR PUSTAKA

- [1] M. P. Bordoloi and S. K. Biswas, "Sentiment Analysis: A Survey on Design Framework, Applications and Future Scopes," *Artificial Intelligence Review*, vol. 56, no. 4, pp. 3505-3560, 2023, doi: 10.1007/s10462-022-10233-3.
- [2] A. Adebisi and M. A. Alshammari, "Innovative Sentiment Analysis and Prediction of Stock Price Using FinBERT, GPT-4 and Logistic Regression: A Data-Driven Approach," *Big Data and Cognitive Computing*, vol. 8, no. 11, Art. 143, 2024, doi: 10.3390/bdcc8110143.
- [3] S. Linhares, N. Oliveira, P. Cortez, and V. Ramos, "Sentiment Trading with Large Language Models," *arXiv preprint arXiv:2412.19245*, 2025. [Online]. Available: <https://arxiv.org/abs/2412.19245>.
- [4] P. He, J. Gao, and W. Chen, "DeBERTaV3: Improving DeBERTa using ELECTRA-Style Pre-Training with Gradient-Disentangled Embedding Sharing," in *Proc. ICLR 2023*, 2023. [Online]. Available: <https://arxiv.org/abs/2111.09543>.
- [5] M. Laurer, W. van Atteveldt, A. Casas, and K. Welbers, "Building Efficient Universal Classifiers with Natural Language Inference," *arXiv preprint arXiv:2312.17543*, 2023. [Online]. Available: <https://arxiv.org/abs/2312.17543>.
- [6] A. Jadhav and V. Mirza, "Large Language Models in Equity Markets: Applications, Techniques, and Insights," *Frontiers in Artificial Intelligence*, vol. 8, 2025, doi: 10.3389/frai.2025.1608365.
- [7] Y. Lin and S. Wang, "Enhancing Stock Market Prediction with Sentiment Analysis Using a BERT-based Model," *Transactions on Computer Science and Intelligent Systems Research*, vol. 7, pp. 309-315, 2024, doi: 10.62051/tq61jb84.
- [8] C. Zhang, "Dynamic Asset Pricing: Integrating FinBERT-Based Sentiment Quantification with the Fama-French Five-Factor Model," *arXiv preprint arXiv:2505.01432*, 2025. [Online]. Available: <https://arxiv.org/abs/2505.01432>.
- [9] X. Liu, Y. Zhao, H. Wang, and J. Chen, "Stock Price Prediction Using FinBERT-Enhanced Sentiment with SHAP Explainability and Differential Privacy," *Mathematics*, vol. 13, no. 17, Art. 2747, 2025, doi: 10.3390/math13172747.
- [10] C. Zhang, B. Yang, H. Zhou, T. Ali Babar, and X. Y. Liu, "Enhancing Financial Sentiment Analysis via Retrieval Augmented Large Language Models," in *Proc. 4th ACM Int. Conf. AI in Finance*, pp. 349-356, 2023, doi: 10.1145/3604237.3626870.
- [11] S. Linhares et al., "Impact of LLMs News Sentiment Analysis on Stock Price Movement Prediction," *arXiv preprint arXiv:2602.00086*, 2026. [Online]. Available: <https://arxiv.org/abs/2602.00086>.
- [12] T. Gangopadhyay and S. Majumder, "Stock Market Forecasting: From Traditional Predictive Models to Large Language Models," *Computational Economics*, 2025, doi: 10.1007/s10614-025-11024-w.
- [13] A. Huang, H. Wang, and Y. Yang, "FinBERT: A Large Language Model for Extracting Information from Financial Text," *Contemporary Accounting Research*, vol. 40, no. 2, pp. 806-841, 2023, doi: 10.1111/1911-3846.12832.

- [14] Q. Zheng, C. Yu, J. Cao, Y. Zhang, Y. Jin, and M. Zhu, "Predicting Stock Prices with FinBERT-LSTM: Integrating News Sentiment Analysis," in Proc. 8th Int. Conf. Cloud and Big Data Computing, pp. 1-8, 2024, doi: 10.1145/3694860.3694870.
- [15] K. Kirtac and G. Germano, "Enhanced Financial Sentiment Analysis and Trading Strategy Development Using Large Language Models," in Proceedings of the 14th Workshop on Computational Approaches to Subjectivity, Sentiment & Social Media Analysis, 2024, pp. 225–236.