

ANALISIS SENTIMENT KEUANGAN MENGGUNAKAN FINE-TUNED FINBERT

(FINANCIAL SENTIMENT ANALYSIS USING FINBERT FINE-TUNED)

Heroe Santoso¹, Raisul Azhar², Suriyati³, Melati Rosanensi⁴, I Made Yadi Dharma⁵, Husain⁶, Ahmat Adil⁷, Muhamad Azwar⁸, I Putu Hariyadi⁹

^{1,2,3,4,5,6,7,8} Universitas Bumigora

e-mail : heroe.santoso@universitasbumigora.ac.id¹, raisulazhar@universitasbumigora.ac.id², suriyati@universitasbumigora.ac.id³, melati.rn@universitasbumigora.ac.id⁴, yadi_dharma@universitasbumigora.ac.id⁵, husain@universitasbumigora.ac.id⁶, ahmat.adil@universitasbumigora.ac.id⁷, muha.azwar@gmail.com⁸, putu.hariyadi@universitasbumigora.ac.id⁹

ABSTRAK

Informasi keuangan merupakan salah satu jenis informasi yang penting untuk dianalisis. Namun, karena banyak informasi teks keuangan yang bersifat tidak terstruktur dan tersebar luas, diperlukan suatu metode analisis yang tepat, salah satunya adalah melalui analisis sentimen. Dalam konteks keuangan, analisis sentimen digunakan industri keuangan untuk menilai persepsi publik tentang perusahaan atau kondisi pasar. Maka penelitian ini menggunakan model FinBERT yang telah dilakukan fine-tuning untuk melakukan analisis sentimen di bidang keuangan. Dataset yang digunakan yaitu gabungan dari FiQA (Financial Question Answering) dan The Financial Phrase Bank, terdiri dari kalimat berbahasa Inggris dengan label sentimen negative, neutral, dan positive. Proses penelitian meliputi pengumpulan data, tahapan pra-pemrosesan data, pelatihan model, dan evaluasi menggunakan metrik accuracy, precision, recall, F1-score, serta confusion matrix. Hasil penelitian menghasilkan model mampu mencapai akurasi 82%, dengan performa terbaik pada kelas positive yaitu F1-score 0.88 dan neutral F1-score 0.85, namun masih lemah dalam mendeteksi kelas negative F1-score 0.49. Temuan ini menunjukkan bahwa fine-tuned FinBERT efektif digunakan untuk analisis sentimen keuangan, khususnya pada sentimen positive dan neutral, diperlukan peningkatan pada deteksi sentimen negative, dengan menambah variasi data latih atau teknik data augmentation.

Kata Kunci: Analisis Sentimen, FinBERT Fine-Tuning, Teks Keuangan

ABSTRACT

Financial information is a critical type of data for analysis. However, because much of it is unstructured and widely dispersed, an appropriate analytical method is required, one of which is sentiment analysis. In the financial context, sentiment analysis is employed by the industry to assess public perceptions of companies or market conditions. This study implements a fine-tuned FinBERT model to perform sentiment analysis in the financial sector. The dataset used is a combination of FiQA (Financial Question Answering) and The Financial Phrase Bank, consisting of English sentences labeled with negative, neutral, and positive sentiments. The research process involved data preprocessing, tokenization, data splitting, model training, and evaluation utilizing criteria such as F1-score, confusion matrix, recall, accuracy, and precision. The findings demonstrate that the model accomplished 82% accuracy, with its best performance in the positive class (F1-score 0.88) and the neutral class (F1-score 0.85), but weaker performance in detecting the negative class (F1-score 0.49). These findings indicate that the fine-tuned FinBERT is effective for financial sentiment analysis, particularly for positive and neutral sentiments, though improvements are needed in negative sentiment detection, potentially through expanding training data diversity or applying data augmentation techniques

Keywords: Sentiment Analysis, FinBERT Fine-Tuning, Financial Text

I. PENDAHULUAN

Informasi tentang keuangan tersebar luas dalam bentuk teks tidak terstruktur seperti berita, ulasan, dan media sosial yang dapat digunakan oleh investor, analis, dan lembaga keuangan untuk pengambilan keputusan. Salah satu pendekatan

yang digunakan untuk menggali informasi ini yakni *sentiment analysis* atau analisis sentimen[1]. Analisis sentimen adalah metode dalam pemrosesan bahasa alami (*Natural Language Processing*) dimaksudkan untuk mengidentifikasi dan mengklasifikasikan opini seseorang tentang suatu subjek atau entitas [2]. Dalam konteks

keuangan, analisis sentimen digunakan industri keuangan untuk menilai persepsi publik tentang perusahaan atau kondisi pasar. Meskipun pendekatan konvensional cukup efektif dalam tugas dasar klasifikasi, pendekatan tersebut masih memiliki keterbatasan dalam memahami konteks semantik yang kompleks pada teks keuangan. [3].

Seiring kemajuan teknologi NLP, model berbasis transformer seperti BERT (*Bidirectional Encoder Representations from Transformers*) telah terbukti meningkatkan pemahaman dan pemrosesan bahasa [4]. Salah satu varian dari model ini adalah FinBERT, yaitu sebuah model bahasa yang telah dilatih khusus untuk domain keuangan, sehingga lebih relevan untuk tugas analisis sentimen dalam sektor tersebut [5].

Dalam penelitian ini, dilakukan proses *fine-tuning* terhadap model FinBERT menggunakan data dari dua sumber utama, yaitu FiQA (*Financial Question Answering*) dan *The Financial PhraseBank*. Kalimat-kalimat berbahasa Inggris dari sumber keuangan di kedua dataset ini telah dilabeli dengan sentimen negatif, netral, dan positif. Tujuan dari penggunaan kombinasi kedua dataset ini untuk meningkatkan keragaman dan tampilan data dalam pelatihan model.

Dengan menggunakan FinBERT yang telah disesuaikan dengan data keuangan ini, diharapkan model akan menghasilkan akurasi dan kinerja yang lebih baik dalam mengklasifikasikan sentimen kategori negatif, netral, dan positif. Selain itu, penelitian ini membantu pembangunan sistem analisis sentimen di bidang keuangan dalam hal pengolahan data teks berbasis domain.

II. STUDI PUSTAKA

Penelitian yang oleh A. M. Priyatno and F. I. Firmanand [6] membandingkan beberapa metode *machine learning* seperti *Naïve Bayes*, *Logistic Regression*, SVM, MLP, dan *Decision Tree* dalam analisis sentimen judul berita keuangan menggunakan fitur n-gram. Fitur n-gram digunakan untuk mengekstraksi informasi dari teks dengan membagi kalimat menjadi satuan kata, yang kemudian diberi bobot berdasarkan frekuensinya. Hasil penelitian menunjukkan bahwa *Logistic Regression* merupakan metode terbaik dengan nilai *f1-measure* sekitar 74%. Penggunaan fitur n-gram dan *machine learning* terbukti efektif dalam analisis sentimen untuk

memahami pengaruh berita keuangan terhadap reputasi perusahaan

Lain halnya dengan J. S. Hutagalung and Rasiban [7] membahas pengembangan model analisis sentimen keuangan dengan dataset gabungan dari FiQA dan *Financial PhraseBank*, dianalisis dengan *Logistic Regression* dan SVM. Hasil menunjukkan model SVM tanpa teknik *Smote* mencapai akurasi tertinggi sebesar 70,70%. Penelitian ini bertujuan untuk membantu pengambilan kebijakan di sektor keuangan, dengan menekankan pentingnya teknik pra-pemrosesan data, *tuning hyperparameter*, dan evaluasi model untuk mendapatkan hasil yang optimal.

Selanjutnya pada penelitian J. Y. Huang, C. L. Tung, and W. Z. Lin [8] menganalisis sentimen di media sosial terkait TSMC dengan 2.177 ulasan dan sekitar 30.000 balasan di PTT. Menggunakan metode kombinasi algoritma genetik (HGA) dan model, dengan tiga level yaitu over, very, dan extreme. Hasil akhirnya menunjukkan bahwa setelah menambahkan indikator sentimen, akurasi mencapai 62,22%, meningkat 2,22% dibandingkan model tanpa variabel sentimen, sehingga menunjukkan bahwa penambahan variabel sentimen dapat meningkatkan akurasi prediksi.

Sedangkan L. Nurina, S. H. Hairuddin, A. A. Bakri, and A. Pilua [9] melakukan tinjauan bibliometrik mengenai tren perkembangan pajak menggunakan Big Data, analisis sentimen, dan kriptokurensi dari tahun 1903 - 2023. Temuan utama menunjukkan pertumbuhan signifikan dalam penelitian terkait teknologi canggih, serta kesenjangan penelitian kolaborasi interdisipliner dan dampak etika. Penelitian ini menekankan perlunya eksplorasi lebih lanjut mengenai pengaruh kriptokurensi terhadap perpajakan untuk meningkatkan pemahaman dan efektivitas sistem pajak.

Penelitian berikutnya E. A. Marwa and A. B. Kristanto [10] memeriksa kesan dari laporan tahunan perusahaan yang terdaftar di Bursa Efek Indonesia dari tahun 2016 hingga 2020. Metode VADER (*Dictionary and Sentiment Aware Reasoner*) dan aplikasi *Orange Data Mining*. Metode VADER digunakan untuk menempatkan sentimen dalam kategori positif, negatif, dan netral. Hasil akhir menunjukkan bahwa rata-rata sentimen positif berkisar antara 11,20% - 12,65%, sementara sentimen negatif berkisar antara 3% - 4%. Dimana menunjukkan sentimen positif lebih

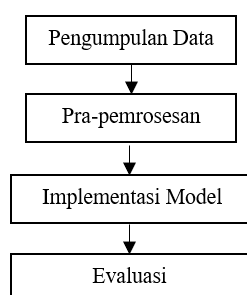
dominan dibandingkan sentimen negatif, dan karakteristik perusahaan tidak mempengaruhi pilihan kata sentimen, sehingga investor perlu waspada terhadap potensi bias dalam pengungkapan informasi.

Berbeda dengan penelitian sebelumnya yang masih menggunakan metode *machine learning* klasik seperti Naïve Bayes, Logistic Regression, atau SVM, penelitian ini menggunakan pendekatan berbasis transformer FinBERT fine tuned yang secara khusus megolah teks keuangan. Pendekatan ini memberikan kemampuan yang lebih baik dalam memahami konteks semantik keuangan, sehingga menghasilkan performa klasifikasi yang lebih tinggi dan relevan.

III. METODE PENELITIAN

Penelitian ini menggunakan penelitian kuantitatif dengan metode eksperimental. Tujuan utama penelitian ini adalah untuk mengetahui seberapa baik model FinBERT melakukan klasifikasi tiga kategori sentimen teks keuangan: negatif, netral, dan positif. Pendekatan eksperimen dilakukan dengan melatih model menggunakan data berlabel, kemudian mengukur performanya menggunakan metrik evaluasi.

Metode penelitian yang dilakukan yaitu, pengumpulan data, pra- pemrosesan, implemtasi model hingga tahap evaluasi model, yang ditampilkan pada Gambar 1.



Gambar 1. Metode penelitian

A. Pengumpulan Data

Menggunakan data gabungan dari dua *dataset open-source* yang berasal dari situs Kaggle.com, yaitu FiQA (*Financial Question Answering*) adalah *dataset* berisi kutipan teks dari berita atau konten keuangan yang bertujuan untuk investasi pengembangan analisis NLP yang terus berkembang [2]. *The Financial PhraseBank* merupakan dataset analisis sentimen di bidang keuangan yang telah dikategorikan berdasarkan

polaritas sentimen[11]. Label pada data terdiri dari tiga kelas, yaitu negatif (0), netral (1), dan positif (2).

B. Pra-pemrosesan

Metode pra-pemrosesan digunakan untuk menyiapkan data dan mendapatkan hasil yang terbaik berdasarkan matriks evaluasi yang digunakan [12]. Proses pra-pemrosesan dilakukan beberapa tahapan yaitu, pembersihan teks (*text cleaning*) dengan menghapus karakter khusus, tanda baca, angka, dan elemen lain yang tidak diperlukan dalam proses analisis sentimen[13]. Tokenisasi dilakukan untuk mengubah kalimat menjadi token yang sesuai dengan arsitektur FinBERT. Kemudian label teks sentimen dikonversi ke bentuk numerik (0, 1, 2) sesuai kebutuhan klasifikasi multi-kelas. Dilanjutkan dengan pembagian data dengan persentase 80% training dan 20% testing.

C. Implementasi Model

Menggunakan model FinBERT, yaitu model *BERT-base* yang telah dilatih khusus pada korpus keuangan. *Trainer API* digunakan untuk menyempurnakan proses dengan menyesuaikan jumlah label 3 kelas. *Batch training* dilakukan pada data latih dengan ukuran tertentu yang disesuaikan dengan kapasitas perangkat. Model menggunakan representasi semantik kalimat keuangan untuk mempelajari pola distribusi sentimen.

D. Evaluasi

Dalam proses evaluasi, data uji digunakan untuk mengevaluasi keakuratan dan keefektifan model dalam klasifikasi sentimen keuangan dengan membandingkan hasil prediksi dengan label sebenarnya [14]. Untuk menilai kinerja model menggunakan metrik *accuracy*, *precision*, *recall*, *f1-score* dan *confusion matrix*. Evaluasi ini memberikan gambaran menyeluruh mengenai kemampuan model dalam mengklasifikasikan sentimen negatif, netral, dan positif dalam teks keuangan.

IV. HASIL DAN PEMBAHASAN

Setelah melalui tahapan metode penelitian yang telah dijelaskan, diperoleh hasil dan pembahasan sebagai berikut.

A. Pengumpulan Data

Dataset yang digunakan hasil penggabungan data FiQA (*Financial Question Answering*) dan *Financial PhraseBank*. Terdiri dari 5.843 kalimat berbahasa Inggris yang diambil dari berbagai sumber keuangan, seperti berita pasar, laporan analisis, dan pernyataan publik. Setiap kalimat telah diberi label sentimen dalam tiga kelas: negatif (0), netral (1), dan positif (2). “Pelabelan data tidak dilakukan manual, melainkan menggunakan label yang sudah tersedia pada dataset FiQA dan Financial PhraseBank, yang sebelumnya telah dianotasi oleh para ahli di bidang keuangan. Dengan demikian, penelitian ini langsung menggunakan label yang ada untuk proses pelatihan dan evaluasi model

Tabel 1. Dataset.

No	Kalimat	Sentiment
1	The GeoSolutions technology will leverage Bene...	positive
2	\$ESI on lows, down \$1.50 to \$2.50 BK a real po...	negative
3	For the last quarter of 2010, Componenta 's n...	positive
4	According to the Finnish-Russian Chamber of Co...	positive
5	The Swedish buyout firm has sold its remaining...	neutral

B. Pra-pemrosesan

Tahap pra-pemrosesan dilakukan untuk meningkatkan kualitas teks sebelum diproses oleh model FinBERT. Ini mencakup proses pembersihan teks dengan menghapus simbol, URL, karakter non-alfabet, dan mengubah semua huruf menjadi huruf kecil, yang ditunjukkan pada Gambar 2. Setelah itu, kalimat diubah menjadi token, yang sesuai dengan arsitektur FinBERT. Data tersebut dilakukan pembagian dengan komposisi 80% data latih dan 20% data uji.

Sentence \	
0	The GeoSolutions technology will leverage Bene...
1	\$ESI on lows, down \$1.50 to \$2.50 BK a real po...
2	For the last quarter of 2010 , Componenta 's n...
3	According to the Finnish-Russian Chamber of Co...
4	The Swedish buyout firm has sold its remaining...
Cleaned	
0	the geosolutions technology will leverage bene...
1	esi on lows down 150 to 250 bk a real possibility
2	for the last quarter of 2010 componenta s net ...
3	according to the finnishrussian chamber of com...
4	the swedish buyout firm has sold its remaining...

Gambar 2. Pembersihan data

C. Implementasi Model

Dengan menggunakan dua *epoch*, model FinBERT melakukan proses *fine-tune* dengan mempertimbangkan metrik *training loss*, *validation loss*, dan *accuracy*. Pada epoch 1, *training loss* bernilai 0.3108, *validation loss* 0.4024 dan *accuracy* 81,27%. Pada epoch 2, *training loss* menurun menjadi 0.2842, yang menunjukkan bahwa kemampuan model semakin efektif dalam mempelajari pola data latih. Namun, *validation loss* sedikit meningkat menjadi 0.4300, yang dapat mengindikasikan adanya gejala *overfitting*. Sementara itu, akurasi pada data validasi meningkat menjadi 81,52%, menandakan performa model relatif stabil

Secara keseluruhan, nilai akurasi di atas 80% menunjukkan bahwa model memiliki kemampuan klasifikasi yang baik, tetapi peningkatan kehilangan validasi menunjukkan potensi *overfitting* dan kemungkinan perbaikan kinerja melalui penyesuaian *hyperparameter*.

Tabel 2. Hasil Pelatihan Model FinBERT.

Epoch	Training Loss	Validation Loss	Accuracy
1	0.3108	0.402396	0.812660
2	0.284200	0.429987	0.815227

Pelatihan tidak dilanjutkan ke epoch berikutnya karena pada epoch ke-2 sudah terlihat indikasi *overfitting* yaitu *validation loss* meningkat meskipun *training loss* menurun. Jika dilanjutkan ke epoch berikutnya, besar kemungkinan model akan semakin menghafal data latih dan menurunkan generalisasi terhadap data uji. Selain itu, akurasi pada epoch kedua sudah mencapai 81,52%, yang relatif stabil dan cukup tinggi untuk dataset yang digunakan. Oleh karena itu, penelitian ini memilih menggunakan hasil terbaik sampai epoch ke-2 untuk menghindari kompleksitas tambahan tanpa peningkatan signifikan pada performa.

D. Evaluasi

Hasil evaluasi *fine-tuning* FinBERT menunjukkan hasil bervariasi pada setiap kelas sentimen. Dalam kelas *negative*, model menghasilkan nilai *precision* sebesar 0.58, *recall* 0.43, dan nilai *F1-score* 0.49, yang dianggap rendah dan menunjukkan bahwa model mengalami kesulitan untuk mendeteksi sentimen *negative*, mungkin dikarenakan jumlah data *negative* yang relatif sedikit atau variasi bahasa

yang lebih kompleks. Sebaliknya, pada kelas *neutral*, model menunjukkan akurasi dan keseimbangan yang tinggi dengan *precision* sebesar 0.83, *recall* 0.87 dan *F1-score* 0.85. Kinerja terbaik ditunjukkan pada kelas *positive* dengan *precision* 0.87, *recall* 0.90, dan *F1-score* 0.88, menandakan kemampuan model yang sangat baik dalam mengidentifikasi sentimen positif. Secara keseluruhan, akurasi model mencapai 82%, dengan nilai *macro average F1-score* sebesar 0,74 dan *weighted average F1-score* sebesar 0,81. Pada Gambar 3 telah ditampilkan hasil yang memuat *classification report* dari model FinBERT. Hasil menunjukkan kemampuan model untuk mengklasifikasikan sentimen keuangan dengan cukup baik, terutama pada sentimen *neutral* dan *positive*, namun masih perlu peningkatan pada deteksi sentimen *negative* melalui penambahan data latih atau penerapan teknik *data augmentation*.

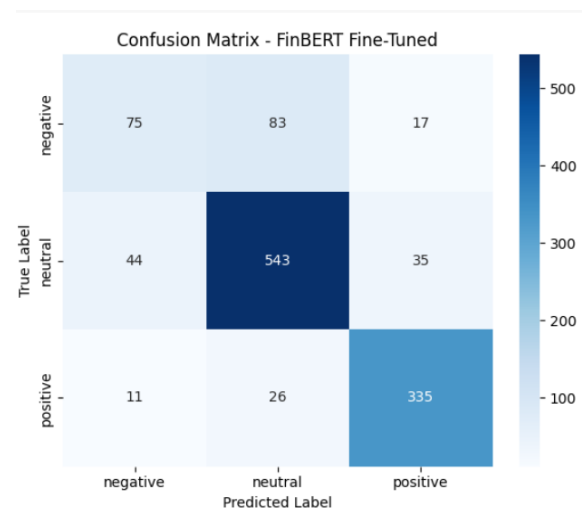
Classification Report (FinBERT Fine-Tuned):				
	precision	recall	f1-score	support
negative	0.58	0.43	0.49	175
neutral	0.83	0.87	0.85	622
positive	0.87	0.90	0.88	372
accuracy			0.82	1169
macro avg	0.76	0.73	0.74	1169
weighted avg	0.80	0.82	0.81	1169

Gambar 3. Hasil evaluasi model *FinBERT*

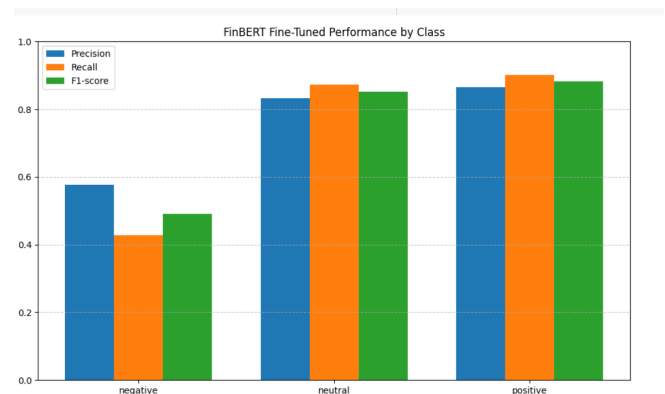
Hasil *confusion matrix fine-tuning* FinBERT, terlihat pada Gambar 4 menunjukkan model mampu mengklasifikasikan kelas *neutral* dan *positive* dengan baik. Dari 622 data *neutral*, sebanyak 543 diprediksi benar, sedangkan 44 diklasifikasikan sebagai *negative* atau *positive* 35. Pada kelas *positive*, dari 372 data, sebanyak 335 diprediksi benar dan hanya 26 salah diklasifikasikan sebagai *neutral* atau *negative* 11 data. Namun, pada kelas *negative*, dari 175 data, hanya 75 yang diprediksi benar, sementara 83 salah diklasifikasikan sebagai *neutral* dan 17 sebagai *positive*. Hal ini memperkuat temuan sebelumnya bahwa model masih mengalami kesulitan dalam mendeteksi sentimen negatif, sedangkan performa pada kelas *neutral* dan *positive* tergolong sangat baik.

Hasil dari *confusion matrix* semakin jelas pada grafik diagram batang yang ditampilkan pada Gambar 5. Untuk kelas *positive*, nilai *precision*, *recall*, dan *F1-score* berada di antara nilai 0.87 -

0.90, yang menunjukkan bahwa model mendeteksi sentimen *positive* dengan sangat baik. Sebaliknya, kelas *negative* memiliki kinerja yang lebih rendah, dengan nilai *recall* sekitar 0.43, yang menunjukkan bahwa banyak sampel yang salah diklasifikasikan. *Matrix* kelas *neutral* stabil dengan nilai *precision*, *recall*, dan *F1-score* antara 0.83 - 0.87. Secara keseluruhan, hasil ini menunjukkan bahwa model mendeteksi sentimen *neutral* dan *positive* dengan baik, tetapi perlu ada peningkatan untuk mendeteksi sentimen *negative*.



Gambar 4. *Confusion matrix*



Gambar 5. Grafik Batang *Confusion matrix*

V. KESIMPULAN

Dengan menggunakan kombinasi *dataset* FiQA dan *The Financial PhraseBank*, penelitian ini berhasil menerapkan FinBERT untuk analisis sentimen teks keuangan. Hasil evaluasi menunjukkan bahwa model memiliki akurasi sebesar 82% dengan formasi terbaik untuk sentimen *positive* dan *neutral*, dengan nilai *F1-score* masing-masing 0.88 dan 0.85. Namun, kinerja model untuk kelas *negative* masih rendah,

dengan nilai *F1-score* 0.49. Ini menunjukkan bahwa ada upaya untuk meningkatkan kinerja melalui penambahan data latihan, penerapan teknik peningkatan data, atau penyesuaian *hyperparameter*. Secara keseluruhan, temuan ini menunjukkan bahwa FinBERT yang disesuaikan dengan baik berhasil mengklasifikasikan sentimen keuangan, serta menunjukkan bahwa model dapat digunakan dalam aplikasi analisis sentimen di sektor keuangan. Penelitian selanjutnya disarankan untuk mengeksplorasi teknik peningkatan kinerja pada kelas minoritas dan membandingkan performa FinBERT dengan model *transformer*-based lainnya.

DAFTAR PUSTAKA

- [1] H. I. Naibaho, K. Jeremi, D. Saragih, R. C. Nasution, S. H. Zebua, and A. N. Wulandari, "Use of Indonesian in Financial Reports for Investor Understanding," *As-Salam J. Islam. Soc. Sci. Humanit.*, vol. 2, no. 4, pp. 95–105, 2024, [Online]. Available: <https://ejournal.as-salam.org/index.php/assalam/article/view/74>
- [2] M. N. Brilianto, Y. Susanti, and E. Zukhronah, "Analisis Sentimen terhadap Kalimat Finansial pada FiQA dan The Financial PhraseBank," *PYTHAGORAS J. Pendidik. Mat.*, vol. 18, no. 1, pp. 48–55, 2023, doi: 10.21831/pythagoras.v18i1.59760.
- [3] N. Ima Nafila and S. Sulisetijono, "Melampaui Pembelajaran Konvensional: Mengintegrasikan Canva Dan Pembelajaran Berbasis Game Dalam Lkpd Untuk Motivasi Optimal Dalam Pendidikan Digital," *J. Inov. Teknol. dan Edukasi Tek.*, vol. 4, no. 1, p. 3, 2024, doi: 10.17977/um068.v4.i1.2024.3.
- [4] M. Amien and G. Frendi Gunawan, "BERT dan Bahasa Indonesia: Studi tentang Efektivitas Model NLP Berbasis Transformer," *ELANG J. Interdiscip.*, vol. 1, pp. 132–140, 2024, doi: 10.32664/elang.v1i02.
- [5] N. Anggraini, D. Arman Prasetya, and T. Trimono, "Prediksi Harga Saham Sektor Energi Menggunakan Metode Spatial Temporal Attention-Based Convolutional Network Berdasarkan Data Teks Dan Numerik," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 9, no. 3, pp. 3872–3880, 2025, doi: 10.36040/jati.v9i3.13443.
- [6] A. M. Priyatno and F. I. Firmanand, "Fitur n-gram untuk perbandingan metode machine learning pada sentimen judul berita keuangan," *J. Artif. Intell. Digit. Bus.*, vol. 1, no. 1, pp. 1–6, 2022, doi: 10.31004/riggs.v1i1.4.
- [7] J. S. Hutagalung and Rasiban, "ANALISIS SENTIMEN KEUANGAN (DATA FIQA AND FINANCIAL PHRASEBANK) MENGGUNAKAN ALGORITMA LOGISTIC REGRESSION DAN SUPPORT VECTOR MACHINE," *J. Indones. Manaj. Inform. dan Komun.*, vol. 4, no. 3, pp. 1654–1669, 2023, doi: 10.35870/jimik.v4i3.404.
- [8] J. Y. Huang, C. L. Tung, and W. Z. Lin, "Using Social Network Sentiment Analysis and Genetic Algorithm to Improve the Stock Prediction Accuracy of the Deep Learning-Based Approach," *Int. J. Comput. Intell. Syst.*, vol. 16, no. 1, 2023, doi: 10.1007/s44196-023-00276-9.
- [9] L. Nurina, S. H. Hairuddin, A. A. Bakri, and A. Pilua, "Tinjauan Bibliometrik Terhadap Pemanfaatan Big Data, Analisis Sentimen, dan Kriptokurensi dalam Analisis Pajak," *Sanskara Akunt. dan Keuang.*, vol. 2, no. 01, pp. 66–76, 2023, doi: 10.58812/sak.v2i01.257.
- [10] E. A. Marwa and A. B. Kristanto, "Analisis Sentimen Pengungkapan Informasi Manajemen: Text Mining Berbasis Metode VADER," *Own. Ris. J. Akunt.*, vol. 6, no. 3, pp. 2853–2864, 2022, doi: 10.33395/owner.v6i3.895.
- [11] N. P. I. Maharani, A. Purwarianti, Y. Yustiawan, and F. C. Rochim, "Domain-Specific Language Model Post-Training for Indonesian Financial NLP," in *Proceedings of the International Conference on Electrical Engineering and Informatics*, Bandung, Indonesia: 2023 International Conference on Electrical Engineering and Informatics (ICEEI), 2023. doi: 10.1109/ICEEI59426.2023.10346625.
- [12] B. Nugroho and A. Denih, "Perbandingan Kinerja Metode Pra-Pemrosesan Dalam Pengklasifikasian Otomatis Dokumen Paten," *Komputasi J. Ilm. Ilmu Komput. dan Mat.*, vol. 17, no. 2, pp. 381–387, 2020, doi: 10.33751/komputasi.v17i2.2148.
- [13] M. Saputra and Sri Wahyuni, "Analisis Sentimen Pengguna Pada Aplikasi Bank Digital Krom Dengan Algoritma Support Vector Machine," *INFOTECH J.*, vol. 10, no. 2, pp. 327–332, 2024, doi: 10.31949/infotech.v10i2.11801.
- [14] L. Palupi, E. Ihsanto, and F. Nugroho, "Analisis Validasi dan Evaluasi Model Deteksi Objek Varian Jahe Menggunakan Algoritma Yolov5," *J. Inf. Syst. Res.*, vol. 5, no. 1, pp. 234–241, 2023, doi: 10.47065/josh.v5i1.4380.