

# Analisis Sentimen Reddit untuk Memprediksi Tren Harga *Meme Coin Chillguy* Menggunakan *Machine Learning*

<sup>1\*</sup> Deasy Indayanti, <sup>2</sup> Muhammad Irfan Tri Atsal, <sup>3</sup> Siti Chodidjah

<sup>1\*</sup> Prodi Sistem Informasi, Fakultas Ilmu Komputer dan Teknologi Informasi,  
Universitas Gunadarma, Depok, Indonesia

<sup>2</sup> Prodi Informatika, Fakultas Teknologi Industri, Universitas Gunadarma, Depok, Indonesia

<sup>3</sup> Prodi Teknik Komputer, Direktorat Diploma Tiga Teknologi Informasi  
Universitas Gunadarma, Depok, Indonesia

Email : <sup>1\*</sup> [deasy@staff.gunadarma.ac.id](mailto:deasy@staff.gunadarma.ac.id), <sup>2</sup> [irfan.basiru@gmail.com](mailto:irfan.basiru@gmail.com),  
<sup>3</sup> [chodi@staff.gunadarma.ac.id](mailto:chodi@staff.gunadarma.ac.id)

Submit : 22 Mar 26 | Diterima : 28 April 2026 | Terbit : 06 Mei 2026

## ABSTRAK

Pergerakan harga *meme coin* cenderung dipengaruhi oleh dinamika sentimen komunitas dibandingkan faktor fundamental yang umum digunakan pada aset keuangan konvensional. Kondisi tersebut mendorong pemanfaatan data media sosial sebagai sumber informasi alternatif dalam analisis pasar kripto. Penelitian ini bertujuan mengembangkan model prediksi tren harga *meme coin* CHILLGUY dengan memanfaatkan sentimen komunitas Reddit melalui pendekatan *Natural Language Processing* (NLP) dan *machine learning*. Data penelitian terdiri atas unggahan Reddit terkait CHILLGUY yang dikumpulkan selama periode November 2024 hingga April 2025 serta data harga harian yang diperoleh dari CoinGecko API. Analisis sentimen dilakukan menggunakan model *RoBERTa* yang telah di-*fine-tune* pada data Reddit untuk mengklasifikasikan opini pengguna ke dalam kategori positif, netral, dan negatif. Nilai sentimen harian kemudian diintegrasikan dengan data harga sebagai variabel prediktor dalam proses klasifikasi arah pergerakan harga. Model yang digunakan meliputi *Logistic Regression* dan *Random Forest*. Hasil penelitian menunjukkan bahwa *Random Forest* menghasilkan performa terbaik dengan tingkat akurasi sebesar 70,97%, sedangkan *Logistic Regression* menunjukkan kinerja yang lebih rendah akibat ketidakseimbangan distribusi kelas pada data sentimen. Temuan ini menunjukkan bahwa sentimen komunitas media sosial memiliki potensi sebagai indikator pendukung dalam memprediksi tren harga aset kripto berbasis komunitas.

**Kata Kunci:** *Machine Learning*, *Natural Language Processing*, Prediksi Tren Harga, Reddit

## PENDAHULUAN

Perkembangan industri aset kripto dalam beberapa tahun terakhir telah menghasilkan berbagai instrumen digital dengan karakteristik yang berbeda dari pasar keuangan tradisional. Salah satu fenomena yang menonjol adalah munculnya *meme coin*, yaitu aset kripto yang memperoleh perhatian pasar terutama melalui aktivitas komunitas dan viralitas media sosial. Berbeda dengan aset yang memiliki fundamental ekonomi yang jelas, nilai *meme coin* lebih banyak dipengaruhi oleh persepsi publik, sentimen investor, serta interaksi pengguna pada platform digital (Ante, 2023).

Meningkatnya peran media sosial dalam ekosistem kripto menjadikan informasi yang dihasilkan pengguna (*user-generated content*) sebagai sumber data yang relevan untuk memahami perilaku pasar. Berbagai penelitian menunjukkan bahwa sentimen yang terbentuk melalui media sosial memiliki hubungan dengan perubahan harga dan volatilitas aset kripto. Informasi yang disebarkan secara cepat melalui komunitas daring dapat memengaruhi keputusan investasi serta menciptakan tekanan beli maupun jual dalam waktu singkat (Nasir et al., 2021).

Salah satu *meme coin* yang menarik perhatian komunitas kripto adalah CHILLGUY. Aset digital ini memperoleh popularitas melalui penyebaran konten berbasis karakter *Just a Chill Guy* yang viral di berbagai platform media sosial. Tingginya eksposur komunitas menyebabkan peningkatan aktivitas perdagangan dalam waktu singkat, namun juga diikuti fluktuasi harga yang signifikan. Kondisi tersebut menunjukkan bahwa pergerakan harga CHILLGUY lebih dipengaruhi oleh faktor psikologis dan sentimen pasar dibandingkan nilai intrinsik aset itu sendiri (Yukun et al., 2022).

Di antara berbagai platform media sosial, Reddit merupakan salah satu media yang banyak digunakan oleh komunitas kripto untuk berbagi informasi, opini, dan prediksi pasar. Diskusi yang terjadi dalam berbagai *subreddit* menghasilkan data tekstual dalam jumlah besar yang dapat dimanfaatkan untuk mengidentifikasi sentimen komunitas secara real time. Menurut Sundarasen dan Saleem (2025), analisis sentimen berbasis media sosial telah berkembang menjadi salah satu pendekatan penting dalam penelitian terkait prediksi harga aset kripto karena mampu merepresentasikan persepsi kolektif investor secara langsung.

Kemajuan teknologi *Natural Language Processing* (NLP), khususnya model berbasis *transformer*, memungkinkan ekstraksi informasi sentimen dari teks media sosial dengan tingkat akurasi yang semakin tinggi. Model *RoBERTa* telah terbukti efektif dalam memahami konteks bahasa informal yang umum digunakan pada platform digital dan mampu meningkatkan kualitas klasifikasi sentimen dibandingkan pendekatan tradisional (Koszweski et al., 2024). Hasil analisis sentimen tersebut selanjutnya dapat digunakan sebagai fitur dalam model *machine learning* untuk memprediksi arah pergerakan harga aset kripto.

Dalam penelitian prediksi harga aset digital, algoritma klasifikasi sering digunakan untuk menentukan arah pergerakan harga berdasarkan variabel masukan tertentu. Salah satu algoritma yang umum digunakan adalah *Logistic Regression* yang memiliki interpretasi sederhana dan efektif untuk hubungan linear. Namun, ketika pola data bersifat kompleks dan non-linear, model berbasis *ensemble* seperti *Random Forest* cenderung memberikan performa yang lebih baik karena mampu menggabungkan banyak pohon keputusan untuk menghasilkan prediksi yang lebih stabil dan akurat (Naseem et al., 2023).

Berdasarkan latar belakang tersebut, penelitian ini bertujuan menerapkan teknik NLP pada data Reddit untuk mengidentifikasi sentimen komunitas terhadap CHILLGUY dan mengintegrasikan hasilnya ke dalam model prediksi tren harga berbasis *machine learning*. Penelitian ini diharapkan dapat memberikan kontribusi terhadap pengembangan metode prediksi harga aset digital dengan memanfaatkan indikator non-tradisional yang bersumber dari media sosial, sekaligus memperluas pemanfaatan analisis sentimen dalam kajian pasar kripto modern.

## TINJAUAN PUSTAKA

### Natural Language Processing (NLP)

*Natural Language Processing* (NLP) merupakan cabang kecerdasan buatan yang berfokus pada kemampuan komputer dalam memahami, mengolah, dan menghasilkan bahasa manusia secara otomatis. Dalam konteks analisis media sosial, NLP digunakan untuk mengekstraksi informasi, opini, dan sentimen dari data teks yang tidak terstruktur. Penerapan NLP pada media sosial menjadi semakin penting karena tingginya volume data yang dihasilkan pengguna setiap hari. Melalui teknik NLP, informasi yang tersembunyi dalam unggahan dan komentar dapat dikonversi menjadi data terstruktur yang mendukung proses analisis dan pengambilan keputusan berbasis data (Liu et al., 2023). Kemampuan NLP dalam mengidentifikasi pola linguistik menjadikannya pendekatan yang relevan untuk mendukung analisis perilaku pasar berbasis opini publik (Koszweski et al., 2024).

### Analisis Sentimen

Analisis sentimen merupakan teknik dalam NLP yang digunakan untuk mengidentifikasi dan mengklasifikasikan opini pengguna ke dalam kategori positif, negatif, atau netral (Lupu & Donoiu, 2025). Pada lingkungan media sosial, analisis sentimen berperan penting dalam memahami persepsi masyarakat terhadap suatu topik, produk, maupun aset digital. Penelitian terkini menunjukkan bahwa sentimen yang diekstraksi dari media sosial memiliki hubungan dengan

perilaku pasar dan dapat digunakan sebagai indikator tambahan dalam proses pengambilan keputusan. Oleh karena itu, analisis sentimen banyak dimanfaatkan dalam bidang keuangan untuk mendeteksi perubahan persepsi investor secara real-time (Sundarasen & Saleem, 2025).

### Reddit

Reddit merupakan salah satu platform diskusi daring yang memiliki komunitas aktif dan terorganisasi berdasarkan topik tertentu melalui fitur *subreddit*. Dalam ekosistem aset kripto, Reddit menjadi media yang banyak digunakan untuk berbagi informasi, opini, serta spekulasi mengenai perkembangan pasar. Karakteristik Reddit yang berbasis diskusi terbuka menjadikannya sumber data yang kaya untuk penelitian analisis sentimen karena mencerminkan persepsi komunitas secara lebih mendalam dibandingkan media sosial lain yang cenderung berorientasi pada interaksi singkat (Medvedev et al., 2022). Informasi yang dihasilkan dari interaksi pengguna dapat mencerminkan kondisi psikologis pasar dan sering kali muncul sebelum terjadinya perubahan harga yang signifikan (Sundarasen & Saleem, 2025).

### Meme Coin

*Meme coin* merupakan aset kripto yang popularitas dan nilainya sebagian besar dipengaruhi oleh fenomena media sosial, tren internet, serta aktivitas komunitas daring. Berbeda dengan aset kripto yang memiliki utilitas atau fundamental teknologi yang kuat, pergerakan harga *meme coin* lebih banyak dipengaruhi oleh sentimen kolektif dan aktivitas spekulatif investor (Corbet et al., 2022). Kondisi tersebut menyebabkan harga *meme coin* cenderung lebih volatil dibandingkan aset kripto konvensional. Perubahan sentimen komunitas dapat memicu lonjakan maupun penurunan harga dalam waktu singkat. Oleh karena itu, pendekatan berbasis sentimen menjadi salah satu metode yang relevan untuk memahami pola pergerakan harga pada aset digital jenis ini (Ante, 2023).

### Model Transformer RoBERTa

Model *transformer* telah menjadi pendekatan dominan dalam berbagai tugas pemrosesan bahasa alami karena kemampuannya memahami hubungan kontekstual antar kata secara lebih efektif dibandingkan metode tradisional. Salah satu pengembangan penting dari arsitektur tersebut adalah *Robustly Optimized Bidirectional Encoder Representations from Transformers* (RoBERTa). Model ini dirancang untuk meningkatkan kualitas representasi bahasa melalui optimasi proses pelatihan dan penggunaan data yang lebih besar (Koszweski et al., 2024). Pada tugas analisis sentimen, RoBERTa terbukti mampu menghasilkan performa yang tinggi, terutama ketika diterapkan pada data media sosial yang mengandung bahasa informal, singkatan, maupun variasi ekspresi pengguna (Zhang et al., 2022).

### Machine Learning

*Machine learning* merupakan pendekatan komputasional yang memungkinkan sistem mempelajari pola dari data historis untuk menghasilkan prediksi terhadap data baru. Dalam bidang keuangan dan aset digital, teknik ini banyak digunakan untuk memprediksi arah pergerakan harga berdasarkan berbagai faktor, termasuk indikator teknikal, fundamental, dan sentimen pasar (Jabeur et al., 2023). Algoritma klasifikasi seperti *Logistic Regression* dan *Random Forest* sering digunakan untuk memprediksi tren harga karena mampu mengidentifikasi hubungan antara variabel masukan dan target prediksi. *Logistic Regression* efektif untuk memodelkan hubungan linear, sedangkan *Random Forest* memiliki kemampuan lebih baik dalam menangani pola non-linear dan data yang kompleks melalui mekanisme *ensemble learning* (Kumar & Mishra, 2024). Dengan mengintegrasikan hasil analisis sentimen ke dalam model *machine learning*, proses prediksi tren harga dapat dilakukan secara lebih komprehensif karena mempertimbangkan aspek psikologis pasar selain data historis harga.

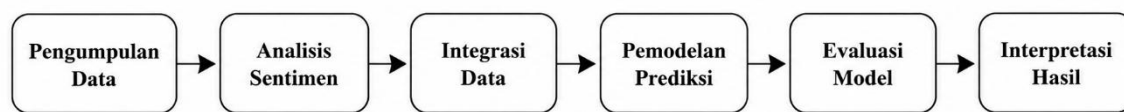
### Sentimen Harga Kripto

Berbagai penelitian menunjukkan bahwa aktivitas dan sentimen pengguna media sosial memiliki hubungan dengan pergerakan harga aset kripto. Informasi yang dibagikan dalam

komunitas daring dapat memengaruhi ekspektasi investor serta membentuk perilaku pasar. Hubungan tersebut tidak selalu bersifat langsung, tetapi sering menunjukkan efek keterlambatan (*lagged effect*) antara perubahan sentimen dan pergerakan harga (Nasir et al., 2021). Oleh karena itu, sentimen media sosial dipandang sebagai indikator pelengkap yang dapat meningkatkan kemampuan model prediksi ketika dikombinasikan dengan data pasar lainnya (Yaijing Zhi et al., 2024).

## METODE PENELITIAN

Penelitian ini menerapkan pendekatan *data science* untuk membangun model prediksi tren harga *meme coin* CHILLGUY berdasarkan sentimen komunitas pada platform Reddit. Tahapan penelitian meliputi pengumpulan data, analisis sentimen, integrasi data, pemodelan, dan evaluasi model. Seluruh proses dilakukan menggunakan bahasa pemrograman *Python* pada lingkungan *Google Colaboratory* dengan dukungan pustaka *PRAW*, *Transformers*, *Pandas*, *Scikit-learn*, dan *Matplotlib*.



Gambar 1. Alur Penelitian

### Pengumpulan Data

Data penelitian terdiri atas dua sumber utama, yaitu data teks dari Reddit dan data harga historis aset kripto. Data teks diperoleh melalui *Application Programming Interface* (API) Reddit menggunakan pustaka *Python Reddit API Wrapper* (*PRAW*) pada subreddit yang membahas *meme coin* CHILLGUY. Pengambilan data dilakukan pada periode 15 November 2024 hingga 30 April 2025. Data yang dikumpulkan berupa judul unggahan (*post title*) yang merepresentasikan opini dan persepsi pengguna terhadap aset yang diteliti.

### Analisis Sentimen

Analisis sentimen dilakukan untuk mengidentifikasi kecenderungan opini pengguna Reddit terhadap *meme coin* CHILLGUY. Proses ini memanfaatkan model *RoBERTa* yang telah melalui proses *fine-tuning* pada data Reddit sehingga mampu menangani karakteristik bahasa informal yang umum ditemukan pada media sosial. Setiap teks diubah menjadi representasi token menggunakan *AutoTokenizer* dan diproses untuk menghasilkan probabilitas sentimen positif, netral, dan negatif. Label sentimen ditentukan berdasarkan probabilitas tertinggi menggunakan fungsi *softmax*. Selanjutnya, setiap label dikonversi ke bentuk numerik dengan nilai 1 untuk sentimen positif, 0 untuk sentimen netral, dan -1 untuk sentimen negatif.

### Integrasi dan Pembentukan Dataset

Nilai sentimen yang diperoleh kemudian diintegrasikan berdasarkan tanggal untuk menghasilkan rata-rata sentimen harian (*average sentiment score*). Data sentimen harian tersebut digabungkan dengan data harga penutupan harian menggunakan atribut tanggal sebagai kunci penghubung. Setelah proses integrasi, dibentuk variabel target berupa arah pergerakan harga (*price movement*), yang diberi nilai 1 apabila harga mengalami kenaikan dibandingkan hari sebelumnya dan nilai 0 apabila harga menurun atau tidak mengalami perubahan. Tahap ini menghasilkan dataset akhir yang siap digunakan dalam proses pemodelan.

### Pemodelan Prediksi

Pemodelan dilakukan menggunakan pendekatan klasifikasi biner untuk memprediksi arah pergerakan harga harian. Dataset dibagi menjadi data latih sebesar 70% dan data uji sebesar 30% menggunakan metode *train-test split* dengan nilai *random state* 42 untuk menjaga konsistensi hasil eksperimen. Variabel masukan yang digunakan adalah rata-rata sentimen harian, sedangkan

variabel keluaran berupa arah pergerakan harga. Dua algoritma *machine learning* diterapkan pada penelitian ini, yaitu *Logistic Regression* dan *Random Forest*. *Logistic Regression* digunakan sebagai model dasar (*baseline model*) karena memiliki mekanisme klasifikasi yang sederhana dan mudah diinterpretasikan. Sementara itu, *Random Forest* dipilih karena kemampuannya dalam menangani hubungan nonlinier serta mengurangi risiko *overfitting* melalui mekanisme agregasi banyak pohon keputusan (*ensemble learning*).

### Evaluasi Model

Kinerja model dievaluasi menggunakan beberapa metrik klasifikasi, yaitu akurasi (*accuracy*), presisi (*precision*), *recall*, dan *F1-score*. Selain itu, *confusion matrix* digunakan untuk menganalisis distribusi prediksi benar dan salah pada masing-masing kelas. Evaluasi dilakukan terhadap data uji untuk mengukur kemampuan model dalam mengidentifikasi arah pergerakan harga berdasarkan informasi sentimen yang diperoleh dari Reddit. Visualisasi hasil evaluasi disajikan dalam bentuk *heatmap* guna mempermudah interpretasi performa model.

## HASIL DAN PEMBAHASAN

### Hasil Analisis Sentimen Data Reddit

Proses pengumpulan data menghasilkan 720 unggahan dari subreddit *r/CHILLGUYCOIN* selama periode 15 November 2024 hingga 30 April 2025. Analisis sentimen menggunakan model *RoBERTa* menunjukkan bahwa sentimen netral mendominasi dengan proporsi 68,89% (496 unggahan), diikuti sentimen positif sebesar 21,94% (158 unggahan) dan sentimen negatif sebesar 9,17% (66 unggahan).

Dominasi sentimen netral menunjukkan bahwa sebagian besar diskusi komunitas berisi informasi, pengamatan pasar, dan pembaruan perkembangan proyek tanpa ekspresi emosional yang kuat. Kondisi ini mencerminkan karakteristik komunitas kripto yang tidak selalu mengekspresikan optimisme atau pesimisme secara eksplisit dalam setiap unggahan. Namun demikian, kemunculan sentimen positif dan negatif tetap memberikan indikasi terhadap perubahan persepsi investor terhadap aset yang diamati.

Distribusi sentimen yang tidak seimbang berpotensi memengaruhi proses pembelajaran model klasifikasi. Ketika mayoritas data berada pada satu kategori tertentu, model cenderung mempelajari pola dominan tersebut dan mengabaikan variasi pada kelas minoritas. Akibatnya, kemampuan model dalam mengenali perubahan harga yang dipengaruhi oleh sentimen ekstrem menjadi berkurang.

### Hubungan Sentimen dan Pergerakan Harga

Hasil penggabungan data sentimen harian dengan harga penutupan harian menghasilkan 103 data observasi yang digunakan dalam proses pemodelan. Analisis visual menunjukkan bahwa harga *CHILLGUY* mengalami kenaikan signifikan pada awal periode pengamatan, kemudian mengalami tren penurunan hingga akhir April 2025. Di sisi lain, rata-rata sentimen harian menunjukkan fluktuasi yang lebih dinamis dan tidak selalu mengikuti arah pergerakan harga secara langsung.

Temuan ini mengindikasikan bahwa hubungan antara sentimen komunitas dan harga aset kripto tidak bersifat linier. Pada beberapa periode tertentu, peningkatan sentimen positif diikuti oleh kenaikan harga dalam jangka pendek. Namun, pada periode lain, perubahan sentimen tidak menghasilkan respons harga yang signifikan. Fenomena tersebut menunjukkan adanya kemungkinan *lag effect*, yaitu pengaruh sentimen yang baru tercermin pada harga setelah jeda waktu tertentu.

Karakteristik tersebut umum ditemukan pada aset kripto, khususnya *meme coin*, yang pergerakan nilainya lebih dipengaruhi oleh dinamika psikologis komunitas dibandingkan faktor fundamental. Oleh karena itu, sentimen media sosial lebih tepat digunakan sebagai indikator pendukung dibandingkan sebagai satu-satunya variabel prediksi.

## Evaluasi Model Prediksi

Kinerja model dievaluasi menggunakan metrik akurasi, *precision*, *recall*, dan *F1-score*. Hasil pengujian menunjukkan bahwa model *Random Forest* menghasilkan performa yang lebih baik dibandingkan *Logistic Regression*.

**Tabel 1.** Hasil Evaluasi Model

Metrik	Logistic Regression	Random Forest
Akurasi	51,61%	70,97%
Precision Kelas 0	0,52	0,67
Precision Kelas 1	0,00	0,80
Recall Kelas 0	1,00	0,88
Recall Kelas 1	0,00	0,53
F1-Score Kelas 0	0,68	0,76
F1-Score Kelas 1	0,00	0,64
F1-Score Rata-rata	0,35	0,70

Model *Logistic Regression* hanya mencapai akurasi 51,61%, yang menunjukkan kemampuan prediksi yang relatif rendah. Seluruh data uji diklasifikasikan sebagai kelas harga turun atau stagnan sehingga model gagal mengenali satu pun kejadian kenaikan harga. Kondisi ini menyebabkan nilai *precision*, *recall*, dan *F1-score* untuk kelas harga naik bernilai nol. Hasil tersebut menunjukkan bahwa model mengalami bias terhadap kelas mayoritas dan tidak mampu menangkap pola yang membedakan kedua kategori pergerakan harga.

Sebaliknya, model *Random Forest* menghasilkan akurasi 70,97% dengan kemampuan klasifikasi yang lebih seimbang. Model berhasil mengidentifikasi 14 data harga turun dan 8 data harga naik secara benar. Nilai *precision* kelas harga naik mencapai 0,80, menunjukkan bahwa sebagian besar prediksi kenaikan harga yang dihasilkan model sesuai dengan kondisi aktual. Selain itu, nilai *F1-score* rata-rata sebesar 0,70 mengindikasikan bahwa model memiliki keseimbangan yang baik antara ketepatan dan kemampuan deteksi kelas.

Perbedaan performa kedua model menunjukkan bahwa hubungan antara sentimen dan arah harga tidak dapat direpresentasikan secara optimal menggunakan pendekatan linier. *Random Forest* mampu membangun banyak pohon keputusan yang menangkap pola kompleks dan interaksi nonlinier dalam data. Kemampuan tersebut menjadikan model lebih adaptif terhadap karakteristik pasar kripto yang memiliki volatilitas tinggi dan pola pergerakan yang sulit diprediksi.

## Pembahasan

Hasil penelitian menunjukkan bahwa sentimen komunitas Reddit memiliki nilai informatif dalam proses prediksi tren harga *meme coin*. Meskipun tidak menjadi faktor utama yang menentukan arah harga, informasi sentimen terbukti mampu meningkatkan kemampuan model dalam mengidentifikasi pola pergerakan pasar. Hal ini dibuktikan oleh performa *Random Forest* yang mencapai tingkat akurasi di atas 70%.

Namun demikian, penggunaan rata-rata sentimen harian sebagai satu-satunya fitur prediksi masih memiliki keterbatasan. Pergerakan harga aset kripto dipengaruhi oleh berbagai faktor lain, seperti volume perdagangan, kapitalisasi pasar, likuiditas, kondisi pasar global, serta aktivitas investor besar. Ketidakhadiran variabel-variabel tersebut menyebabkan sebagian variasi harga tidak dapat dijelaskan oleh model.

Selain itu, dominasi sentimen netral sebesar 68,89% mengurangi variasi informasi yang tersedia untuk proses klasifikasi. Akibatnya, model menghadapi kesulitan dalam membedakan kondisi pasar yang benar-benar mencerminkan optimisme atau pesimisme komunitas. Kondisi ini juga menjelaskan rendahnya performa *Logistic Regression* yang sangat sensitif terhadap ketidakseimbangan distribusi kelas.

Secara keseluruhan, hasil penelitian mengonfirmasi bahwa data sentimen dari media sosial dapat dimanfaatkan sebagai indikator pendukung dalam analisis pasar kripto. Model *Random Forest* menunjukkan kemampuan yang lebih baik dalam memanfaatkan informasi sentimen dibandingkan *Logistic Regression*, sehingga lebih sesuai diterapkan pada permasalahan prediksi harga aset digital yang memiliki karakteristik nonlinier dan volatilitas tinggi. Dengan demikian,

integrasi analisis sentimen dan *machine learning* berpotensi menjadi pendekatan alternatif dalam mendukung pengambilan keputusan investasi pada aset *meme coin*.

### KESIMPULAN

Penelitian ini menunjukkan bahwa penerapan *Natural Language Processing* (NLP) pada data Reddit dapat dimanfaatkan untuk mendukung prediksi tren harga *meme coin* CHILLGUY melalui pendekatan *machine learning*. Hasil analisis membuktikan bahwa sentimen komunitas yang diekstraksi menggunakan model berbasis *transformer* memiliki keterkaitan dengan arah pergerakan harga, sehingga berpotensi menjadi indikator tambahan dalam analisis pasar aset kripto. Dari dua model yang diuji, *Random Forest* memberikan kinerja terbaik dengan akurasi sebesar 70,97% serta kemampuan yang lebih baik dalam mengidentifikasi pergerakan harga naik maupun turun. Sebaliknya, *Logistic Regression* menunjukkan performa yang kurang optimal akibat keterbatasannya dalam menangani hubungan nonlinier dan ketidakseimbangan kelas pada data. Temuan ini mengindikasikan bahwa integrasi analisis sentimen media sosial dengan metode *machine learning* dapat meningkatkan pemahaman terhadap dinamika harga aset digital yang dipengaruhi oleh perilaku komunitas, khususnya pada kategori *meme coin* yang memiliki volatilitas tinggi.

### REFERENSI

- Ante, L. (2023). Social media and cryptocurrency markets: A review of empirical literature. *Journal of Risk and Financial Management*, 16(1), 21. <https://doi.org/10.3390/jrfm16010021>
- Liu, P., Yuan, W., Fu, J., Jiang, Z., Hayashi, H., & Neubig, G. (2023). Pre-train, prompt, and predict: A systematic survey of prompting methods in natural language processing. *ACM Computing Surveys*, 55(9), 1–35. <https://doi.org/10.1145/3560815>
- Koszweski, M., Kocoń, J., & Wróblewska, A. (2024). Investigating transformer-based sentiment analysis models for noisy social media text. *Expert Systems with Applications*, 233, 121089. <https://doi.org/10.1016/j.eswa.2023.121089>
- Sundarasan, S., & Saleem, F. (2025). From tweets to trades: A bibliometric and systematic review of social media's influence on cryptocurrency. *International Journal of Financial Studies*, 13(2), 87. <https://doi.org/10.3390/ijfs13020087>
- Zhi, Y., Kong, L., Atouati, S., Chan, T. H. H., & Sozio, M. (2024). Explaining cryptocurrency price trends: Statistical analysis of social media posts versus market prices. *ACM Transactions on the Web*, 18(4), 1–28. <https://doi.org/10.1145/3652526>
- Medvedev, A. N., Lambiotte, R., & Delvenne, J. C. (2022). The anatomy of Reddit: An overview of academic research. *Dynamics on and of Complex Networks*, 183–204. [https://doi.org/10.1007/978-3-030-65351-4\\_9](https://doi.org/10.1007/978-3-030-65351-4_9)
- Corbet, S., Hou, Y., Hu, Y., Oxley, L., & Xu, D. (2022). Cryptocurrency liquidity and volatility interrelationships during the COVID-19 pandemic. *Finance Research Letters*, 43, 101937. <https://doi.org/10.1016/j.frl.2021.101937>
- Yukun, H., Huang, X., & Ye, Y. (2022). Investor sentiment and cryptocurrency price dynamics: Evidence from social media activities. *Finance Research Letters*, 49, 103093. <https://doi.org/10.1016/j.frl.2022.103093>
- Noori, I. K. B., & Putra, P. S. (2025). *Sentiment analysis and classification of user reviews on the Redbus application using Logistic Regression and SVM*. INOVTEK Polbeng – Seri Informatika, 10(2), 877–887. <https://doi.org/10.35314/k6k6m469>
- Arora, N., Patil, V., & Rathi, V. (2021). Performance analysis of classification algorithms for social media sentiment analysis. *Materials Today: Proceedings*, 47, 4774–4780. <https://doi.org/10.1016/j.matpr.2021.05.520>
- Lupu, R., & Donoiu, P. C. (2025). Sentiment matters for cryptocurrencies: Evidence from tweets. *Data*, 10(4), 50. <https://doi.org/10.3390/data10040050>
- Nasir, M. A., Huynh, T. L. D., Nguyen, S. P., & Duong, D. (2021). Role of media coverage in predicting cryptocurrency returns and volatility. *International Review of Financial Analysis*, 74, 101670. <https://doi.org/10.1016/j.irfa.2021.101670>

- Jabeur, S. B., Mefteh-Wali, S., & Viviani, J. L. (2023). Forecasting cryptocurrency returns using machine learning techniques: A systematic review. *Research in International Business and Finance*, 64, 101889. <https://doi.org/10.1016/j.ribaf.2022.101889>
- Nasir, R., & Mishra, R. (2024). Comparative analysis of machine learning algorithms for financial market prediction. *Expert Systems with Applications*, 237, 121453. <https://doi.org/10.1016/j.eswa.2023.121453>
- Naseem, U., Razzak, I., Musial, K., & Imran, M. (2023). Transformer-based deep intelligent contextual embedding for Twitter sentiment analysis. *Future Generation Computer Systems*, 139, 64–76. <https://doi.org/10.1016/j.future.2022.09.021>