

Analisis Sentimen dan Data Market Untuk Memprediksi Pergerakan Kapitalisasi Pasar *Altcoin* Menggunakan *CryptoBERT*

Tugas Akhir

Diajukan untuk memenuhi persyaratan mencapai derajat

Sarjana Sistem Informasi



Marcellino Natalio Widanarto

NPM: 211711178

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNOLOGI INDUSTRI
UNIVERSITAS ATMA JAYA YOGYAKARTA
2025**

HALAMAN PENGESAHAN

Tugas Akhir Berjudul

ANALISIS SENTIMEN DAN DATA MARKET UNTUK MEMPREDIKSI PERGERAKAN TOTAL2ES
MENGUNAKAN CRYPTOBERT

yang disusun oleh

Marcellino Natalio Widanarto

211711178

dinyatakan telah memenuhi syarat pada tanggal 15 Juli 2025

		Keterangan
Dosen Pembimbing 1	: Generosa Lukhayu Pritalia, S.T., M.Eng.	Telah Menyetujui
Dosen Pembimbing 2	: Putri Nastiti, S.Kom., M.Eng.	Telah Menyetujui
Tim Penguji		
Penguji 1	: Generosa Lukhayu Pritalia, S.T., M.Eng.	Telah Menyetujui
Penguji 2	: Herlina, S.Kom., M.Eng.	Telah Menyetujui
Penguji 3	: Hendro Gunawan, S.Si., M.T.	Telah Menyetujui

Yogyakarta, 15 Juli 2025

Universitas Atma Jaya Yogyakarta

Teknologi Industri

Dekan

ttd.

Dr. Ir. Parama Kartika Dewa SP., S.T., M.T.

Dokumen ini merupakan dokumen resmi UAJY yang tidak memerlukan tanda tangan karena dihasilkan secara elektronik oleh Sistem Bimbingan UAJY. UAJY bertanggung jawab penuh atas informasi yang tertera di dalam dokumen ini

LEMBAR PERNYATAAN KEASLIAN PENELITIAN

Saya yang bertanda tangan di bawah ini:

Nama Lengkap : Marcellino Natalio Widanarto
NPM : 211711178
Program Studi : Sistem Informasi
Fakultas : Teknologi Industri
Judul Penelitian : Analisis Sentimen dan Data Market Untuk Memprediksi
Pergerakan Kapitalisasi Pasar *Altcoin* Menggunakan *CryptoBERT*

Menyatakan dengan ini:

1. Skripsi ini adalah benar merupakan hasil karya sendiri dan tidak merupakan salinan sebagian atau keseluruhan dari karya orang lain.
2. Memberikan kepada Universitas Atma Jaya Yogyakarta, berupa Hak Bebas Royalti non eksklusif (*Non-Exclusive-Royalty-Free Right*) atas Penelitian ini, dan berhak menyimpan, mengelola dalam pangkalan data, mendistribusikan, serta menampilkan untuk kepentingan akademis, tanpa perlu meminta izin selama tetap mencantumkan nama penulis.
3. Bersedia menanggung secara pribadi segala bentuk tuntutan hukum yang mengikuti atas pelanggaran Hak Cipta dalam pembuatan skripsi ini.

Demikianlah pernyataan ini dibuat dan dapat dipergunakan sebagaimana mestinya.

Yogyakarta, 6 Juli 2025
Yang menyatakan,

Marcellino Natalio Widanarto
211711178

PRAKATA

Puji dan Syukur kehadiran Tuhan Yang Maha Esa atas Rahmat dan karunia-Nya sehingga peneliti dapat menyelesaikan tugas akhir sebagai syarat penuntasan gelar Sarjana Sistem Informasi dengan penelitian berjudul “Analisis Sentimen dan Data Market Untuk Memprediksi Pergerakan Kapitalisasi Pasar *Altcoin* Menggunakan *CryptoBERT*” Pada kesempatan ini, peneliti ingin menyampaikan rasa terima kasih yang mendalam kepada berbagai pihak yang telah memberikan dukungan, bimbingan, dan bantuan selama proses penyelesaian tugas akhir ini, yaitu kepada:

1. Tuhan Yesus Kristus dan Bunda Maria atas segala bantuan, berkat, dan kemudahan yang diberikan kepada peneliti selama mengerjakan tugas akhir ini.
2. Papa, Mama, Mbak Cloudia, dan Mbak Sandy yang telah memberikan doa, dukungan moral dan modal, motivasi, dan kasih sayang yang tak terhingga kepada peneliti dari awal perkuliahan hingga penyelesaian tugas akhir.
3. Ibu Generosa Lukhayu Pritalia, S.T., M.Eng., selaku dosen pembimbing 1 dan dosen pembimbing akademik yang telah memberikan arahan, bimbingan, dan motivasi dengan penuh kesabaran dalam mengarahkan peneliti dalam menyelesaikan tugas akhir maupun permasalahan akademik selama perkuliahan.
4. Ibu Putri Nastiti, S.Kom., M.Eng., selaku dosen pembimbing 2 dan tempat curhat, yang telah memberikan masukan, bimbingan, dan arahan yang sangat berharga untuk perbaikan diri dan penelitian ini.
5. Ivana Angelica, yang telah memberikan dukungan emosional, motivasi, dan semangat yang luar biasa, serta selalu setia menemani dan menyemangati peneliti dalam setiap fase pengerjaan tugas akhir ini.
6. Teman-teman SMA Kolese De Britto dan tambahan, seperti Adi, Mikhael, Salomo, David, Bryan, Enggar, Abi, Lenkky, Yogi, Bayu, Zent, dan teman-teman lainnya yang telah saling mendukung, berbagi ilmu, mengingatkan, dan memberikan semangat dalam menghadapi berbagai tantangan selama proses perkuliahan dan penyelesaian tugas akhir.

7. Teman-teman Sistem Informasi Atma Jaya Yogyakarta, yang telah menjadi rekan seperjuangan, saling berbagi pengalaman, dan memberikan dukungan selama masa perkuliahan.
8. Seluruh dosen dan staf Program Studi Sistem Informasi yang telah memberikan ilmu, pengalaman, dan dukungan selama masa perkuliahan.
9. Semua pihak yang tidak dapat peneliti sebutkan satu per satu, yang telah membantu dan mendukung peneliti dalam menyelesaikan tugas akhir ini.

Peneliti berharap hasil penelitian tugas akhir ini dapat memberikan kontribusi positif dalam bidang keilmuan sistem informasi, khususnya dalam cakupan *machine learning*, serta dapat menjadi referensi untuk penelitian dan pengembangan akademik selanjutnya. Semoga penelitian ini dapat bermanfaat bagi semua pihak yang membutuhkan.

INTISARI

Cryptocurrency tidak hanya memberikan kontribusi terhadap perubahan paradigma keuangan global tetapi juga kesempatan baru dalam dunia investasi. *Bitcoin* sebagai pionir dan *cryptocurrency* dengan kapitalisasi pasar terbesar memiliki fenomena vital bernama *halving*. Halving merupakan mekanisme pembatasan *supply* yang menciptakan siklus lonjakan harga berulang dan secara historis memicu *Altseason* melalui pergeseran dana dari *Bitcoin* ke *Altcoin*. Penelitian ini bertujuan menganalisis sentimen publik terhadap *Altseason* pada periode *post-halving* 2024 yang menunjukkan pola anomali dibanding siklus sebelumnya, serta menganalisis performa prediksi pergerakan kapitalisasi pasar *Altcoin* menggunakan model *LSTM* dengan data *OHLCV* dan dengan data *OHLCV* yang diintegrasikan sentimen. Data penelitian mencakup 3.996 *Tweet* dari platform *X* dengan kata kunci "*Altcoin*" dan "*Altseason*" yang diklasifikasikan menggunakan *CryptoBERT*, serta data *OHLCV* indeks *TOTAL2ES*. Hasil analisis menunjukkan dominasi sentimen *bullish* sebesar 80% terhadap *Altseason* yang kontradiktif dengan penurunan kapitalisasi *Altcoin* sebesar 42%, mengindikasikan ketidaksinambungan antara sentimen publik dan realitas pasar. Evaluasi performa menunjukkan *baseline model* superior dengan R^2 0.8338 dibanding *multimodal model* dengan R^2 0.8074, membuktikan bahwa dalam periode anomali pasar, sentimen media sosial justru menjadi *noise* yang mengurangi akurasi prediksi jangka pendek dibandingkan analisis teknikal murni.

Kata Kunci: Analisis Sentimen, Time-Series Analysis, Price Prediction, BERT, CryptoBERT, LSTM, Cryptocurrency, Altcoin, Altseason.

ABSTRACT

Cryptocurrency not only contributes to the global financial paradigm shift but also provides new opportunities in the investment world. Bitcoin, as the pioneer and cryptocurrency with the largest market capitalization, has a vital phenomenon called halving. Halving is a supply limitation mechanism that creates recurring price surge cycles and historically triggers Altseason through fund shifts from Bitcoin to Altcoins. This research aims to analyze public sentiment towards Altseason during the post-halving 2024 period, which shows anomalous patterns compared to previous cycles, and to analyze the predictive performance of Altcoin market capitalization movements using LSTM models with OHLCV data and with OHLCV data integrated with sentiment. The research data includes 3,996 Tweets from platform X with keywords "Altcoin" and "Altseason" classified using CryptoBERT, as well as OHLCV data from the TOTAL2ES index. The analysis results show a dominance of bullish sentiment of 80% towards Altseason that contradicts the 42% decline in Altcoin capitalization, indicating a disconnect between public sentiment and market reality. Performance evaluation shows the baseline model is superior with R^2 0.8338 compared to the multimodal model with R^2 0.8074, proving that during market anomaly periods, social media sentiment becomes noise that reduces short-term prediction accuracy compared to pure technical analysis.

Keywords: Sentiment Analysis, Time-Series Analysis, Price Prediction, BERT, CryptoBERT, LSTM, Cryptocurrency, Altcoin, Altseason.

DAFTAR ISI

PRAKATA.....	IV
INTISARI	VI
ABSTRACT	VII
DAFTAR ISI	VIII
DAFTAR GAMBAR.....	X
DAFTAR TABEL	XIII
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah.....	3
1.3. Pertanyaan Penelitian	3
1.4. Tujuan Penelitian	4
1.5. Batasan Masalah	4
1.6. Manfaat Penelitian.....	4
1.7. Bagan Keterkaitan.....	5
BAB II TINJAUAN PUSTAKA	6
2.1. Studi Sebelumnya	6
2.2. Dasar Teori.....	11
2.2.1. Analisis Sentimen	11
2.2.2. Pendekatan dalam Tahapan Prediksi dan Klasifikasi Sentimen.....	12
2.2.3. Bidirectional <i>Encoder Representations from Transformers (BERT)</i>	13
2.2.4. Kapitalisasi Pasar	14
2.3. Tabel Perbandingan	15
BAB III METODOLOGI PENELITIAN.....	16
3.1. Metode Penelitian	16
3.2. Tahapan Penelitian	18
3.2.1. Data Extraction	18
3.2.2. Textual Data Processing	19
3.2.3. Sentiment Classification with <i>CryptoBERT</i>	21
3.2.4. Output Presentation	21
3.2.5. Numerical Data Processing	22
3.2.6. Next 4-Hour Close Prediction.....	23
3.2.7. Evaluation Metrics	25
3.3. Tools yang Digunakan	26
3.3.1. Google Colab	26
3.3.2. Tweet Harvest	26
3.3.3. TradingView	26
BAB IV HASIL DAN PEMBAHASAN.....	27

4.1.	Hasil.....	27
4.1.1.	Data Extraction.....	27
4.1.2.	Textual Data Processing	35
4.1.3.	Sentiment Classification with CryptoBERT	43
4.1.4.	Output Presentation	47
4.1.5.	Numerical Data Processing	48
4.1.2.	Next 4-Hour Close Prediction.....	59
4.1.3.	Evaluation Metrics	67
4.2.	Pembahasan	68
4.2.1.	Analisis Sentimen dengan CryptoBERT.....	68
4.2.2.	Baseline Model.....	75
4.2.3.	Multimodal Model	76
BAB V	KESIMPULAN DAN SARAN.....	79
5.1.	Kesimpulan	79
5.2.	Saran.....	80
GLOSARIUM		82
DAFTAR PUSTAKA		86
TABEL REVISI		94
Lampiran.....		95

DAFTAR GAMBAR

Gambar 1. 1. Bagan Keterkaitan Penelitian.....	5
Gambar 2. 1. Kerangka Kerja Umum Analisis Sentimen	11
Gambar 2. 2. Arsitektur <i>BERT</i>	13
Gambar 3. 1. Arsitektur <i>LSTM</i>	17
Gambar 3. 2. Tahapan Penelitian	18
Gambar 4. 1. <i>Code Setup</i> dan Konfigurasi <i>Tweet Harvest</i>	27
Gambar 4. 2. Bagian Pertama <i>Code</i> Pembuatan <i>Class Twitter Scraper</i>	28
Gambar 4. 3. Bagian Kedua <i>Code</i> Pembuatan <i>Class Twitter Scraper</i>	28
Gambar 4. 4. <i>Twitter Authentication Token</i> pada <i>Secrets</i>	29
Gambar 4. 5. <i>Code</i> Konfigurasi Kriteria Penelitian dan Proses Pengumpulan Data	30
Gambar 4. 6. Kesalahan Hasil dalam <i>Data Extraction</i>	31
Gambar 4. 7. Parameter <i>Textual Data Extraction</i>	32
Gambar 4. 8. Hasil <i>Textual Data Extraction</i> dengan <i>Tweet Harvest</i>	32
Gambar 4. 9. <i>Chart</i> <i>TOTAL2ES</i> dalam <i>Timeframe 4H</i>	33
Gambar 4. 10. Data Numerik <i>TOTAL2ES</i>	35
Gambar 4. 11. Hasil Penggabungan <i>CSV Dataset</i>	36
Gambar 4. 12. <i>Code Load and Combine CSV Dataset</i>	36
Gambar 4. 13. Hasil Penghapusan Kolom Irrelevan dan <i>Data Cleaning</i>	37
Gambar 4. 14. <i>Code</i> Penghapusan Kolom Irrelevan	37
Gambar 4. 15. Tipe Data Kolom.....	38
Gambar 4. 16. <i>Code</i> Penyesuaian Tipe Data.....	38
Gambar 4. 17. Hasil Identifikasi dan Penghapusan Data Terduplikat.....	39
Gambar 4. 18. <i>Code</i> Identifikasi dan Penghapusan Data Terduplikat	39
Gambar 4. 19. <i>Code Data Cleaning</i>	40
Gambar 4. 20. <i>Tweet</i> Tidak Relevan Hasil <i>Data Cleaning</i>	41
Gambar 4. 21. Hasil Keseluruhan <i>Data Cleaning and adjustment</i>	41
Gambar 4. 22. <i>Code Sorting</i> dan <i>Save Dataset</i>	42
Gambar 4. 23. Hasil <i>Word Cloud</i>	42
Gambar 4. 24. <i>Code Word Cloud</i>	43
Gambar 4. 25. <i>Code Function</i> Pembatasan <i>Max Sequence Length</i>	44

Gambar 4. 26. Hasil Identifikasi Jumlah Maksimal <i>Token</i>	44
Gambar 4. 27. <i>Code</i> Pembatasan Jumlah Maksimal <i>Token</i>	45
Gambar 4. 28. Hasil <i>Sentiment Classification</i>	45
Gambar 4. 29. <i>Code Sentiment Classification</i>	46
Gambar 4. 30. Bagian Pertama <i>Code Output Presentation</i>	47
Gambar 4. 31. Bagian Kedua <i>Code Output Presentation</i>	48
Gambar 4. 32. Bagian Ketiga <i>Code Output Presentation</i>	48
Gambar 4. 33. Hasil Identifikasi Tipe Data Kolom <i>time</i>	49
Gambar 4. 34. Hasil Filtrasi Rentang Waktu Data Numerik.....	50
Gambar 4. 35. <i>Code Load</i> dan <i>Adjustment</i> Data Numerik.....	50
Gambar 4. 36. Hasil <i>Load</i> dan <i>Adjustment</i> Data Hasil Analisis Sentimen	51
Gambar 4. 37. <i>Code Load</i> and <i>Adjustment</i> Data Sentimen	51
Gambar 4. 38. Hasil Agregasi <i>Timeframe</i>	52
Gambar 4. 39. Bagian Pertama <i>Code Agregasi 4H</i>	53
Gambar 4. 40. Bagian Kedua <i>Code Agregasi 4H</i>	53
Gambar 4. 41. Hasil <i>Merging</i> Data Sentimen dengan Data Numerik	54
Gambar 4. 42. <i>Code Merging</i> Data Sentimen dengan Data Numerik.....	55
Gambar 4. 43. Hasil <i>Data Splitting</i>	56
Gambar 4. 44. <i>Code Data Splitting</i>	57
Gambar 4. 45. Hasil <i>Normalization</i> pada <i>OHLCV</i>	58
Gambar 4. 46. <i>Code Normalization</i>	59
Gambar 4. 47. <i>Code Class Config</i> untuk Konfigurasi.....	60
Gambar 4. 48. <i>Code Class Flow</i> dan Arsitektur Model <i>LSTM</i>	61
Gambar 4. 49. <i>Code Class Dataset</i>	61
Gambar 4. 50. <i>Code Function Create Model</i>	62
Gambar 4. 51. Bagian Pertama <i>Code Function Train Model</i>	63
Gambar 4. 52. Bagian Kedua <i>Code Function Train Model</i>	64
Gambar 4. 53. <i>Code Training Baseline Model</i>	65
Gambar 4. 54. Visualisasi Hasil <i>Loss</i> pada Kedua Model	67
Gambar 4. 55. <i>Code Training Multimodal Model</i>	67
Gambar 4. 56. Visualisasi Distribusi Sentimen pada Dataset	69
Gambar 4. 57. Pola Distribusi Skor Probabilitas Sentimen <i>Bullish</i>	69

Gambar 4. 58. Distribusi Volume <i>Tweet</i> dan <i>Tweet</i> dengan Sentimen <i>Bullish</i>	70
Gambar 4. 59. Distribusi Rata-Rata Probabilitas <i>Bullish</i> per Bulan.....	71
Gambar 4. 60. <i>Chart</i> TOTAL2ES <i>Post-Halving</i> 2024	71
Gambar 4. 61. Pergerakan <i>BTC.D</i> 230-550 Hari setelah <i>Bitcoin Halving</i>	72
Gambar 4. 62. Pergerakan TOTAL2ES 230-550 Hari setelah <i>Bitcoin Halving</i>	72
Gambar 4. 63. Hasil Identifikasi <i>Tweet</i> Pertama	73
Gambar 4. 64. Hasil Identifikasi <i>Tweet</i> Kedua	73
Gambar 4. 65. Hasil Identifikasi <i>Tweet</i> Ketiga	74
Gambar 4. 66. Hasil Identifikasi <i>Tweet</i> Keempat.....	74
Gambar 4. 67. Hasil Perbandingan Performa Kedua Model.....	75
Gambar 4. 68. Visualisasi Distribusi Sentimen Timeframe	76

DAFTAR TABEL

Tabel 2. 1. Tabel Penelitian Terdahulu	8
Tabel 2. 2. Tabel Perbandingan Teknik Pendekatan Machine Learning	15
Tabel 4. 1. Kolom Hasil <i>Data Extraction</i> dengan <i>Tweet Harvest</i>	30
Tabel 4. 2. Kolom Hasil <i>Export TradingView</i>	34
Tabel 4. 3. Hasil <i>Training Baseline Model</i>	64
Tabel 4. 4. Hasil <i>Training Multimodal Model</i>	65
Tabel 4. 5. Hasil <i>Evaluation Metrics</i> pada Kedua Model	67