

ANALISIS KLASTERISASI FASE PASAR BITCOIN BERDASARKAN DATA HARGA HISTORIS MENGGUNAKAN ALGORITMA K-MEANS

Laurentina Luisa Br Lase¹, Eka Ramadanta Sitepu², Muhammad Fahri Rinanda³, Isnani Yudi Kurniawan⁴

Sekolah Tinggi Manajemen Informatika dan Komputer KAPUTAMA, Sumatera Utara

E-mail : louisasembiring4@gmail.com¹, dhanta07@gmail.com², mfahririnanda03@gmail.com³, isnanyudi22@gmail.com⁴

ABSTRAK

Penelitian ini menggunakan data historis harga Bitcoin periode 2017 hingga 2024 yang terdiri dari variabel open, high, low, close, dan volume (OHLCV). Pesatnya perkembangan pasar keuangan digital telah mendorong peningkatan signifikan dalam aktivitas perdagangan kripto, dengan Bitcoin tetap menjadi aset dominan dari sisi kapitalisasi pasar dan volume transaksi. Pertumbuhan ini menghasilkan data harga historis berskala besar yang memiliki volatilitas tinggi dan dinamika yang kompleks, sehingga identifikasi pola secara manual menjadi tidak efisien dan cenderung subjektif. Oleh karena itu, diperlukan pendekatan data mining untuk mengubah data pasar mentah menjadi informasi yang terstruktur dan objektif sebagai dasar pengambilan keputusan. Penelitian ini bertujuan untuk menganalisis fase-fase pasar Bitcoin dengan menerapkan algoritma K-Means clustering pada data historis harga yang mencakup variabel OHLCV. Tahap prapemrosesan data dilakukan untuk memastikan konsistensi dan keterbandingan antarvariabel, kemudian analisis kluster diimplementasikan menggunakan bahasa pemrograman Python melalui Google Colaboratory. Penentuan jumlah kluster optimal dilakukan dengan metode Elbow. Hasil penelitian menunjukkan bahwa pergerakan pasar Bitcoin dapat dikelompokkan secara efektif ke dalam tiga kluster utama yang merepresentasikan fase pasar dengan aktivitas tinggi, aktivitas sedang, dan kondisi pasar yang relatif stabil. Temuan ini membuktikan bahwa teknik data mining berbasis klusterisasi mampu mengungkap struktur laten pasar dan dapat dijadikan landasan analitis yang sistematis bagi pengembangan sistem informasi pendukung analisis pasar kripto.

Kata kunci

Bitcoin, Pasar Mata Uang Kripto, Data Mining, Klusterisasi K-Means, Analisis Fase Pasar

ABSTRACT

This study uses historical Bitcoin price data from 2017 to 2024, consisting of open, high, low, close, and volume (OHLCV) variables. The rapid expansion of digital financial markets has significantly increased cryptocurrency trading activity, with Bitcoin remaining the dominant asset in terms of market capitalization and transaction volume. This growth produces large-scale historical price data characterized by high volatility and complex dynamics, making manual pattern identification inefficient and subjective. Therefore, a data mining approach is required to transform raw market data into structured and objective information for decision support. This study aims to analyze Bitcoin market phases by applying the K-Means clustering algorithm to historical price data consisting of open, high, low, close, and volume (OHLCV). Data preprocessing was conducted to ensure consistency and comparability among variables, followed by clustering analysis implemented in Python using Google Colaboratory. The optimal number of clusters was determined using the Elbow method. The results indicate that Bitcoin market movements can be effectively grouped into three distinct clusters representing high-activity, moderate, and stable market phases. These findings demonstrate that clustering-based data mining techniques are capable of revealing latent market structures and can serve as a structured analytical foundation for information systems supporting cryptocurrency market analysis.

Keywords | *Bitcoin, Cryptocurrency Market, Data Mining, K-Means Clustering, Market Phase Analysis*

1. PENDAHULUAN

Perkembangan teknologi informasi telah mendorong transformasi signifikan pada sektor keuangan melalui munculnya cryptocurrency sebagai aset digital global. Bitcoin sebagai cryptocurrency pertama masih mendominasi pasar baik dari sisi kapitalisasi maupun volume transaksi, serta sering dijadikan indikator utama dalam analisis dinamika pasar aset kripto (Corbet et al., 2019; Chan et al., 2017).

Meskipun data historis Bitcoin tersedia secara luas, karakteristik pasar aset kripto yang kompleks dan volatil menyulitkan proses identifikasi pola secara manual, sehingga diperlukan pendekatan analitis berbasis data mining untuk mengekstraksi struktur data secara objektif (Song et al., 2019; Corbet et al., 2019).

Dalam konteks sistem informasi, teknik data mining berperan penting dalam mentransformasikan data mentah menjadi informasi bermakna melalui proses penggalian pola dan struktur tersembunyi pada data berskala besar. Dalam konteks sistem informasi, teknik clustering merupakan salah satu metode data mining yang digunakan untuk mengelompokkan data berdasarkan tingkat kemiripan karakteristik. Metode ini banyak diterapkan dalam analisis data pasar keuangan untuk mengidentifikasi struktur dan pola tersembunyi dalam data berskala besar (Aghabozorgi & Teh, 2014; Aghabozorgi et al., 2015).

Meskipun berbagai penelitian sebelumnya telah menerapkan teknik clustering untuk menganalisis struktur pasar keuangan dan cryptocurrency, Berbagai penelitian sebelumnya telah menerapkan teknik clustering dalam analisis pasar keuangan dan aset kripto, terutama untuk mengkaji hubungan antar aset atau struktur pasar secara agregat (Bonanno et al., 2004; Liao & Chou, 2013; Song et al., 2019). Namun, kajian yang secara khusus memanfaatkan variabel harga historis Bitcoin berupa open, high, low, close, dan volume transaksi (OHLCV) untuk mengidentifikasi fase pasar secara langsung masih terbatas. Oleh karena itu, penelitian ini difokuskan pada pemetaan fase pasar Bitcoin berbasis data historis harga sebagai dasar penyediaan informasi pendukung keputusan dalam sistem informasi.

2. METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan memanfaatkan data sekunder berupa data historis harga Bitcoin. Pendekatan ini dipilih karena memungkinkan analisis pola pasar secara objektif melalui penerapan teknik data mining berbasis clustering.

a. Jenis dan Pendekatan Penelitian

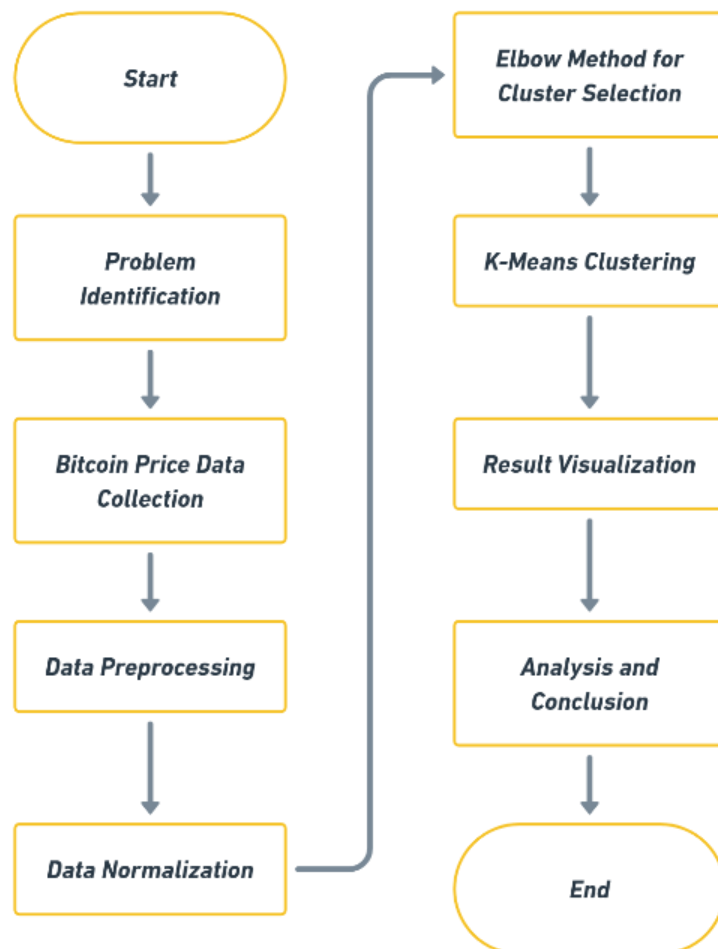
Penelitian ini menggunakan pendekatan kuantitatif dengan jenis penelitian deskriptif-analitis. Penelitian difokuskan pada pengelompokan data historis harga Bitcoin untuk mengidentifikasi pola dan fase pasar tanpa melakukan proses prediksi harga.

b. Sumber dan Teknik Pengumpulan Data

Data yang digunakan dalam penelitian ini merupakan data sekunder berupa data historis harga Bitcoin yang diperoleh dari Yahoo Finance. Data mencakup variabel open, high, low, close, dan volume transaksi, serta disimpan dalam format CSV untuk memudahkan proses pengolahan dan analisis.

c. Tahapan Penelitian

Tahapan penelitian disusun secara sistematis untuk memastikan proses analisis berjalan secara terstruktur dan dapat direplikasi. Tahapan penelitian meliputi pengumpulan data, praproses data, analisis clustering menggunakan algoritma K-Means, serta visualisasi hasil pengelompokan. Alur tahapan penelitian secara keseluruhan ditampilkan pada Gambar 1.



Gambar 1. Flowchart Tahapan Penelitian

d. Proses Pengolahan Data

Data historis yang telah diperoleh selanjutnya melalui beberapa tahapan pengolahan data. Proses pengolahan data diawali dengan tahap data cleaning untuk menghilangkan data yang tidak lengkap, data duplikat, serta nilai yang tidak valid. Selanjutnya dilakukan normalisasi data untuk menyamakan skala antar variabel, sehingga setiap atribut memiliki kontribusi yang seimbang dalam proses clustering.

Data yang telah melalui proses normalisasi kemudian dianalisis menggunakan metode *K-Means Clustering* untuk mengelompokkan data

berdasarkan tingkat kemiripan karakteristik. Alur proses pengolahan data secara lebih rinci disajikan pada Gambar 2.

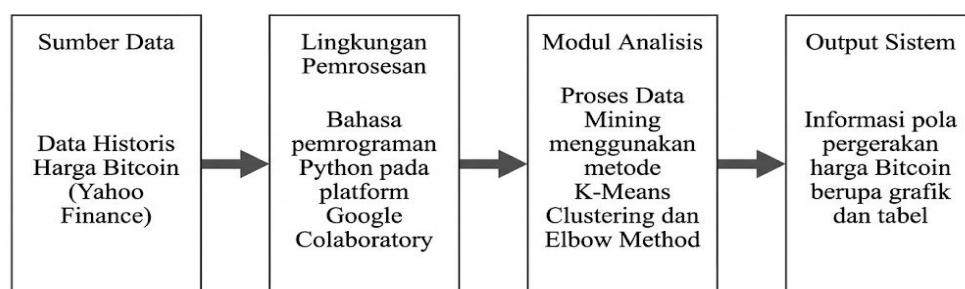


Gambar 2. Diagram Alur Proses Pengolahan Data

e. Arsitektur Sistem Penelitian

Penelitian ini dirancang menggunakan arsitektur sistem sederhana yang menggambarkan alur pemrosesan data mulai dari sumber data hingga menghasilkan informasi. Data historis harga Bitcoin digunakan sebagai input, kemudian diproses menggunakan bahasa pemrograman Python pada lingkungan Google Colaboratory. Output sistem berupa informasi pola pergerakan harga Bitcoin disajikan dalam bentuk tabel dan visualisasi grafik.

Gambaran arsitektur sistem penelitian secara keseluruhan disajikan pada Gambar 3.



Gambar 3. Arsitektur Sistem Analisis Pola Pergerakan Harga Bitcoin

f. Tools dan Lingkungan Penelitian

Proses pengolahan dan analisis data dilakukan menggunakan bahasa pemrograman Python dengan memanfaatkan platform Google Colaboratory sebagai lingkungan komputasi berbasis cloud. Pustaka yang digunakan meliputi Pandas untuk pengolahan data, NumPy untuk komputasi numerik, Scikit-learn untuk penerapan algoritma K-Means Clustering, serta Matplotlib untuk visualisasi data.

g. Metode Analisis Data

Metode analisis data yang digunakan dalam penelitian ini adalah K-Means Clustering yang termasuk ke dalam kategori unsupervised learning pada teknik data mining. K-Means bekerja dengan mengelompokkan data ke dalam sejumlah cluster berdasarkan kedekatan jarak antara setiap data dan pusat cluster (centroid). Metode analisis data yang digunakan dalam penelitian ini adalah K-Means Clustering, yang termasuk ke dalam kategori unsupervised learning dalam teknik data mining. Metode ini bertujuan untuk mengelompokkan data ke dalam sejumlah cluster berdasarkan tingkat kemiripan karakteristik numerik yang dimiliki setiap data (Aghabozorgi et al., 2015).

Penentuan jumlah cluster optimal dilakukan menggunakan metode Elbow dengan menganalisis perubahan nilai Within Cluster Sum of Squares (WCSS) terhadap variasi jumlah cluster. Pendekatan ini digunakan untuk memperoleh jumlah cluster yang mampu merepresentasikan struktur data secara optimal tanpa bersifat subjektif (Arbelaitz et al., 2013; Charrad et al., 2014).

3. HASIL DAN PEMBAHASAN

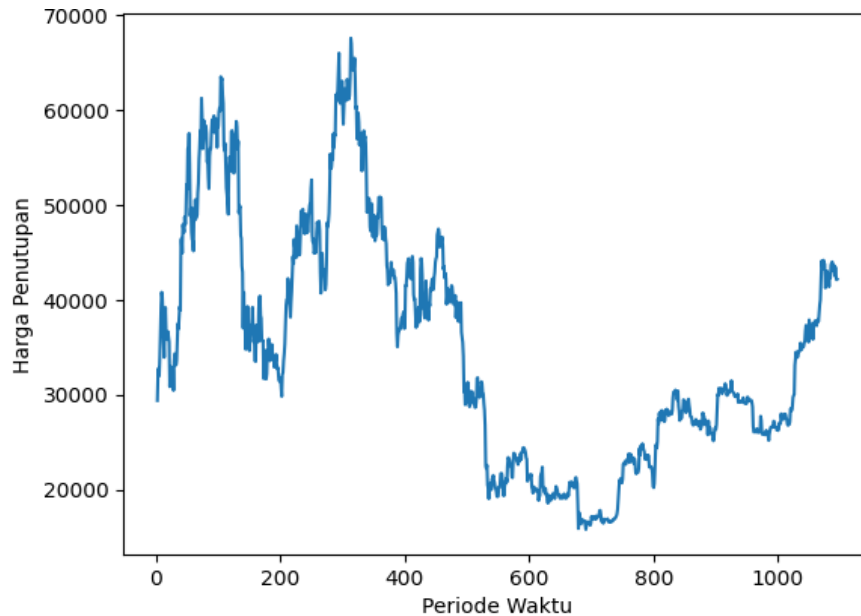
3.1 Hasil Penelitian

Bagian ini menyajikan hasil pengolahan dan analisis data historis harga Bitcoin menggunakan algoritma K-Means Clustering. Seluruh hasil yang disampaikan merupakan output langsung dari proses pengolahan data tanpa interpretasi teoritis atau perbandingan dengan penelitian sebelumnya.

a. Hasil Visualisasi Pergerakan Harga Bitcoin

Berdasarkan data historis harga Bitcoin yang digunakan dalam penelitian ini, dilakukan visualisasi terhadap harga penutupan (close price) untuk mengamati pola pergerakan harga secara umum. Grafik pergerakan harga menunjukkan fluktuasi yang signifikan selama periode pengamatan, dengan pola kenaikan dan penurunan yang tidak bersifat linier.

Pergerakan harga yang ditampilkan mencerminkan dinamika pasar Bitcoin yang berubah-ubah, di mana terdapat periode dengan volatilitas tinggi serta periode dengan pergerakan harga yang relatif stabil. Visualisasi ini digunakan sebagai gambaran awal karakteristik data sebelum dilakukan proses pengelompokan menggunakan metode clustering.

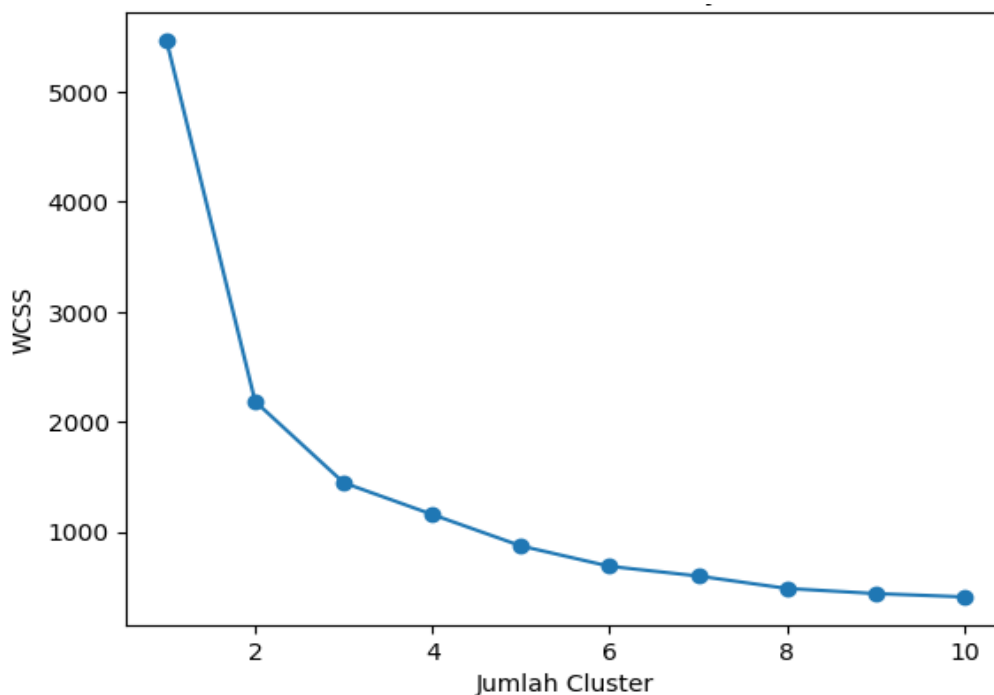


Gambar 1. Grafik Pergerakan Harga Penutupan (*Close Price*) Bitcoin

b. Penentuan Jumlah Cluster Menggunakan Metode Elbow

Penentuan jumlah cluster optimal dilakukan menggunakan metode Elbow dengan menganalisis perubahan nilai Within Cluster Sum of Squares (WCSS) terhadap variasi jumlah cluster. Grafik Elbow menunjukkan penurunan nilai WCSS yang cukup signifikan pada jumlah cluster satu hingga tiga, kemudian mengalami penurunan yang lebih landai pada penambahan cluster selanjutnya.

Berdasarkan pola tersebut, titik siku (elbow point) teridentifikasi pada jumlah cluster sebanyak tiga. Oleh karena itu, jumlah cluster yang digunakan dalam proses K-Means Clustering pada penelitian ini ditetapkan sebanyak tiga cluster.

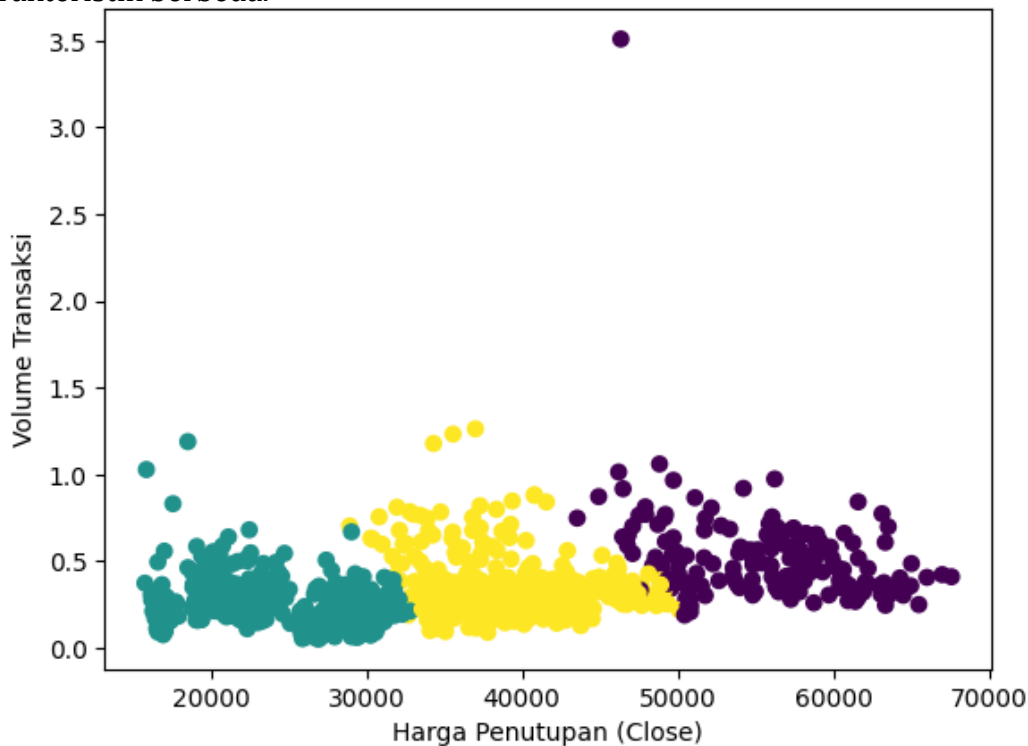


Gambar 2. Grafik Metode *Elbow* untuk Penentuan Jumlah Cluster Optimal

c. Hasil K-Means Clustering Data Bitcoin

Proses pengelompokan data dilakukan menggunakan algoritma K-Means Clustering dengan jumlah cluster sebanyak tiga. Variabel yang digunakan dalam proses clustering meliputi harga open, high, low, close, serta volume transaksi.

Hasil clustering kemudian divisualisasikan menggunakan dua variabel utama, yaitu harga penutupan (close) dan volume transaksi, untuk mempermudah interpretasi pola pengelompokan. Visualisasi menunjukkan bahwa data historis harga Bitcoin dapat dikelompokkan ke dalam tiga kelompok yang memiliki karakteristik berbeda.



Gambar 3. Visualisasi Hasil K-Means Clustering Data Harga Bitcoin Berdasarkan Harga Penutupan dan Volume Transaksi

d. Ringkasan Karakteristik Setiap Cluster

Untuk mengetahui karakteristik masing-masing cluster yang terbentuk, dilakukan perhitungan nilai rata-rata setiap variabel pada masing-masing cluster. Hasil perhitungan tersebut disajikan dalam bentuk tabel untuk memberikan gambaran kuantitatif mengenai perbedaan karakteristik antar cluster.

Tabel 1. Ringkasan Nilai Rata-rata Setiap Cluster Hasil K-Means Clustering

Cluster	Open	High	Low	Close	Volume
0	55.808	57.270	54.114	55.814	5.23E+10
1	24.243	24.654	23.828	24.246	2.34E+10
2	40.145	41.172	39.008	40.163	3.42E+10

Berdasarkan nilai rata-rata tersebut, setiap cluster menunjukkan perbedaan tingkat harga dan volume transaksi yang cukup jelas, sehingga mengindikasikan adanya variasi kondisi pasar dalam data historis yang dianalisis.

3.2 Pembahasan

Bagian pembahasan bertujuan untuk menginterpretasikan hasil clustering yang diperoleh dengan mengaitkannya pada konteks dinamika pasar Bitcoin serta temuan penelitian sebelumnya.

Hasil pengelompokan menunjukkan bahwa cluster dengan nilai rata-rata harga dan volume transaksi tertinggi merepresentasikan kondisi pasar dengan tingkat aktivitas perdagangan yang tinggi. Kondisi ini mencerminkan fase pasar yang bersifat spekulatif, di mana pergerakan harga cenderung lebih fluktuatif. Temuan ini sejalan dengan penelitian Balcilar et al. (2017) yang menyatakan bahwa peningkatan volume transaksi berkorelasi dengan tingginya volatilitas pasar Bitcoin.

Sebaliknya, cluster dengan nilai rata-rata harga dan volume transaksi terendah menggambarkan kondisi pasar yang relatif stabil. Pada fase ini, aktivitas perdagangan cenderung menurun dan pergerakan harga tidak mengalami fluktuasi yang signifikan. Karakteristik tersebut sesuai dengan temuan Drożdż et al. (2018) yang mengidentifikasi adanya fase stabil dalam dinamika pasar keuangan berbasis aset digital.

Cluster dengan nilai rata-rata berada di antara kedua kondisi tersebut dapat diinterpretasikan sebagai fase transisi, di mana pasar berada pada kondisi normal dengan aktivitas perdagangan yang tidak terlalu tinggi maupun terlalu rendah. Fase ini mencerminkan peralihan kondisi pasar dari keadaan stabil menuju aktif, atau sebaliknya.

a. Implikasi dalam Sistem Informasi

Dalam konteks sistem informasi, hasil clustering yang diperoleh menunjukkan bahwa algoritma K-Means Clustering mampu mengelompokkan data historis harga Bitcoin ke dalam beberapa kondisi pasar yang berbeda secara objektif. Informasi mengenai fase pasar ini dapat dimanfaatkan sebagai dasar dalam pengembangan sistem pendukung keputusan (Decision Support System) yang menyajikan informasi kondisi pasar secara lebih terstruktur.

Integrasi hasil clustering ke dalam sistem informasi memungkinkan pengguna untuk memahami kondisi pasar Bitcoin berdasarkan data historis, sehingga proses analisis pasar dapat dilakukan secara lebih sistematis dan berbasis data.

b. Keterbatasan Penelitian

Penelitian ini memiliki beberapa keterbatasan, di antaranya penggunaan data historis harga Bitcoin tanpa mempertimbangkan faktor eksternal seperti sentimen pasar, kondisi ekonomi makro, maupun kebijakan regulasi. Selain itu, algoritma K-Means Clustering bersifat sensitif terhadap penentuan jumlah cluster dan kurang optimal dalam menangani data dengan distribusi yang tidak seimbang.

Meskipun demikian, penelitian ini tetap memberikan gambaran awal mengenai pemanfaatan teknik data mining berbasis clustering dalam analisis fase pasar Bitcoin.

4. KESIMPULAN DAN SARAN

4.1 Kesimpulan

Berdasarkan hasil penelitian dan pembahasan yang telah dilakukan, dapat disimpulkan bahwa metode *K-Means Clustering* mampu mengelompokkan data pergerakan harga Bitcoin ke dalam beberapa kelompok yang memiliki karakteristik berbeda berdasarkan variabel harga dan volume transaksi. Penelitian ini memanfaatkan data historis harga Bitcoin yang diproses melalui tahapan *data preprocessing*, penentuan jumlah cluster optimal menggunakan metode *Elbow*, serta proses pengelompokan data dengan algoritma *K-Means*.

Hasil pengelompokan menunjukkan bahwa data pergerakan harga Bitcoin dapat diklasifikasikan ke dalam tiga cluster utama, yaitu kondisi pasar dengan tingkat aktivitas tinggi, tingkat aktivitas sedang, dan tingkat aktivitas rendah. Masing-masing cluster memiliki karakteristik yang berbeda dari sisi nilai rata-rata harga dan volume transaksi, sehingga mampu memberikan gambaran pola pergerakan pasar kripto secara lebih terstruktur.

Selain itu, hasil visualisasi data dan proses *clustering* menunjukkan bahwa pendekatan *data mining* dapat dimanfaatkan sebagai alat bantu analisis untuk memahami perilaku pasar aset kripto. Dalam konteks sistem informasi, hasil penelitian ini dapat digunakan sebagai dasar dalam pengembangan sistem pendukung keputusan yang menyajikan informasi kondisi pasar berdasarkan data historis.

4.2 Saran

Berdasarkan hasil penelitian yang telah dilakukan, terdapat beberapa saran yang dapat diajukan untuk pengembangan penelitian selanjutnya. Penelitian berikutnya disarankan untuk menggunakan rentang data yang lebih panjang atau menambahkan variabel lain, seperti indikator teknikal maupun faktor eksternal, sehingga hasil analisis yang diperoleh menjadi lebih komprehensif.

Selain itu, metode *clustering* lain dapat digunakan sebagai pembanding, seperti *Hierarchical Clustering* atau *DBSCAN*, untuk mengamati perbedaan hasil pengelompokan data. Penelitian selanjutnya juga dapat mengintegrasikan hasil *clustering* ke dalam sistem informasi atau aplikasi analisis pasar kripto agar dapat dimanfaatkan secara langsung oleh pengguna sebagai alat bantu dalam pengambilan keputusan.

5. DAFTAR PUSTAKA

- Acharya, J.G.A. (2011) 'Cluster ensembles', *WIREs Data Mining and Knowledge Discovery*, 1. <https://doi.org/10.1002/widm.32>
- Aghabozorgi, S. and Teh, Y.W. (2014) 'Stock market co-movement assessment using a three-phase clustering method', *Expert Systems with Applications*, 41. <https://doi.org/10.1016/j.eswa.2013.08.028>
- Aghabozorgi, S., Shirkhorshidi, A.S. and Wah, T.Y. (2015) 'Time-series clustering—a decade review', *Information Systems*, 53. <https://doi.org/10.1016/j.is.2015.04.007>

- Arbelaitz, O., Gurrutxaga, I., Muguerza, J., Pérez, J.M. and Perona, I. (2013) 'An extensive comparative study of cluster validity indices', *Pattern Recognition*, 46. <https://doi.org/10.1016/j.patcog.2012.07.021>
- Arroyo, J. and Maté, C. (2009) 'Forecasting histogram time series with k-nearest neighbours methods', *International Journal of Forecasting*, 25. <https://doi.org/10.1016/j.ijforecast.2008.07.003>
- Bonanno, G., Caldarelli, G., Lillo, F., Micciche, S., Vandewalle, N. and Mantegna, R.N. (2004) 'Networks of equities in financial markets', *European Physical Journal B*, 38. <https://doi.org/10.1140/epjb/e2004-00129-6>
- Brauneis, A. and Mestel, R. (2019) 'Cryptocurrency-portfolios in a mean-variance framework', *Finance Research Letters*, 28. <https://doi.org/10.1016/j.frl.2018.05.008>
- Chan, S., Chu, J., Nadarajah, S. and Osterrieder, J. (2017) 'A statistical analysis of cryptocurrencies', *Journal of Risk and Financial Management*, 10. <https://doi.org/10.3390/jrfm10020012>
- Charrad, M., Ghazzali, N., Boiteau, V. and Niknafs, A. (2014) 'NbClust: an R package for determining the relevant number of clusters in a data set', *Journal of Statistical Software*, 61. <https://doi.org/10.18637/jss.v061.i06>
- Corbet, S., Lucey, B., Urquhart, A. and Yarovaya, L. (2019) 'Cryptocurrencies as a financial asset: a systematic analysis', *International Review of Financial Analysis*, 62. <https://doi.org/10.1016/j.irfa.2018.09.003>
- D'Urso, P., Giovanni, L. and Massari, R. (2016) 'Garch-based robust clustering of time series', *Fuzzy Sets and Systems*, 305. <https://doi.org/10.1016/j.fss.2016.01.010>
- González-Rivera, G. and Arroyo, J. (2012) 'Time series modeling of histogram-valued data: the daily histogram time series of S&P500 intradaily returns', *International Journal of Forecasting*, 28. <https://doi.org/10.1016/j.ijforecast.2011.02.007>
- Irpino, A. and Verde, R. (2015) 'Basic statistics for distributional symbolic variables: a new metric-based approach', *Advances in Data Analysis and Classification*, 9, pp.241–260. <https://doi.org/10.1007/s11634-014-0176-4>
- Liao, S.-H. and Chou, S.-Y. (2013) 'Data mining investigation of co-movements on the Taiwan and China stock markets for future investment portfolio', *Expert Systems with Applications*, 40. <https://doi.org/10.1016/j.eswa.2012.08.075>
- Song, J., Chang, W. and Song, J. (2019) 'Cluster analysis on the structure of the cryptocurrency market via bitcoin-ethereum filtering', *Physica A: Statistical Mechanics and its Applications*. <https://doi.org/10.1016/j.physa.2019.121339>