# DAFTAR ISI

[DAFTAR ISI i](#_Toc133985358)

[BAB I. PENDAHULUAN 1](#_Toc133985359)

[1.1 Latar Belakang 1](#_Toc133985360)

[1.2 Rumusan Masalah dan Pertanyaan Penelitian 2](#_Toc133985361)

[1.1 Batasan Masalah 2](#_Toc133985362)

[1.4 Tujuan Perancangan 3](#_Toc133985367)

[4.5 Manfaat Perancangan 3](#_Toc133985373)

[BAB II. TINAJAUAN PUSTAKA 4](#_Toc133985374)

[2.1 Tinjauan Pustaka 4](#_Toc133985375)

[2.2 Landasan Teori 5](#_Toc133985376)

[BAB III. KONSEP PERANCANGAN 7](#_Toc133985377)

[2.3 Diagram Alir Perancangan 7](#_Toc133985378)

[2.4 Pertimbangan Perancangan 7](#_Toc133985379)

[2.5 Analisis Teknis 7](#_Toc133985380)

[2.3.1 Bahasa Pemrograman 8](#_Toc133985386)

[2.3.2 Rentang Data 8](#_Toc133985387)

[2.3.3 Format Data 8](#_Toc133985388)

[2.3.4 Algoritma K-Means 8](#_Toc133985389)

[2.6 Peralatan dan Bahan 8](#_Toc133985390)

[DAFTAR PUSTAKA 10](#_Toc133985391)

# BAB I. PENDAHULUAN

## Latar Belakang

Berdasarkan data dari Badan Pengawas Perdagangan Berjangka Komoditi (Bappebti), terdapat sekitar 16,1 juta investor aset kripto di Indonesia pada bulan Agustus 2022. Jumlah total investor aset kripto pada akhir tahun 2021 hanya sebesar 11,2 juta, yang menunjukkan bahwa terjadi peningkatan sekitar 43,75 persen jumlah investor aset kripto selama periode Januari hingga Agustus 2022. Fakta ini menjadi semakin menarik jika dibandingkan dengan data investor pasar modal per Juli 2022, di mana jumlah investor aset kripto pada bulan tersebut melebihi 15,57 juta orang. Jumlah ini lebih besar sebesar 6,27 juta orang dibandingkan dengan jumlah investor pasar modal Indonesia yang hanya sebesar 9,3 juta orang. Selain itu, Bappebti juga melaporkan bahwa volume transaksi aset kripto pada periode Januari hingga Agustus 2022 mencapai setidaknya Rp. 249,3 triliun.

Cryptocurrency merupakan mata uang virtual yang digunakan sebagai mata uang alternatif dimana mata uang tersebut dihasilkan dan diperdagangkan melalui proses kriptografi (Nugraha, 2018). Cryptocurrency merupakan salah satu bentuk mata uang digital yang semakin populer belakangan ini dan menarik perhatian banyak orang karena potensi keuntungannya. Namun, Cryptocurrency juga memiliki nilai yang sangat volatil dan sering mengalami fluktuasi harga yang cukup besar, sehingga tidak selalu menjadi pilihan yang aman bagi para investor (Investopedia, 2022). Oleh karena itu, terdapat banyak penelitian yang dilakukan untuk memahami tren dan melakukan prediksi pada suatu aset kripto.

Penelitian terhadap Cryptocurrency telah banyak dilakukan. Beberapa diantaranya yaitu, Andi (2021), Hua (2020) dan Kwon et al (2019) menggunakan model LSTM untuk memprediksi harga Bitcoin. Lalu, Pant et al (2018) menggunakan Recurrent Neural Network (RNN) dan Huang et al (2021) menggunakan Long Short-Term Memory Network (LSTM) untuk memprediksi harga Cryptocurrency berdasarkan analisis sentimen Twitter. Sementara itu, dalam penggunaan Natural Language Processing (NLP) terdapat penelitian yang melakukan pemodelan pada sentimen Twitter terhadap Cryptoccurency oleh Citien et al (2022), Wong (2021), dan Deebadi (2022).

Berdasarkan penelitian sejenis mengenai prediksi, perbandingan, maupun pembelajaran dalam memahami Bitcoin sebagai mata uang digital. Bitcoin cenderung memiliki pola dalam transaksi dan data historikal-nya. Hal ini sering dimanfaatkan untuk mengurangi resiko investasi pada Bitcoin. Seringnya, prediksi Bitcoin hanya memanfaatkan data historikal harga tutup berdasarkan waktu, sehingga pola yang dimiliki cenderung bergantung pada kondisi pasar yang stabil atau terdistribusi secara normal. Beberapa penelitian cenderung memiliki nilai galat yang besar saat memprediksi harga. Salah satu di antaranya terjadi pada kondisi pasar dengan sentimen negatif yang cukup besar di media sosial terhadap aset kripto.

Berdasarkan latar belakang tersebut, terbuka peluang untuk melakukan penelitian yang dapat mengatasi permasalahan dalam memprediksi harga Cryptocurrency. Diperlukan analisis yang tepat untuk menambahkan variabel dalam pemodelan prediksi harga Cryptocurrency. Salah satu pendekatan yang dapat dilakukan adalah menggunakan sentimen analisis pada media sosial Twitter. Namun, penelitian yang telah melibatkan sentimen analisis tersebut masih memiliki kekurangan. Hal ini terlihat dari pengambilan data teks dalam pemodelan sentimen yang tidak memperhatikan kualitas data. Pengambilan data teks cenderung acak, sesuai dengan hasil scrapping dari API yang digunakan. Padahal, popularitas suatu tweet tentunya akan mempengaruhi psikologi pasar. Oleh karena itu, dengan memanfaatkan interaksi yang banyak terjadi pada sebuah tweet, diharapkan dapat menghasilkan model dengan akurasi yang lebih baik.

## Rumusan Masalah dan Pertanyaan Penelitian

Berdasarkan penjelasan yang telah disampaikan pada latar belakang, Secara umum, transaksi Cryptocurrency dianggap cepat dan aman karena sistemnya yang terdesentralisasi. Namun, terdapat risiko volatilitas harga Cryptocurrency yang signifikan akibat sentimen negatif pada sosial media. Oleh karena itu, perlu dilakukan penelitian untuk memprediksi harga suatu Cryptocurrency berdasarkan sentimen yang sedang terjadi di sosial media.

## Batasan Masalah

Berdasarkan latar belakang yang telah diuraikan di atas, diperoleh batasan masalah yang dalam penelitian ini adalah:

1. Data yang digunakan merupakan data 10 teratas Coinmarketcap.
2. Data sosial media yang digunakan merupakan data API Twitter
3. Rentang data yang digunakan dari Januari 2019 hingga Desember 2022

## Tujuan Perancangan

Berdasarkan rumusan masalah yang telah diuraikan di atas, maka tujuan penelitian yang ingin dicapai adalah :

1. Untuk mengetahui bahwa sosial media memiliki pengaruh terhadap harga Cryptocurrency.
2. Untuk mengetahui bahwa 10 teratas Coinmarketcap memiliki pengaruh terhadap harga Cryptocurrency lainnya.
3. Untuk mengetahui akurasi dari hasil prediksi harga suatu Cryptocurrency dengan sentimen sosial media.

## Manfaat Perancangan

Berdasarkan latar belakang yang telah diuraikan di atas, diperoleh manfaat perancangan yang ingin dicapai adalah:

1. Manfaat Teoritis

Sebagai sarana untuk menambah ilmu pengetahuan di bidang machine learning, Cryptocurrency, dan sentimen sosial media.

1. Manfaat Praktis

Hasil penelitian diharapkan dapat memberikan wawasan lebih lanjut mengenai supply and demand yang terjadi pada transaksi Cryptocurrency. Sehingga, dapat menjadikan hasil penelitian ini sebagai bahan pertimbangan dalam pengambilan keputusan saat melakukan perdagangan Cryptocurrency.

# BAB II. TINAJAUAN PUSTAKA

## Tinjauan Pustaka

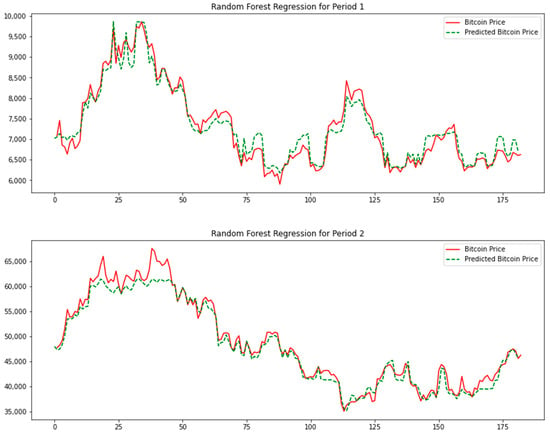
Chilamkurti, dkk dari La Trobe University (2020) melakukan penelitian untuk memprediksi harga Bitcoin menggunakan ARIMA Model. Penelitian ini menggunakan data historis dari harga sebelumnya untuk memprediksi harga Bitcoin di masa depan dan bertujuan membantu mengurangi risiko investasi pada Bitcoin. Dalam penelitiannya, dijelaskan bahwa Bitcoin mampu menciptakan sistem transaksi yang lebih cepat dengan mengeliminasi otoritas pusat atau yang saat ini dikenal dengan sistem terdesentralisasi. Sistem ini juga menyebabkan biaya transaksi menjadi semakin murah karena semua sumber daya dilakukan secara langsung tanpa perlu melibatkan pihak ketiga. Kondisi ini juga membuat likuiditas pada supply dan demand Bitcoin menjadi tinggi.

Debasmita, dkk (2023) melakukan penelitian yang kemudian diterbitkan oleh Elsevier Inc. dalam Decision Analitycs Journal 6. Penelitian ini menggunakan klaster untuk menganalisis nilai ekstrim dari Bitcoin. Dalam literatur penelitiannya, disebutkan bahwa Bitcoin memiliki nilai volatilitas yang tinggi, namun pada saat yang sama sangat responsif dalam memperbarui informasi harga. Hal ini seringkali menyebabkan terjadinya "overreaction hypothesis", yang merupakan salah satu anomali pasar saham yang banyak dipelajari dalam studi De Bondt dan Thaler. Studi tersebut menunjukkan bahwa perubahan harga pada hari berikutnya cenderung lebih besar daripada setelah hari yang normal.

Dalam sebuah penelitian yang dipublikasikan di Journal of Soft Computing Paradigm (JSCP) pada tahun 2021, Andri H. membahas penggunaan regresi logistik dengan model machine learning LSTM untuk memprediksi harga Bitcoin. Pada pendahuluan penelitiannya, Andri H. mengemukakan bahwa Bitcoin sebagai salah satu mata uang digital memiliki identitas, struktur, dan fungsi yang konsisten sehingga memiliki masa depan yang cerah. Bitcoin juga diakui memiliki keunggulan dalam mengurangi keterlibatan pihak ketiga seperti bank, kartu kredit, dan fungsi pemerintahan dalam transaksi, yang pada akhirnya dapat mengurangi waktu dan biaya transaksi.

Berdasarkan penelitian sejenis mengenai prediksi, perbandingan, dan pembelajaran dalam memahami Bitcoin sebagai mata uang digital, Bitcoin cenderung memiliki pola dalam transaksi dan data historisnya. Hal ini sering dimanfaatkan untuk mengurangi risiko investasi pada Bitcoin. Umumnya, prediksi harga Bitcoin menggunakan data historis harga penutupan berdasarkan waktu, sehingga pola yang dimiliki cenderung bergantung pada kondisi pasar yang stabil. Namun, beberapa penelitian menunjukkan nilai galat yang besar saat memprediksi harga pada kondisi pasar dengan transaksi yang tidak wajar, yang kemudian disebut sebagai "transaksi paus".

Hasil penelitian Chen J. (2023) membahas mengenai Analisis Prediksi Harga Bitcoin Menggunakan Machine Learning, yang menunjukkan hasil yang cukup stabil ketika tren harga Bitcoin juga stabil. Namun, ketika harga Bitcoin mengalami transaksi yang tidak wajar (disebut sebagai transaksi paus), galat antara prediksi dan harga sebenarnya cenderung besar.



Gambar 2.1. Grafik perbandingan hasil prediksi Bitcoin terhadap harga sebenarnya menggunakan Random Forest (Chen, 2023)

## Landasan Teori

### Data Mining

Data Mining adalah…

### Bitcoin

Bitcoin adalah

### Clustering

Clustering adalah

### K-Means

K-Means adalah…

K-Means keunggulan dibanding algoritma lain…

K-Means sesuai dengan data Bitcoin…

K-Means sensitif terhadap data outlier sehingga…

# BAB III. KONSEP PERANCANGAN

## Diagram Alir Perancangan

Berikut diagram alir pada pemodelan transaksi paus pada Bitcoin:

Gambar 1. Diagram Alir Perancangan

## Pertimbangan Perancangan

Dalam pelaksanaannya, terdapat pertimbangan yang perlu diperhatikan dalam membuat pemodelan yang akan dilakukan. Seperti membuat Pipeline yang dapat menyesuaikan data kedalam model training yang akan di jalankan. Bersamaan dengan hal tersebut data harus di normalisasi sehingga tidak terdapat data yang kosong, null, atau memiliki satuan ukur yang berbeda. Hal ini karena Algoritma K-Means sangat sensitif terhadap data outlier.

Selain itu, terdapat pertimbangan dalam menentukan spesifikasi sarana dalam pembuatan model yang akan dilakukan. Terdapat 2 pilihan utama dalam melakukan pelatihan yakni dengan menggunakan hardware dari aplikasi berbasis cloud atau menggunakan hardware lokal dengan aplikasi yang dijalankan secara On-Premises. Jika pemodelan akan dijalankan menggunakan basis cloud maka perlu diperhatikan penggunaan akselerasi hardware. Sebaiknya digunakan akselerasi GPU untuk memberikan performa optimal terhadap waktu pelatihan model. Selain itu, pemodelan dapat dilakukan secara On-Premises pada komputer pribadi, diperlukan setidaknya spesifikasi minimal yang tertera pada paltform aplikasi yang akan digunakan.

## Analisis Teknis

Beberapa analisis teknis yang dipertimbangkan pada pemodelan menggunakan algoritma K-Means ini adalah sebagai berikut:



### Bahasa Pemrograman

Bahasa pemrograman yang digunakan dalam pemodelan ini adalah Python. Penggunaan Python dalam bidang machine learning merupakan hal yang populer. Hal ini karena bahasa pemrograman Python yang open source memudahkan pengembangan library yang berguna bagi komunitasnya. Syntax yang sederhana juga menjadi alasan penggunaan bahasa pemrogram ini.

### Rentang Data

Bitcoin pertama kali diperdagangkan pada tahun 2008. Pada dasarnya semua transaksi Bitcoin tercatat dalam block chain Bitcoin itu sendiri. Namun, saat itu perdagangan bitcoin masih tabu dan belum dapat dilihat sebagai aset yang berharga sehingga membuat data transaksi kebanyakan berasal dari penambangan dan airdrop. Sehingga, untuk menyesuaikan tujuan pemodelan dalam mengelompokkan transaksi paus agar dapat digunakan trader sebagai wawasan dalam bertransaksi. Akhirnya, diputuskan pengambilan data transaksi Bitcoin pada 3 tahun sebelumnya.

### Format Data

Sebagai bentuk penyajian hasil pemodelan yang baik. Diperlukan sebuah format data yang mudah dipahami. Digunakan format excel yang *readable* untuk memberikan kemudahan dalam pengambilan manfaat pada pemodelan ini.

### Algoritma K-Means

Penggunaan Algoritma K-Means merupakan inisialisasi awal untuk mengetahui kemampuan metode cluster dalam mengelompokkan data transaksi pada Bitcoin sehingga memperoleh wawasan mengenai transaksi paus. Kompleksitas algoritma yang sederhana dengan tujuan yang sama dengan algoritma clustering lainnya menjadikan algoritma ini sesuai untuk pemodelan yang akan dilakukan dan memberi ruang untuk penelitian sejenis jika metode clustering pada umumnya dapat menemukan pola dari transaksi paus pada transaksi Bitcoin.

## Peralatan dan Bahan

Terdapat beberapa sarana yang diperlukan untuk melakukan clustering pada transaksi Bitcoin ini sebagai berikut:

1. Server atau komputer dengan sistem operasi yang mendukung Python, seperti Linux, macOS, atau Windows.
2. Browser web yang mendukung Python, seperti Google Chrome, Mozilla Firefox, atau Microsoft Edge.
3. Data Formatting yang mudah dipahami dan mendukung pengaksesan menggunakan Python seperti Excel .
4. Data source atau sumber data yang akan diterima dan disimpan kedalam bentuk Excel, seperti Web downloader atau API yang terkoneksi ke data transaksi Bitcoin pada platform Binance.
5. Library tambahan atau add-on yang dapat memperluas fitur dan kemampuan Python. Seperti library untuk visualisasi hasil, Library yang dapat mengolah data, serta library untuk menjalankan algoritma K-Means.

# DAFTAR PUSTAKA

Nugraha, G. (2018) "PENGARUH RESPONS IMF TERKAIT PENGGUNAAN CRYPTOCURRENCY TERHADAP KEBIJAKAN PEMERINTAH INDONESIA," eprints.undip

Chen, J. (2023). Analysis of Bitcoin Price Prediction Using Machine Learning. Journal of Risk and Financial Management, 16(1), 51. https://doi.org/10.3390/jrfm16010051

Andi, H. K. (2021). An accurate bitcoin price prediction using logistic regression with LSTM machine learning model. Journal of Soft Computing Paradigm, 3(3), 205-217.

Poongodi, M., Vijayakumar, V., & Chilamkurti, N. (2020). Bitcoin price prediction using ARIMA model. International Journal of Internet Technology and Secured Transactions, 10(4), 396-406.

Das, D., Kayal, P., & Maiti, M. (2023). A K-means clustering model for analyzing the Bitcoin extreme value returns. Decision Analytics Journal, 6, 100152.

Nakamoto, S. (2018) Bitcoin: A Peer-to-Peer Electronic Cash System. *bitcoin.org*

Investopedia, Cryptocurrency Explained with Pros and Cons for Investment. Available: https://www.investopedia.com/terms/c/Cryptocurrency.asp 26-Sep-2022 [Online]. [Diakses: 22-Sep-2021].

Hua, Y. (2020). Bitcoin price prediction using ARIMA and LSTM. In E3S Web of Conferences (Vol. 218, p. 01050). EDP Sciences.

Kwon, D. H., Kim, J. B., Heo, J. S., Kim, C. M., & Han, Y. H. (2019). Time series classification of Cryptocurrency price trend based on a recurrent LSTM neural network. Journal of Information Processing Systems, 15(3), 694-706.

FITRIANI, F., Stiawan, D., & Heryanto, A. (2021). *DETEKSI SERANGAN MALWARE RANSOMWARE PADA BITCOIN MINING DENGAN METODE K-MEANS CLUSTERING* (Doctoral dissertation, Sriwijaya University).

Barradas, A., Tejeda-Gil, A., & Cantón-Croda, R. M. (2022). Real-Time Big Data Architecture for Processing Cryptocurrency and Social Media Data: A Clustering Approach Based on k-Means. *Algorithms*, *15*(5), 140.

-----------------------------------------------------------------------------------------------------------------------

Pant, D. R., Neupane, P., Poudel, A., Pokhrel, A. K., & Lama, B. K. (2018, October). Recurrent neural network based bitcoin price prediction by Twitter sentiment analysis. In *2018 IEEE 3rd International Conference on Computing, Communication and Security (ICCCS)* (pp. 128-132). IEEE.

Critien, J. V., Gatt, A., & Ellul, J. (2022). Bitcoin price change and trend prediction through Twitter sentiment and data volume. Financial Innovation, 8(1), 1-20.

Wong, E. L. X. (2021). Prediction of Bitcoin prices using Twitter Data and Natural Language Processing.

Huang, X., Zhang, W., Tang, X., Zhang, M., Surbiryala, J., Iosifidis, V., ... & Zhang, J. (2021). Lstm based sentiment analysis for cryptocurrency prediction. In Database Systems for Advanced Applications: 26th International Conference, DASFAA 2021, Taipei, Taiwan, April 11–14, 2021, Proceedings, Part III 26 (pp. 617-621). Springer International Publishing.

Deebadi, A. (2020). Understanding Impact of Twitter Feed on Bitcoin Price and Trading Patterns.