

BAB II

KAJIAN PUSTAKA

2.1 Tinjauan Pustaka

Pesatnya pertumbuhan Pasar digital *Cryptocurrency* dengan volatilitas harganya yang tinggi serta penerapannya dalam berbagai transaksi komersial telah menarik perhatian akademisi dan investor. *Cryptocurrency* semakin meluas perkembangannya sebagai alat tukar dan sebagai komoditas yang diperdagangkan 24 jam nonstop di berbagai bursa di seluruh dunia (Rizky Parlika and Atmaja 2018a).

Cryptocurrency terdiri atas *Bitcoin* dan *Altcoin* (*Alternative coin*). *Bitcoin* adalah bentuk Sistem Uang Elektronik *Peer-to-Peer* yang memungkinkan pembayaran secara daring dan dapat dikirim langsung dari satu pihak ke pihak lain tanpa melalui perantara lembaga keuangan (Nakamoto n.d.). Bila *Bitcoin* hanya berjumlah 1 jenis saja dan merupakan pioneer dari *Cryptocurrency*, maka *Altcoin* memiliki jumlah ribuan jenis (<https://Coinmarketcap.com>, 2023). Di antara seluruh *Cryptocurrency*, *Bitcoin* hampir setiap saat memiliki volume perdagangan yang paling tinggi dibandingkan seluruh *Cryptocurrency* yang bereputasi (<https://Coinmarketcap.com>, 2023).

Bitcoin sebagai *Cryptocurrency* terbesar dan pioner semakin populer dan telah menarik perhatian investor individu dan institusional, sehingga penting bagi investor dan perusahaan untuk dapat memprediksi volatilitasnya dengan lebih akurat (Tapia and Kristjanpoller 2022). Hal ini menyebabkan pentingnya perkiraan arah harga *Bitcoin* yang akurat, untuk menentukan tren harga dan alokasi aset *Bitcoin* (Basher and Sadorsky 2022).

2.1.1 Volatilitas *Cryptocurrency*

Karena harga *Bitcoin* sangat fluktuatif, memperkirakan volatilitasnya sangat penting untuk banyak aplikasi, seperti manajemen risiko atau lindung nilai (Bergsli et al. 2022). Untuk memperkuat kemampuan Dataset Terpola dalam

melakukan Prediksi Fluktuasi hingga Pemetaan Pola Volatilitas harga *Cryptocurrency* terutama terhadap harga *Bitcoin* sebagai *Cryptocurrency* yang paling berpengaruh, Dataset Terpolat dapat dijadikan sebagai Data *Training* dan Data *Testing* pada beberapa metode prediksi baik berbasis *Statistical Learning* (D. Kumar, Kumar, and Verma 2022), *Machine Learning* (Yijun Wang, Andreeva, and Martin-barragan 2023), *Deep Learning* (D'Amato, Levantesi, and Piscopo 2022), Analisis Teknikal (Goutte et al. 2023), dan Analisis Sentimen berbasis *Fear and Greed Index* (Barua and Sharma 2023).

2.1.2 Model Prediksi *Cryptocurrency*

Model prediksi pada pasar perdagangan yang akurat dapat memberi investor alat untuk membantu membuat keputusan berbasis dataset yang lebih baik. Model ini dapat membantu pedagang mengurangi risiko investasi dan memilih aset yang berpotensi paling menguntungkan. Selain itu, membuat model lanjutan memungkinkan penggunaan data non-tradisional seperti historis harga aset dan berita terkait aset tersebut. Kemajuan terbaru dalam teknik pembelajaran mesin, seperti Pembelajaran Mendalam, Teknik Penambangan Teks, dan Teknik Ensemble, menimbulkan kebutuhan untuk melakukan penelitian yang berkelanjutan (Bustos and Pomares-Quimbaya 2020). Contoh Penelitian terkait model pembelajaran mesin diantaranya menggunakan model *Long Short-Term Memory* untuk memprediksi harga *Bitcoin* dengan interval 5 menit. Hasilnya model pembelajaran mesin lebih unggul daripada metode statistik, dengan akurasi mencapai 67,2% (Z. Chen, Li, and Sun 2020b). Investigasi mereka terhadap prediksi harga *Bitcoin* dapat dianggap sebagai studi percontohan tentang pentingnya dimensi sampel dalam teknik pembelajaran mesin. Dengan menggunakan Model Dataset Terpolat diharapkan akan diperoleh capaian akurasi yang lebih baik dalam memprediksi pergerakan harga *Cryptocurrency* di masa depan, serta mampu menunjukkan Pola terbentuknya fluktuasi hingga pola berakhirnya fluktuasi, sehingga secara keseluruhan volatilitas harga *Cryptocurrency* dapat diprediksi dengan Model Dataset Terpolat.

2.1.3 *Cryptocurrency, Bitcoin, dan Altcoin*

Perdagangan Cryptocurrency telah menjadi jenis investasi yang makin diminati Masyarakat Indonesia dan Dunia. Mengapa diminati, salah satunya karena Cryptocurrency bekerja mirip dengan mata uang standar, namun mekanisme pembayaran virtual dilakukan tanpa campur tangan otoritas pusat mana pun (Aslam et al. 2022). Konsep mengenai Cryptocurrency pertama diperkenalkan oleh (Nakamoto n.d.), dengan nama Bitcoin. Seiring meningkatnya kepercayaan masyarakat pengguna teknologi Blockchain dengan produk utamanya yakni Cryptocurrency, bertumbuhlah berbagai vendor dan organisasi dan lembaga yang mengembangkan Cryptocurrency dengan tujuan yang spesifik. Makin banyaknya pilihan komoditas Cryptocurrency, muncullah pengindeks dan peranking Cryptocurrency. Coinmarketcap adalah organisasi pengindeks bereputasi yang sering menjadi acuan utama pada skala internasional (<https://Coinmarketcap.com/> 2023).

Pada tahun-tahun berikutnya bermunculan beraneka jenis dan varian Cryptocurrency (<https://Coinmarketcap.com/> 2023). Bila Bitcoin hanya terdiri dari 1 jenis saja, maka bentuk Cryptocurrency lainnya adalah Alternative coin yang disingkat dengan Altcoin. Altcoin terdiri dari berbagai ragam, varian, jenis, dan tipe. Cryptocurrency diperdagangkan pada berbagai Pasar digital bereputasi yang tersebar di seluruh dunia (Coinmarketcap.com 2023b).

Indodax adalah salah satu Pasar digital perdagangan Cryptocurrency yang terkemuka di Indonesia (Coinmarketcap.com 2023a). Indodax menyediakan akses terhadap data seluruh perdagangan aset Cryptocurrency secara realtime melalui koneksi berbasis API (Indodax.com 2023). Dari data-data inilah dikembangkanlah Pada Penelitian ini Model Dataset Terpola untuk memprediksi volatilitas harga Cryptocurrency.

2.1.4 Pengaruh *Bitcoin* terhadap *Altcoin*

David dkk, mempertimbangkan tiga metode, yaitu *Auto-Regressive Integrated Moving Average* (ARIMA), *Auto-Regressive Fractionally Integrated Moving Average* (ARFIMA) dan Analisis *Fluktuasi Detrended*, serta tiga indeks

yang diberikan oleh eksponen *Hurst* dan *Lyapunov* atau Dimensi Fraktal. Informasi ini memungkinkan menilai perilaku deret waktu, seperti persistensi, keacakan, prediktabilitasnya dan kekacauan data. Hasilnya menunjukkan bahwa, kecuali *Bitcoin*, mata uang *Cryptocurrency* lainnya menunjukkan prediktabilitas yang lebih rendah bila dibandingkan dengan *Bitcoin* (David et al. 2021). *Bitcoin* adalah Pioneer dan tidak mudah untuk ditiru. Hal ini dibuktikan oleh Cahil dan Frank, menggunakan pengaturan unik dalam mata uang *Cryptocurrency* dengan *Bitcoin* menjadi fokus utamanya, mereka mendefinisikan peniru sebagai mata uang *Cryptocurrency* yang memiliki nama yang mirip dengan *Bitcoin*. Hasilnya menunjukkan peniru mendapatkan pengembalian yang lebih tinggi di hari pertama dan minggu pertama perdagangan, tetapi memiliki tingkat kelangsungan hidup yang lebih rendah (Cahill and Liu 2021).

Bitcoin hingga saat ini adalah *Cryptocurrency* yang paling dikenal dan disebarluaskan, dengan volume transaksi, nilai pasar, dan penerimaan yang lebih besar dalam layanan pertukaran (<https://Coinmarketcap.com>, 2023). Qiu, dkk, mempelajari apakah efek limpahan volatilitas di antara mata uang *Cryptocurrency* penting untuk meramalkan volatilitas realisasi *Bitcoin*. Hasil penelitian mereka menunjukkan bahwa model volatilitas *Bitcoin* menunjukkan efek keterkaitan yang kuat di antara masing-masing *Cryptocurrency*, namun penggunaan informasi ini untuk prediksi volatilitas *Bitcoin* masih belum maksimal (Qiu et al. 2021).

Bitcoin merupakan Aset *Cryptocurrency* pertama dan sering menjadi yang terbesar kapitalisasi pasarnya serta mempengaruhi banyak *Cryptocurrency* lainnya yang transaksi perdagngannya tercatat secara periodik di *Coinmarketcap* (<https://Coinmarketcap.com>, 2023). Pada detail dari laman *Coinmarketcap* (<https://Coinmarketcap.com>, 2023), ditunjukkan kapitalisasi pasar *Bitcoin* seringkali lebih besar dibandingkan dengan total keseluruhan kapitalisasi pasar *Altcoin* yang tercatat. *Bitcoin* juga banyak digunakan sebagai pasangan perdagangan aset *Cryptocurrency* pada banyak sekali Pasar digital *Cryptocurrency* bereputasi dari seluruh penjuru dunia yang tercatat pada laman *Coinmarketcap* (*Coinmarketcap.com* 2023b). Sebagai contoh pasangan perdagangan ETH BTC berarti memperdagangkan Aset *Altcoin Ethereum* (ETH)

dengan satuan harga *Bitcoin* (BTC). Hal inilah yang memungkinkan *Bitcoin* memiliki kekuatan untuk mempengaruhi *Altcoin* baik secara parsial maupun keseluruhan.

2.1.5 Penelitian pendahuluan

Pada penelitian sebelumnya diusulkan persamaan sederhana yang disebut rizubot untuk mendapatkan cara pembacaan pasar perdagangan *Cryptocurrency* (Rizky Parlika and Atmaja 2018a). Data-data perdagangan *Cryptocurrency* dapat diperoleh dari koneksi API yang dibagikan oleh berbagai Pasar digital perdagangan *Cryptocurrency*. Penelitian selanjutnya berhasil mengekstraksi data harga *Bitcoin* terbaru dari 20 pasar *Cryptocurrency* (Rizky Parlika and Atmaja 2020). Data-data awal ini kemudian disimpan dalam database *MySQL* di *hosting* dan ditambah secara terus menerus (Rizky Parlika and Pratama 2021a). 2 tahun berselang mulailah dikembangkan ide terkait Model Dataset Terpola. Pada Penelitian ini dijelaskan pembentukan hingga Validasi dari Model Dataset Terpola yang dibangun menggunakan sumber data dari akses API pada pasar *Indodax* (*Indodax.com* 2023).

2.1.6 Pasar *Indodax* sebagai penyedia data-data transaksi perdagangan

Pasar *Indodax* merupakan Pasar digital *Cryptocurrency* yang memiliki platform marketplace terbesar di Indonesia untuk membeli dan menjual aset digital yakni *Bitcoin* dan *Altcoin* serta telah terdaftar resmi di Badan Pengawas Perdagangan Berjangka Komoditi di Indonesia (*Coinmarketcap.com* 2023a). Pada Penelitian ini digunakan data-data pergerakan harga *Bitcoin* dan *Altcoin* dari pasar perdagangan *Cryptocurrency Indodax* melalui akses API pada pasar *Indodax* (*Indodax.com* 2023).

Meskipun Model dataset terpola ini awalnya dibangun menggunakan data-data pergerakan *Cryptocurrency* dari pasar perdagangan *Indodax*, namun Model ini juga dapat secara mudah diimplementasikan pada berbagai pasar perdagangan *cryptocurrency* lainnya yang terdaftar di *Coinmarketcap* (Rizky Parlika and Atmaja 2018a), (*Coinmarketcap.com* 2023b), selama pasar-pasar tersebut

memberikan akses pembacaan data-data transaksi perdagangannya secara *Realtime* melalui koneksi API.

2.1.7 Prediksi *Cryptocurrency* menggunakan *Statistical Learning*

Penelitian Terkait *Statistical Learning* pada tahun 2019 menyelidiki karakteristik distribusi dan dinamika return dan volatilitas *Bitcoin* menggunakan dataset dari 2011 hingga 2017, serta menilai pengaruh sentimen investor terhadap volatilitas *Bitcoin*. Model autoregressive sederhana digunakan untuk prediksi, dengan Google Trend Index sebagai proksi sentimen investor. Hasilnya menunjukkan bahwa volatilitas *Bitcoin* tidak memiliki sifat persistence yang biasa ditemukan dalam *Time Series* keuangan, tetapi sentimen investor signifikan dalam menjelaskan perubahan volatilitas di masa depan (Eom et al. 2019).

Pada tahun 2020, penelitian terkait yang bertujuan untuk mengidentifikasi dan memprediksi keberadaan gelembung dan keruntuhan di pasar *Bitcoin* menggunakan dataset harga *Bitcoin* harian dari 2017 hingga 2019. Metode yang digunakan adalah model LPPLS dengan deteksi *Time Series* multilevel adaptif yang bekerja pada skala waktu satu jam dan 30 menit. Hasil utama menunjukkan bahwa model ini mampu mendeteksi gelembung dan memprediksi keruntuhan dengan akurasi tinggi, terutama dalam menangani fluktuasi harga yang ekstrem dalam jangka pendek (Shu and Zhu 2020).

Penelitian terkait pada tahun 2022 bertujuan untuk menganalisis dinamika korelasi di pasar *Cryptocurrency* dengan menggunakan data dari *Coinmarketcap* untuk periode 2017-2021. Metode yang digunakan termasuk analisis berbasis Influence Strength (IS) dan berbagai teknik jaringan seperti MST dan PMFG. Hasilnya menunjukkan bahwa terdapat event-event kritis dalam dinamika korelasi yang berkaitan dengan fluktuasi pasar yang besar, dan struktur jaringan cenderung mengalami perubahan signifikan selama event-event tersebut (Nie, 2022). Penelitian lainnya pada tahun yang sama bertujuan untuk mengeksplorasi sumber-sumber munculnya peluang arbitrase statistik di pasar *Bitcoin* dengan menggunakan data harga harian dari lima bursa besar selama periode 2017-2022. Metode yang digunakan termasuk Grey correlation dan DCC-GARCH untuk

mengukur korelasi antara return *Bitcoin* di berbagai bursa. Hasil utama menunjukkan bahwa peluang arbitrase muncul terutama saat terjadi kongesti jaringan dan volatilitas tinggi, sementara volume perdagangan yang tinggi dan aktivitas on-chain cenderung mengurangi peluang tersebut (Kristoufek and Bouri 2022).

Penelitian terkait pada tahun 2023 menganalisis pertumbuhan dan dampak *Bitcoin* menggunakan data historis dari 2009 hingga 2021. Dengan menggunakan analisis regresi linear berganda, penelitian ini menemukan bahwa beberapa variabel seperti ukuran blok rata-rata dan kapitalisasi pasar memiliki pengaruh signifikan terhadap waktu penyelesaian akhir harian aset. Hasil menunjukkan model memiliki tingkat prediksi yang baik, namun terdapat masalah multikolinearitas yang perlu diatasi dalam penelitian lanjutan (Kumarasamy and Vijayalakshmi 2024).

Metode *Autoregressive Integrated Moving Average* (ARIMA) adalah metode peramalan deret waktu yang digunakan untuk menganalisis dan memprediksi data time-series dengan mempertimbangkan hubungan antara nilai historis dan pola fluktuasi data. ARIMA merupakan model statistik yang banyak digunakan dalam prediksi keuangan, termasuk pergerakan harga *Cryptocurrency*.

Model ARIMA terdiri dari tiga komponen utama yang direpresentasikan dalam bentuk ARIMA(p, d, q):

a. (p) *Autoregressive* (AR):

Komponen p menunjukkan hubungan antara nilai saat ini dengan nilai sebelumnya dalam data time-series. Model ini menggunakan regresi linier dari data masa lalu untuk memperkirakan nilai di masa depan.

b. (d) Differencing (I - Integrated):

Komponen d Digunakan untuk membuat data menjadi stasioner, yaitu menghilangkan tren atau pola musiman. *Differencing* dilakukan dengan menghitung perubahan antara nilai saat ini dan nilai sebelumnya.

c. (q) Moving Average (MA):

Komponen q memperhitungkan kesalahan prediksi masa lalu dalam peramalan nilai saat ini. Model MA digunakan untuk mengurangi noise dan menangkap pola residual dalam data.

Pada laporan disertasi, ARIMA digunakan sebagai *model baseline* untuk membandingkan performa dengan model *Deep Learning* (LSTM). Berikut langkah-langkah implementasi ARIMA dalam penelitian:

1) Praproses Data

Pada tahap Praproses data, harga *Cryptocurrency* dikumpulkan dalam bentuk time-series berdasarkan harga harian (misalnya, harga *Bitcoin* dalam Rupiah). Data diperiksa untuk mengetahui apakah sudah stasioner (tidak memiliki tren atau pola musiman). Jika data tidak stasioner, maka dilakukan differencing (d) untuk menghilangkan tren dan membuat data lebih sesuai dengan model ARIMA.

2) Pemilihan Parameter ARIMA (p, d, q)

a) Identifikasi Stasioneritas:

Uji *Dickey-Fuller* dilakukan untuk memastikan apakah data sudah stasioner atau perlu dilakukan differencing.

b) Menentukan Nilai p dan q:

Autocorrelation Function (ACF) digunakan untuk menentukan nilai q (*Moving Average*). Partial Autocorrelation Function (PACF) digunakan untuk menentukan nilai p (*Autoregressive*).

c) Grid Search untuk Optimasi (p, d, q):

Dicoba beberapa kombinasi nilai p, d, q untuk mendapatkan kombinasi terbaik yang menghasilkan error paling rendah.

3) Pelatihan dan Evaluasi Model ARIMA

- a) Model ARIMA dilatih menggunakan data historis harga *Cryptocurrency*. Setelah model dipelajari, dilakukan prediksi harga untuk periode tertentu ke depan.

- b) Evaluasi dilakukan menggunakan metrik seperti *Mean Absolute Percentage Error* (MAPE) dan *Root Mean Squared Error* (RMSE) untuk mengukur akurasi prediksi.

4) Hasil dan Perbandingan dengan Model LSTM

Kelebihan ARIMA adalah Cocok untuk dataset dengan tren yang dapat diprediksi secara linier. Model sederhana dan tidak memerlukan daya komputasi tinggi. Kekurangan ARIMA kurang mampu menangkap pola non-linear dan volatilitas tinggi di pasar *Cryptocurrency*. Model ARIMA Tidak efektif jika data memiliki pola kompleks seperti pump and dump, flash *Crash*, atau bullish trend mendadak. ARIMA memiliki performa lebih buruk dibandingkan LSTM dalam memprediksi harga *Bitcoin* karena tidak mampu menangkap pola kompleks dalam data time-series *Cryptocurrency*.

2.1.8 Prediksi *Cryptocurrency* menggunakan *Machine Learning*

Penelitian sebelumnya berfokus pada pengembangan model prediksi arah harga, harga maksimum, minimum, dan penutupan harian *Bitcoin* menggunakan kombinasi teknik *Machine Learning* seperti ANN, SVM, dan ensemble RNN. Data yang digunakan meliputi data harga harian dari tahun 2013 hingga 2017, dengan hasil menunjukkan bahwa kombinasi model ini dapat mencapai akurasi yang lebih tinggi dibandingkan studi sebelumnya, dengan MAPE serendah 1.28% untuk prediksi harga maksimum (Mallqui and Fernandes 2019).

Penelitian lainnya memprediksi harga *Bitcoin* pada interval harian dan 5 menit menggunakan teknik *Machine Learning* dengan mempertimbangkan dimensi sampel. Dataset harian diambil dari *Coinmarketcap* dan dataset interval 5 menit dari pasar *Binance*. Model yang digunakan termasuk *Logistic Regression* (LR) dan *Linear Discriminant Analysis* (LDA) untuk harga harian, serta *Random Forest* (RF), *XGBoost* (XGB), *Quadratic Discriminant Analysis* (QDA), *Support Vector Machine* (SVM), dan *Long Short-Term Memory* (LSTM) untuk harga interval 5 menit. Hasil menunjukkan bahwa LR memberikan akurasi terbaik untuk

prediksi harga harian (66%), sementara LSTM paling akurat untuk interval 5 menit (67.2%) (Z. Chen, Li, and Sun 2020b).

Selanjutnya, dilakukan penelitian tentang prediksi pergerakan jangka pendek dari pasar *Bitcoin*, yang meliputi prediksi harga dalam rentang waktu 1 hingga 60 menit. Mengingat volatilitas pasar *Bitcoin* dan tantangan dalam memprediksi pergerakan harga secara akurat dalam jangka pendek, penelitian ini bertujuan untuk mengevaluasi berbagai model pembelajaran mesin dan menentukan model mana yang paling efektif dalam memprediksi pergerakan pasar *Bitcoin* dalam jangka waktu yang singkat. Hasil penelitian menunjukkan bahwa model RNN (terutama LSTM dan GRU) serta model GBC adalah yang paling efektif dalam memprediksi pergerakan pasar *Bitcoin* dalam jangka pendek, dengan akurasi prediksi yang lebih tinggi dibandingkan model lainnya. Secara khusus, LSTM menunjukkan performa terbaik untuk prediksi 1 menit, sementara GBC unggul dalam prediksi 5 menit. Meskipun demikian, hasil trading berdasarkan prediksi ini menunjukkan bahwa meskipun prediksi dapat memberikan keuntungan sebelum biaya transaksi, setelah memperhitungkan biaya transaksi, hasilnya menjadi negatif (Jaquart, Dann, and Weinhardt 2021).

Penelitian berikutnya bertujuan memprediksi tren harga *Bitcoin* dengan menggabungkan analisis sentimen dari media sosial dan indikator teknis. Dataset terdiri dari lebih dari 26 juta komentar dan 26 ribu berita antara 2017-2022. Enam model *Machine Learning* diuji, dengan XGBoost memberikan akurasi tertinggi sebesar 90,57%. Studi ini menunjukkan bahwa kombinasi indikator teknis dan sentimen meningkatkan prediksi harga *Bitcoin* (Jung et al. 2023)

Penelitian lainnya mengembangkan indikator perdagangan *Bitcoin* berbasis *Machine Learning* dengan menggunakan data harga *Bitcoin* dari yfinance dan data sentimen dari Twitter. Metode yang digunakan meliputi analisis sentimen dengan TextBlob dan prediksi harga dengan FBProphet. Hasil menunjukkan bahwa indikator ini mampu menghasilkan keuntungan yang signifikan dengan RMSE yang rendah, menjadikannya alat yang efektif untuk perdagangan *Bitcoin* di pasar *Cryptocurrency* (Rahaman et al. 2023).

Penelitian terkait memprediksi *Crash* pasar *Bitcoin* yang sering terjadi secara mendadak akibat volatilitas tinggi. Meskipun berbagai model statistik dan pembelajaran mesin telah dikembangkan untuk memprediksi harga *Bitcoin*, tantangan utama terletak pada akurasi dalam memprediksi event *Crash*, terutama karena fenomena gelembung harga yang sering muncul sebelum *Crash*. Oleh karena itu, penelitian ini bertujuan untuk meningkatkan akurasi prediksi *Crash* pasar *Bitcoin* dengan menggabungkan analisis fenomena gelembung menggunakan uji *Generalized Supremum Augmented Dickey-Fuller (GSADF)* dengan berbagai model pembelajaran mesin. Hasil penelitian menunjukkan bahwa pendekatan hibrida yang menggabungkan uji GSADF dengan model pembelajaran mesin secara signifikan meningkatkan akurasi prediksi *Crash* pasar *Bitcoin*. Penggunaan indeks gelembung yang dihasilkan dari uji GSADF dalam model pembelajaran mesin menghasilkan performa prediksi yang lebih baik, terutama untuk prediksi jangka menengah hingga panjang (7 dan 14 hari). Peningkatan signifikan tercapai dalam akurasi prediksi dibandingkan model yang tidak menggunakan indeks gelembung (Park and Yang 2024).

2.1.9 Prediksi *Cryptocurrency* menggunakan *Deep Learning*

Penelitian sebelumnya untuk membandingkan berbagai model *Deep Learning* seperti DNN, LSTM, CNN, dan ResNet untuk memprediksi harga *Bitcoin* menggunakan data harian dari 2011 hingga 2018. LSTM terbukti unggul dalam prediksi regresi, sedangkan DNN lebih baik untuk klasifikasi. Metrik yang digunakan meliputi MAPE untuk regresi dan akurasi untuk klasifikasi. Kesimpulannya, meskipun semua model memiliki performa yang sebanding, klasifikasi memberikan hasil yang lebih baik untuk strategi perdagangan (Ji, Kim, and Im 2019).

Pada tahun 2020 terdapat penelitian untuk memprediksi harga Litecoin dan Monero menggunakan dataset dari tahun 2015 hingga 2020. Model hybrid LSTM-GRU. Data uji menggunakan data harga harian yang telah dinormalisasi. Hasil menunjukkan bahwa model ini memberikan prediksi yang lebih akurat

dibandingkan model LSTM standar, terutama untuk jendela prediksi 1 hari dan 3 hari (Patel et al. 2020).

Pada tahun 2021 terdapat Penelitian yang bertujuan untuk meningkatkan akurasi prediksi harga Litecoin dan Zcash dengan menggunakan model hibrida LSTM-GRU yang mempertimbangkan arah dari *Bitcoin* sebagai parent coin. Dataset yang digunakan mencakup data harga harian yang telah dinormalisasi, dan hasilnya menunjukkan bahwa model ini berhasil memberikan prediksi yang lebih akurat dibandingkan dengan model LSTM dan GRU standar, terutama untuk jendela prediksi 1 hari dan 3 hari (Tanwar et al. 2021).

Pada tahun 2022 terdapat penelitian untuk mengembangkan model hybrid DL-GuesS untuk memprediksi harga *Cryptocurrency* dengan menggabungkan data harga historis dan analisis sentimen dari tweet. Dataset mencakup data harga harian dari Dash, *Bitcoin*, dan Litecoin, serta sentimen yang diambil dari Twitter. Model ini menggunakan kombinasi LSTM dan GRU serta analisis sentimen dengan VADER. Hasilnya menunjukkan bahwa DL-GuesS lebih akurat dibandingkan dengan model prediksi tradisional, dengan MSE lebih rendah (Parekh et al. 2022). Pada tahun yang sama terdapat penelitian yang bertujuan untuk meningkatkan akurasi klasifikasi pergerakan harga *Bitcoin* dan Ethereum dengan menggunakan kombinasi indikator teknikal, trading, dan sosial media dari Reddit dan GitHub. Dataset yang digunakan mencakup data harga per jam dari 2017 hingga 2021. Model *Deep Learning* yang digunakan termasuk MLP, CNN, LSTM, dan MALSTM-FCN. Hasilnya menunjukkan bahwa model unrestricted dengan kombinasi indikator tersebut memberikan akurasi yang lebih tinggi dibandingkan model yang hanya menggunakan indikator teknikal (Ortu et al. 2022).

Pada tahun 2023, terdapat penelitian Penelitian yang bertujuan untuk menguji potensi penggunaan informasi teknikal sebagai input untuk model *Deep Learning* dalam menghasilkan sinyal perdagangan di pasar *Cryptocurrency*, khususnya *Bitcoin*. Dataset yang digunakan adalah data historis perdagangan *Bitcoin* dari *Binance* antara 2017-2022 yang telah diproses dan dipolakan menggunakan berbagai indikator teknikal. Metode yang digunakan termasuk Logistic

Regression, XGBoost, LightGBM, MLP, GRU, LSTM, dan CNN. Hasil utama menunjukkan bahwa model GRU memberikan hasil terbaik dengan precision 58,36% dan Sharpe ratio 1,85, mengungguli strategi Buy-and-Hold (Goutte et al. 2023).

2.1.10 Prediksi *Cryptocurrency* menggunakan LSTM

Model *Long Short-Term Memory* (LSTM) yang termasuk dalam Model *Deep Learning* akan banyak digunakan pada penelitian ini, sehingga dibahas tersendiri. *Long Short-Term Memory* (LSTM) adalah jenis jaringan saraf tiruan (*Artificial Neural Network*) yang dirancang khusus untuk memproses dan menganalisis data *Time Series* atau data dengan ketergantungan temporal jangka panjang. LSTM adalah varian dari *Recurrent Neural Network* (RNN), yang dilengkapi dengan mekanisme untuk mengatasi masalah *vanishing gradient*, yang sering terjadi pada RNN standar ketika mencoba belajar dari data dengan ketergantungan waktu yang panjang. Beberapa Penelitian terdahulu terkait prediksi *Cryptocurrency* menggunakan LSTM dapat dilihat pada bagian Lampiran 3 Tabel 6.3.

LSTM terdiri dari unit dasar yang disebut *LSTM cell*, yang memiliki tiga gerbang utama untuk mengatur informasi yang mengalir melalui sel tersebut. Ketiga gerbang ini adalah:

1. *Forget Gate*: Gerbang ini menentukan informasi mana yang akan dihapus dari sel. Input ke gerbang ini adalah keluaran dari sel sebelumnya (*hidden state*) dan input saat ini, yang kemudian dikombinasikan dengan fungsi sigmoid untuk menghasilkan nilai antara 0 dan 1. Nilai 0 berarti informasi sepenuhnya dilupakan, sedangkan nilai 1 berarti informasi sepenuhnya disimpan.
2. *Input Gate*: Gerbang ini menentukan informasi baru yang akan disimpan dalam sel. Input gate juga memiliki dua bagian utama:
 - a. *Candidate Memory Cell*: Menghitung nilai-nilai baru yang berpotensi ditambahkan ke sel, menggunakan fungsi tanh.

- b. *Input Gate Layer*: Memutuskan informasi mana dari *candidate memory cell* yang akan dimasukkan ke sel dengan mengalikan fungsi *sigmoid* pada nilai tersebut.
3. *Output Gate*: Gerbang ini memutuskan bagian dari *cell state* yang akan dikeluarkan ke *hidden state* berikutnya. Output dari sel saat ini dihitung dengan mengalikan nilai dari *cell state* yang telah dimodifikasi dengan fungsi *tanh*, dan output gate layer yang memiliki fungsi *sigmoid*.
4. *Cell State*: Ini adalah inti dari LSTM yang menyimpan informasi yang mengalir melalui jaringan dengan gangguan minimal. Cell state diperbarui berdasarkan informasi dari forget gate dan input gate.

LSTM memiliki mekanisme unik dalam menangani dependensi jangka panjang. LSTM dirancang untuk mempertahankan informasi yang relevan dalam jangka waktu yang panjang sambil membuang informasi yang tidak relevan. Dengan adanya tiga gerbang utama ini, LSTM mampu:

1. *Mempertahankan Informasi Penting*: Forget gate memungkinkan jaringan untuk membuang informasi yang tidak relevan dari *cell state*.
2. *Menambah Informasi Baru*: Input gate menentukan informasi baru yang akan ditambahkan ke *cell state*, sesuai dengan data baru yang diterima.
3. *Mengeluarkan Informasi Relevan*: Output gate mengontrol seberapa banyak informasi dari *cell state* yang perlu diteruskan ke *hidden state* berikutnya, memungkinkan jaringan untuk menyesuaikan output berdasarkan data terkini.

LSTM sangat sesuai untuk diterapkan dalam Prediksi *Time Series Cryptocurrency*. LSTM sangat berguna dalam memprediksi harga *Cryptocurrency* karena kemampuannya dalam menangani data yang memiliki ketergantungan waktu yang panjang, fluktuasi harga yang signifikan, dan pola yang mungkin berulang. Berikut adalah beberapa langkah yang biasa dilakukan:

1. *Pra-Pemrosesan Data*: Langkah ini meliputi normalisasi data untuk memastikan semua nilai berada dalam rentang yang sama. Langkah ini membantu dalam mempercepat proses pelatihan model.

2. Pelatihan Model: Data dilatih dalam rentang waktu tertentu (misalnya, beberapa minggu atau bulan), di mana LSTM akan mengidentifikasi pola dan belajar dari fluktuasi harga *Cryptocurrency*.
3. Evaluasi Model: Metrik seperti Mean Absolute Error (MAE) atau Root Mean Square Error (RMSE) digunakan untuk mengukur akurasi model.
4. Prediksi: Model yang telah dilatih dapat digunakan untuk memprediksi harga di masa depan, baik dalam jangka pendek (misalnya, hari berikutnya) atau jangka panjang (minggu atau bulan mendatang).

Kelebihan Model LSTM antara lain:

1. Kemampuan Menangani Dependensi Jangka Panjang: LSTM efektif dalam mengingat informasi yang relevan untuk periode waktu yang panjang.
2. Mengatasi *Vanishing Gradient Problem*: Struktur LSTM yang kompleks mengatasi masalah yang sering dihadapi oleh RNN tradisional.

Kekurangan Model LSTM antara lain:

1. Komputasi yang Lambat: Karena kompleksitas arsitektur, LSTM bisa memakan waktu lebih lama dalam pelatihan dibandingkan dengan model sederhana.
2. Kebutuhan Data yang Besar: LSTM membutuhkan jumlah data yang cukup banyak untuk pelatihan, agar bisa mengenali pola dengan baik.

LSTM adalah alat yang sangat berguna dalam prediksi data *Time Series*, terutama dalam konteks *Cryptocurrency* yang memerlukan pemahaman ketergantungan data jangka panjang. Namun, untuk performa optimal, diperlukan tuning parameter yang baik dan data yang berkualitas tinggi.

2.1.11 Prediksi *Cryptocurrency* menggunakan *Technical Analysis*

Penelitian terkait *Technical Analysis* tahun 2019 bertujuan memprediksi return *Bitcoin* menggunakan data harga dari tahun 2010 hingga 2018. Data diolah menggunakan lima indikator teknikal seperti momentum dan *Moving Average*. Regresi prediktif diterapkan untuk menghasilkan prediksi return. Hasil menunjukkan bahwa kombinasi dari semua indikator teknikal memberikan

prediksi yang lebih akurat dibandingkan dengan strategi buy-and-hold atau benchmark historis (L. Liu 2019). Penelitian lainnya tahun 2019 bertujuan meningkatkan prediksi performa strategi trading teknikal dengan menggunakan data aset harian S&P 500 dari tahun 2014 hingga 2015. Metode yang digunakan adalah *Long Short-Term Memory* (LSTM) Neural Network yang dilatih dengan dataset yang diformulasikan dari SMA, RSI, dan MACD. Hasilnya menunjukkan bahwa prediksi berbasis LSTM dapat meningkatkan keuntungan kumulatif dibandingkan dengan strategi teknikal saja (Sang and Pierro 2019). Penelitian lainnya mencoba memprediksi return harian *Bitcoin* menggunakan data harga dari 2012 hingga 2017. Dataset diformulasikan dengan 124 indikator teknikal. Metode yang digunakan adalah pohon keputusan berbasis CART, dengan hasil menunjukkan prediksi rentang return harian yang lebih akurat dibandingkan strategi buy-and-hold (Huang, Huang, and Ni 2019).

Penelitian terkait pada tahun 2020 bertujuan untuk memprediksi tren harga enam *Cryptocurrency* (*Bitcoin*, Dash, Ether, Litecoin, Monero, Ripple) terhadap USD dalam interval satu menit menggunakan model *Deep Learning*. Dataset yang digunakan mencakup data harga dari Q3 2018 hingga Q2 2019 yang diproses menggunakan 18 indikator teknis. Berbagai model neural network, termasuk CNN, CLSTM, MLP, dan RBFNN, digunakan untuk prediksi. Hasil menunjukkan bahwa CLSTM memberikan akurasi prediksi terbaik, terutama untuk *Bitcoin*, Ether, dan Litecoin. Penelitian lainnya pada tahun yang sama bertujuan untuk menguji apakah aturan perdagangan teknikal bisa menghasilkan keuntungan yang lebih besar daripada strategi buy-and-hold di pasar *Bitcoin*. Dataset yang digunakan adalah harga harian *Bitcoin* dari tahun 2010 hingga 2018, dengan tujuh aturan teknikal diterapkan, termasuk *Moving Averages* dan trading range breakout. Metode yang digunakan adalah evaluasi berdasarkan rasio Sharpe melalui bootstrapping. Hasilnya menunjukkan bahwa aturan trading range breakout secara konsisten lebih unggul dari buy-and-hold dalam pasar yang sedang trending (Gerritsen et al. 2020).

Penelitian Terkait pada tahun 2021 mengkaji penggunaan deep neural networks untuk memprediksi arah harga aset dengan menggunakan 124 indikator

teknikal yang dipilih melalui tiga metode seleksi fitur. Dataset yang digunakan mencakup data harian dari tujuh indeks pasar global selama periode 2008-2019. Hasil penelitian menunjukkan bahwa akurasi model berkisar antara 50% hingga 65%, tetapi strategi perdagangan yang dihasilkan tidak dapat mengungguli strategi Buy-and-Hold (Y. Peng et al. 2021). Penelitian pada tahun yang sama bertujuan memprediksi return harian *Bitcoin* menggunakan analisis teknikal dan model non-linear seperti Random Forest dan SVM, dengan dataset dari 2015 hingga 2019. Hasil menunjukkan bahwa model Random Forest memberikan prediksi paling akurat, dibandingkan model lainnya (Gradojevic et al. 2021).

Penelitian tahun 2022 menggunakan kombinasi indikator teknikal, perdagangan, dan media sosial untuk memprediksi pergerakan harga harian *Bitcoin* dan *Ethereum*. Data diambil dari 1 Januari 2017 hingga 1 Januari 2021 dan dianalisis menggunakan berbagai algoritma *Deep Learning* seperti MLP, LSTM, CNN, dan MALSTM-FCN. Hasilnya menunjukkan bahwa menggabungkan indikator sosial dan perdagangan secara signifikan meningkatkan akurasi klasifikasi, dengan peningkatan akurasi dari 51%-55% menjadi 67%-84% (Ortu et al. 2022). Penelitian terkait pada tahun yang sama bertujuan untuk menguji apakah *Cryptocurrency*, terutama *Bitcoin* dan *Ethereum*, dapat dianalisis menggunakan metode teknikal yang biasa digunakan di pasar aset atau forex. Data dari platform trading digunakan bersama alat teknikal seperti candlestick pattern, MACD, RSI, dan *Bollinger Bands*. Hasil menunjukkan Indikator teknikal yang digunakan dapat memberikan sinyal yang cukup akurat untuk mengambil keputusan trading di pasar *Cryptocurrency*, menghasilkan keuntungan signifikan jika diikuti dengan benar (Jain, Bhardwaj, and Soni 2022).

Penelitian tahun 2023 menggunakan data perdagangan *Bitcoin* dari *Binance* (2017–2022) dan menerapkan algoritma seperti regresi logistik, XGBoost, LightGBM, serta model jaringan saraf MLP, GRU, LSTM, dan CNN. Dataset dibagi menjadi pelatihan, validasi, dan pengujian, dengan evaluasi menggunakan precision, recall, Sharpe ratio, dan rata-rata pengembalian. Hasilnya, model GRU menunjukkan kinerja terbaik dengan precision 58,39% dan Sharpe ratio 1,85. Penggunaan indikator teknikal tambahan tidak meningkatkan akurasi secara

signifikan. GRU secara konsisten mengungguli model lain, sementara CNN tidak menunjukkan hasil optimal dalam prediksi *Time Series*. (Goutte et al. 2023). Pada tahun yang sama Penelitian tahun 2023 menggunakan data harga aset Indonesia (IDX) dari Yahoo Finance, meliputi harga penutupan, tertinggi, terendah, volume, serta indikator teknikal (SMA, EMA, SMMA, LWMA) dengan jendela waktu 7 dan 30 hari. Model Multilayer Perceptron (MLP) dibangun dan dioptimalkan menggunakan Grid Search. Prediksi harga aset tujuh hari ke depan dibandingkan dengan harga aktual untuk mengevaluasi kinerja model. Hasilnya menunjukkan akurasi tinggi, dengan nilai R^2 sebesar 0,995 dan RMSE 657,18. Prediksi cenderung tertinggal di belakang harga aktual, dan diperkenalkan metrik MEMPR untuk analisis toleransi kesalahan prediksi (Julian et al. 2023). Penelitian lainnya bertujuan untuk memprediksi harga *Bitcoin* menggunakan analisis teknikal. Dataset terdiri dari data harga *Bitcoin* per jam dan per hari dari 2015-2019 yang diformulasikan menjadi indikator teknikal. Penulis menggunakan model non-linear seperti Random Forest, ANN, dan SVM. Hasilnya, Random Forest mengungguli model lainnya dalam prediksi data harian, sementara data frekuensi jam menunjukkan efisiensi pasar yang lemah. (Gradojevic et al. 2023).

Penelitian Tahun 2024 Penelitian terkait mengembangkan model TechSupportLSTM yang menggabungkan LSTM dengan SVM/SVR untuk memprediksi pergerakan harga dan tren *Cryptocurrency* menggunakan 24 indikator teknikal. Dataset diambil dari data harian *Bitcoin*, Ethereum, Litecoin, dan Ripple Credits. Hasilnya menunjukkan peningkatan akurasi prediksi dan profitabilitas dibandingkan metode lain, serta strategi trading dengan unsur acak yang lebih efektif dibandingkan strategi *buy-and-hold* (Y. H. Liu and Huang 2024). Penelitian pada tahun yang sama bertujuan untuk menguji apakah aturan perdagangan teknis sederhana dapat mengungguli strategi *buy-and-hold* di pasar *Bitcoin* setelah memperhitungkan biaya transaksi dan perilaku investor yang realistis. Dataset yang digunakan mencakup data transaksi *tick-by-tick* BTC-USD dari *Bitstamp* antara 2012-2022. Metodologi melibatkan pengujian 75,360 aturan perdagangan dari berbagai frekuensi perdagangan menggunakan uji *multiple hypothesis* untuk menghindari hasil bias dari *Data Mining*. Hasil menunjukkan

bahwa beberapa aturan, terutama pada frekuensi intraday, mampu mengungguli strategi *buy-and-hold* dalam hal *risk-return* (Deprez and Frömmel 2024).

2.1.12 Prediksi *Cryptocurrency* menggunakan *Descriptive Statistics*

Penelitian terkait penggunaan Statistik Deskriptive dalam Prediksi *Cryptocurrency* pada tahun 2023 menyimpulkan bahwa penggunaan dataset berpola dengan fungsi minimum dan maksimum merupakan metode yang efektif untuk memprediksi pergerakan harga Bitcoin dan Ethereum. Pola yang dihasilkan dapat membantu trader dalam mengambil keputusan perdagangan yang lebih baik dengan memanfaatkan titik-titik harga tertinggi dan terendah yang diidentifikasi. Metode ini juga memberikan alternatif yang lebih sederhana dan cepat dibandingkan dengan teknik prediksi yang lebih kompleks (Rizky Parlika, Mustafid 2023).

Pada tahun 2024 terkait tema yang sama, hasil penelitian menyimpulkan bahwa penggunaan fungsi rata-rata (Average) pada dataset terpola adalah metode yang efektif untuk memprediksi fluktuasi harga Bitcoin. Pola yang dihasilkan dari fungsi ini memungkinkan prediksi yang lebih akurat dan dapat membantu dalam mengurangi risiko serta meningkatkan potensi keuntungan dalam perdagangan *Cryptocurrency*. Penelitian ini juga membuka peluang untuk pengembangan lebih lanjut dalam penerapan dataset terpola pada berbagai jenis *Cryptocurrency* lainnya (Rizky Parlika, Mustafid, and Rahmat 2024a).

2.1.13 *Return on Investment (ROI)* dalam Prediksi *Cryptocurrency*

Penelitian terkait penggunaan perhitungan Return on Investment (ROI) pada tahun 2021 bertujuan untuk mengembangkan sistem perdagangan otomatis di pasar Forex dengan menggunakan kombinasi EmcSVM dan fuzzy NSGA-II untuk memprediksi tren pasar. Dataset terdiri dari pasangan mata uang EUR/USD dari 2014 hingga 2019. Hasil menunjukkan peningkatan ROI dan pengurangan *drawdown* yang signifikan dibandingkan metode sebelumnya (Sadeghi, Daneshvar, and Madanchi 2021). Penelitian terkait pada tahun yang sama mengkaji kemampuan belajar pedagang mata uang *Cryptocurrency* dalam

lingkungan yang ambigu, menggunakan dataset transaksi dari CryptEx antara 2016 dan 2019. Metode yang digunakan meliputi dekomposisi profit dan regresi panel untuk mengevaluasi performa berdasarkan trading dan komponen directional. Hasilnya menunjukkan bahwa meskipun pedagang tidak secara signifikan meningkatkan pengembalian, mereka dapat beradaptasi dan merealisasikan profit lebih baik seiring waktu (Gemayel and Preda 2021).

2.1.14 Penentuan kondisi Kenaikan (*Bullish*) dan Penurunan (*Bearish*) dalam Prediksi *Cryptocurrency*

Penelitian terkait kondisi Kenaikan dan Penurunan dalam Prediksi *Cryptocurrency* pada tahun 2020 11 mengkaji dampak kecenderungan *Bullish-Bearish* pengguna di komunitas finansial online terhadap volatilitas dan pengembalian pasar aset di China. Dataset terdiri dari lebih dari 5 juta komentar yang diproses menggunakan CNN. Hasilnya menunjukkan bahwa kecenderungan *Bearish* pengguna meningkatkan volatilitas pasar dan pengembalian yang lebih tinggi, dengan konsistensi kecenderungan yang berdampak positif pada volatilitas (Qian, Li, and Yuan 2020).

Penelitian terkait pada tahun 2021 bertujuan untuk mengeksplorasi bagaimana volatilitas *Bitcoin*, baik upside maupun downside, mempengaruhi reaksi pasar aset internasional dengan menggunakan analisis kuantil. Dataset yang digunakan terdiri dari data harga *Bitcoin* harian dan data pasar aset dari 11 pasar aset internasional selama periode 2014 hingga 2020. Metode kuantil-regresi diterapkan untuk menganalisis hubungan tersebut. Hasilnya menunjukkan bahwa volatilitas downside dari *Bitcoin* memiliki dampak lebih besar pada kuantil return pasar aset yang lebih rendah, sementara volatilitas upside lebih mempengaruhi kuantil yang lebih tinggi (Ahmed 2021). Penelitian terkait pada tahun yang sama mendeteksi skema pump-and-dump dalam *Cryptocurrency* dengan menggunakan sinyal pasar dan sosial. Dataset terdiri dari 355 *Cryptocurrency* antara 2018-2019. CNN dan BLSTM digunakan untuk mendeteksi target koin dan memprediksi harga puncak. Model terbaik mencapai F1-score tertinggi untuk prediksi koin target dan MAPE 6-11% untuk prediksi harga puncak pump (Nghiem et al. 2021).

Penelitian terkait pada tahun 2022 menganalisis prediksi ahli *Cryptocurrency* untuk menentukan apakah prediksi tersebut memberikan keuntungan bagi investor *Bitcoin*. Dataset terdiri dari 222 prediksi antara 2012-2020. Menggunakan metode cumulative abnormal returns (CAR), ditemukan bahwa prediksi *Bearish* dan *neutral* menghasilkan return negatif yang signifikan, sementara prediksi *Bullish* tidak signifikan. Hasil ini menunjukkan bahwa ahli *Cryptocurrency* dapat membantu efisiensi pasar, khususnya dengan prediksi non-*Bullish* (Gerritsen, Lugtigheid, and Walther 2022).

Penelitian terkait pada tahun 2024 menentukan apakah risiko penurunan dihargai dalam pasar *Cryptocurrency*. Dataset mencakup sekitar 2000 *Cryptocurrency* dari 2014 hingga 2021. Metode yang digunakan adalah model multifactor pricing dengan penambahan risiko penurunan. Hasil utama menunjukkan bahwa risiko penurunan dihargai dengan premi yang signifikan, memperkuat kekuatan penjelasan model (Dobrynskaya 2024). Penelitian pada tahun yang sama bertujuan untuk mengungkap efek asimetris antara risiko naik dan turun di pasar *Cryptocurrency* dengan menggunakan data volatilitas *Cryptocurrency* frekuensi tinggi dari 2020 hingga 2024, serta data dari krisis LUNA dan FTX. Metode TVP-VAR digunakan untuk memetakan jaringan spillover risiko. Hasilnya menunjukkan bahwa volatilitas buruk lebih sering menyebar, tetapi volatilitas baik yang sering memicu krisis pada periode risiko tinggi (Aibai, Julaiti, and Gou 2024). Penelitian yang lain dengan tema yang sama menggunakan analisis event study yang menggabungkan grafik dan regresi panel untuk mengukur hubungan antara aktivitas Twitter dan pergerakan harga selama acara pump-and-dump. Hasil menunjukkan bahwa promosi di Twitter meningkatkan perhatian terhadap mata uang *Cryptocurrency* yang dipromosikan sebelum acara, tetapi menyebabkan investor yang mengandalkan informasi dari Twitter mengalami kerugian karena keterlambatan dalam menjual (Ardia and Bluteau 2024).

2.1.15 Implementasi Data *Time Series* dalam memprediksi *Cryptocurrency*

Data Time Series adalah serangkaian data yang diambil atau dicatat dalam interval waktu tertentu secara berurutan. Dalam konteks prediksi harga *Cryptocurrency*, *data Time Series* mencakup harga *Cryptocurrency* yang dicatat pada interval waktu tertentu, seperti harian, per jam, atau per menit. Karena harga *Cryptocurrency* berfluktuasi secara dinamis, analisis *Time Series* menjadi salah satu metode yang efektif untuk memahami tren, pola musiman, dan perilaku data yang berulang. Beberapa Penelitian pendahuluan terkait Pengelolaan Data *Time Series* dalam memprediksi harga *Cryptocurrency* dapat dilihat pada bagian Lampiran 2 pada Tabel 6.2.

Data Time Series memiliki tiga komponen utama yang perlu diperhatikan dalam konteks analisis harga *Cryptocurrency*:

1. Tren: Pergerakan harga yang menunjukkan kecenderungan naik atau turun dalam jangka panjang.
2. Musiman: Pola berulang dalam data yang terjadi secara periodik, seperti pola harian atau mingguan.
3. Siklus: Perubahan yang berulang namun tidak bersifat tetap atau reguler, biasanya terkait dengan perubahan kondisi ekonomi atau sentimen pasar.

Pengelolaan data *Time Series* melibatkan beberapa langkah utama untuk menyiapkan, memproses, dan menganalisis data tersebut. Langkah-langkah ini termasuk:

1. Pengumpulan Data: Data harga *Cryptocurrency* bisa diperoleh dari berbagai sumber, seperti API platform perdagangan (misalnya, *Binance*, *CoinGecko*, atau *Coinmarketcap*). Data yang diambil biasanya mencakup harga pembukaan, penutupan, tertinggi, terendah, dan volume perdagangan.
2. Pra-Pemrosesan Data: Tahap ini melibatkan pembersihan data dari anomali atau nilai hilang (*missing values*). Selain itu, data dapat di-*resample* agar sesuai dengan interval waktu tertentu, seperti mengubah data per menit menjadi data harian. Normalisasi atau standarisasi data juga dilakukan untuk memperlancar proses analisis.

3. Visualisasi Data: Langkah ini berguna untuk mengidentifikasi pola tren atau musiman yang mungkin ada pada data. Visualisasi umum meliputi plot garis untuk melihat tren, serta plot korelasi atau *autocorrelation* untuk mengeksplorasi ketergantungan data terhadap waktu.
4. Pemilihan dan Penggunaan Model: Model yang dipilih bergantung pada pola dan karakteristik data. Beberapa model yang umum digunakan dalam prediksi harga *Cryptocurrency* meliputi:
 - a. *ARIMA (Auto-Regressive Integrated Moving Average)*: Digunakan ketika data menunjukkan hubungan linier. ARIMA cocok untuk data yang memiliki pola tren dan musiman yang bisa diidentifikasi.
 - b. *LSTM (Long Short-Term Memory)*: Salah satu jenis RNN yang populer untuk data *Time Series* yang kompleks. LSTM bisa menangani pola jangka panjang serta ketergantungan temporal, yang umum pada harga *Cryptocurrency*.
5. Validasi dan Evaluasi Model: Untuk mengevaluasi performa model, data biasanya dibagi menjadi data latih (*training set*) dan data uji (*testing set*). Beberapa metrik yang digunakan untuk mengukur akurasi model prediksi adalah *Mean Absolute Error (MAE)*, *Root Mean Square Error (RMSE)*, atau *Mean Absolute Percentage Error (MAPE)*.
6. Implementasi dan Pemantauan: Setelah model dipilih dan dievaluasi, model bisa digunakan untuk membuat prediksi harga di masa depan. Prediksi ini kemudian bisa digunakan untuk keputusan perdagangan atau strategi investasi.

Tantangan dalam prediksi harga *Cryptocurrency* menggunakan Data *Time Series*

1. Volatilitas Tinggi: Harga *Cryptocurrency* sering kali berubah secara drastis dalam waktu singkat, sehingga mempersulit pembuatan model yang akurat.
2. Ketergantungan pada Faktor Eksternal: Harga *Cryptocurrency* juga dipengaruhi oleh berita, regulasi, dan sentimen pasar, yang tidak selalu tercermin pada data historis.

3. Overfitting: Karena fluktuasi yang ekstrem, ada risiko model menjadi terlalu menyesuaikan (*overfit*) terhadap data latih dan tidak bekerja baik pada data baru.

Secara keseluruhan, data *Time Series* merupakan alat yang kuat untuk menganalisis dan memprediksi harga *Cryptocurrency*. Namun, untuk mendapatkan hasil yang optimal, penting untuk memilih model yang tepat dan memperhitungkan faktor volatilitas serta ketidakpastian pasar.

2.1.16 Perbedaan penggunaan Dataset pada Penelitian *Cryptocurrency*

Terdapat perbedaan mendasar antara sumber data awal yang digunakan pada Model Dataset Terpola dengan sumber data awal pada penelitian-penelitian terdahulu dalam kaitannya dengan prediksi harga *Bitcoin* dan aset *Cryptocurrency* lainnya dalam suatu pasar perdagangan *Cryptocurrency*. Apakah data-data yang digunakan selama ini sebagai dataset pada penelitian-penelitian terdahulu sudah tepat formulasi maupun variasinya, ataukah bila dataset yang dibangun diolah dengan diprogram terlebih dahulu, berpotensi memberikan hasil akurasi prediksi yang lebih baik juga menjadi pertanyaan penelitian yang akan dijawab.

Dari berbagai penelitian yang dijelaskan pada tinjauan Pustaka serta pada Tabel 6.1 hingga 6.3 pada bagian lampiran, tentang perbedaan penggunaan dataset pada penelitian-penelitian terdahulu, tampak pada penelitian dengan karakteristik yang sama, sering peneliti sebelumnya menggunakan data-data seperti *High*, *Low*, *Last*, dan *Current (HLOC)* yang secara langsung digunakan dalam periode tertentu, tanpa dipolakan terlebih dahulu. Pada penelitian ini, diduga ketika suatu dataset dipolakan terlebih dahulu akan memberikan cara pembacaan posisi harga dan pergerakan harga transaksi *Cryptocurrency* dengan lebih baik dalam memprediksi pergerakan harga *Cryptocurrency*, fluktuasinya serta volatilitasnya dibandingkan dengan melakukan pembacaan pada grafik pasar secara langsung maupun dari penyedia data mentah pihak ketiga seperti *Coinmarketcap*, yahoo finance, dan investing.com.

Penelitian pendahuluan terkait perbedaan penggunaan dataset pada penelitian *Cryptocurrency* dapat dilihat pada bagian Lampiran 1 Tabel 6.1.

2.1.17 Perbedaan Perdagangan *Cryptocurrency* dengan Perdagangan Saham

Perdagangan *Cryptocurrency* memiliki beberapa perbedaan dengan Perdagangan Saham. Perbedaan mendasar antara lain:

1. Sifat Aset

Cryptocurrency adalah aset digital yang berfungsi sebagai alat pembayaran, investasi, dan teknologi blockchain. Saham adalah bukti kepemilikan suatu perusahaan, memberikan hak atas keuntungan (dividen) dan suara dalam keputusan perusahaan.

2. Volatilitas

Harga *Cryptocurrency* lebih tidak stabil dibandingkan saham, sering dipengaruhi oleh sentimen pasar dan spekulasi. Harga saham cenderung lebih stabil dibandingkan *Cryptocurrency* karena didukung fundamental perusahaan dan regulasi ketat. *Cryptocurrency* lebih volatil dan spekulatif dibandingkan saham.

3. Regulasi Minim

Cryptocurrency beroperasi dalam ekosistem yang kurang diatur dibandingkan pasar saham, memberikan lebih banyak kebebasan namun risiko yang lebih tinggi. Bursa saham diatur oleh badan pengawas (misalnya OJK di Indonesia), sehingga lebih memberikan perlindungan kepada investor. Pasar saham lebih diatur dibandingkan *Cryptocurrency*, yang cenderung bebas dan kurang transparan dalam regulasi.

4. Waktu Perdagangan

Perdagangan *Cryptocurrency* berlangsung 24/7 tanpa batasan waktu, tidak seperti saham yang memiliki jam bursa tertentu. Perdagangan saham hanya berlangsung pada jam operasional bursa tertentu. Perdagangan *Cryptocurrency* terjadi sepanjang waktu, sementara saham hanya diperdagangkan pada jam bursa.

5. Dividen

Pada Perdagangan *Cryptocurrency* tidak ada pembagian keuntungan seperti dividen pada saham, dan keuntungan murni berasal dari apresiasi harga.

6. Teknologi Blockchain

Pada Perdagangan *Cryptocurrency*, transaksi dicatat di buku besar digital yang terdesentralisasi, sehingga memberikan transparansi tinggi namun lebih rawan serangan teknologi. Pada perdagangan Saham, Histori harga dicatat namun tidak menggunakan sistem Blockchain.

7. Fokus Fundamental

Penilaian saham sering kali didasarkan pada kinerja perusahaan, laporan keuangan, dan proyeksi bisnis. Sedangkan Penilaian Aset *Cryptocurrency* selain didasarkan pada Portofolio Vendor yang mencetak aset kripto, seringkali juga dipengaruhi sentiment terhadap berbagai berita baik dari kanal bereputasi hingga media sosial influencer kelas dunia. Nilai saham didukung oleh fundamental perusahaan, sedangkan nilai *Cryptocurrency* lebih bergantung pada adopsi teknologi dan sentimen pasar.

2.2 Keaslian Penelitian

Karena tujuan dasar perdagangan mata uang *Cryptocurrency* adalah cara mendapatkan titik beli/harga terendah kemudian menjualnya dengan harga lebih tinggi (Rizky Parlika and Atmaja 2018b), maka sebagai kelanjutan penelitian sebelumnya, pada penelitian ini akan dibangun Model Dataset Terpola yang dilandasi dua pertanyaan awal yakni kapan saat yang lebih baik untuk melakukan pembelian *cryptocurrency* pada pasar *Spot* atau bersiap mengambil posisi *Long* pada pasar *Future* (Hipotesis 1), dan kapan saat yang lebih baik untuk melakukan penjualan *cryptocurrency* pada pasar *Spot* atau bersiap mengambil posisi *Short* pada pasar *Future* (Hipotesis 2).

Novelty pada Penelitian ini adalah Model Dataset Terpola yang diformulasikan menjadi hipotesis 1, 2, 3, 4 dan pembuktian masing-masing hipotesis. Sehingga, mampu menghasilkan susunan informasi yang secara periodik bisa digunakan pedagang *Cryptocurrency* baik di level perusahaan maupun level retail sebagai pendukung keputusan sebelum mengambil langkah pembelian ataupun penjualan. Hasil penelitian ini akan menjadi kebaruan pada penelitian pergerakan harga *Bitcoin* khususnya pada sisi penyajian dataset terpola.

Untuk membuktikan Hipotesis 1 dan Hipotesis 2 pada Model Dataset Terpola yang diusulkan, dapat digunakan perbandingan Posisi harga *Bitcoin* dan *Altcoin* pada saat munculnya informasi penanda bahwa harga sudah relatif rendah. Maka, pada saat yang sama harga asli *Bitcoin* dan kebanyakan aset *Cryptocurrency* pada pasar *Indodax* akan mengalami penurunan harga yang signifikan dan begitupula sebaliknya. Penelitian di masa depan, dataset terpola ini akan dikombinasikan dengan berbagai algoritma yang lebih canggih dan Penelitian ini akan menjadi dasar teori pijakan bagi penelitian-penelitian terkait pergerakan harga *Cryptocurrency* di masa depan.

Beberapa Penelitian pendahuluan yang terkait hingga mendekati tujuan pada penelitian disertasi ini telah dirangkum dalam 3 tabel pada bagian Lampiran, yakni Tabel 6.1 yang menjelaskan mengenai Perbedaan penggunaan dataset pada penelitian-penelitian terdahulu, Tabel 6.2 mengenai Penelitian-penelitian terkait pengelolaan data *Time Series* dalam memprediksi harga *Cryptocurrency*, dan Tabel 6.3 yang menjelaskan mengenai Varian Model LSTM dalam memprediksi harga *Cryptocurrency* (*Bitcoin* dan *Altcoin*). LSTM dibuatkan tabel tersendiri karena hasil pengujian Dataset Terpola pada bagian akhir menunjukkan implementasi Model LSTM sebagai upaya optimasi Model Dataset Terpola memberikan hasil terbaik.

2.3 Landasan Teori

Dasar Teori yang digunakan dijelaskan lebih dalam pada sub bagian 2.3 dari 2.3.1 hingga 2.3.7

2.3.1 Formula dasar Model Dataset Terpola

Pada bagian ini akan dijelaskan teori-teori yang digunakan pada Penelitian ini. Parlita dan Atmaja (2018) mengusulkan algoritma sederhana yang disebut rizubot untuk mendapatkan posisi pembelian yang relatif lebih baik. Rizubot diKodekan dari dokumen API yang dibagikan oleh Pasar digital *Cryptocurrency* digital. Terdapat lebih dari 100 Kode akses API dari 200 Pasar digital

Cryptocurrency di seluruh dunia yang masih aktif tahun 2018. Pada penelitian tersebut dihasilkan 6 Formula awal :

$$\text{Highest Price} - \text{Lowest Price} = \text{Range}, \quad (2.1)$$

$$\text{Highest Price} - \text{Current Price} = \text{Top Range}, \quad (2.2)$$

$$\text{Current Price} - \text{Lowest Price} = \text{Lower Range}, \quad (2.3)$$

$$\text{Range} / \text{Lower Range} * 100\% = \text{Percent of Range}, \quad (2.4)$$

$$\text{Percent of Top Range} / \text{Range} * 100\% = \text{Percent of Top Range}, \text{ dan} \quad (2.5)$$

$$\text{Percent of Low Range} / \text{Range} * 100\% = \text{Percent of Lower Range} \quad (2.6)$$

Parlika dan Atmaja (2020), menguji ratusan data API yang dibagikan oleh ratusan Pasar digital *Cryptocurrency* untuk memudahkan pembacaan data yang telah dikumpulkan, Pada penelitian ini juga digunakan bot telegram serta aplikasi pemantauan harga sederhana secara *Realtime* berbasis web. Langkah ini diharapkan dapat membantu dalam upaya memantau pergerakan harga *Cryptocurrency* serta memprediksi pergerakan harga *Cryptocurrency* di masa depan. Pada penelitian tersebut, berhasil diekstraksi data harga *Bitcoin* terbaru dari 20 *Cryptocurrency* Markets melalui Akses API bersama. Data ini kemudian disimpan dalam Database *MySQL* pada hosting secara otomatis. Data ini kemudian dapat ditambah terus menerus menggunakan fitur *Cron Jobs* yang disediakan oleh Layanan hosting yang telah disewa sebelumnya. Percobaan penyimpanan data berhasil dilakukan secara *Realtime* dan Simultan, namun masih belum ada penerapan pelevelan, dan belum ditemukan gagasan untuk membangun Dataset Terpol. Model yang baru ini diharapkan dapat membantu dalam upaya memantau pergerakan harga *Bitcoin* dan memprediksi pergerakan harga kedepannya (Rizky Parlika and Atmaja 2018a).

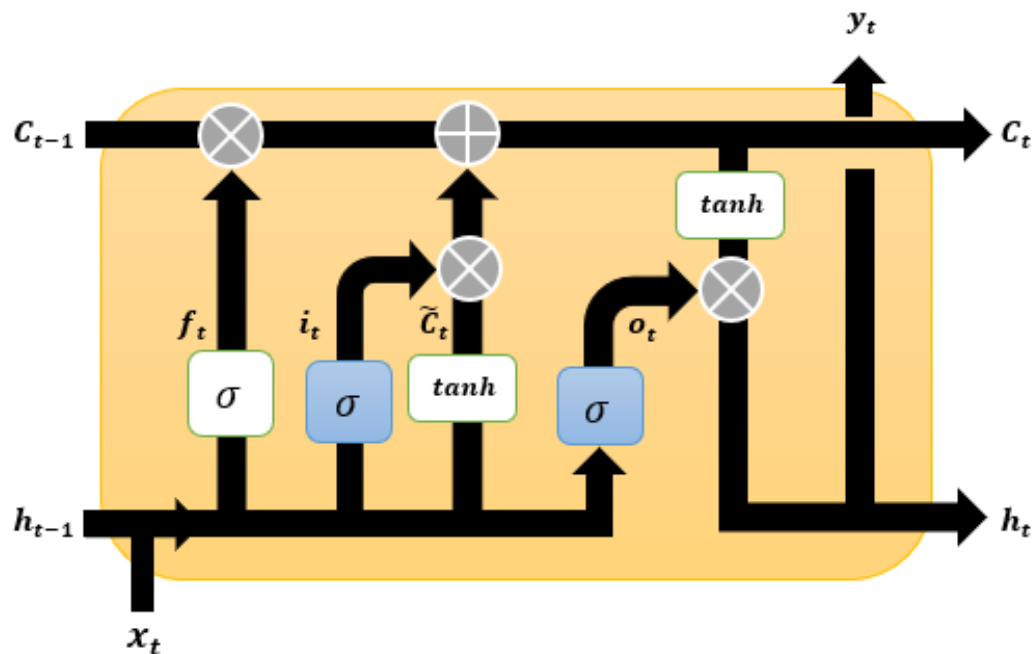
Parlika dan Pratama (2021) menggunakan bahasa pemrograman PHP dan implementasi fungsi CURL untuk pembacaan JSON. Dari hasil penelitian, data pergerakan *Bitcoin* secara real time dapat ditambah secara *realtime* dan simultan. Pada Penelitian tersebut, *Bitcoin* diamati secara khusus karena merupakan komoditas *Cryptocurrency* utama dan penentu pergerakan harga *Altcoin*. Pada penelitian tersebut, belum diterapkan pelevelan, dan belum ada gagasan untuk

membangun Dataset Terpola yang mampu melakukan prediksi fluktuasi dan volatilitas harga *Cryptocurrency*.

Dengan hasil-hasil dari beberapa penelitian pendahuluan, hasil penelitian yang akan dilakukan diharapkan dapat memberikan wawasan baru untuk investor *Cryptocurrency*, pedagang, dan analis keuangan sebagaimana juga pada penelitian lain yang mendekati seperti pada Lampiran 1, Lampiran 2, dan Lampiran 3.

2.3.2 Model *Deep Learning Long Short-Term Memory* (LSTM)

Model LSTM adalah jenis jaringan saraf tiruan yang digunakan dalam bidang pembelajaran mendalam (*Deep Learning*) untuk memproses dan memprediksi data yang memiliki urutan (*sequential data*). Berikut adalah persamaan dasar yang digunakan dalam LSTM



Gambar 2.1 Arsitektur Model LSTM (Yuliyono and Girsang 2019).

Berdasar Gambar 2.1, dijabarkanlah kedalam beberapa persamaan

Forget Gate:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (2.7)$$

Input Gate:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (2.8)$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \quad (2.9)$$

Cell State Update:

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (2.10)$$

Output Gate:

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (2.11)$$

$$h_t = o_t * \tanh(C_t) \quad (2.12)$$

Dengan:

(σ) adalah fungsi aktivasi sigmoid.

($\tanh$) adalah fungsi aktivasi tanh.

(W_f, W_i, W_C, W_o) adalah Level untuk forget gate, input gate, candidate cell state, dan output gate.

(b_f, b_i, b_C, b_o) adalah bias untuk masing-masing gate.

(h_{t-1}) adalah output dari langkah waktu sebelumnya.

(x_t) adalah input pada langkah waktu saat ini.

(C_t) adalah cell state pada langkah waktu saat ini.

(f_t, i_t, o_t) adalah output dari forget gate, input gate, dan output gate.

Beberapa Penelitian terdahulu terkait prediksi *Cryptocurrency* menggunakan LSTM dapat dilihat pada bagian Lampiran 3 Tabel 6.3.

2.3.3 Metrik Evaluasi Kinerja

Pada penelitian sebelumnya, digunakanlah Metrik evaluasi seperti, MAE, MSE, RMSE, dan MAPE saat merancang model untuk memprediksi perbedaan harga *Cryptocurrency* dengan menerapkan pengembangan *Reinforcement Learning* (Shahbazi and Byun 2021). Adapun formula *Mean Absolute Error* (Scikit-learn.org (mae) 2024), *Mean Absolute Percentage Error* (Scikit-learn.org (mape) 2024), *Mean Square Error* (Scikit-learn.org (mse) 2024), dan *Root Mean Square Error* (Scikit-learn.org (rmse) 2024) :

Untuk mengevaluasi hasil penerapan model prediksi dataset terpola secara *Hybrid*, digunakan 4 metrik berikut:

1. MAE (Mean Absolute Error) mengukur rata-rata selisih absolut antara nilai prediksi dan nilai aktual, dengan MAE yang lebih kecil menunjukkan model yang lebih baik (P. L. Seabe, et. Al, 2023).

$$MAE(y, \hat{y}) = \frac{1}{n_{samples}} \sum_{i=0}^{n_{samples}-1} |y_i - \hat{y}_i| \quad (2.13)$$

2. MSE (Mean Squared Error) mengukur rata-rata selisih kuadrat antara nilai prediksi dan nilai aktual, dengan MSE yang lebih kecil menunjukkan model yang lebih baik (P. L. Seabe, et. Al, 2023).

$$MSE(y, \hat{y}) = \frac{1}{n_{samples}} \sum_{i=0}^{n_{samples}-1} (y_i - \hat{y}_i)^2 \quad (2.14)$$

3. RMSE (Root Mean Squared Error) adalah akar kuadrat dari MSE, memberikan ukuran kesalahan dalam satuan yang sama dengan data asli, sehingga lebih mudah diinterpretasikan (P. L. Seabe, et. Al, 2023).

$$RMSE = \sqrt{MSE} \quad (2.15)$$

4. MAPE (Mean Absolute Percentage Error) mengukur rata-rata kesalahan persentase absolut antara nilai prediksi dan nilai aktual, dengan MAPE yang lebih kecil menunjukkan model yang lebih baik (P. L. Seabe, et. Al, 2023).

$$MAPE(y, \hat{y}) = \frac{1}{n_{samples}} \sum_{i=0}^{n_{samples}-1} \frac{|y_i - \hat{y}_i|}{\max(\epsilon, |y_i|)} \quad (2.16)$$

Root Mean Square Error (RMSE) atau Kesalahan kuadrat rata-rata akar atau deviasi kuadrat rata-rata akar adalah salah satu ukuran yang paling umum digunakan untuk mengevaluasi kualitas prediksi. Ukuran ini menunjukkan seberapa jauh perbedaan antara prediksi dengan nilai sebenarnya yang diukur menggunakan jarak *Euclidean*. Untuk menghitung RMSE, dihitung selisih perbedaan antara prediksi dengan fakta yang muncul untuk setiap titik data. *Root mean square error* dapat dinyatakan sebagai:

$$RMSE = \sqrt{\frac{1}{n_{samples}} \sum_{i=0}^{n_{samples}-1} (y_i - \hat{y}_i)^2} \quad (2.17)$$

Dengan N adalah jumlah titik data, $y(i)$ adalah pengukuran ke-i, dan $\hat{y}(i)$ adalah prediksi yang sesuai. RMSE bukan merupakan bentuk invarian dari skala sehingga perbandingan model yang menggunakan ukuran ini dipengaruhi oleh

skala data. Untuk alasan ini, RMSE umumnya digunakan pada data standar (C3.AI, 2023).

Penelitian pendahuluan lainnya menggunakan ukuran RSME pada prediksi harga *Bitcoin* dan mencapainya menggunakan implementasi jaringan saraf rekuren yang dioptimalkan secara *Bayesian* (RNN) dan Jaringan Memori Jangka Pendek (LSTM), serta dibandingkan dengan ARIMA yang merupakan model deret waktu populer (Roche and McNally n.d.). Seperti yang diharapkan, model pembelajaran mendalam (*Deep Learning*) memiliki kinerja yang lebih baik dibandingkan ARIMA dengan akurasi sebesar 52% dan RSME sebesar 8%.

2.3.4 Klasterisasi Dataset Terpola menggunakan *K-Means Clustering*

K-Means Clustering (Scikit-learn.org (KMeans) 2024) merupakan algoritma pembelajaran mesin yang termasuk dalam kategori *unsupervised learning* dan digunakan untuk melakukan analisis Klaster atau pengelompokan data. Algoritma ini bertujuan untuk membagi data ke dalam sejumlah Klaster (*k Clusterings*) berdasarkan kemiripan atau jarak antar data. Tiap Klaster berisi data yang memiliki karakteristik serupa, sementara data antar Klaster akan memiliki perbedaan yang signifikan. *K-Means* bekerja dengan membagi data ke dalam *k* Klaster dengan langkah-langkah:

1. Menentukan Jumlah Klaster (*k*): Sebelum memulai proses Klastering, jumlah Klaster (*k*) yang diinginkan harus ditentukan terlebih dahulu. Ini bisa dilakukan dengan beberapa metode seperti *Elbow Method* atau *Silhouette Score* untuk menentukan nilai *k* yang optimal.
2. Menginisialisasi *Centroid*: Algoritma secara acak memilih *k* titik awal sebagai centroid awal untuk masing-masing Klaster. *Centroid* adalah titik pusat dari sebuah Klaster, yang merupakan rata-rata dari semua titik dalam Klaster tersebut.
3. Menentukan Klaster untuk Setiap Titik Data: Algoritma mengukur jarak setiap titik data ke centroid yang telah ditentukan. Titik data akan dikelompokkan ke Klaster dengan centroid terdekat. Biasanya, jarak yang

digunakan adalah jarak *Euclidean*, meskipun metrik lain juga bisa digunakan.

4. Memperbarui Centroid: Setelah semua titik data dikelompokkan ke Klaster, centroid dari setiap Klaster diperbarui. Centroid baru dihitung dengan cara mengambil rata-rata dari semua titik data yang berada di Klaster tersebut.
5. Mengulangi Langkah 3 dan 4: Proses pengelompokan (Langkah 3) dan pembaruan *centroid* (Langkah 4) diulangi hingga tidak ada perubahan pada Klaster (konvergen), atau jumlah iterasi maksimal tercapai. Konvergensi terjadi ketika anggota Klaster tidak berubah lagi, atau perpindahan centroid di bawah ambang batas yang ditentukan.

Pemilihan Nilai k yang Optimal sangat penting dalam penerapan *K-Means Clustering*. Memilih jumlah Klaster yang optimal adalah salah satu aspek krusial dalam *K-Means*. Beberapa metode yang umum digunakan meliputi:

1. *Elbow Method*: Dalam metode ini, jumlah Klaster dipilih berdasarkan plot dari *sum of squared errors* (SSE) terhadap jumlah Klaster. Titik dimana penurunan nilai SSE mulai melambat (membentuk sudut seperti siku) dianggap sebagai nilai k yang optimal.
2. *Silhouette Score*: Metode ini mengevaluasi seberapa baik setiap titik data cocok dengan Klaster yang di dalamnya, dibandingkan dengan Klaster lainnya. Skor ini berkisar dari -1 hingga 1, di mana nilai lebih tinggi menunjukkan Klaster yang lebih baik.

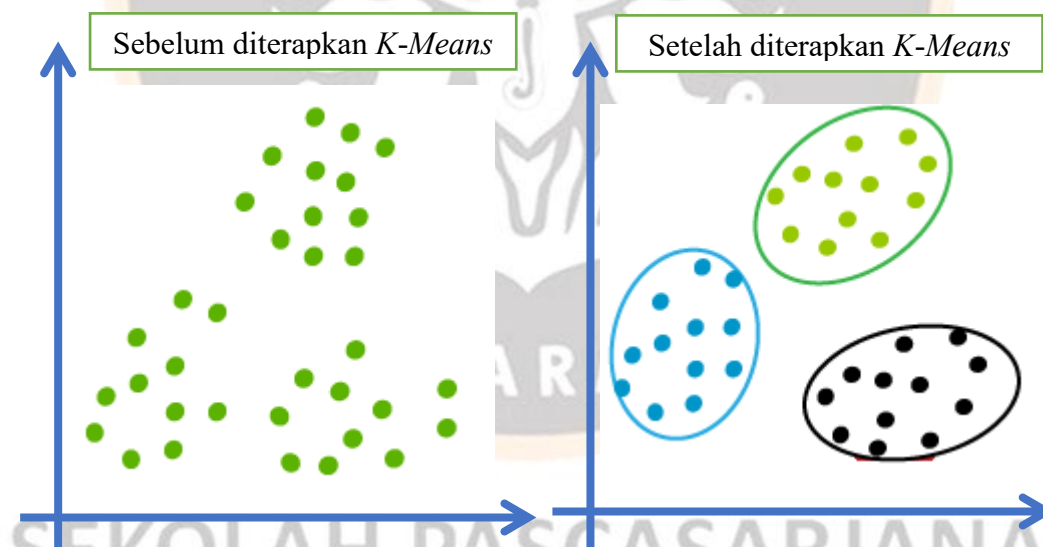
Kelebihan *K-Means Clustering* antara lain :

1. Sempel dan Cepat: *K-Means* relatif sederhana dan mudah diimplementasikan, serta memiliki waktu komputasi yang cepat untuk dataset besar.
2. Efisien untuk Klaster yang Berbentuk Bulat: *K-Means* bekerja dengan baik ketika Klaster yang terbentuk berbentuk bulat dan memiliki ukuran yang kurang lebih seragam.

Kekurangan *K-Means Clustering* antara lain :

1. Sensitif Terhadap *Outliers*: *K-Means* tidak tahan terhadap outlier, karena data yang jauh dari Klaster bisa menarik centroid sehingga menghasilkan Klaster yang kurang akurat.
2. Hanya Mendeteksi Klaster Berbentuk Bulat: Jika data memiliki Klaster yang tidak berbentuk bulat atau memiliki variasi ukuran yang signifikan, *K-Means* dapat memberikan hasil yang kurang baik.
3. Pemilihan k Harus Diketahui: Algoritma ini memerlukan jumlah Klaster yang ditentukan terlebih dahulu, yang bisa jadi sulit untuk ditentukan pada data yang kompleks.

Secara keseluruhan, *K-Means Clustering* adalah metode yang kuat dan serbaguna, terutama ketika digunakan dalam konteks yang sesuai dengan karakteristik datanya. Namun, penting untuk mempertimbangkan batasan dan pemilihan jumlah Klaster yang tepat agar mendapatkan hasil yang optimal (Mallqui and Fernandes 2019), (Das, Kayal, and Maiti 2023). Lebih jelasnya dapat dilihat pada Gambar 2.2.



Gambar 2.2 Visualisasi Implementasi Metode *K-Means Clustering* (dqlab.id, 2023)

K-Means bekerja dengan membagi data ke dalam beberapa kelompok berdasarkan seberapa mirip kualitasnya. Sasaran utama *K-Means* adalah memisahkan data ke dalam beberapa "kelompok" sehingga data dalam setiap kelompok memiliki ciri yang sama (dqlab.id, 2023).

2.3.5 Implementasi Bahasa Pemrograman berbasis *Python* beserta librari Pendukung

Beberapa *Library Python scikit-learn* yang digunakan dalam penelitian ini

Tabel 2.1 *Library Python scikit-learn*.

No	Modul <i>scikit-learn.org</i>	Fungsi	Sumber Pustaka
1.	<i>mean_absolute_percentage_error</i>	MAPE	(Scikit-learn.org (mape) 2024)
2.	<i>mean_absolute_error</i>	MAE	(Scikit-learn.org (mae) 2024)
3.	<i>mean_squared_error</i>	MSE	(Scikit-learn.org (mse) 2024)
4.	<i>root_mean_squared_error</i>	RMSE	(Scikit-learn.org (rmse) 2024)
5.	<i>KMeans</i>	KMeans	(Scikit-learn.org (KMeans) 2024)
6.	<i>mutual_info_score</i>	MIS	(Scikit-learn.org (mis) 2024)
7.	<i>adjusted_mutual_info_score</i>	AMIS	(Scikit-learn.org (ami) 2024)
8.	<i>normalized_mutual_info_score</i>	NMIS	(Scikit-learn.org (nmis) 2024)
9.	<i>fowlkes_mallows_score</i>	FMS	(Scikit-learn.org (fms) 2024)
10.	<i>homogeneity_score</i>	HS	(Scikit-learn.org (hs) 2024)
11.	<i>rand_score</i>	RS	(Scikit-learn.org (rs) 2024)
12.	<i>adjusted_rand_score</i>	ARS	(Scikit-learn.org (ars) 2024)
13.	<i>v_measure_score</i>	VMS	(Scikit-learn.org (vms) 2024)
14.	<i>MLP Classifier</i>	MLP C	(Scikit-learn.org (MLP C) 2024)
15.	<i>RandomForestRegressor</i>	RFR	(Scikit-learn.org (rfr) 2024)
16.	<i>Support Vector Machines</i>	SVM	(Scikit-learn.org (svm) 2024)
17.	<i>GradientBoostingRegressor</i>	GBR	(Scikit-learn.org (gbr) 2024)

Beberapa *Library Python TensorFlow.Keras* yang digunakan dalam penelitian :

Tabel 2.2 *Library Python TensorFlow.Keras*.

No.	Modul <i>tensorflow.org</i>	Fungsi	Sumber Pustaka
1.	<i>tf.keras.layers.LSTM</i>	LSTM	(Keras.LSTM 2024)
2.	<i>tf.keras.layers.Bidirectional</i>	<i>Bidirectional</i>	(Keras.Bidirectional 2024)
3.	<i>tf.keras.layers.GRU</i>	GRU	(Keras.GRU 2024)
4.	<i>tf.keras.layers.Conv1D</i>	Conv1D	(Keras.Conv1D 2024)
5.	<i>tf.keras.layers.Conv2D</i>	Conv2D	(Keras.Conv2D 2024)
6.	<i>tf.keras.layers.ConvLSTM1D</i>	ConvLSTM1D	(Keras.ConvLSTM1D 2024)
7.	<i>tf.keras.layers.ConvLSTM2D</i>	ConvLSTM2D	(Keras.ConvLSTM2D 2024)

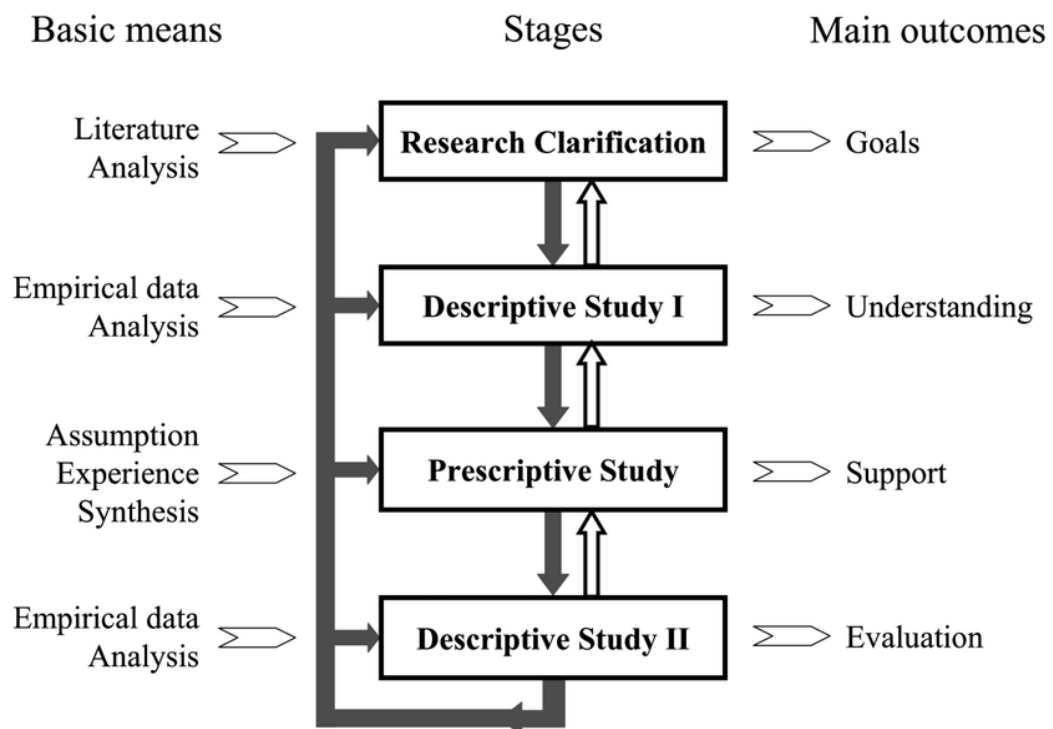
Beberapa *Library Python* lainnya yang digunakan dalam penelitian ini

Tabel 2.3 *Library Python* lainnya.

No.	Library <i>Python</i>	Modul lainnya	Fungsi	Sumber Pustaka
1.	<i>statsmodels0.14.4</i>	<i>statsmodels.tsa.arima.model.ARIMA</i>	ARIMA	(Statsmodels.org (ARIMA) 2024)
2.	<i>pypi.org</i>	<i>pypi.org/project/transformers</i>	<i>Transformers</i>	(Pypi.org (transformers) 2024)
3.	<i>pypi.org</i>	<i>pypi.org/project/xgboost</i>	<i>Xgboost</i>	(Pypi.org (xgboost) 2024)
4.	<i>pypi.org</i>	<i>pypi.org/project/lightgbm</i>	<i>Light GBM</i>	(Pypi.org (lightgbm) 2024)

2.3.6 Design Research Methodology (DRM)

Design Research Methodology (DRM) adalah pendekatan sistematis dalam penelitian desain yang digunakan untuk mengembangkan solusi berbasis metode ilmiah. DRM sering digunakan dalam penelitian berbasis sistem informasi, teknik, dan ilmu komputer untuk memastikan bahwa pengembangan solusi dilakukan dengan pendekatan yang terstruktur dan berbasis bukti.



Gambar 2.3 Framework *Design Research Methodology* (Blessing dan Chakrabarti 2009)

Konsep Dasar *Design Research Methodology* DRM dikembangkan oleh Blessing dan Chakrabarti (2009) sebagai metode penelitian desain yang menekankan pendekatan berbasis bukti dalam membangun dan mengevaluasi solusi. DRM terdiri dari empat tahap utama:

1. *Research Clarification* (RC): Mengidentifikasi masalah penelitian, merumuskan tujuan, dan memahami batasan masalah.
2. *Descriptive Study 1* (DS1): Mengkaji literatur dan melakukan analisis awal untuk memahami tantangan dalam domain penelitian.

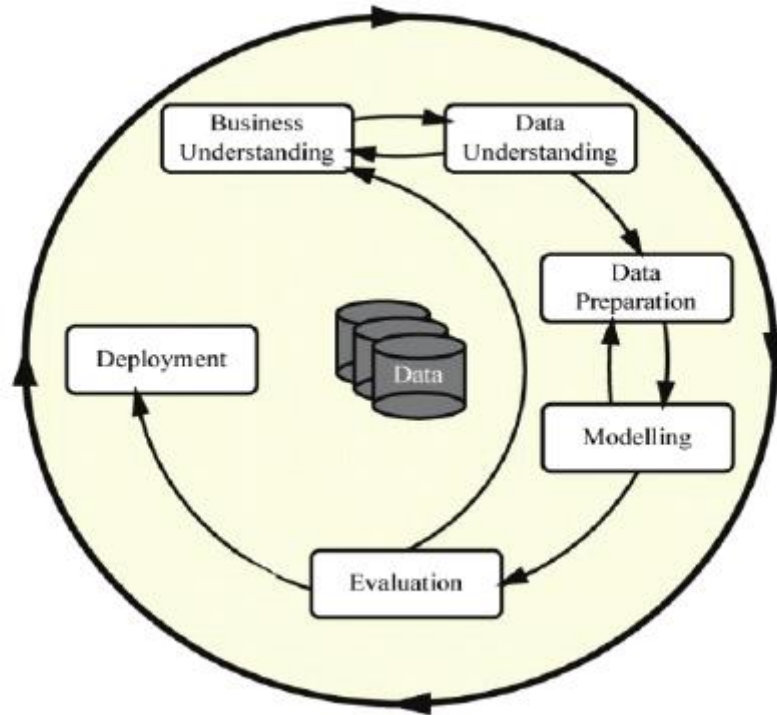
3. *Prescriptive Study* (PS): Merancang dan mengembangkan model atau solusi berbasis teori dan hasil studi literatur.
4. *Descriptive Study 2* (DS2): Mengevaluasi model atau solusi yang dikembangkan berdasarkan data empiris untuk mengukur efektivitasnya.

DRM telah banyak diterapkan dalam penelitian berbasis sistem informasi untuk mengembangkan model berbasis data, sistem pendukung keputusan (*Decision Support System* - DSS), dan kecerdasan buatan (*Artificial Intelligence* - AI). Gregor dan Hevner (2013) menyatakan bahwa DRM memberikan pendekatan sistematis dalam mengembangkan dan mengevaluasi sistem informasi berbasis data. Peffers et al. (2007) mengusulkan Design Science Research Methodology (DSRM) yang beririsan dengan DRM, terutama dalam hal iterasi desain dan evaluasi berbasis empiris.

Studi lain oleh Tönnessen dan Teuteberg (2020) menunjukkan bahwa DRM efektif dalam merancang model berbasis *Deep Learning* untuk memprediksi pergerakan harga pasar. DRM sering dikombinasikan dengan CRISP-DM (*Cross-Industry Standard Process for Data Mining*) untuk memastikan pendekatan yang lebih sistematis dalam eksplorasi data. Menurut Wirth dan Hipp (2000), CRISP-DM digunakan sebagai framework standar untuk proses *Data Mining* yang melibatkan siklus Business Understanding, *Data Understanding*, *Data Preparation*, Modeling, *Evaluation*, dan *Deployment*. Kombinasi DRM dan CRISP-DM terbukti efektif dalam penelitian berbasis kecerdasan buatan dan analisis data pasar keuangan.

2.3.7 Cross-Industry Standard Process for Data Mining (CRISP-DM)

Cross-Industry Standard Process for Data Mining (CRISP-DM) adalah metodologi yang terstruktur dan sistematis untuk proyek *Data Mining*, mencakup enam tahapan utama, yakni pemahaman bisnis, pemahaman data, persiapan data, pemodelan, evaluasi, dan penerapan (Chapman, dkk., 2000).



Gambar 2.4 6 Tahapan CRISP-DM (Chapman, dkk., 2000).

Pada Gambar 2.4 terdapat Enam Tahapan CRISP-DM:

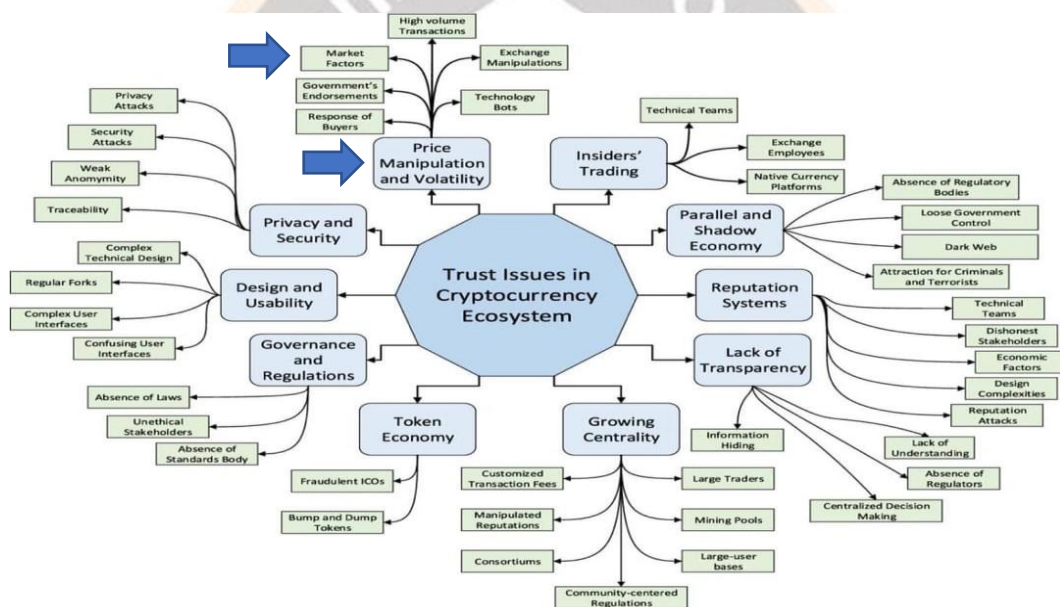
1. *Business Understanding* (Pemahaman Bisnis):
Menentukan tujuan dan persyaratan bisnis, menerjemahkannya ke dalam tujuan *Data Mining*, dan mempersiapkan strategi awal.
2. *Data Understanding* (Pemahaman Data):
Memahami data yang tersedia, mengidentifikasi masalah seperti data yang hilang, outlier, dan distribusi data.
3. *Data Preparation* (Persiapan Data):
Memperbaiki masalah data, melakukan transformasi data, dan menyiapkan data untuk pemodelan.
4. *Modelling* (Pemodelan):
Membangun model prediktif atau deskriptif menggunakan metode statistik dan *Machine Learning*, serta memilih algoritma yang tepat.
5. *Evaluation* (Evaluasi):
Mengevaluasi kinerja model dan memastikan bahwa model yang dipilih sesuai dengan tujuan bisnis yang telah ditetapkan.

6. Deployment (Penerapan):

Merencanakan penggunaan model di dunia nyata, termasuk bagaimana model akan digunakan dan bagaimana hasil model akan digunakan untuk pengambilan keputusan.

2.3.8 Area Penelitian *Cryptocurrency*

Berikut adalah area penelitian *Cryptocurrency* menurut M. H. U. Rehman, dkk., 2020.



Gambar 2.5 Area penelitian *Cryptocurrency* (M. H. U. Rehman, dkk., 2020)

Dari Gambar 2.5 telah diberikan penanda bahwa fokus area Penelitian *Cryptocurrency* yang lebih dekat dengan tujuan disertasi ini dapat dilihat pada diagram alir Gambar 2.6.



Gambar 2.6 Detail Area penelitian *Cryptocurrency* (M. H. U. Rehman, dkk., 2020)

Dari Gambar 2.6 tampak bahwa salah satu masalah terkait kepercayaan dalam ekosistem mata uang kripto adalah adanya tindakan manipulasi harga dan

volatilitas. Untuk melakukan tindakan ini tentu para pelaku pasar yang memiliki modal besar akan sangat memperhatikan faktor-faktor yang merupakan bagian dari transaksi perdagangan dalam suatu Pasar digital mata uang kripto. Model Dataset terpola akan dengan cepat mengetahui adanya perilaku ini dengan mengevaluasi secara realtime dan terus menerus terkait pergerakan harga dari waktu ke waktu dan menampilkannya dalam informasi yang matang, terpola, dan akan sangat membantu Pedagang retail dalam mengambil keputusan perdagangan aset *Cryptocurrency*.

2.4 Hipotesis Model Dataset Terpola

Hipotesis adalah pernyataan yang dapat diuji dengan penelitian ilmiah. Jika Peneliti ingin menguji hubungan antara dua variabel atau lebih, Peneliti perlu menulis hipotesis sebelum memulai eksperimen atau pengumpulan data. Hipotesis menyatakan prediksi Peneliti tentang apa yang akan ditemukan dalam penelitian Peneliti. Ini adalah jawaban tentatif atas pertanyaan penelitian Peneliti yang belum diuji. Untuk beberapa proyek penelitian, Peneliti mungkin harus menulis beberapa hipotesis yang membahas berbagai aspek pertanyaan penelitian Peneliti. Hipotesis bukan sekedar dugaan – hipotesis harus didasarkan pada teori dan pengetahuan yang ada. Hal ini juga harus dapat diuji, artinya Peneliti dapat mendukung atau menyangkalnya melalui metode penelitian ilmiah (seperti eksperimen, observasi, dan analisis data statistik).

Peneliti harus mempunyai Gambaran tentang apa yang ingin ditemukan. Tulis jawaban awal Peneliti atas pertanyaan tersebut dalam kalimat yang jelas dan ringkas. Contoh merumuskan hipotesis Peneliti, “Menghadiri lebih banyak perkuliahan menghasilkan hasil ujian yang lebih baik”(McCombes 2022).

Hipotesis yang telah dibuat selanjutnya disempurnakan. Peneliti perlu memastikan hipotesis Peneliti spesifik dan dapat diuji. Ada berbagai cara untuk mengungkapkan hipotesis, namun semua istilah yang Peneliti gunakan harus memiliki definisi yang jelas. Hipotesis harus berisi Variabel yang relevan, Kelompok tertentu yang sedang dipelajari, dan Hasil prediksi percobaan atau

analisis (Dias, Fernando, and Fernando 2022), (Rizky Parlika, Mustafid, and Rahmat 2024a), (Al-Omoush, Gomez-Olmedo, and Funes 2024).

Penelitian ini menggunakan 4 Hipotesis yang saling berkaitan yang berhasil dibuktikan kebenarannya, sehingga semakin memperkuat validasi dari novelty Model Dataset Terpola, yakni:

1. Hipotesis 1, menunjukkan mekanisme memperoleh posisi beli aset *Cryptocurrency* pada harga yang relatif lebih rendah dan berpotensi menghasilkan profit saat dijual pada harga yang lebih tinggi, serta sebagai penanda bahwa harga mayoritas aset *Cryptocurrency* sedang berada di area support.
2. Hipotesis 2, menunjukkan mekanisme memperoleh posisi jual aset *Cryptocurrency* pada harga yang relatif lebih tinggi dan berpotensi menghasilkan profit bila sebelumnya telah melakukan pembelian pada harga yang lebih rendah, serta sebagai penanda bahwa harga mayoritas aset *Cryptocurrency* sedang berada di area resisten.
3. Hipotesis 3, menunjukkan mekanisme untuk menandai dimulainya fluktuasi penurunan pada area harga dengan skala mingguan hingga bulanan untuk mendapatkan posisi harga beli yang relatif lebih rendah, menandai posisi harga yang mendekati titik paling rendah, hingga berakhirnya fase fluktuasi penurunan.
4. Hipotesis 4 menunjukkan mekanisme untuk menandai dimulainya fluktuasi kenaikan pada area harga dengan skala mingguan hingga bulanan untuk mendapatkan posisi harga jual yang relatif lebih tinggi, menandai posisi harga yang mendekati titik paling tinggi, hingga berakhirnya fase fluktuasi kenaikan.

Pada penelitian ini keempat hipotesis utama, masing-masing diuji menggunakan dataset terpola dan metode yang relevan. Berikut penjelasan status setiap hipotesis:

1. Hipotesis 1: Zona harga rendah merupakan titik beli yang optimal.

Status: Terbukti, Analisis data menunjukkan bahwa zona harga rendah yang teridentifikasi oleh model secara konsisten memberikan peluang keuntungan tinggi ketika digunakan sebagai titik beli.

2. Hipotesis 2: Zona harga tinggi merupakan titik jual yang optimal.

Status: Terbukti, Hasil validasi menunjukkan bahwa zona harga tinggi yang ditentukan oleh model merupakan area optimal untuk menjual aset, dengan korelasi tinggi terhadap potensi keuntungan maksimum.

3. Hipotesis 3: Fluktuasi harga turun dapat dideteksi dan diprediksi.

Status: Terbukti, Model berhasil mendeteksi awal dan akhir fluktuasi harga turun dengan akurasi tinggi, terutama menggunakan teknik *Clustering* dan LSTM untuk pola temporal.

4. Hipotesis 4: Fluktuasi harga naik dapat dideteksi dan diprediksi.

Status: Terbukti, Tren harga naik dapat dipetakan secara akurat menggunakan model, menunjukkan bahwa pola temporal dan dataset terpola dapat memberikan prediksi yang andal.

Dengan hasil pembuktian Hipotesis 1-4 diharapkan dapat memberikan hasil prediksi volatilitas *Cryptocurrency* yang lebih sederhana dan mudah dipahami menggunakan Model Dataset Terpola, untuk digunakan sebagai bahan pertimbangan sebelum mengambil keputusan untuk melakukan pembelian dan penjualan aset *Cryptocurrency*.