

Prediksi Harga Mata Uang Kripto Menggunakan LSTM dan MLR

^{1st} Samuel Erlangga
Fakultas Informatika
Universitas Telkom
Bandung, Indonesia
samuelerlangga@student.telkomuni-
versity.ac.id

^{2nd} Indwiarti
Fakultas Informatika
Universitas Telkom
Bandung, Indonesia
indwiarti@telkomuniversity.ac.id

^{3rd} Annisa Aditsania
Fakultas Informatika
Universitas Telkom
Bandung, Indonesia
aaditsania@telkomuniversity.ac.id

Abstrak—Perkembangan mata uang kripto terjadi begitu pesat dalam beberapa tahun ini. Hal ini membuat orang-orang semakin tertarik untuk menginvestasikan uangnya ke dalam mata uang kripto. Namun untuk mempelajari pergerakan pasar dan ilmu untuk memprediksi harga menggunakan analisa teknikal tidaklah mudah, seringkali terjadinya perubahan yang begitu besar pada harga membuat mata uang kripto menjadi aset yang cukup diminati namun beresiko tinggi. Tidak jarang para investor mengalami kerugian dikarenakan hasil analisa yang kurang tepat dan karena transaksi jual beli mata uang kripto yang tidak dipatok oleh waktu memungkinkan harga berubah kapan saja. Pada penelitian ini melakukan prediksi harga mata uang kripto berupa harga penutupan mata uang kripto tersebut. Prediksi mata uang kripto ini dilakukan dengan metode *Long Short-Term Memory (LSTM)* dan *Multiple Linear Regression (MLR)*. Dengan menggunakan histori harga tutup sebelumnya (*Close*) dan jumlah penambahan kasus positif Covid-19 untuk menjadi data latih. Menggunakan dua mata uang kripto untuk memastikan mesin dapat bekerja untuk lebih dari satu mata uang kripto. Hasil dari kedua metode ini adalah prediksi harga mata uang kripto yang dibandingkan dengan menghitung nilai *Root Mean Squared Error (RMSE)* untuk memastikan metode apa yang lebih baik digunakan dalam memprediksi harga mata uang kripto. Hasil terbaik dihasilkan oleh model *LSTM* dengan nilai *RMSE* 0.000240.

Kata kunci—LSTM, MLR, mata uang kripto

Abstract—The development of cryptocurrencies has occurred so rapidly in recent years. This makes people more and more interested in investing their money in cryptocurrencies. However, studying market movements and the science of predicting prices using technical analysis is not easy, often such large changes in prices make cryptocurrencies an asset that is quite attractive but high risk. Investors frequently experience losses due to inaccurate analysis results. Buying and selling cryptocurrency transactions that are not pegged by time allow prices to change at any time. This study predicted the price of the cryptocurrency in the form of the closing price of the cryptocurrency. This cryptocurrency prediction made using the *Long Short-Term Memory (LSTM)* and *Multiple Linear Regression (MLR)* methods. Using the history of previous closing prices (*Closing*) and the number of additional positive cases of Covid-19 as training data.

Using two cryptocurrencies to ensure the machine can work for more than one cryptocurrency. The results of these two methods are cryptocurrency price predictions which compared by calculating the Root Mean Squared Error (RMSE) value to determine which method is better for predicting cryptocurrency prices. The best results were produced by the LSTM model with an RMSE value of 0.000240.

Keywords—LSTM, MLR, mata uang kripto

I. PENDAHULUAN

A. Latar Belakang

Perkembangan minat berinvestasi yang cukup besar di Indonesia bahkan diseluruh dunia, membuat dampak positif pada pasar saham maupun pasar *cryptocurrency*. Salah satu dampak positifnya adalah meningkatnya harga dan volume pasar karena mulai bermunculannya investor-investor baru atau kemauan investor untuk menanamkan uangnya pada pasar *cryptocurrency*. Bisa dilihat perkembangan pasar *cryptocurrency* mengalami perubahan yang cukup pesat pada tahun 2017 sampai tahun 2021 saat pandemi terjadi [1].

Dikarenakan pandemi yang terjadi mempengaruhi harga saham dimana ekonomi dunia terganggu dan banyak kegiatan menjadi terhenti yang membuat perusahaan-perusahaan berjalan tidak semestinya. Namun berbanding terbalik dengan *cryptocurrency*, pandemi memberi pengaruh besar dalam perubahan harga *cryptocurrency* ini dengan meningkatnya jual beli *cryptocurrency* karena banyak investor mulai memilih *cryptocurrency* sebagai salah satu cara untuk berinvestasi dan karena naik turunnya harga suatu koin *cryptocurrency* hanya dipengaruhi dari minat pembeli (teori *Supply and Demand*), walaupun terjadi pandemi dan minat pembeli masih baik harga tidak akan mengalami perubahan yang drastis [2].

Pada umumnya perkembangan *cryptocurrency* tidak terpengaruh oleh Covid-19, namun disebabkan juga karena penurunan besar-besaran pada pasar saham yang disebabkan oleh pandemi. Terlihat jelas pada awal masa pandemi para investor mengalami

Covid-19 Shock dan mengalami kepanikan sehingga menjual asset investasi miliknya [3]. Oleh karena itu banyak investor mulai tertarik dan berminat untuk berinvestasi di pasar *cryptocurrency*.

Adanya perkembangan minat berinvestasi pada mata uang kripto membuat harga semakin bergerak dengan lincah/*volatile* dan membuat seseorang lebih sulit dalam melakukan *Technical Analysis* untuk memprediksi harga selanjutnya. Pergerakan harga saham ini yang membuat para investor pemula menjadi kesulitan dalam membaca arah gerak harga, karena sewaktu-waktu harga bisa saja berubah cukup besar karena psikologi investor secara keseluruhan ingin membeli atau menjual koin tersebut.

Menganalisa pasar menggunakan *Technical Analysis* tidak hanya melihat pada data *Open*, *Close*, *High*, dan *Low* namun dari tren yang sedang terjadi dan dapat menggunakan indikator seperti *Moving Average (MA)* dengan berbagai jenis *MA* yang tersedia, *Relative Strength Index (RSI)* dan beberapa indikator lainnya untuk lebih memastikan prediksi harga saham selanjutnya [4].

Melihat peluang mata uang kripto akan digunakan secara umum untuk bertransaksi dikemudian hari dan dengan kondisi pasar yang seperti ini tidak sedikit dari investor mengalami kerugian karena salah memprediksi pergerakan harga dan mengalami kepanikan. Dengan menggunakan *Machine Learning* sebagai salah satu solusi dalam memperkecil kerugian dan memaksimalkan keuntungan, diharapkan dalam penelitian ini mesin mampu memprediksi harga koin dalam kurun waktu tertentu dengan akurasi yang baik.

Dari penelitian yang dilakukan oleh Yadav, Jha dan Sharan pada tahun 2020 [5], mereka menggunakan *LSTM* sebagai metode untuk memprediksi saham di India dengan menggunakan *RMSE* sebagai evaluasi hasil. Dari kedua eksperimen yang dilakukan ditemukan bahwa *LSTM* dapat memprediksi harga saham dengan cukup akurat, diikuti dengan pengaturan jumlah *hidden layer*, *batch* dan *epoch*-nya, namun dari eksperimen pertamanya ditemukan bahwa *stateless LSTM* memiliki hasil yang lebih stabil dibanding *stateful LSTM*. *RMSE* *statless LSTM* lebih rendah dibandingkan *stateful LSTM* pada keempat perusahaan dan *RMSE* paling rendah ditemukan pada 1 *hidden layer*, namun ada satu saham yang dimana semakin banyak jumlah *hidden layer* maka nilai *RMSE* semakin baik.

Oleh karena itu prediksi pada penelitian ini dilakukan tidak hanya menggunakan machine learning dengan metode Long Short-Term Memory (*LSTM*) namun dengan membandingkannya dengan metode Multiple Linear Regression (*MLR*). Pada penelitian yang dilakukan oleh S. Shakhla, B. Shah, N. Shah, V. Unadkat, and P. Kanani pada tahun 2018 [6], mereka

melakukan prediksi harga saham sebuah perusahaan teknologi ternama yaitu Apple dengan kode saham AAPL. Pada penelitiannya mereka menggunakan *RMSE* sebagai evaluasi performansi dari model yang mereka buat, dan hasilnya cukup menjanjikan namun masih dapat dikembangkan dengan lebih memperhatikan parameternya.

M. Patel, S. Tanwar, R. Gupta dan N. Kumah [7] pada tahun 2020, melakukan sebuah penelitian untuk memprediksi *cryptocurrency* dengan menggunakan dua metode yaitu *LSTM* dan Gated Recurrent Unit (*GRU*). Dengan menggunakan data historikal dari dua mata uang kripto yaitu Ethereum dan Litecoin dan menggunakan harga rata-rata harian untuk menjadi fitur utamanya. Pada penelitiannya mereka menghasilkan skema yang akurat dalam memprediksi harga mata uang kripto tersebut, dan mengatakan bahwa skema ini dapat digunakan untuk memprediksi harga mata uang kripto lainnya dengan membandingkan beberapa metrik evaluasi seperti *Mean Squared Error (MSE)*, *Root Mean Squared Error (RMSE)*, *Mean Average Error (MAE)* dan *Mean Absolute Percentage Error (MAPE)*, dan melakukan pengujian dengan membedakan *window* dalam model sebagai parameter. Perbedaan dengan penelitian ini adalah tidak menggunakan perbedaan jumlah *window* sebagai pengujian karena metode *MLR* tidak memiliki *window*, namun memiliki persamaan yaitu melihat nilai *RMSE* untuk dibandingkan.

Dengan hasil penelitian sebelumnya yang menyatakan bahwa *LSTM* dan *MLR* dapat digunakan untuk prediksi dan prediksi pun dapat dilakukan dengan menggunakan data mata uang kripto maka mesin dapat dibuat untuk belajar dari data yang telah ada dari waktu sebelumnya dan memprediksi harga pada waktu berikutnya. *LSTM* dan *MLR* akan menggunakan data harga *cryptocurrency* sebelumnya dan jumlah kasus penambahan positif Covid-19 diseluruh dunia lalu memprediksi berapakah harga *cryptocurrency* pada hari berikutnya.

B. Topik dan Batasannya

Topik utama dalam penelitian ini adalah memprediksi harga *cryptocurrency* atau mata uang kripto dengan menggunakan dua metode yaitu Long Short-Term Memory (*LSTM*) dan Multiple Linear Regression (*MLR*). Lalu membandingkan hasil dari kedua metode yang digunakan untuk mengetahui metode mana yang memberikan hasil yang lebih baik ataupun akurat.

Batasan yang digunakan dalam penelitian ini adalah data yang digunakan mungkin saja memiliki data yang terbatas karena data hanya tersedia mulai dari tanggal 22 Januari 2020saja hingga saat ini, dan menggunakan data hingga tanggal 20 September 2022.

Dikarenakan harga *cryptocurrency* pada beberapa waktu kebelakang yaitu mulai dari pertengahan 2022 hingga September 2022 mengalami perubahan yang cukup besar dapat membuat model lebih sulit dalam melakukan prediksi.

C. Tujuan

Tujuan dari penelitian ini adalah memprediksi harga *cryptocurrency* atau mata uang kripto menggunakan metode pembelajaran mesin *Long Short-Term Memory* dan *Multiple Linear Regression* dengan menggunakan data histori *cryptocurrency* tersebut dan jumlah penambahan kasus positif *Covid-19* di dunia. Menggunakan model *LSTM* dan *MLR* yang optimal untuk mendapatkan prediksi harga yang paling optimal dan membandingkan performansi hasil dari kedua metode tersebut.

D. Organisasi Tulisan

Pada bagian berikutnya yaitu bagian kedua membahas tentang metode yang akan digunakan dan studi-studi terkait. Pada bagian ketiga membahas tentang sistem dan rancangannya, lalu membahas dan menganalisa hasil dari sistem yang telah dibangun pada bagian keempat dan kesimpulan dari keseluruhan penelitian pada bagian kelima.

II. KAJIAN TEORI

A. Long Short-Term Memory

Long Short-Term Memory (LSTM) merupakan salah satu metode modifikasi dari *RNN*, *LSTM* juga mampu untuk menggunakan data dari hasil pengamatan sebelumnya untuk dipakai dalam memprediksi. *LSTM* pertama kali dibuat pada tahun sekitar 1990 untuk menyelesaikan salah satu kelemahan *RNN*. Dalam pengaplikasiannya *RNN* sering kali menemui permasalahan yang disebut *Vanishing*

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

$$f_t = \sigma(w_f[h_{t-1}, x_t] + b_f) \quad (2)$$

$$o_t = \sigma(w_o[h_{t-1}, x_t] + b_o) \quad (3)$$

i_t adalah *input gate*, f_t adalah *forget gate*, o_t adalah *output gate*, σ adalah fungsi *sigmoid*, w_x adalah *weight* untuk setiap *gate* yang ada, h_{t-1} adalah *output* dari perhitungan *LSTM* sebelumnya saat $t-1$, x_t adalah *input LSTM* saat ini dan b_x adalah bias untuk setiap *gate*.

Dari perhitungan pertama yaitu *forget gate* yang membuat data apa yang harus dibuang dari *cell state*, perhitungan kedua yaitu *input gate* yang dimana memberitau data baru yang mana yang harus disimpan di

Gradient dimana model pembelajaran menjadi begitu lambat atau bahkan berhenti. Namun *LSTM* memiliki ingatan yang lebih panjang dan dapat belajar dari input yang terpisah dari yang lainnya dalam jangka waktu yang cukup lama [5].

Karena *RNN* tidak memiliki memori jangka panjang, *LSTM* digunakan untuk menangani masalah ini yang dimana memiliki memori jangka panjang yang berguna dalam memprediksi data. Sebuah *LSTM* memiliki tiga gerbang atau *gates* yaitu:

1. **Input Gate** : Menentukan apakah *input* baru akan diperbolehkan masuk atau tidak.
2. **Forget Gate** : Menghapus data yang sudah dianggap tidak penting.
3. **Output Gate** : Menentukan informasi apakah yang akan dikeluarkan.

Ketiga *gate* ini bekerja karena fungsi aktivasi yang menentukan perilaku apa yang akan dilakukan kepada data-data yang ada pada setiap *layer*-nya [5].

Pada setiap *gate* akan digunakan fungsi aktivasi, ada beberapa fungsi aktivasi yang diantaranya adalah *Linear*, *ReLU*, *Tanh* dan *Sigmoid* [8]. Pada *LSTM* akan digunakan fungsi aktivasi *sigmoid* yang mengeluarkan nilai 0-1, karena membutuhkan nilai positif dari setiap gerbang untuk mengetahui apakah fitur tersebut harus disimpan atau dihapus. Nilai "0" mengartikan bahwa gerbang memblokir fitur untuk lewat dan nilai "1" [9] untuk gerbang mengizinkan fitur untuk lewat, seperti pada rumus dibawah ini.

dalam *cell state*, dan perhitungan ketiga yaitu *output gate* yang menyediakan fungsi aktivasi untuk hasil akhir untuk *LSTM* saat waktu " t ".

Cell state digunakan sebagai memori dalam *LSTM* dan menampung data dari *LSTM* sebelumnya lalu membawa data untuk *LSTM* selanjutnya. *Cell State* ini menjadi bagian penting dalam *LSTM* karena dapat mengingat data dari proses sebelumnya.

$$\tilde{c}_t = \tanh(w_c[h_{t-1}, x_t] + b_c) \quad (4)$$

$$c_t = f_t * c_{t-1} + i_t * \tilde{c}_t \quad (5)$$

$$h_t = o_t * \tanh(c_t) \quad (6)$$

c_t adalah *cell state* atau memori saat waktu “t” dan $\tilde{c}_t$ adalah *cell state candidate* atau kandidat memori yang akan disatukan dengan *cell state*. Dari persamaan di atas dapat terlihat bahwa pada waktu kapanpun, *cell state* mengetahui data apa yang harus dilupakan dari proses sebelumnya dan data apa yang harus dilanjutkan dari proses saat ini. Setelah melakukan penyaringan untuk *cell state* lalu *cell state* akan melalui fungsi aktivasi dan *softmax layer* untuk mendapatkan output.

Output layer akan terus terhubung dengan *hidden layer* sampai pelatihan atau pembelajaran selesai dilakukan. Seberapa banyak perputaran ini dilakukan tergantung dari jumlah *epoch* yang diberikan, dengan jumlah *epoch* yang tepat dapat membuat hasil akhir lebih akurat.

B. Multiple Linear Regression

Multiple Linear Regression (MLR) merupakan salah satu metode yang kerap digunakan dalam melakukan analisis pasar saham [6]. *MLR* digunakan agar dapat memprediksi menggunakan beberapa variabel yang mempengaruhi hasil yang akan diprediksi. Model ini secara matematis menghitung hubungan antara variabel independen (prediktor) dengan variabel dependen atau target prediksi [10]. Variabel dependen akan berubah-ubah tergantung dari variabel independen, karena variabel independen mempengaruhi variabel dependen. Persamaan matematika umum yang sering digunakan untuk mewakili *MLR* adalah:

$$y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n + e_n \quad (7)$$

Dimana β adalah koefisien dari setiap variabel, x merupakan variabel independen (prediktor), e_n adalah ukuran parameter kesalahan dan y_i merupakan variabel dependen yang akan dicari hasilnya [11].

C. Metrik Pengukuran

Pada penelitian ini melakukan perbandingan hasil prediksi dari kedua model

regresi yang dibangun. Untuk mengetahui dan mengukur hasil dari penelitian ini agar dapat membandingkan kedua hasil model yang dibangun dengan menggunakan *Root Mean Squared Error (RMSE)*. *RMSE* digunakan untuk mengetahui deviasi antara dua nilai yang didalam penelitian ini adalah nilai yang digunakan model untuk belajar dan nilai prediksi yang dihasilkan oleh model.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n}} \quad (8)$$

Dengan perhitungan matematika seperti ini, dapat menghasilkan nilai deviasi antar kedua variabelnya yaitu, y_i sebagai nilai observasi (nilai yang digunakan model untuk belajar), $\hat{y}_i$ sebagai nilai prediksi yang dikeluarkan oleh model, i menunjukkan iterasi ke- i yang sedang berjalan dan n jumlah iterasi yang dilakukan.

III. METODE

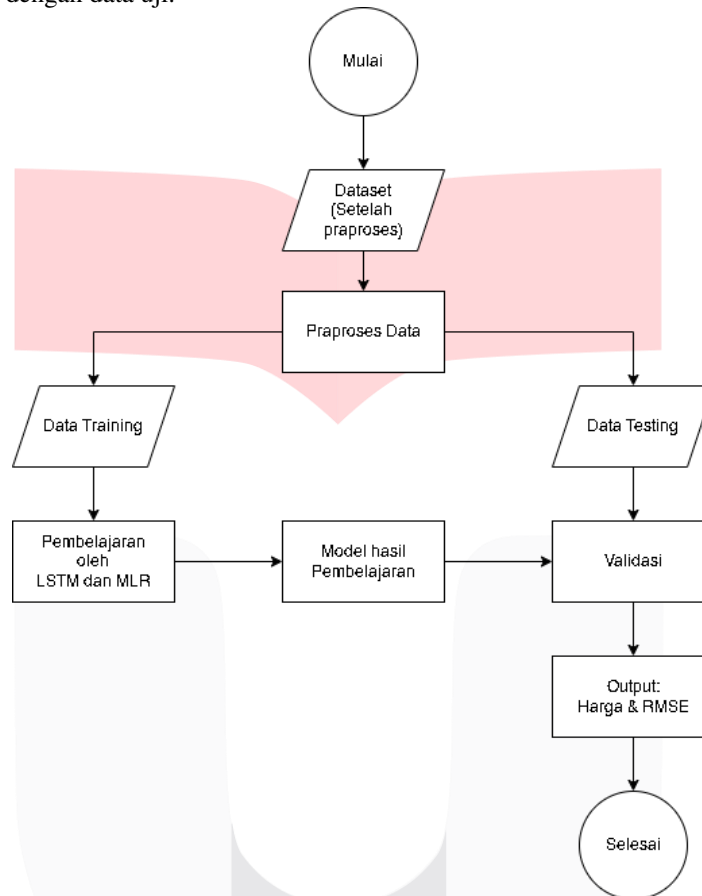
A. Sistem Secara Umum

Sistem yang dibangun dalam penelitian ini adalah model *Machine Learning* yang

digunakan untuk memprediksi harga mata uang kripto. Data yang digunakan didapat dari dua sumber yaitu *Yahoo! Finance* dan *Our World in Data*, dan data yang didapat sudah dalam bentuk data deret waktu. Lalu merubah data dalam bentuk *Numpy array* agar dapat lebih mudah diproses. Sebelum data di proses, dilakukan praproses data untuk melihat apakah data sudah dalam tipe data yang sesuai dan apakah data tidak ada yang kosong, lalu dilakukan proses *scaling* data agar data berada dalam skala angka yang sama.

Setelah melakukan praproses data, data tersebut sudah siap untuk digunakan model untuk berlatih, kedua model *Long Short-Term Memory (LSTM)* dan *Multiple Linear Regression (MLR)* menggunakan data latih yang sama. Dilakukan pelatihan berulang kali agar hasil bisa didapat sebaik mungkin dan dibandingkan dengan data uji.

Evaluasi dilakukan dengan menggunakan metrik *Root Mean Squared Error (RMSE)* setelah model selesai untuk mengetahui model mana yang memberikan hasil yang lebih baik dalam memprediksi harga mata uang kripto dengan kedua data yang diberikan.



GAMBAR 1
GAMBARAN UMUM SISTEM

B. Dataset

Dataset yang digunakan dalam penelitian ini adalah dataset mata uang kripto dan data set jumlah kasus positif baru perharinya di dunia. Dataset mata uang kripto diambil dari *Yahoo! Finance* dengan menggunakan API yang dibuka secara umum. *Yahoo! Finance* adalah sebuah media bagian dari *Yahoo!* yang menyediakan berita finansial, harga saham, laporan keuangan dan konten pribadinya. Namun hanya mengambil dua mata uang kripto yang dianggap cukup stabil di pasar uang yaitu *Ethereum (ETH)* dan *Ripple (XRP)*. Karena kedua *cryptocurrency* ini memiliki perbedaan yaitu *Ethereum* yang merupakan *stable coin* atau

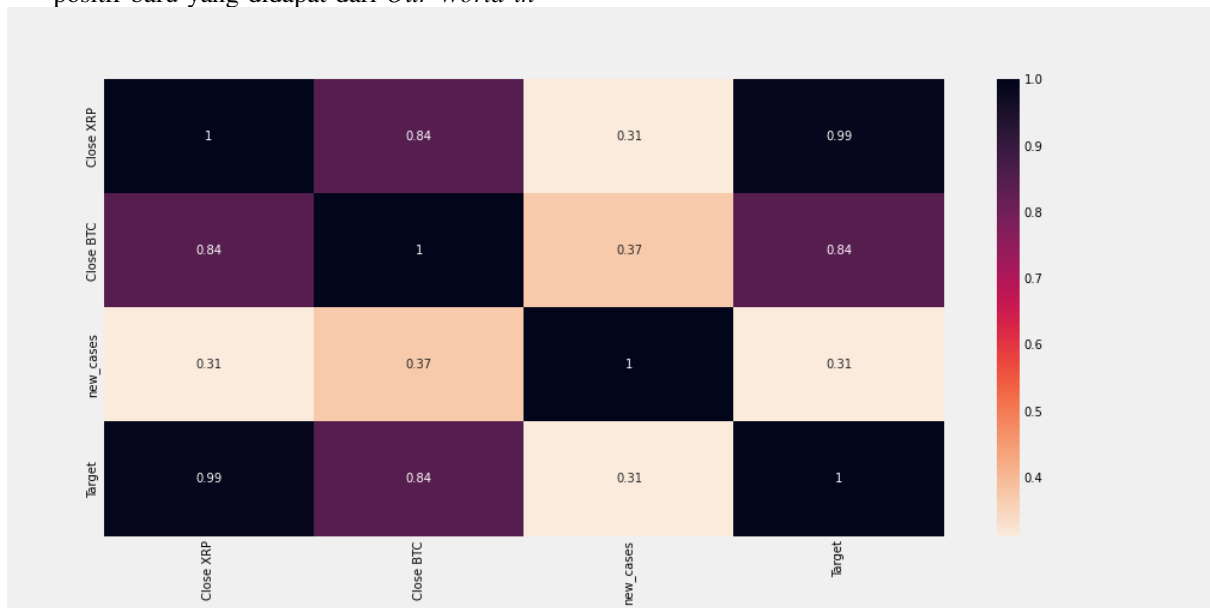
salah satu koin yang dianggap stabil oleh pasar, karena *Ethereum* sendiri menjadi patokan banyak harga dan standarisasi nilai tukar pada *cryptocurrency* sebagai contoh yaitu penggunaan dollar pada perdagangan internasional dan *Ripple* merupakan salah satu *cryptocurrency* yang cukup populer namun tidak termasuk *stable coin*. Dan data harga mata uang kripto *Bitcoin (BTC)* sebagai salah satu variabel dependen pada skenario di dalam pengujian.

Lalu untuk data jumlah kasus positif baru di dunia diambil dari *website Our World in Data* yang menyediakan banyak data yang biasa digunakan dalam sains seperti data

kemiskinan, penyakit, kelaparan, perubahan iklim, resiko eksistensial dan ketidakrataan.

Data yang didapat dari *Yahoo! Finance* berupa harga *High*, *Low*, *Open*, *Close*, *Adj Close* dan *Volume*, namun hanya menggunakan harga *Close* pada hari sebelumnya sebagai variable independennya dan *Close* hari berikutnya sebagai variable dependennya atau *Target*. Data jumlah kasus positif baru yang didapat dari *Our World in*

Data akan menjadi variable independen Bersama dengan harga *Close* mata uang kripto tersebut. Korelasi antar variabelnya cukup baik, *Close* dan *Target* sangat berkorelasi dengan nilai 0.99, untuk total kasus positif baru (*new_cases*) dan *Target* memiliki korelasi yang cukup kecil yaitu 0.31 dan untuk harga *Close Bitcoin* dengan target yaitu 0.84.



GAMBAR 2
TABEL KORELASI ANTAR VARIABEL

Dataset yang digunakan merupakan data deret waktu yang diambil dari tanggal 22 Januari 2020 dimana *Our World in Data*

hanya menyajikan data kasus positif Covid-19 dimulai dari tanggal tersebut, sampai tanggal 20 September 2022.

TABEL 1
SAMPEL DATA

Tanggal	Close (\$)	Close BTC(\$)	New Case	Target (\$)
22/01/2020	168.2941	8680.8759	0	162.9285
23/01/2020	162.9285	8406.5156	100	163.0511
24/01/2020	163.0511	8445.4345	287	161.2839
25/01/2020	161.2839	8367.8476	493	168.0771
26/01/2020	168.0771	8596.8300	683	170.9308

C. Praproses

Praproses data dilakukan agar data yang akan diproses dipastikan merupakan data yang baik dan mempermudah dalam membangun model. Dalam praproses data melakukan pengecekan apakah data yang diterima memiliki data yang kosong atau tidak lengkap, lalu melihat info data secara

menyeluruh untuk memastikan tipe data yang digunakan dan jumlah data sesuai.

Lalu data diubah ke dalam *numpy array* agar data lebih mudah diproses. Setelah itu melakukan *scaling* pada data menggunakan *Min Max Scaler* agar data berada dalam jangka yang sama yaitu 0 untuk nilai minimalnya dan 1 untuk nilai maksimalnya.

$$X' = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (9)$$

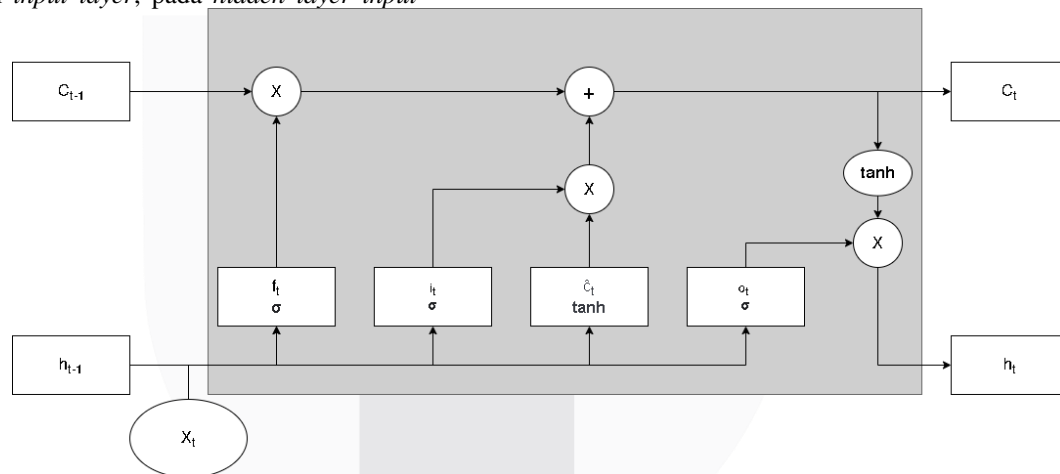
Min Max Scaler digunakan karena data yang digunakan memiliki rentang maksimal dan minimal yang berbeda-beda, oleh karena itu *Min Max Scaler* digunakan agar semua data yang dipakai dilakukan normalisasi terlebih dahulu agar semua datanya memiliki rentang maksimal dan minimal yang sama yaitu nol sampai 1 agar model dapat lebih baik melakukan pembelajaran. Lalu membagi data dengan perbandingan data latih dan data uji sebesar 8:2, dengan variabel dependen yaitu harga koin yang diprediksi, dan variabel independen yaitu harga coin sebelumnya dan jumlah kasus positif baru Covid-19 di dunia.

D. Long Short-Term Memory (LSTM)

Menggunakan data yang telah melalui praproses dan dibagi dua mejadi data *training*, dan data *testing*. Data akan dilatih menggunakan metode *LSTM* yang dimana metode ini sangat bergantung pada *hidden layer*nya. Data *training* akan dimasukkan ke dalam *input layer*, pada *hidden layer input*

data akan dilakukan proses aktivasi untuk mengetahui perilaku apa yang akan dikenakan pada data tersebut. Dengan menggunakan 3 *hidden layer* yang pada 2 layer pertamanya menggunakan layer *LSTM* dengan 50 neuron pada dua unit pertama, 25 neuron *Dense layer* pada *hidden layer* terakhir dan 1 neuron *Dense layer* pada unit terakhir sebagai *Output layer*, menggunakan *batch size* sebesar 64, menggunakan *optimizer* Adam dan dilakukan analisis terhadap nilai *Root Mean Squared Error* (RMSE) yang didapat.

Eksperimen dilakukan dengan tujuan mengetahui nilai *epoch* yang paling baik. Jumlah *epoch* akan diinisiasi 10 dan akan dilakukan eksperimen sampai jumlah *epoch* 100 (contoh: 10, 20, 30, 50, 70, 100). Setiap perubahan *epoch* dan *hidden layer* dilakukan pengulangan sebanyak 10 kali untuk memastikan kekonsistenan data. Lalu akan diambil nilai *RMSE* paling kecil untuk mendapatkan harga prediksi yang paling mendekati harga aktual.

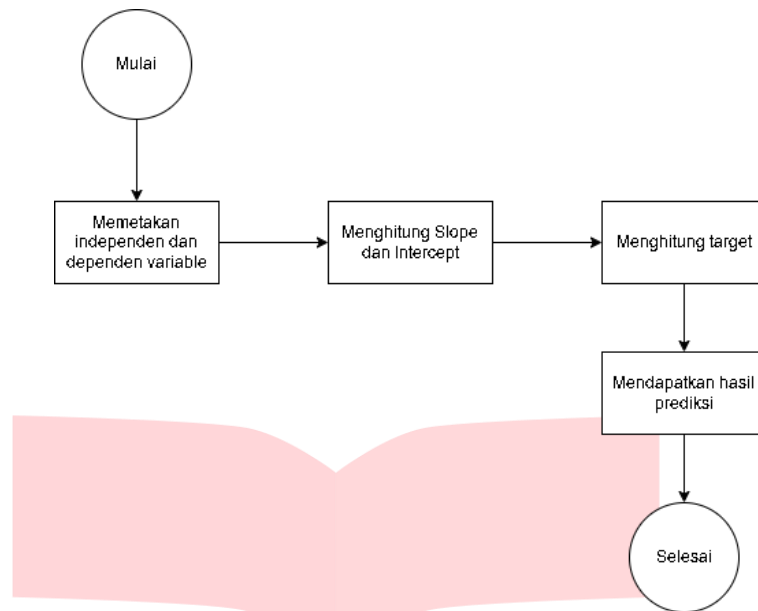


GAMBAR 3
ALUR SISTEM MODEL LSTM

i_t adalah *input gate*, f_t adalah *forget gate*, o_t adalah *output gate*, σ adalah fungsi sigmoid, h_{t-1} adalah *output* dari perhitungan *LSTM* sebelumnya saat $t-1$, x_t adalah *input LSTM* saat ini, c_t adalah *cell state* atau memori saat waktu " t " dan $\hat{c}_t$ adalah *cell state candidate* atau kandidat memori yang akan disatukan dengan *cell state*.

E. Multiple Linear Regression (MLR)

Pada metode *Multiple Linear Regression (MLR)* ini menggunakan data yang sama yang digunakan pada metode *Long Short-Term Memory (LSTM)*. Data juga dibagi ke dalam bagian data *training* dan data *testing*, lalu data *training* digunakan oleh model *MLR* untuk belajar. Model ini menghitung nilai hubungan antara variabel independen dan dependennya satu persatu.



GAMBAR 4
ALUR SISTEM MODEL MLR

IV. HASIL DAN PEMBAHASAN

A. Skenario Pengujian

Dalam pengujian yang dilakukan menggunakan beberapa data seperti yang sudah di siapkan yaitu data harga mata uang kripto dan data *Covid-19*. Namun akan dilakukan pengujian dengan dua skenario agar mengetahui apakah data yang digunakan mempengaruhi perubahan harga dan prediksi harga.

Skenario pertama akan menggunakan data *Close* mata uang kripto XRP/ETH dan data penambahan kasus positif *Covid-19* di dunia menjadi variabel independennya. Untuk variabel dependennya akan menggunakan data *Close* mata uang kripto XRP/ETH dan memprediksi harga di hari berikutnya.

Skenario kedua akan menggunakan data *Close* mata uang kripto XRP/ETH dan BTC sebagai variabel independen, dan menggunakan data *Close* mata uang kripto XRP/ETH sebagai variabel dependen. Menggunakan data BTC sebagai variabel prediktornya disebabkan karena kenaikan atau penurunan harga dari BTC diasumsikan mempengaruhi harga dari mata uang kripto yang lainnya karena BTC adalah mata uang kripto paling populer dan diasumsikan sebagai perubahan musim mata uang kripto bila harga BTC mengalami penurunan atau kenaikan yang cukup besar.

TABEL 2
HASIL PENGUJIAN LSTM PADA DATA XRP

Skenario 1		Skenario 2	
Epoch	RMSE	Epoch	RMSE
10	0.006095	10	0.000628
20	0.003969	20	0.000240
30	0.003433	30	0.005350
50	0.011134	50	0.001249
70	0.002425	70	0.001565
100	0.001327	100	0.003619

TABEL 3
HASIL PENGUJIAN LSTM PADA DATA ETH

Skenario 1		Skenario 2	
Epoch	RMSE	Epoch	RMSE
10	19.2526	10	0.1286
20	8.5906	20	0.7609
30	3.4347	30	1.9531
50	1.8454	50	1.1743
70	13.9909	70	3.4198
100	0.5543	100	1.7479

B. Analisis Hasil LSTM

Dari hasil pengujian yang dilakukan menggunakan *LSTM* terhadap data *cryptocurrency XRP* dan *ETH*, terlihat bahwa pada setiap skenario pengujian jumlah *epoch* yang menghasilkan error terkecil berada pada jumlah *epoch* yang berbeda. Hal ini dapat disebabkan oleh pembelajaran mesin itu sendiri, karena hasil dari pembelajaran mesin dapat berubah-ubah pada setiap pembelajaran yang dilakukan.

C. Analisis Hasil MLR

TABEL 4
HASIL PENGUJIAN MLR PADA DATA XRP

RMSE Skenario 1	RMSE Skenario 2
0.02013	0.02072

TABEL 5
HASIL PENGUJIAN MLR PADA DATA ETH

RMSE Skenario 1	RMSE Skenario 2
88.6579	87.6505

Dari hasil pengujian yang dilakukan menggunakan *MLR* terhadap data *cryptocurrency XRP* dan *ETH*, dapat terlihat pada hasil pengujiannya bahwa nilai *error* dari setiap skenario tidak begitu jauh berbeda, untuk data mata uang kripto *XRP* hasil terbaik didapat pada skenario 1 dan untuk

mata uang kripto *ETH* didapat pada skenario 2.

D. Perbandingan Analisis Performa Metode LSTM dan MLR

TABEL 6
HASIL RMSE OPTIMAL DARI METODE LSTM DAN MLR

Metode	Data	RMSE Skenario 1	RMSE Skenario 2
LSTM	XRP	0.002425	0.000240
LSTM	ETH	0.5543	0.1286
MLR	XRP	0.02013	0.02072
MLR	ETH	88.6579	87.6505

Nilai RMSE yang dihasilkan oleh kedua model pada mata uang kripto *XRP* dan *ETH* memiliki perbedaan yang cukup jauh, ini disebabkan oleh kedua harga mata uang kripto itu tersendiri. Harga mata uang kripto *XRP* berada dibawah 2 dollar dan harga mata uang kripto *ETH* berada dikisaran ribuan dollar, inilah yang menyebabkan nilainya berbeda jauh karena rentang harga kedua mata uang kripto sudah berbeda.

TABEL 7
HASIL PREDIKSI HARGA OPTIMAL DARI METODE LSTM DAN MLR

Metode	Data	Tanggal	Aktual (\$)	Prediksi S1 (\$)	Prediksi S2 (\$)
LSTM	XRP	22/09/20	0.41518	0.34793	0.35335
LSTM	ETH	22/09/20	1324.38818	1297.25817	1293.89965
MLR	XRP	22/09/20	0.41518	0.38970	0.38791
MLR	ETH	22/09/20	1324.38818	1382.23112	1372.02121

Terlihat dari Tabel 6, bahwa *RMSE* terkecil dihasilkan oleh *LSTM* pada data *XRP* yaitu dengan nilai error 0.000240 karena perbedaan skala harga kedua mata uang kripto tersebut, namun bila kita bandingkan harga aktual di hari terakhir prediksi dengan hasil *RMSE* untuk kedua mata uang kripto untuk *XRP* didapat perbandingan 1/1729.92

dan untuk *ETH* 1/15.11. Dan dapat terlihat juga pada Tabel 6, bahwa Skenario 2 dengan menggunakan data *XRP/ETH* dan *BTC* sebagai variabel prediktor membuktikan bahwa model menghasilkan hasil yang paling optimal.

V. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan menggunakan skenario yang telah dijelaskan, prediksi cryptocurrency dengan menggunakan metode *Long Short-Term Memory (LSTM)* dapat dilakukan dengan baik. Hasil nilai evaluasi *Root Mean Squared Error (RMSE)* terbaik pada model *LSTM* yaitu sebesar 0.000240 dan menyatakan bahwa skenario 2 merupakan skenario terbaik pada pengujian sehingga dapat disimpulkan bahwa data pertambahan kasus positif *Covid-19* tidak berpengaruh besar untuk hasil dari kedua model dan *Covid-19* tidak berpengaruh pada kenaikan maupun penurunan harga mata uang kripto. Tetapi harga mata uang kripto *Bitcoin (BTC)* cukup mempengaruhi harga mata uang kripto yang lain dan hasil dari model itu sendiri, karena *BTC* merupakan koin yang paling populer dan diasumsikan mempengaruhi trend pada pasar mata uang kripto.

Namun hasil ini tidak menutup kemungkinan untuk metode *Multiple Linear Regression* tetap digunakan untuk prediksi atau bahkan menghasilkan nilai yang lebih baik dari *LSTM*. Maka dapat disimpulkan bahwa *Long Short-Term Memory* menghasilkan nilai yang lebih baik dibandingkan dengan *Multiple Linear Regression* untuk prediksi harga *cryptocurrency* pada penelitian ini. Mungkin ini disebabkan karena metode *LSTM* menggunakan algoritma dan proses yang lebih kompleks dibandingkan dengan *MLR* walaupun tidak menutup kemungkinan bahwa *MLR* dapat memberikan performansi yang lebih baik dibandingkan *LSTM*.

Walaupun hasil pada penelitian ini menghasilkan nilai seperti demikian, kedepannya penelitian ini masih dapat dikembangkan dengan menggunakan data yang lebih relevan lagi dengan target penelitian. Melakukan pengaturan performansi yang lebih baik lagi untuk kedua model yaitu *LSTM* dan *MLR* dan tidak menutup kemungkinan juga untuk menggunakan metode *Machine Learning* lainnya untuk melakukan prediksi maupun regresi.

REFERENSI

- [1] S. Corbet, Y. (Greg) Hou, Y. Hu, C. Larkin, B. Lucey, and L. Oxley, "Cryptocurrency

liquidity and volatility interrelationships during the COVID-19 pandemic," *Financ. Res. Lett.*, no. April, p. 102137, 2021, doi: 10.1016/j.frl.2021.102137.

- [2] F. Fang *et al.*, "Cryptocurrency trading: A comprehensive survey," *arXiv*, 2020.
- [3] A. Mahata, A. Rai, O. Prakash, and M. Nurujjaman, "Modeling and analysis of the effect of Covid-19 on the stock price: v and l-shape recovery," *arXiv*, vol. 574, p. 126008, 2020, doi: 10.1016/j.physa.2021.126008.
- [4] J. Patel, S. Shah, P. Thakkar, and K. Kotecha, "Predicting stock and stock price index movement using Trend Deterministic Data Preparation and machine learning techniques," *Expert Syst. Appl.*, vol. 42, no. 1, pp. 259–268, 2015, doi: 10.1016/j.eswa.2014.07.040.
- [5] A. Yadav, C. K. Jha, and A. Sharan, "Optimizing LSTM for time series prediction in Indian stock market," *Procedia Comput. Sci.*, vol. 167, no. 2019, pp. 2091–2100, 2020, doi: 10.1016/j.procs.2020.03.257.
- [6] S. Shakhla, B. Shah, N. Shah, V. Unadkat, and P. Kanani, "Stock price trend prediction using multiple linear regression," *Int. J. Eng. Sci. Invent.*, vol. 7, no. 10, pp. 29–33, 2018, [Online]. Available: www.ijesi.org.
- [7] M. M. Patel, S. Tanwar, R. Gupta, and N. Kumar, "A Deep Learning-based Cryptocurrency Price Prediction Scheme for Financial Institutions," *J. Inf. Secur. Appl.*, vol. 55, no. August, p. 102583, 2020, doi: 10.1016/j.jisa.2020.102583.
- [8] C. Chalvatzis and D. Hristu-Varsakelis, "High-performance stock index trading via neural networks and trees," *Appl. Soft Comput. J.*, vol. 96, p. 106567, 2020, doi: 10.1016/j.asoc.2020.106567.
- [9] A. Moghar and M. Hamiche, "Stock Market Prediction Using LSTM Recurrent Neural Network," *Procedia Comput. Sci.*, vol. 170, pp. 1168–1173, 2020, doi: 10.1016/j.procs.2020.03.049.
- [10] S. K. Prion and K. A. Haerling, "Making Sense of Methods and Measurements: Simple Linear Regression," *Clin. Simul. Nurs.*, vol. 48, pp. 94–95, 2020, doi: 10.1016/j.ecns.2020.07.004.
- [11] L. S. Aiken, S. G. West, and S. C. Pitts, "Multiple Linear Regression," in *Handbook of Psychology*, American Cancer Society, 2003, pp. 481–507.
- [12] D. Enke, M. Grauer, and N. Mehdiyev, "Stock market prediction with Multiple Regression, Fuzzy type-2 clustering and neural networks,"

Procedia Comput. Sci., vol. 6, pp. 201–206,
2011, doi: 10.1016/j.procs.2011.08.038.

