

## Prediksi Harga pada Trading Forex Pair USDCHF Menggunakan Regresi Linear

### *Price Prediction on USDCHF Pair Forex Trading Using Linear Regression*

Mohammad Edi<sup>1\*</sup>, Ema Utami<sup>2</sup>, Ainul Yaqin<sup>3</sup>

Program Studi Magister Teknik Informatika, Universitas Amikom Yogyakarta, Yogyakarta, Indonesia

\*E-mail: mohammad.edi100@students.amikom.ac.id

#### Abstrak

Dalam era globalisasi, perdagangan bebas tumbuh pesat dan teknologi berkembang mempengaruhi persaingan ekonomi. Forex, perdagangan valuta asing, menjadi salah satu investasi yang digunakan untuk menghadapi tantangan ini. Analisis teknis dan fundamental digunakan untuk memprediksi pergerakan harga dalam trading forex. Penelitian sebelumnya menggunakan algoritma regresi linear dan teknik lainnya untuk prediksi harga dalam forex. Dalam penelitian ini, algoritma regresi linear digunakan untuk memprediksi harga close dalam trading forex karena algoritma tersebut telah banyak digunakan dalam prediksi, kelebihanannya dalam menaksirkan parameter model yang sederhana, dan data yang berbasis runtun waktu. Selain itu, algoritma ini dapat melakukan analisis dengan menggunakan banyak variabel bebas sehingga hasil prediksi dapat lebih akurat. Tujuan penelitian ini adalah membuat model prediksi harga forex untuk mempermudah trader melakukan prediksi harga. Dataset sebanyak 2066 data diperoleh melalui software metatrader dan diproses melalui tahap preprocessing. Model regresi linear dibuat menggunakan 5 skenario dan evaluasi dilakukan menggunakan nilai Mean Squared Error (MSE) dan Root Mean Square Error (RMSE) untuk memilih model terbaik. Hasilnya menunjukkan bahwa regresi linear mampu memprediksi harga penutupan pada pair USDCHF. Model regresi linear terbaik diperoleh menggunakan variabel bebas pada skenario 1, yaitu variabel Open, dengan persamaan regresi linear  $y=0,0145+0,9849x$ , MSE terbaik sebesar 0,0000328509 dan RMSE terbaik sebesar 0,0057315705.

**Kata kunci:** Prediksi; Forex; Regresi Linear; MSE; RMSE.

#### Abstract

In the era of globalization, free trade grows rapidly and technology develops, affecting economic competition. Forex, foreign exchange trading, is one of the investments used to face this challenge. Technical and fundamental analysis is used to predict price movements in forex trading. Previous studies have used linear regression algorithms and other techniques for price prediction in forex. In this study, the linear regression algorithm is used to predict closing prices in forex trading because the linear regression algorithm is an algorithm that has been widely used in predictions, its strengths are in estimating simple model parameters and data based on time series. In addition, the linear regression algorithm can perform analysis using several independent variables so that the prediction results can be more accurate. The purpose of this study is to create a forex price prediction model, to make it easier for traders to make price predictions. A dataset of 2066 data was obtained through the metatrader software and processed through the preprocessing stage. The linear regression model was created using 5 scenarios, and the evaluation was carried out using the Mean Squared Error (MSE) and Root Mean Square Error (RMSE) values to select the best model. The results show that linear regression is able to predict the closing price of the USDCHF pair. The best linear regression model is obtained using the independent variable in scenario 1, namely the Open variable, with a linear regression equation of  $y=0.0145+0.9849x$ , the best MSE is 0.0000328509 and the best RMSE is 0.0057315705.

**Keywords:** Prediction; Forex; Linear Regression; MSE; RMSE.

Naskah diterima 29 Mei. 2023; direvisi 02 Sept. 2023; dipublikasikan 1 Okt. 2023.

JAMIKA is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.



## I. PENDAHULUAN

Dalam era globalisasi yang saat ini terjadi, perdagangan bebas berkembang pesat dan berdampak pada persaingan ekonomi yang semakin kompetitif. Hal ini diperkuat dengan terus berkembangnya teknologi yang mampu menyaingi kemampuan manusia. Perkembangan teknologi ini terutama terjadi dalam sistem informasi yang telah bertransformasi menjadi media internet. Dalam konteks ini, manusia dituntut untuk selalu berusaha, berpikir, bersaing, dan meningkatkan kompetensinya agar tidak tertinggal dalam segi ekonomi, berpikir, dan akses informasi.

Salah satu cara untuk mengatasi tantangan ini adalah melalui investasi. Investasi melibatkan menempatkan modal pada suatu perusahaan atau aset dengan harapan menghasilkan keuntungan dalam jangka

waktu tertentu. Salah satu bidang investasi yang cukup populer saat ini adalah perdagangan valuta asing atau *Foreign Exchange (Forex)* [1].

*Forex*, juga dikenal sebagai pasar valuta asing atau valas, adalah transaksi perdagangan mata uang suatu negara terhadap negara lain melalui perantara seperti broker atau dealer [2]. Contohnya adalah membandingkan Euro (EUR) dengan Dolar Amerika Serikat (USD) atau Dolar Amerika Serikat (USD) dengan Franc Swiss (CHF). Pasar *forex* beroperasi secara global, sehingga transaksi dapat terjadi selama 24 jam sehari. Menurut Bank International for Settlement (BIS), rata-rata perputaran harian dalam transaksi *forex* diperkirakan mencapai \$3,21 triliun[3].

Dalam menghadapi tantangan globalisasi dan persaingan ekonomi yang semakin ketat, investasi dalam perdagangan *forex* dapat menjadi salah satu pilihan. Namun, penting untuk memahami risiko yang terkait dengan perdagangan ini dan melengkapi diri dengan pengetahuan dan strategi yang tepat. Sebelum terlibat dalam perdagangan *forex*, disarankan untuk melakukan riset yang cermat, mengikuti pelatihan yang relevan, dan berhubungan dengan profesional keuangan yang berpengalaman guna membantu mengurangi risiko dan meningkatkan peluang kesuksesan dalam investasi ini.

Ada dua jenis analisis saat menganalisis pasar *forex* [4][5]: (1) analisis teknis dan (2) analisis fundamental. Analisis ini penting untuk memprediksi pergerakan harga mata uang dan mengambil keputusan investasi yang lebih baik.

Analisis teknis melibatkan pengamatan terhadap grafik dan pola pergerakan harga historis untuk memprediksi kemana arah pergerakan harga selanjutnya. Dalam analisis ini, pedagang menggunakan alat dan indikator teknis seperti garis tren, pola candlestick, indikator momentum, dan lainnya. Tujuan dari analisis teknis adalah untuk mengidentifikasi tren, level support dan resistance, serta peluang untuk masuk atau keluar dari perdagangan. Sementara itu, analisis fundamental berfokus pada faktor-faktor ekonomi, sosial, dan politik yang dapat mempengaruhi nilai mata uang. Analisis ini melibatkan pemantauan berita dan peristiwa yang dapat mempengaruhi pasar, seperti kebijakan moneter, data ekonomi, peristiwa politik, dan lain sebagainya. Analisis fundamental membantu trader memahami faktor-faktor yang dapat mempengaruhi penawaran dan permintaan mata uang, sehingga mereka dapat mengambil keputusan yang lebih informasi.

Mengetahui harga *forex* di masa depan adalah suatu hal yang sangat penting bagi *trader* [6], namun tidak semua *trader* mampu untuk melakukan prediksi harga. Maka dibutuhkan teknik untuk memprediksi harga [7]. Proses prediksi harus mengupayakan keakuratan hasil prediksi, meskipun terdapat ketidakpastian dari pergerakan harga dalam pasar *forex* karena faktor tertentu. Tujuan penelitian ini, yaitu pembuatan model untuk prediksi harga pada pasar *forex*, untuk mempermudah *trader* dalam melakukan prediksi harga.

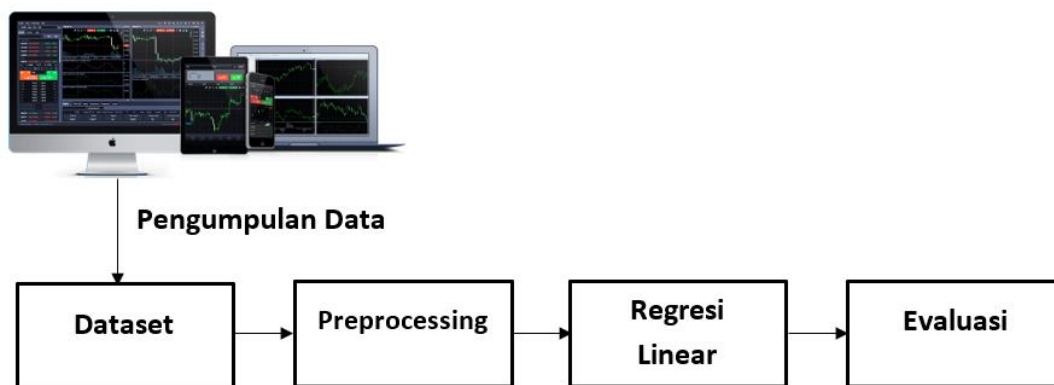
Salah satu cara untuk prediksi nilai harga dalam *trading forex* adalah dengan memanfaatkan teknologi dibidang *Artificial Intelligence*. Penelitian tentang prediksi harga telah banyak dilakukan dengan implementasi beberapa algoritma untuk dibandingkan tingkat akurasi prediksi. Adapun penelitian sebelumnya terkait prediksi harga, yaitu dilakukan oleh Wijaya[1], penelitian prediksi harga pada *pair EURUSD* dan *GBPUSD* dengan metode *Long Short Term Memory (LSTM)*. Pada *pair EUR/USD* menggunakan model 5 input (*open* dan *close* 4 hari). Dan *pair GBP/USD* menggunakan model 3 input (*open* dan *close* 2 hari). Jumlah profit yang diperoleh, *loss* MSE dan hasil akurasi MSE semuanya adalah yang terbaik jika dibandingkan model dengan jumlah input lainnya. Penelitian selanjutnya oleh Falentino Sembiring [8], tentang analisis tingkat akurasi *moving average* dalam memprediksi pergerakan uang elektronik Bitcoin. Penelitian ini bertujuan untuk mengidentifikasi dan menganalisis data perubahan nilai mata uang elektronik dengan menggunakan metode prediksi pergerakan harga rata-rata atau *moving average*. Hasil dari penelitian ini dan hasil pengujian menunjukkan bahwa dengan teknik *Buy*/beli persentase prediksi dengan benar sebesar 62.86%, sedangkan dengan teknik *Sell*/jual persentase prediksi dengan benar hanya sebesar 25% pada periode yang sama. Penelitian lainnya oleh Imam Priyadi [9] yang menggunakan regresi linear. Penelitian ini untuk prediksi harga emas dengan pengolahan data menggunakan Rapidminer. Penelitian ini merekomendasikan regresi linear untuk melakukan prediksi yang dapat dijadikan acuan untuk trading pada metatrader. Hasil rekomendasi tersebut menghasilkan pencapaian akurasi sebesar 85%. Penelitian selanjutnya oleh Benaya Pradito [10], tentang prediksi nilai kurs mata uang dengan membandingkan algoritma regresi linear dan *neural network*. Hasil penelitian disimpulkan bahwa algoritma regresi linear merupakan algoritma yang paling sesuai untuk melakukan prediksi kurs mata uang. Penelitian selanjutnya oleh Janur Syah Putra [11], tentang prediksi harga saham Bank BRI menggunakan algoritma regresi linear. Rasio pembagian data yang digunakan 80:20 menghasilkan akurasi terbaik 0.8992 untuk data training dan 0.9125 untuk data testing. prediksi dari model yang digunakan penulis bisa dikatakan masuk kategori akurat karena menghasilkan nilai *error* yang kecil. Penelitian berikutnya oleh Jevin Putra Halawa [12], tentang prediksi harga saham dengan algoritma regresi linear. Hasil penelitian ini menghasilkan model prediksi harga saham dengan baik, dengan nilai RMSE 0.898.

Penelitian selanjutnya oleh Aziz Sarkar [13] menggunakan *machine learning* algoritma regresi linear pada EUR/USD untuk memprediksi pergerakan masa depan. Hasil penelitian ini, persentase akurasi ramalan data harian lebih tinggi dari perkiraan data per jam pada tahap pengujian. Penelitian selanjutnya oleh Hilman Winnos [14], membandingkan metode regresi linear berganda dan *Autoregressive Integrated Moving Average* (ARIMA) untuk prediksi saham PT. BSI, Tbk. Hasil perbandingan antara model ARIMA dan regresi linear berganda menunjukkan bahwa model persamaan dengan menggunakan metode regresi linear berganda mendapatkan nilai akurasi tertinggi, yaitu 98,9%.

Berdasarkan uraian permasalahan dan pemaparan beberapa penelitian yang telah dilakukan sebelumnya, maka penelitian ini menggunakan algoritma regresi linear dalam memprediksi harga *close* pada *trading forex*, karena algoritma regresi linear merupakan algoritma yang telah banyak digunakan dalam prediksi [9]–[14], kelebihanannya dalam menaksirkan parameter model yang sederhana dan data yang berbasis runtun waktu. Selain itu, algoritma regresi linear dapat melakukan analisis dengan menggunakan beberapa variabel bebas sehingga hasil prediksi bisa lebih akurat [15]. Dalam pembentukan model algoritma regresi linear dalam penelitian ini, dilakukan variasi variabel bebas untuk mendapatkan model regresi linear terbaik berdasarkan evaluasi dengan menggunakan nilai MSE dan *Root Mean Squared Error* (RMSE). Dataset yang digunakan dalam penelitian ini dikumpulkan melalui software metatrader, berbeda dengan penelitian sebelumnya yang memperoleh dataset dari [www.investing.com](http://www.investing.com) [9], [11], [12], website Yahoo Finance [14].

## II. METODE PENELITIAN

Tahapan-tahapan yang dilakukan dalam penelitian ini, yaitu dimulai dari pengumpulan dataset, tahapan preprocessing, regresi linear dan tahap evaluasi. Tahapan metode penelitian yang digunakan dalam penelitian ini ditunjukkan pada Gambar 1.



Gambar 1. Metode Penelitian

### Dataset

Penelitian ini menggunakan data harian trading forex pair USD/CHF yang diambil dari perangkat lunak metatrader, sampel dataset ditunjukkan pada Tabel 1. metatrader adalah platform/aplikasi untuk akses perdagangan valas yang sering digunakan oleh para trader [16]. Untuk download software metatrader bisa melalui [www.metatrader4.com](http://www.metatrader4.com). Jumlah data yang berhasil dikumpulkan sejumlah 2066 data histori harian USDCHF mulai tanggal 1 Desember 2014-16 November 2022, yang berisi *Date*, *Time*, *Open*, *High*, *Low*, *Close*, dan *Volume*. Sabtu dan Minggu adalah hari libur di pasar *forex* [13]. Jumlah data pada dataset terdiri sebanyak 2066 data. Tipe data pada setiap variabel adalah *object* (*Date*, *Time*), *float64* atau bilangan pecahan (*Open*, *High*, *Low*, *Close*) dan *int64* atau bilangan bulat (*Volume*) seperti ditunjukkan dalam Gambar 2.

TABEL 1  
SAMPEL DATASET

Date	Time	Open	High	Low	Close	Volume
2014.12.01	00.00	0.9654	0.9687	0.9615	0.9645	4964
2014.12.02	00.00	0.9646	0.9729	0.9641	0.9720	3655
2014.12.03	00.00	0.9721	0.9782	0.9714	0.9774	4244
2014.12.04	00.00	0.9775	0.9793	0.9648	0.9712	6139

Date	Time	Open	High	Low	Close	Volume
2014.12.05	00.00	0.9713	0.9799	0.9702	0.9789	4988
...	...	...	....	...	...	...
2022.11.10	00.00	0.9840	0.9896	0.9629	0.963	8146
2022.11.11	00.00	0.9631	0.968	0.9396	0.9415	8720
2022.11.14	00.00	0.9418	0.9488	0.9404	0.9435	6817
2022.11.15	00.00	0.9436	0.9474	0.9356	0.943	9743
2022.11.16	00.00	0.9429	0.9468	0.9385	0.9443	7312

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2066 entries, 0 to 2065
Data columns (total 7 columns):
#   Column  Non-Null Count  Dtype
---  -
0   Date    2066 non-null   object
1   Time    2066 non-null   object
2   Open    2066 non-null   float64
3   High    2066 non-null   float64
4   Low     2066 non-null   float64
5   Close   2066 non-null   float64
6   Volume  2066 non-null   int64
dtypes: float64(4), int64(1), object(2)
memory usage: 113.1+ KB
```

Gambar 2. Tipe Data Pada Variabel Dalam Dataset

### Preprocessing

Setelah dataset terkumpul, tahap selanjutnya adalah tahap *preprocessing* data, yaitu melakukan pembersihan data terhadap *missing value*. Data yang akan dibersihkan adalah data ganda dan data yang bernilai kosong (bukan nol). Gambar 3 menampilkan hasil pengecekan *missing value*, menunjukkan bahwa dalam dataset tidak ada yang bernilai kosong, sehingga data bisa diolah pada tahapan selanjutnya.

```
: Date      0
   Time      0
   Open      0
   High      0
   Low       0
   Close     0
   Volume    0
dtype: int64
```

Gambar 3. Source code dan hasil cek *missing value*

### Regresi Linear

Regresi linear merupakan metode yang digunakan untuk membangun model hubungan antara 1 variabel dependen dan 1 atau lebih variabel independen. Regresi linear ada 2 macam, yaitu: regresi linear sederhana dan regresi linear berganda.

Regresi linear sederhana adalah teknik yang digunakan untuk membangun model hubungan antara 1 variabel terikat dengan 1 variabel bebas [17], [18]. Persamaan (1) merupakan rumus regresi linear sederhana, dan untuk garis kemiringan  $b$  dapat ditemukan dengan persamaan (2).  $a$  merupakan konstanta regresi linear, yang dihitung menggunakan persamaan (3), dimana  $y$  merupakan variabel terikat yang akan diprediksi,  $x$  merupakan variabel bebas,  $\bar{x}$  merupakan rata-rata nilai variabel bebas ( $x$ ), dan  $\bar{y}$  merupakan rata-rata nilai variabel terikat ( $y$ ). Persamaan (4) dan (5), merupakan persamaan untuk mencari nilai rata-rata nilai variabel  $x$  dan  $y$ , dimana hasilnya akan diterapkan pada persamaan (2). Sebagai ilustrasi penerapan regresi linear sederhana dalam penelitian ini menggunakan variabel bebas ‘*Open*’ untuk memprediksi “*Close*” sebagai variabel terikat seperti ditunjukkan pada Tabel 2.

$$y = a + bx \tag{1}$$

$$b = \frac{\sum xy - n\bar{x}\bar{y}}{\sum x^2 - n\bar{x}^2} \tag{2}$$

$$a = \bar{y} - b\bar{x} \tag{3}$$

$$\bar{x} = \frac{\sum x}{n} \tag{4}$$

$$\bar{y} = \frac{\sum y}{n} \tag{5}$$

TABEL 2  
 CONTOH DATA UNTUK REGRESI LINEAR SEDERHANA

Date	Open (x)	Close (y)	$x^2$	$y^2$	xy
2014.12.18	0,9728	0,9798	0,9463	0,9600	0,9531
2014.12.19	0,9799	0,9843	0,9602	0,9688	0,9645
2014.12.22	0,9840	0,9837	0,9683	0,9677	0,9680
2014.12.23	0,9836	0,9877	0,9675	0,9756	0,9715
2014.12.24	0,9876	0,9863	0,9754	0,9728	0,9741
2014.12.26	0,9862	0,9874	0,9726	0,9750	0,9738
2014.12.29	0,9870	0,9897	0,9742	0,9795	0,9768
2014.12.30	0,9896	0,9888	0,9793	0,9777	0,9785
2014.12.31	0,9887	0,9936	0,9775	0,9872	0,9824
2015.01.02	0,9943	1,0016	0,9886	1,0032	0,9959
Total	9,8537	9,8829	9,7099	9,7675	9,7386

$$b = \frac{9,7386 - 10\left(\frac{9,8537}{10}\right)\left(\frac{9,8829}{10}\right)}{9,7099 - 10\left(\frac{9,8537}{10}\right)^2} = 0,8765$$

$$a = \left(\frac{9,8829}{10}\right) - 0,8765\left(\frac{9,8537}{10}\right) = 0,1246$$

Berdasarkan data pada Tabel 2, nilai  $a$  dan  $b$  dapat dihitung menggunakan persamaan (2) dan (3). Nilai  $a$  diperoleh sebesar 0,9883 dan  $b$  sebesar 0,1246. Berdasarkan perhitungan diatas, model persamaan regresi linear adalah  $y = 0,1246 + 0,8765x$ . selain menggunakan regresi linear sederhana dalam penelitian ini juga diterapkan pembuatan model regresi linear berganda.

Regresi linear berganda adalah metode yang digunakan untuk membangun model hubungan antara 2 atau lebih variabel bebas dengan 1 variabel terikat [19]–[21]. Pada umumnya persamaan regresi linear berganda dapat dituliskan seperti persamaan (6). Bila terdapat 2 variabel bebas, yaitu  $X_1$  dan  $X_2$ , persamaan regresi linear berganda dapat dituliskan seperti persamaan (7).  $a$  merupakan konstanta regresi linear. Garis kemiringan  $b$  pada regresi linear berganda. Untuk membuat persamaan regresi linear berganda, dapat menerapkan persamaan (8)–(10). sebagai ilustrasi penggunaan pembentukan persamaan regresi linear dalam penelitian ini dapat menggunakan tabel bantu, dengan variabel bebas *Open* sebagai  $X_1$ , *Close-1* merupakan harga penutupan 1 hari sebelumnya yang dijadikan  $X_2$ , *Close* sebagai variabel terikat ( $Y$ ) seperti ditunjukkan pada Tabel 3.

$$Y = a + b_1X_1 + b_2X_2 + \dots + b_nX_n \tag{6}$$

$$Y = a + b_1X_1 + b_2X_2 \tag{7}$$

$$\sum Y = an + b_1 \sum X_1 + b_2 \sum X_2 \tag{8}$$

$$\sum X_1Y = a \sum X_1 + b_1 \sum X_1^2 + b_2 \sum X_1X_2 \tag{9}$$

$$\sum X_2 Y = a \sum X_2 + b_1 \sum X_1 X_2 + b_2 \sum X_2^2 \quad (10)$$

TABEL 3  
 CONTOH DATA UNTUK REGRESI LINEAR BERGANDA

Date	Open (X <sub>1</sub> )	Close-1 (X <sub>2</sub> )	Close (Y)	X <sub>1</sub> <sup>2</sup>	X <sub>2</sub> <sup>2</sup>	X <sub>1</sub> Y	X <sub>2</sub> Y	X <sub>1</sub> X <sub>2</sub>
2014.12.18	0,9728	0,9729	0,9798	0,9463	0,9465	0,9531	0,9532	0,9464
2014.12.19	0,9799	0,9798	0,9843	0,9602	0,9600	0,9645	0,9644	0,9601
2014.12.22	0,9840	0,9843	0,9837	0,9683	0,9688	0,9680	0,9683	0,9686
2014.12.23	0,9836	0,9837	0,9877	0,9675	0,9677	0,9715	0,9716	0,9676
2014.12.24	0,9876	0,9877	0,9863	0,9754	0,9756	0,9741	0,9742	0,9755
2014.12.26	0,9862	0,9863	0,9874	0,9726	0,9728	0,9738	0,9739	0,9727
2014.12.29	0,9870	0,9874	0,9897	0,9742	0,9750	0,9768	0,9772	0,9746
2014.12.30	0,9896	0,9897	0,9888	0,9793	0,9795	0,9785	0,9786	0,9794
2014.12.31	0,9887	0,9888	0,9936	0,9775	0,9777	0,9824	0,9825	0,9776
2015.01.02	0,9943	0,9936	1,0016	0,9886	0,9872	0,9959	0,9952	0,9879
Total	9,8537	9,8542	9,8829	9,7099	9,7108	9,7386	9,7391	9,7103

Berdasarkan perhitungan data pada Tabel 3, besaran nilai a, b<sub>1</sub> dan b<sub>2</sub> dapat dihitung menggunakan persamaan (8)-(10). Nilai a yang diperoleh adalah 0,03572 b<sub>1</sub> adalah 0, dan b<sub>2</sub> adalah 0,9667. Ketiga nilai ini diperoleh dari perhitungan:

- Masukkan nilai-nilai variabel yang terdapat pada Tabel 3 kedalam persamaan (8)-(10).

$$9,8829 = 10a + 9,8537b_1 + 9,8542b_2$$

$$9,7386 = 9,8537a + 9,7099b_1 + 9,7103b_2$$

$$9,7391 = 9,8542a + 9,7103b_1 + 9,7108b_2$$

- Persamaan (8) dikali 9,8537 dan persamaan (9) dikali 10.

$$97,3831 = 98,5370a + 97,095b_1 + 97,100b_2$$

$$97,3860 = 98,5370a + 97,099b_1 + 97,103b_2$$

$$\underline{-0,0029 = 0a - 0,0031b_1 - 0,0030b_2}$$

$$-0,0029 = -0,0031b_1 - 0,0030b_2 \quad (11)$$

- Persamaan (8) dikali 9,8537 dan persamaan (10) dikali 10.

$$97,3881 = 98,5420a + 97,100b_1 + 97,105b_2$$

$$97,3910 = 98,5420a + 97,103b_1 + 97,108b_2$$

$$\underline{-0,0029 = 0a - 0,0030b_1 - 0,0030b_2}$$

$$-0,0029 = -0,0030b_1 - 0,0030b_2 \quad (12)$$

- Eliminasi variabel b<sub>2</sub> menggunakan persamaan (11) dan persamaan (12).

$$-0,0029 = -0,0031b_1 - 0,0030b_2$$

$$\underline{-0,0029 = -0,0030b_1 - 0,0030b_2}$$

$$0 = -0,0001b_1 - 0b_2$$

$$b_1 = \frac{0}{-0,001}$$

$$b_1 = 0$$

- Masukkan nilai b<sub>1</sub> ke persamaan (11).

$$-0,0029 = -0,0031 \cdot 0 - 0,0030b_2$$

$$-0,0029 = -0,0030b_2$$

$$b_2 = \frac{-0,0029}{-0,0030} = 0,9667$$

- Masukkan nilai b<sub>1</sub> dan b<sub>2</sub> ke persamaan (8).

$$9,8829 = 10a + 9,8537 \cdot 0 + 9,8542 \cdot 0,9667$$

$$9,8829 = 10a + 0 + 9,5257$$

$$9,8829 = 10a + 9,5257$$

$$0,3572 = 10a$$

$$a = \frac{0,3572}{10} = 0,03572$$

Berdasarkan hasil perhitungan di atas, model persamaan regresi linear yang terbentuk adalah  $Y = 0,03572 + X_1 + 0,9667X_2$ . Penelitian ini membuat beberapa model regresi linear dengan menggunakan variabel bebas yang berbeda disetiap model untuk melakukan variabel terikat *close* ( $y$ ). Variabel bebas yang digunakan dalam penelitian ini ditunjukkan pada Tabel 4. Pembuatan model regresi linear menggunakan bahasa pemrograman *Python* untuk mendapatkan model persamaan dari data train yang telah dikumpulkan. Bahasa pemrograman *Python* dipilih karena bahasanya ini memiliki banyak keuntungan, terutama dalam pemrograman berbasis pembelajaran mesin [22].

Terdapat beberapa skenario untuk membuat model regresi linear untuk prediksi harga *Close* seperti pada Tabel 4, Perbedaan dari masing-masing skenario adalah terletak pada variabel bebas yang digunakan, *Open* merupakan harga pembukaan pada hari yang mau diprediksi, *Close-1* merupakan harga penutupan 1 hari sebelumnya, *Close-2* merupakan harga penutupan 2 hari sebelumnya, *Close-3* merupakan harga penutupan 3 hari sebelumnya, dan *Close-4* merupakan harga penutupan 4 hari sebelumnya. *Pseudocode* dalam membentuk model regresi linear ditunjukkan pada Gambar 4.

```
# Pseudocode for Linear Regression

# Define the dataset
data = [(skenario_1, y, date),
        (skenario_2, y, date),
        (skenario_3, y, date),
        (skenario_4, y, date),
        (skenario_5, y, date)]

# Split data into training and testing sets based on date
train_data = [(xi, yi) for xi, yi, date in data if date < '2022-01-01']
test_data = [(xi, yi) for xi, yi, date in data if date >= '2022-01-01']

# Initialize variables to track best model and its metrics
best_model = None
best_mse = infinity

# Loop through different combinations of independent variables
for variable_combination in possible_combinations:
    # Prepare the data for the selected variables
    selected_train_data = [(xi[var] for var in variable_combination) + (yi,) for xi, yi in train_data]
    selected_test_data = [(xi[var] for var in variable_combination) + (yi,) for xi, yi in test_data]

    # Perform linear regression on the selected variables
    coeff_matrix = calculate_coefficients(selected_train_data)

    # Calculate predictions for the test data
    predictions = [predict(coeff_matrix, xi) for xi, _ in selected_test_data]

    # Calculate MSE and RMSE for the current model
    mse = calculate_mse(selected_test_data, predictions)
    rmse = sqrt(mse)

    # Compare with the best model so far
    if mse < best_mse:
        best_mse = mse
        best_model = {
            'variables': variable_combination,
            'coefficients': coeff_matrix,
            'mse': mse,
            'rmse': rmse
        }

# Output the best model and its metrics
print("Best Model:")
print("Variables:", best_model['variables'])
print("Coefficients:", best_model['coefficients'])
print("MSE:", best_model['mse'])
print("RMSE:", best_model['rmse'])
```

Gambar 4. *Pseudocode* model regresi linear

TABEL 4  
 SKENARIO MODEL REGRESI LINEAR

Skenario	Variabel Bebas (x)
1	<i>Open</i>
2	<i>Open, Close-1</i>
3	<i>Open, Close-1, Close-2</i>
4	<i>Open, Close-1, Close-2, Close-3</i>
5	<i>Open, Close-1, Close-2, Close-3, Close-4</i>

*Pseudocode* yang ditunjukkan pada Gambar 4, adalah skrip yang menerapkan pada data yang mengandung beberapa variabel independen (skenario\_i) dan variabel dependen (y), serta informasi tanggal (date). Tujuan dari skrip pada Gambar 4 adalah untuk mencari model regresi linear terbaik dengan membandingkan beberapa model berdasarkan nilai MSE dan RMSE, menggunakan data uji yang memiliki tanggal setelah 1 Januari 2022.

### Evaluasi

Pada tahap ini dilakukan pengukuran untuk menentukan performa terbaik dari regresi linear dengan menentukan nilai MSE dan nilai RMSE.

### Mean Square Error

*Mean Squared Error* (MSE) adalah rerata nilai kesalahan kuadrat nilai aktual dengan prediksi [23]. Metode MSE pada umumnya digunakan untuk mengetahui estimasi nilai kesalahan nilai prediksi. Nilai MSE yang nilainya mendekati nol, atau yang nilai terendah, berarti hasil prediksi sesuai dengan nilai sebenarnya dan model tersebut dapat digunakan untuk mencari nilai prediksi.

MSE dihitung dengan mengurangi nilai data sebenarnya dengan nilai data hasil prediksi, kemudian dikuadratkan hasil yang diperoleh lalu menjumlahkan dan membaginya dengan banyaknya data yang ada. Persamaan (13) merupakan persamaan untuk menghitung nilai MSE.

$$MSE = \frac{\sum_{t=1}^n (A_t - F_t)^2}{n} \quad (13)$$

### Root Mean Square Error

*Root Mean Square Error* (RMSE) adalah metode perhitungan dengan cara menghitung selisih nilai sebenarnya dengan nilai prediksi [23]. Nilai RMSE adalah hasil perhitungan akar kuadrat perhitungan MSE pada persamaan (11). Nilai RMSE dengan nilai lebih kecil bisa dipastikan lebih baik nilai akuratnya daripada metode prediksi yang bernilai RMSE lebih besar.

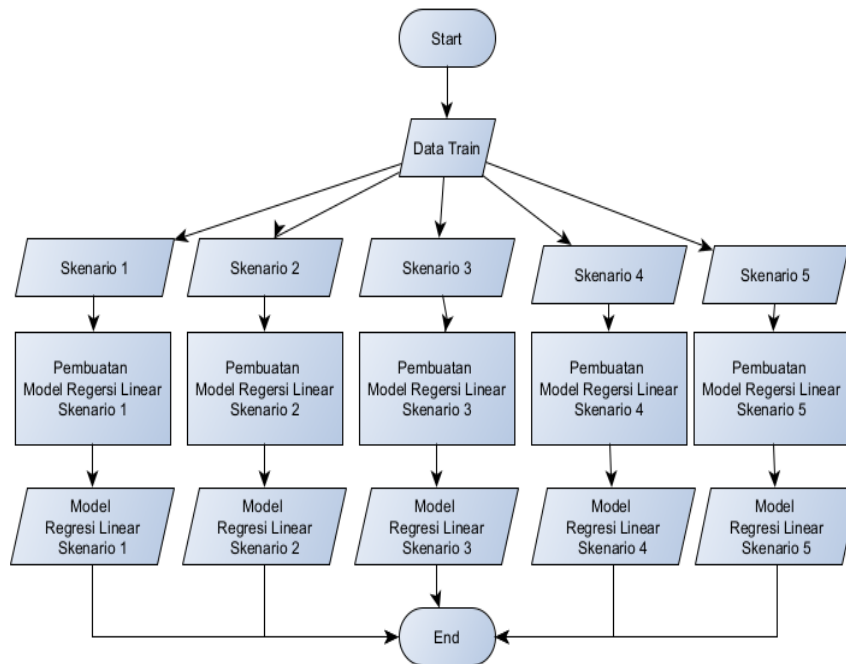
RMSE adalah hasil perhitungan pengurangan nilai sebenarnya dengan nilai hasil prediksi, hasil pengurangan tersebut kemudian dikuadratkan, dan menjumlahkan semuanya, hasil penjumlahan selanjutnya dibagi dengan banyak data. Hasil dari perhitungan tersebut, kemudian dihitung ulang untuk mendapatkan nilai dari akar kuadrat. Persamaan (14) merupakan persamaan untuk perhitungan RMSE.

$$RMSE = \sqrt{\frac{\sum_{t=1}^n (A_t - F_t)^2}{n}} \quad (14)$$

## III. HASIL DAN PEMBAHASAN

Pemodelan regresi linear dan tahapan evaluasi diproses dengan menggunakan bahasa pemrograman python, dengan menggunakan *library sklearn*. Fungsi *library sklearn* untuk membuat model *machine learning* salah satunya adalah regresi linear. Gambar 5 merupakan flowchart dalam pembuatan model regresi linear dalam penelitian ini. Banyaknya model yang dibentuk sebanyak 5 model regresi linear dengan variabel bebas yang berbeda sesuai dengan skenario yang telah ditetapkan pada Tabel 5. Hasil pembentukan model pada penelitian ini ditunjukkan pada Tabel 5. Dimana y adalah *close* yang merupakan variabel terikat pada penelitian ini, x atau x<sub>1</sub> adalah *open*, x<sub>2</sub> adalah *close-1*, x<sub>3</sub> adalah *close-2*, x<sub>4</sub> adalah *close-3*, x<sub>5</sub> adalah *close-4*

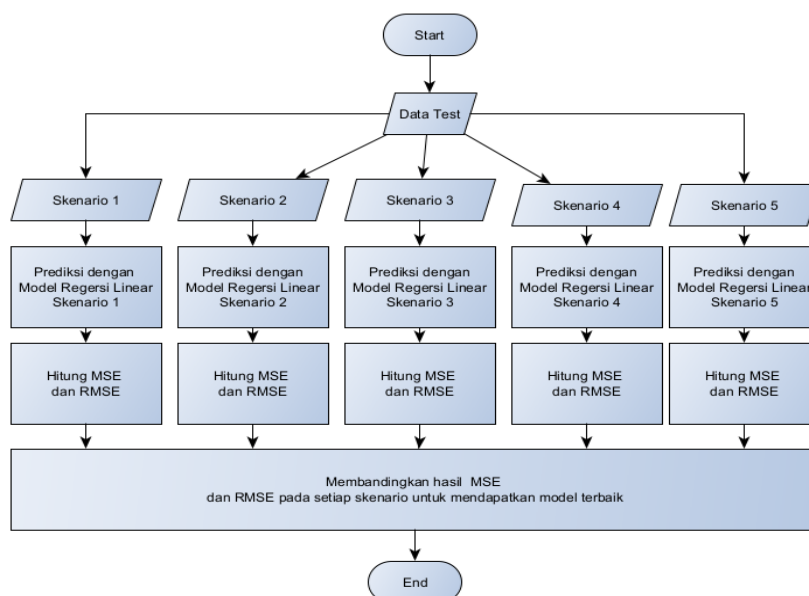




Gambar 5. Flowchart pembentukan model regresi linear

TABEL 5  
HASIL PEMBENTUKAN MODEL PADA TIAP SKENARIO

Skenario	Model Regresi Linear
1	$y = 0,0145 + 0,9849x$
2	$y = 0,0149 + 1,0079x_1 - 0,0230x_2$
3	$y = 0,0141 + 1,0014x_1 - 0,0380x_2 + 0,0219x_3$
4	$y = 0,0137 + 0,9872x_1 - 0,0245x_2 - 0,0029x_3 + 0,0258x_4$
5	$y = 0,0136 + 0,9859x_1 - 0,0234x_2 - 0,0028x_3 + 0,0214x_4 + 0,0046x_5$



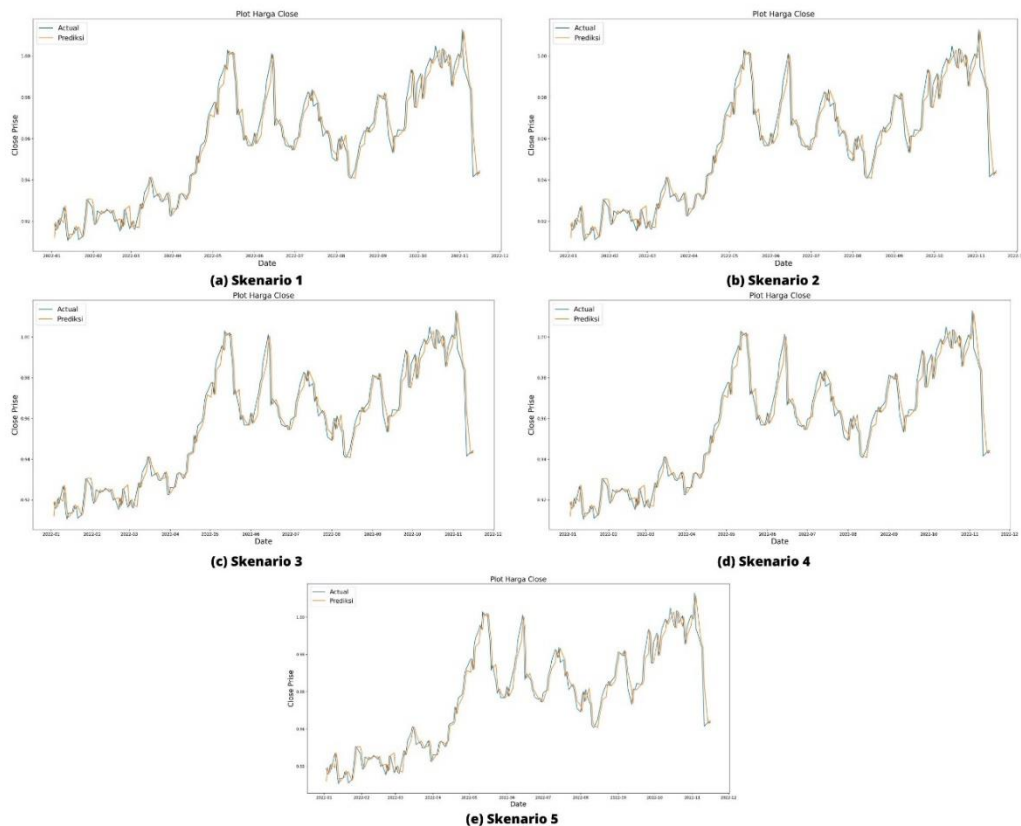
Gambar 6. Flowchart tahapan evaluasi model regresi linear

Setelah pembuatan model regresi linear, tahap berikutnya adalah evaluasi model regresi linear. Gambar 6 merupakan flowchart tahapan evaluasi model regresi linear. Pada tahapan ini menggunakan data test yang telah disiapkan sebelumnya, kemudian ditentukan variabel bebas yang digunakan yang sesuai dengan skenario. Tabel 6 menunjukkan hasil prediksi pada setiap skenario, hasil prediksi diperoleh selisih antara *close* dengan prediksi, dengan rata-rata selisih sebesar 0,0001, hal ini menunjukkan hasil prediksi hampir sama dengan keadaan sebenarnya.

TABEL 6  
HASIL PREDIKSI MODEL REGRESI LINEAR PADA TIAP SKENARIO

Date	Close	Skenario 1		Skenario 2		Skenario 3		Skenario 4		Skenario 5	
		Prediksi	Selisih	Prediksi	Selisih	Prediksi	Selisih	Prediksi	Selisih	Prediksi	Selisih
03/01/2022	0,9188	0,9121	0,0067	0,9121	0,0067	0,9121	0,0067	0,9121	0,0067	0,9121	0,0067
04/01/2022	0,9157	0,9195	-0,0038	0,9195	-0,0038	0,9194	-0,0037	0,9194	-0,0037	0,9194	-0,0037
05/01/2022	0,9163	0,9165	-0,0002	0,9165	-0,0002	0,9165	-0,0002	0,9163	0,0000	0,9163	0,0000
06/01/2022	0,9207	0,9169	0,0038	0,9169	0,0038	0,9168	0,0039	0,9169	0,0038	0,9169	0,0038
07/01/2022	0,9183	0,9214	-0,0031	0,9214	-0,0031	0,9213	-0,0030	0,9213	-0,0030	0,9213	-0,0030
...	...	...	...	...	...	...	...	...	...	...	...
10/11/2022	0,9630	0,9837	-0,0207	0,9837	-0,0207	0,9837	-0,0207	0,9838	-0,0208	0,9838	-0,0208
11/11/2022	0,9415	0,9631	-0,0216	0,9631	-0,0216	0,9635	-0,0220	0,9636	-0,0221	0,9636	-0,0221
14/11/2022	0,9435	0,9421	0,0014	0,9421	0,0014	0,9426	0,0009	0,9431	0,0004	0,9431	0,0004
15/11/2022	0,9430	0,9439	-0,0009	0,9439	-0,0009	0,9438	-0,0008	0,9444	-0,0014	0,9445	-0,0015
16/11/2022	0,9443	0,9432	0,0011	0,9432	0,0011	0,9432	0,0011	0,9431	0,0012	0,9432	0,0011
Rata-rata			0,0001		0,0001		0,0001		0,0001		0,0001

Setelah pembuatan model regresi linear, tahap berikutnya adalah evaluasi model regresi linear. Gambar 6 merupakan flowchart tahapan evaluasi model regresi linear. Pada tahapan ini menggunakan data test yang telah disiapkan sebelumnya, kemudian ditentukan variabel bebas yang digunakan yang sesuai dengan skenario. Tabel 6 menunjukkan hasil prediksi pada setiap skenario, hasil prediksi diperoleh selisih antara *close* dengan prediksi, dengan rata-rata selisih sebesar 0,0001, hal ini menunjukkan hasil prediksi hampir sama dengan keadaan sebenarnya.



Gambar 7. Plot hasil prediksi harga pada setiap skenario (a) Skenario 1, (b) Skenario 2, (c) Skenario 3, (d) Skenario 4, (e) Skenario 5

Hasil prediksi harga *close* pada setiap skenario dalam bentuk grafik ditunjukkan pada Gambar 7. Berdasarkan plot hasil prediksi harga *close* pada setiap skenario pada Gambar 7, regresi linear mampu melakukan prediksi dengan baik dibuktikan grafik *close* dan hasil prediksi hampir sama. Untuk mengetahui kinerja model regresi linear yang telah dibentuk, perlu dilakukan evaluasi pada model regresi linear setiap skenario.

Setelah mengetahui nilai harga prediksi yang dihasilkan oleh model regresi linear yang telah dibentuk, selanjutnya dilakukan evaluasi dengan melakukan perhitungan nilai MSE dan RMSE guna menentukan model terbaik. Semakin kecil MSE dan RMSE menunjukkan metode prediksi yang dihasilkan bisa dikatakan baik [9], [24]. Secara manual perhitungan MSE dapat menggunakan persamaan (13) dan RMSE menggunakan persamaan (14). Secara lengkap hasil perhitungan MSE dan RMSE ditunjukkan pada Tabel 7.

TABEL 7  
 HASIL PERHITUNGAN MSE DAN RMSE PADA SETIAP SKENARIO

Skenario	MSE	RMSE
1	0,0000328509	0,0057315705
2	0,0000328643	0,0057327352
3	0,0000331035	0,0057535656
4	0,0000331819	0,0057603740
5	0,0000332024	0,0057621485

Nilai MSE yang rendah, atau nilainya mendekati nol, berarti hasil prediksi sesuai dengan data sebenarnya. RMSE dengan nilai yang paling mendekati nol bisa dikatakan lebih baik akurasi daripada metode prediksi yang bernilai RMSE lebih besar. Secara umum model yang dihasilkan dalam setiap skenario baik, namun model terbaik dengan menggunakan skenario 1, yaitu dengan variabel bebas yang digunakan *Open*. Berbeda dengan penelitian sebelumnya, Imam Priyadi [9] mendapatkan model terbaik dengan menggunakan variabel bebas *Open*, *High* dan *Low*. Janur Syah Putra [11] yang menggunakan *Date* dan *Close* dengan mendapatkan nilai *error* sebesar 0,0013751. Jevin Putra Halawa [12] yang menggunakan *moving average Open*, *moving average High* dan *moving average Low*, tahap evaluasi menggunakan RMSE, dengan hasil RMSE 0,898. Hilman Winnos [14] yang menggunakan variabel bebas *Open*, *High*, *Low* dan *Volume*, dengan tahap evaluasi MAPE, mendapatkan hasil sebesar sebesar 0,011. Sedangkan penelitian ini, mendapatkan model terbaik dengan menggunakan variabel bebas *Open*, persamaan regresi linear  $y=0,0145+0,9849x$ , dengan menggunakan perhitungan nilai MSE dan RMSE untuk menentukan model regresi linear terbaik, nilai MSE terbaik pada penelitian ini dengan nilai MSE sebesar 0,0000328509 dan nilai RMSE terbaik dengan nilai RMSE sebesar 0,0057315705.

#### IV. KESIMPULAN

Berdasarkan pembahasan yang telah dilakukan, bahwa regresi linear mampu melakukan prediksi harga pada *trading forex* dengan baik, ini dibuktikan dengan hasil rata-rata selisih antara *close* dan hasil prediksi sangat kecil, yaitu sebesar 0,0001. Secara umum model yang dihasilkan dalam setiap skenario baik, karena semua hasil perhitungan MSE dan RMSE mendekati nilai nol, namun model terbaik dengan menggunakan skenario 1, yaitu variabel bebas yang digunakan *Open*, persamaan regresi linear  $y=0,0145+0,9849x$ . Dengan nilai MSE sebesar 0,0000328509 dan nilai RMSE terbaik dengan nilai RMSE sebesar 0,0057315705. Berbeda dengan penelitian sebelumnya, yang menghasilkan nilai MSE dan RMSE lebih besar dari hasil penelitian ini. Saran untuk penelitian berikutnya dapat menggunakan variabel yang lain yang tersedia dalam dataset, serta bisa dikombinasikan dengan nilai indikator analisa teknikal misalnya *moving average*, RSI, dan lain-lain.

#### DAFTAR PUSTAKA

- [1] A. J. Wijaya, W. Swastika, and O. H. Kelana, "Prediksi Harga Foreign Exchange Mata Uang Eur/Usd Dan Gbp/Usd Menggunakan Long Short-Term Memory," *Sainsbertek J. Ilm. Sains Teknol.*, vol. 2, no. 1, pp. 16–31, 2021, doi: 10.33479/sb.v2i1.121.
- [2] A. Sespajayadi, "PREDIKSI NILAI TUKAR VALUTA ASING MENGGUNAKAN METODE GENETIC ALGORITHM-NEURAL NETWORK ( Euro Terhadap US Dollar )," *J. Teknol. Inf. dan Komun.*, vol. 3, pp. 21–28, 2013.
- [3] Rafiqah, "ANALISIS SHARI'AH COMPLIANCE PADA TRADING FOREX ONLINE," *Misykat*, vol. 5, pp. 147–164, 2020.

- [4] M. Ayitey Junior, P. Appiahene, O. Appiah, and C. N. Bombie, "Forex market forecasting using machine learning: Systematic Literature Review and meta-analysis," *J. Big Data*, vol. 10, no. 1, 2023, doi: 10.1186/s40537-022-00676-2.
- [5] T. Titin, "Analisis Pengambilan Keputusan Dalam Transaksi Trading Forex Di Fxindo Regional Lamongan," *J. Ekbis*, vol. 14, no. 2, p. 5, 2015, doi: 10.30736/ekbis.v14i2.121.
- [6] R. R. Elhakim, "Prediksi Nilai Tukar Rupiah Ke Dollar As Menggunakan Metode Arima," *MATHunesa J. Ilm. Mat.*, vol. 8, no. 2, pp. 145–150, 2020, doi: 10.26740/mathunesa.v8n2.p145-150.
- [7] R. Maulana and D. Kumalasari, "Analisis Dan Perbandingan Algoritma Data Mining Dalam Prediksi Harga Saham Ggrm," *J. Inform. Kaputama*, vol. 3, no. 1, pp. 22–28, 2019, [Online]. Available: <https://finance.yahoo.com/quote/GGRM.J>
- [8] F. Sembiring, D. Gustian, A. Erfina, and Y. Vikriansyah, "Analisis Tingkat Akurasi Algoritma Moving Average dalam Prediksi Pergerakan Uang Elektronik Bitcoin," *J. Ilm. Tek. Inform. dan Sist. Inf.*, vol. 10, no. 1, pp. 23–30, 2021, [Online]. Available: <http://ojs.stmik-banjarbaru.ac.id/index.php/jutisi/article/view/577>
- [9] I. Priyadi, J. Santony, and J. Na'am, "Data Mining Predictive Modeling for Prediction of Gold Prices Based on Dollar Exchange Rates, Bi Rates and World Crude Oil Prices," *Indones. J. Artif. Intell. Data Min.*, vol. 2, no. 2, p. 93, 2019, doi: 10.24014/ijaidm.v2i2.6864.
- [10] D. S. Pradito, Benaya. Purnia, "KOMPARASI ALGORITMA LINEAR REGRESSION DAN NEURAL NETWORK UNTUK MEMREDIKSI NILAI KURS," *Evolusi J. Sains dan Manaj.*, vol. 10, no. 2, pp. 64–71, 2022.
- [11] J. S. Putra, R. D. Ramadhani, and A. Burhanuddin, "Prediksi Harga Saham Bank Bri Menggunakan Algoritma Linear Regresion Sebagai Strategi Jual Beli Saham," *J. Dinda Data Sci. Inf. Technol. Data Anal.*, vol. 2, no. 1, pp. 1–10, 2022, doi: 10.20895/dinda.v2i1.273.
- [12] J. P. Halawa and A. Hermawan, "Implementation of Linear Regression Algorithm to Predict Stock Prices Based on Historical Data," *Bit-Tech*, vol. 5, no. 2, pp. 103–112, 2022, doi: 10.32877/bt.v5i2.616.
- [13] M. S. Aziz Sarkar and U. A. M. Ehsan Ali, "EUR/USD Exchange Rate Prediction Using Machine Learning," *Int. J. Math. Sci. Comput.*, vol. 8, no. 1, pp. 44–48, 2022, doi: 10.5815/ijmsc.2022.01.05.
- [14] H. Winnos, R. Septima, and Hu. Gemasih, "Perbandingan Metode Regresi Linier Berganda dan Autoregressive Integrated Moving Average ( ARIMA ) Untuk Prediksi Saham PT. BSI, Tbk.," *Ocean Eng. J. Ilmu Tek. dan Teknol. Marit.*, vol. 1, no. 4, pp. 15–23, 2022.
- [15] T. Indarwati, T. Irawati, and E. Rimawati, "Penggunaan Metode Linear Regression Untuk Prediksi Penjualan Smartphone," *J. Teknol. Inf. dan Komun.*, vol. 6, no. 2, pp. 2–7, 2019, doi: 10.30646/tikomsin.v6i2.369.
- [16] G. S. Firlil, "Sharia Business Ethics Online Forex Trading at Company Signal Providers Volatility 75 Index JR ( VIXJR )," vol. 1, no. 19, pp. 809–818, 2023.
- [17] Harsiti, Z. Muttaqin, and E. Srihartini, "Penerapan Metode Regresi Linier Sederhana Untuk Prediksi Persediaan Obat Jenis Tablet," *JSiI (Jurnal Sist. Informasi)*, vol. 9, no. 1, pp. 12–16, 2022, doi: 10.30656/jsii.v9i1.4426.
- [18] T. N. Putri, A. Yordan, and D. H. Lamkaruna, "Peramalan Penerimaan Mahasiswa Baru Universitas Samudra Menggunakan Metode Regresi Linear Sederhana," *J-TIFA (Jurnal Teknol. Inform.)*, vol. 2, no. 1, pp. 21–27, 2019.
- [19] A. M. Yusuf, *Metode Penelitian Kuantitatif, kualitatif & Penelitian Gabungan*, Edisi Pert. Jakarta: PRENADAMEDIA GROUP, 2014.
- [20] A. F. L. Putra, J. Junaidi, Z. Situmorang, and A. H. Nasyuha, "Regresi Linier Berganda Untuk Memprediksi Jumlah Nasabah," *J. Sci. Soc. Res.*, vol. 5, no. 2, p. 236, 2022, doi: 10.54314/jssr.v5i2.915.
- [21] A. Z. Siregar, "Implementasi Metode Regresi Linier Berganda Dalam Estimasi Tingkat Pendaftaran Mahasiswa Baru," ... *J. Penerapan Sist. Inf. (Komputer dan ...)*, vol. 2, no. 3, pp. 133–137, 2021,

---

[Online]. Available: <http://tunasbangsa.ac.id/pkm/index.php/kesatria/article/view/73>

- [22] A. S. Bayangkari Karno, “Analisis Data Time Series Menggunakan LSTM (Long Short Term Memory) Dan ARIMA (Autocorrelation Integrated Moving Average) Dalam Bahasa Python.,” *Ultim. InfoSys J. Ilmu Sist. Inf.*, vol. 11, no. 1, pp. 1–7, 2020, doi: 10.31937/si.v9i1.1223.
- [23] N. Afrianto, D. H. Fudholi, and S. Rani, “Prediksi Harga Saham Menggunakan BiLSTM dengan Faktor Sentimen Publik,” *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 6, no. 1, pp. 41–46, 2022, doi: 10.29207/resti.v6i1.3676.
- [24] A. Aulia, B. Aprianti, Y. Supriyanto, and C. Rozikin, “Prediksi Harga Emas dengan Menggunakan Algoritma Support Vector Regression (Svr) dan Linear Regression (LR),” *J. Ilm. Wahana Pendidik.*, vol. 8, no. 5, pp. 84–88, 2022, doi: 10.5281/zenodo.6408864.