

Journal of Dinda

**Kelompok Keahlian Rekayasa Data
Institut Teknologi Telkom Purwokerto**

Vol. 2 No. 1 (2022) 1 - 10

ISSN Media Elektronik: 2809-8064

Prediksi Harga Saham Bank Bri Menggunakan Algoritma Linear Regresion Sebagai Strategi Jual Beli Saham

Janur Syahputra¹, Rima Dias Ramadhani², Auliya Burhanudin³

¹Program Studi S1 Teknik Informatika, Fakultas Informatika, Institut Teknologi Telkom Purwokerto

²Program Studi S1 Sains Data, Fakultas Informatika, Institut Teknologi Telkom Purwokerto

³Program Studi S1 Informatika, Fakultas Informatika, Institut Teknologi Telkom Purwokerto

¹17102178@ittelkom-pwt.ac.id, ²rima@ittelkom-pwt.ac.id, ²auliya@ittelkom-pwt.ac.id

Abstract

Shares are securities as proof of ownership of investors in a company. Stocks have a volatile nature, this makes stocks difficult to predict. Stock prediction is an effort to estimate the stock price, especially in the Bank Rakyat Indonesia company that will appear in the future, and to increase investors' profit opportunities in making investment decisions. During the COVID-19 pandemic, Bank BRI's shares experienced significant ups and downs in four months, which illustrates the sensitivity of the stock to an event. Therefore, it is important to predict stock prices to reduce the risk accepted by investors. The prediction itself requires time series data. Time series is data that is collected sequentially from time to time. The method used for time series data is Linear Regression because this method can handle time-series data. Based on these problems, stock prediction research will be conducted at the Bank Rakyat Indonesia company using the Linear Regression method. Bank Rakyat Indonesia share price data were obtained from the investing.com website from the period starting on January 1, 2008, to June 1, 2020. The data is processed starting from preprocessing to determine attributes, remove unnecessary attributes, and change the contents of the data type, then process split data to divide the dataset into training and test data. The attributes used in this study are Date and Price and the distribution of the data used is 60:40, 65:35, 70:30, 75:25, and 80:20. The best ratio is at 80:20 which produces train and test accuracy of 0.89 and 0.91, Then each training data and testing data are entered into the linear regression model for prediction. The error results from the predictions were calculated using MAPE and yielded a percentage of 13.751% for training data, 13.773% for test data, and 13.755% for overall data. The MAPE results indicate that the linear regression method can be used to predict the stock price of BRI Bank.

Keywords: stokcs, linear regression, MAPE

Abstrak

Saham merupakan surat berharga sebagai bukti kepemilikan investor dalam suatu perusahaan. Saham memiliki sifat yang fluktuatif secara tidak menentu, hal ini yang membuat saham sulit diprediksi. Prediksi saham merupakan upaya memperkirakan berapa harga saham khususnya pada perusahaan Bank Rakyat Indonesia yang akan muncul dimasa depan serta untuk meningkatkan peluang keuntungan investor dalam mengambil keputusan investasi. pada saat pandemi covid-19, saham Bank BRI mengalami naik dan turun secara signifikan dalam rentang empat bulan, hal tersebut yang menggambarkan sensitifnya saham terhadap suatu peristiwa. Oleh karena itu pentingnya melakukan prediksi harga saham agar bisa mengurangi resiko yang diterima para investor. Prediksi sendiri membutuhkan data yang berdimensi time series. Time series merupakan data yang dikumpulkan secara berurutan dari waktu ke waktu. Metode yang digunakan untuk data berdimensi time series yaitu salah satunya Linear Regression, karena metode ini bisa menangani data rentetan waktu. Berdasarkan permasalahan tersebut maka akan dilakukan penelitian prediksi saham pada perusahaan Bank Rakyat Indonesia menggunakan metode Linear Regression. Data harga saham Bank Rakyat Indonesia diperoleh dari website investing.com dari periode yang dimulai pada 1 januari 2008 sampai dengan 1 juni 2020. Data tersebut diolah dimulai dari preprocessing untuk menentukan atribut, menghilangkan atribut yang tidak perlu, dan mengubah isi tipe data, lalu proses split data untuk membagi dataset menjadi data latih dan uji. Atribut yang digunakan pada penelitian ini yaitu Date dan Price

Diterima Redaksi : 23-06-2021 | Selesai Revisi : 06-01-2022 | Diterbitkan Online : 23-02-2022

serta pembagian data yang digunakan adalah 60:40, 65:35, 70:30, 75:25, dan 80:20. Rasio terbaik ada pada 80:20 yang menghasilkan akurasi train dan test sebesar 0,89 dan 0,91, Kemudian masing-masing data training dan data testing dimasukkan kedalam model linear regression untuk kemudian dilakukan prediksi. Hasil error dari prediksi dihitung menggunakan MAPE dan menghasilkan persentase sebesar 13,751% untuk data pelatihan, 13,773% untuk data pengujian, dan 13,755% untuk data keseluruhan. Hasil MAPE tersebut mengindikasikan bahwa metode linear regression dapat digunakan untuk prediksi harga saham Bank BRI.

Kata Kunci : Saham, Linear Regression, MAPE

© 2022 Journal of DINDA

1. Pendahuluan

Pasar modal merupakan pasar untuk berbagai instrumen keuangan jangka panjang yang bisa diperjualbelikan seperti Saham [1]. Menurut penelitian [2] Citra Puspa Permata dkk, berdasarkan undang-undang Pasar Modal (UUPM), yaitu Undang-Undang Nomor 8 tahun 1995 Pasal 1 Angka 13 menjelaskan, Pasar modal adalah kegiatan yang bersangkutan dengan penawaran umum dan perdagangan efek, perusahaan publik yang berkaitan dengan efek yang diterbitkannya, serta lembaga dan profesi yang berkaitan dengan efek. Dengan memperjualbelikan saham, investor bertujuan mendapatkan keuntungan yang maksimal dan resiko yang minimal [3]. Saham merupakan surat berharga yang menunjukkan bukti kepemilikan atau penyertaan pasar modal investor dalam suatu perusahaan dan saham mempunyai nilai yang bersifat naik dan turun atau fluktuatif [4].

Beberapa hal yang menyebabkan harga saham naik dan turun yaitu, kondisi dan kinerja perusahaan, resiko, deviden, tingkat suku bunga, kondisi perekonomian, kebijakan pemerintah, isu pemerintah atau isu lain, laju inflasi, penawaran dan permintaan [5], sedangkan menurut Arifin pada penelitian [6] faktor yang menentukan naik dan turunnya harga saham yaitu kondisi fundamental emiten, permintaan dan penawaran, tingkat suku bunga, valuta asing, dana asing, indeks harga saham gabungan, dan rumor.

Investor yang bertujuan untuk investasi jangka panjang perlu melakukan analisis fundamental maupun teknikal, tingkat analisis yang diteliti berdasarkan laporan keuangan perusahaan dapat diketahui kondisi perusahaan, lalu hasil analisis ini berupa laporan keuangan yang diinterpretasikan dengan menggunakan analisis teknikal kemudian menjadi informasi yang berguna dalam mengambil keputusan investasi yang tepat [7]. Semakin tinggi keuntungan atau kerugian (return) yang diperoleh dari investasi yang ditawarkan maka semakin tinggi pula resiko yang harus ditanggung investor, dan sebaliknya semakin rendah keuntungan atau kerugian (return) yang ditawarkan maka semakin rendah pula resiko yang harus ditanggung investor,

dalam hal ini return dan resiko memiliki hubungan yang searah atau bersifat linear [8].

Indeks harga saham adalah nilai atau harga dari saham yang diterbitkan oleh perusahaan dan dikelompokkan dalam klasifikasi tertentu [9]. Setidaknya saat ini ada 34 indeks harga saham yang terdaftar di Bursa Efek Indonesia (BEI), salah satu indeks saham yang diminati oleh investor yaitu LQ45 pada bursa saham Indonesia. Indeks harga saham LQ45 terdiri dari 45 perusahaan-perusahaan besar yang memiliki likuiditas tinggi yang dapat dijadikan prospek bisnis yang menjanjikan bagi para investor [10].

TOP 10 CONSTITUENTS				
Code	Company Name	Market Cap* (in trill. IDR)	Index Weight	Sector
BBCA	Bank Central Asia Tbk.	255.73	15.63%	Finance
BBRI	Bank Rakyat Indonesia (Persero) Tbk.	231.74	14.16%	Finance
TLKM	Telekomunikasi Indonesia (Persero) Tbk.	188.18	11.50%	Infra., Utilities & Transportation
BMRI	Bank Mandiri (Persero) Tbk.	141.83	8.67%	Finance
ASII	Astra International Tbk.	126.41	7.73%	Miscellaneous Industry
BBNI	Bank Negara Indonesia (Persero) Tbk.	57.97	3.54%	Finance
UNVR	Unilever Indonesia Tbk.	45.79	2.80%	Consumer Goods Industry
CPIN	Charoen Pokphand Indonesia Tbk.	42.76	2.61%	Basic Industry & Chemicals
BRPT	Barito Pacific Tbk.	37.90	2.32%	Basic Industry & Chemicals
SMGR	Semen Indonesia (Persero) Tbk.	34.86	2.13%	Basic Industry & Chemicals
Total (Top 10 Constituents)		1,163.18	71.09%	

Data prior to the launch date is back-tested data.
*Adjusted Market Capitalization

Data as of: Dec 30, 2019

Gambar 1 Daftar 10 Teratas Perusahaan Indeks Saham LQ45 Tahun 2019 [10]

Gambar 1 merupakan 10 daftar perusahaan teratas pada indeks saham LQ45 berdasarkan Harga Keseluruhan Saham Perusahaan, peringkat atas ditempati oleh saham dari sektor perbankan. Tentunya saham sektor perbankan sangat digemari investor untuk terus diperdagangkan di lantai bursa karena memiliki kapitalisasi pasar dan likuiditas yang tinggi. Saham sektor perbankan didominasi beberapa saham besar perbankan yaitu saham PT Bank Centra Asia Tbk (BBCA), PT Bank Rakyat Indonesia Tbk (BBRI), PT Bank Mandiri Tbk (BMRI).



Gambar 2 Grafik saham Bank BRI pada tahun 2019 – 2020 [11]

Gambar 2 merupakan grafik saham pada saat pandemi covid-19, terlihat bahwa saham Bank BRI mengalami naik dan turun secara signifikan dalam rentang waktu januari hingga april tahun 2020 nilai saham paling tinggi yaitu tanggal 24 januari 2020 sebesar Rp. 4.740 per lembar saham, sedangkan terendah yaitu pada tanggal 9 april 2020 sebesar Rp. 2.790 per lembar. Hal tersebut yang dapat meningkatkan resiko rugi pada saat investor melakukan investasi.

Berdasarkan permasalahan tersebut, sangat penting dilakukan prediksi pada harga saham untuk mengurangi resiko pada investor saat melakukan keputusan investasi. diperlukan metode analisa prediksi saham yang memiliki hasil lebih akurat serta untuk mengurangi resiko pada investor saat mengambil keputusan investasi.

Prediksi atau peramalan adalah memperkirakan sesuatu yang akan terjadi dimasa mendatang, jadi sesuatu yang di prediksi dari hasil akan ada dimasa yang dating [12]. Prediksi merupakan salah satu cabang machine learning yang metode penyelesaian masalahnya menggunakan supervised learning dan menggunakan time series atau rentetan waktu sebagai datanya [13].

Beberapa metode yang cocok digunakan untuk time series yaitu Linear Regression dan Auto Regressive Integrative Moving Average (ARIMA). Definisi metode Linear Regression itu sendiri merupakan metode yang biasanya untuk menemukan persamaan dari sebuah data yang dimana data tersebut saling berhubungan antara variabel satu dengan variabel lain yang ingin diramalkan dalam satu database yang cukup besar [14]. Sementara definisi dari metode ARIMA merupakan suatu pendekatan pemodelan yang dapat digunakan untuk menghitung probabilitas dari nilai masa depan yang terletak di antara dua batas yang sudah ditentukan [15].

Penelitian sebelumnya sudah pernah membahas yang berjudul “Prediksi Harga Saham PT. BRI, Tbk. Menggunakan metode Auto Regressive Integrative Moving Average (ARIMA) [16]. Pada penelitian tersebut berisi mengenai pembahasan prediksi harga saham menggunakan metode Auto Regressive Integrative Moving Average (ARIMA), namun dalam penerapan metode tersebut dataset harus memiliki nilai yang stabil (stasioner). Berdasarkan penelitian tersebut

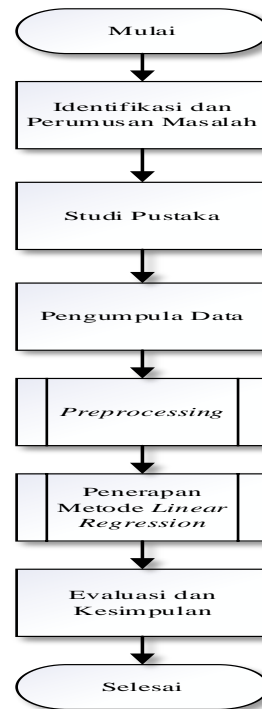
penulis ingin menggunakan metode Linear Regression karena metode ini dapat digunakan pada dataset yang memiliki nilai tidak stabil, serta metode ini dapat diaplikasikan ke dalam data tren seperti saham. Penelitian ini akan memprediksi Harga Saham Pada Bank Rakyat Indonesia (BRI) Prediksi yang akan dilakukan menggunakan metode Linear Regression sehingga diharapkan dapat membantu para investor atau pelaku pasar modal dalam mengambil keputusan investasi saham pada Bursa Efek Indonesia. Penelitian lainnya membahas tentang prediksi harga saham BMRI yang dihubungkan dengan judul berita berbahasa Indonesia dengan model klasifikasi nya yaitu neural network [17].

Berdasarkan dari alasan dan penelitian tersebut, maka pertanyaan yang muncul pada penelitian ini adalah sebagai berikut :

1. Bagaimana menerapkan metode linear regression yang dapat membantu melakukan prediksi pada harga saham Bank Rakyat Indonesia periode Januari 2008 sampai dengan Juni 2020?
2. Berapa presentase error yang diperoleh dari model prediksi linear regression pada harga saham Bank Rakyat Indonesia dari periode Januari 2008 sampai dengan Juni 2020 menggunakan MAPE?

2. Metode Penelitian

Metode penelitian ini dilakukan melalui tahap tahap untuk memprediksi harga saham pada Gambar 3 dibawah ini.



Gambar 3 Flowchart Penelitian

2.1. Identifikasi dan Perumusan Masalah

Tahap awal pada penelitian ini melakukan identifikasi dan perumusan masalah pada penelitian yaitu memprediksi harga saham pada Bank Rakyat Indonesia dengan menentukan metode prediksi yang digunakan linear regression dan atribut yang ada pada data harga saham yaitu Date, Price, Open, High, Low, Volume (M), Change (%) serta atribut digunakan untuk memprediksi saham yaitu Date, Price.

2.2. Studi Pustaka

Pada tahap studi Pustaka ini dilakukan dengan melakukan studi literatur dari berbagai sumber yang tersedia seperti jurnal ilmiah, skripsi, buku, website, maupun sumber lain yang memiliki hubungan dengan permasalahan pada penelitian ini.

2.3. Pengumpulan Data

Tahap pengumpulan data, pada tahap pengumpulan ini menggunakan dataset yang diperoleh dari *website investing.com*, dapat diakses dengan link <https://www.investing.com/equities/bank-rakyat-in-historical-data>. Dataset ini merupakan dataset harga saham Bank Rakyat Indonesia (BRI) pada periode yang dimulai dari tahun 2008 hingga 2020 yang mempunyai tujuh atribut yaitu *Date, Price, Open, High, Low, Vol. Change %*. Adapun atribut yang digunakan yaitu *Date* dan *Price*, untuk atribut *Price* merupakan harga penutupan saham.

2.4. Preprocessing

Tahap preprocessing pada penelitian ini digunakan untuk mengolah dataset yang diperoleh agar membantu memudahkan pengolahan data pada model. Berdasarkan dataset yang telah diperoleh akan dilakukan langkah Pengolahan dataset, Langkah ini bertujuan untuk memudahkan proses pengolahan data selanjutnya. Dalam pengolahan dataset ini memiliki dua tahap yaitu:

a. Penghapusan dan perubahan karakter pada dataset:

Tahap ini melakukan penghapusan dan perubahan setiap karakter yang tidak diperlukan, Hal ini bertujuan untuk mempermudah langkah perubahan tipe data. Adapun proses yang dilakukan pada tahap ini yaitu penghapusan karakter koma (“,”) pada kolom Price, kemudian pengubah karakter koma (“,”) menjadi titik (“.”) pada kolom Open sampai Low, penghapusan karakter “M”, dan penghapusan simbol persen (%).

	Date	Price	Open	High	Low	Vol.	Change %
0	2008-01-02	735	735.0	740.0	730.0	44.77	-0.68
1	2008-01-03	720	735.0	735.0	710.0	102.94	-2.04
2	2008-01-04	750	720.0	755.0	715.0	143.67	4.17
3	2008-01-07	745	735.0	750.0	730.0	63.70	-0.67
4	2008-01-08	750	745.0	755.0	745.0	84.23	0.67
...
3036	2020-06-24	3120	3050.0	3140.0	3010.0	308.48	3.65
3037	2020-06-25	3060	3100.0	3110.0	3040.0	158.21	-1.92
3038	2020-06-26	3030	3100.0	3120.0	3020.0	147.31	-0.98
3039	2020-06-29	3040	3040.0	3060.0	2980.0	118.81	0.33
3040	2020-06-30	3030	3070.0	3110.0	3030.0	134.01	-0.33

Gambar 4 Hasil Proses Pengubahan dan Penghilangan karakter

Berdasarkan Gambar 4 dilihat bahwa penghilangan dan perubahan karakter yang tidak perlu telah berhasil dilakukan

b. Perubahan tipe data pada setiap kolom:

Tahap kedua yaitu perubahan tipe, setiap kolom data masih menggunakan tipe data object sehingga data tidak cocok dengan isi dari tipe data sebenarnya, oleh karena itu agar isi data sesuai dengan tipe data sebenarnya maka akan dilakukan perubahan tipe data. Adapun kolom yang diubah yaitu kolom Date diubah menjadi tipe data datetime, kolom Price, Open, High, Low diubah menjadi tipe data “int”, kemudian kolom Volume dan Change % diubah menjadi tipe data “float”.

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 3041 entries, 0 to 3040
Data columns (total 7 columns):
#   Column      Non-Null Count  Dtype
---  -
0   Date        3041 non-null   datetime64[ns]
1   Price       3041 non-null   int32
2   Open        3041 non-null   int32
3   High        3041 non-null   int32
4   Low         3041 non-null   int32
5   Vol.        3041 non-null   float64
6   Change %    3041 non-null   float64
dtypes: datetime64[ns](1), float64(2), int32(4)
memory usage: 118.9 KB
```

Gambar 5 Hasil Proses perubahan tipe data

Gambar 5 menunjukkan hasil dari proses perubahan tipe data pada setiap kolom menjadi tipe data yang cocok dengan tipe dari isi data.

Berikut pengaplikasian proses preprocessing menggunakan pemrograman python dapat di tuliskan seperti pada program berikut ini:

```
Program Preprocessing
Input: Dataset saham
Output: hasil penghapusan dan konversi tipe data

# Membaca input dataset
Dataset = (pd.read_csv('BBRI Historical
Data.csv'))
```

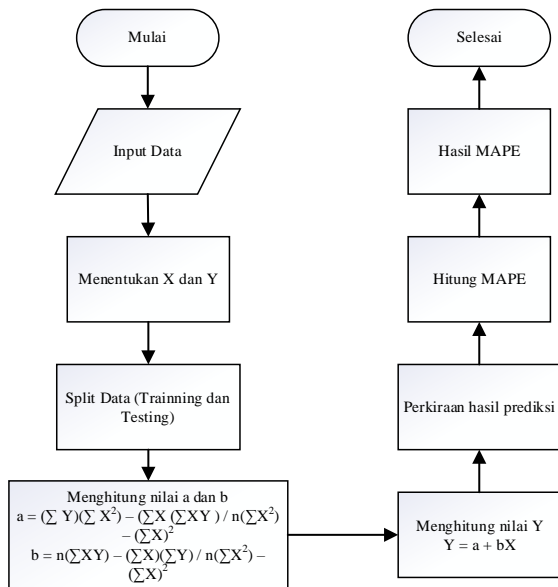


```
# Perubahan tipe data pada setiap kolom dan
penghapusan karakter
dataset['Date'] =
pd.to_datetime(dataset['Date']) dataset[] =
dataset['Price'] =
dataset['Price'].str.replace(',','').astype(int)
dataset['open'] =
dataset['open'].str.replace(',','').astype(int)
dataset['High'] =
dataset['High'].str.replace(',','').astype(int)
dataset['Low'] =
dataset['Low'].str.replace(',','').astype(int)
dataset['Vol.']=
dataset['Vol.'].str.replace('M','').astype(float)
dataset['change %'] =
dataset['change %'].str.replace(',','').astype(int)

# Untuk melihat informasi mengenai dataset
dataset.info()

# Menampilkan tabel dataset
dataset.head()
```

2.5. Penerapan Metode Linear Regression



Gambar 6 Flowchart Metode Linear Regression

Gambar 6 merupakan alur proses dari metode *linear regression* yang dimulai dari input data berupa dataset harga saham yang sudah melalui tahap *preprocessing*, *preprocessing* bertujuan untuk mengubah tipe data yang sesuai dengan bentuk data dan penghilangan karakter yang tidak diperlukan. Kemudian setelah input data, selanjutnya menentukan kolom yang digunakan untuk variabel X dan Y. setelah menentukan variabel X dan Y, selanjutnya masuk kedalam proses split data untuk data latih dan data uji. Selanjutnya masuk kedalam proses perhitungan nilai a (intercept) dan b (slope) menggunakan persamaan rumus [18]:

a. Mencari nilai a (intercept) dengan menggunakan persamaan rumus:

$$a = \frac{(\sum Y)(\sum X^2) - (\sum X)(\sum XY)}{n(\sum X^2) - (\sum X)^2} \quad (1)$$

b. Mencari nilai b (slope) dengan menggunakan persamaan rumus:

$$b = \frac{n(\sum XY) - (\sum X)(\sum Y)}{n(\sum X^2) - (\sum X)^2} \quad (2)$$

Setelah nilai a dan nilai b dihitung, selanjutnya menentukan Y untuk hasil prediksi. Adapun persamaan rumus Y sebagai berikut:

$$Y = a + bX \quad (3)$$

Setelah nilai Y (prediksi) sudah ditentukan, selanjutnya menghitung jumlah *error* pada prediksi, penghitungan jumlah *error* menggunakan MAPE. MAPE adalah suatu ukuran ketetapan relatif yang digunakan untuk menghitung presentase hasil dari penyimpangan peramalan dengan cara menghitung kesalahan absolut pada tiap priode dibagi dengan hasil observasi nyata pada periode tersebut kemudian dihitung rata-ratanya. Adapun cara perhitungan MAPE menggunakan persamaan rumus sebagai berikut [19]:

$$MAPE = \frac{100\%}{n} \sum_{t=1}^n \left| \frac{x_t - \hat{x}_t}{x_t} \right| \quad (4)$$

Keterangan: n = banyak data, x_t = data observasi nyata pada waktu t, \hat{x}_t = data hasil peramalan pada waktu t

2.6. Evaluasi dan Kesimpulan

Tahap selanjutnya adalah evaluasi dan kesimpulan, pada tahap ini dilakukan analisis dari hasil proses linear regression yang berisi uraian dari hasil proses penelitian dan hasil dari pengujian dalam prediksi harga saham pada Bank Rakyat Indonesia. Hasil dari analisis tersebut akan dijadikan sebagai bahan kesimpulan dari penelitian yang dilakukan ini.

3. Hasil dan Pembahasan

3.1. Pengolahan Dataset

Input dataset yang peneliti gunakan diperoleh dari website [investing.com](https://www.investing.com) yang didownload dari <https://www.investing.com/equities/bank-rakyat-in-historical-data>. Dataset tersebut merupakan data harga saham Bank BRI dari tahun 2008 hingga 2020 dengan periode harian yang berjumlah 3041 baris. Ada tujuh kolom yang terdapat pada dataset ini, yaitu Date, Price, Open, High, Low, Vol. dan Change (%) semua kolom tersebut dikonversi tipe datanya supaya lebih memudahkan dalam proses memasukan data kedalam metode. Berikut hasil proses dari konversi tipe data pada setiap kolom ditunjukkan pada Gambar 8.

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 3041 entries, 0 to 3040
Data columns (total 7 columns):
#   Column      Non-Null Count  Dtype
---  ---
0   Date         3041 non-null   datetime64[ns]
1   Price        3041 non-null   int32
2   Open         3041 non-null   int32
3   High         3041 non-null   int32
4   Low          3041 non-null   int32
5   Vol.         3041 non-null   float64
6   Change %     3041 non-null   float64
dtypes: datetime64[ns](1), float64(2), int32(4)
memory usage: 118.9 KB
```

Gambar 7 Hasil proses konversi setiap kolom pada dataset

Selanjutnya, setelah dilakukan proses konversi tipe data adalah penentuan kolom yang akan digunakan seperti pada Gambar 8

	Date	Price
0	2008-01-02	735
1	2008-01-03	720
2	2008-01-04	750
3	2008-01-07	745
4	2008-01-08	750
...
3036	2020-06-24	3120
3037	2020-06-25	3060
3038	2020-06-26	3030
3039	2020-06-29	3040
3040	2020-06-30	3030

Gambar 8 Kolom Date dan Price

Gambar 8 merupakan kolom yang akan digunakan untuk prediksi pada metode Linear Regression yaitu kolom Date dan Price. (penjelasan isi fungsi price dan date) kolom Date dan Price berkaitan karena, pada kolom Price bisa bertambah dan berkurang nilainya berdasarkan kolom Date, alasan tersebut yang membuat kedua kolom tersebut digunakan.



Gambar 9 Grafik plot antara Price dan Date

Dilihat pada Gambar 3.3 alasan tersebut dibuktikan oleh pergerakan harga dari waktu ke waktu yang ditandai dengan garis biru pada grafik dengan memplot Price dan Date, dapat dilihat bahwa harga cenderung meningkat dari hari ke hari. Diperkirakan tren pada gambar tersebut bersifat linier.

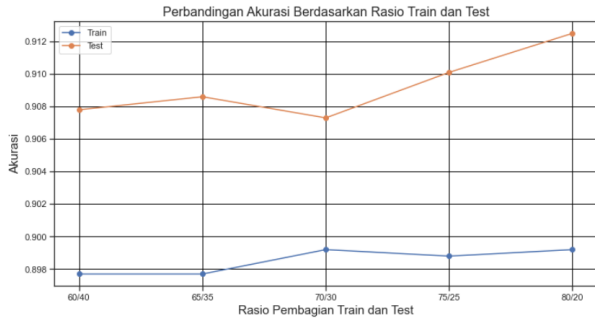
3.2. Perbandingan Hasil Akurasi Metode Berdasarkan Rasio Pembagian Data

Pembagian data antara data latih dan uji bertujuan untuk proses pelatihan dan untuk mengetahui performa dari metode linear regression. Data latih digunakan untuk melatih metode yang nantinya akan menghasilkan model pelatihan, setelah itu untuk mengetahui akurasi dari model yang sudah dilatih penulis menggunakan data uji. Penentuan rasio pembagian antara data latih dan data uji disini penulis mengambil beberapa pembagian untuk diuji mana rasio pembagian yang menghasilkan akurasi latih dan uji tertinggi, sekaligus mengetahui rata rata akurasi latih dan uji pada model yang penulis buat. Adapun beberapa rasio pembagian data latih dan train ditunjukkan pada Tabel 1.

Table 3.1 Hasil akurasi data train dan test berdasarkan rasio pembagian

No	Rasio Data Latih dan Uji	Akurasi Latih	Akurasi Uji
1	60:40	0.8977	0.9078
2	65:35	0.8977	0.9086
3	70:30	0.8992	0.9073
4	72:25	0.8988	0.9101
5	80:20	0.8992	0.9125
Rata Rata		0.8985	0.9092

Berdasarkan Tabel 1 akurasi tertinggi pada pembagian data latih dan uji ada pada rasio pembagian 80:20 dengan akurasi train dan test sebesar 0,89 dan 0,91, serta rata-rata akurasi latih dan uji dari berbagai rasio sebesar yaitu 0,89 dan 0,90. Dari Tabel 3.1 dapat disimpulkan bahwa akurasi terbesar dari rasio pembagian data latih dan uji ada pada 80:20, jadi rasio tersebut yang digunakan pada proses pelatihan model. Adapun grafik yang merepresentasikan Tabel 1 ditunjukkan pada Gambar 10



Gambar 10 Grafik plot akurasi latih dan uji berdasarkan rasio pembagian

Berdasarkan Gambar 10 dijelaskan bahwa semakin besar rasio data pelatihan dan semakin kecil data pengujian akurasi yang didapat pada kedua data tersebut akan semakin bagus, hal tersebut ditunjukkan pada grafik Gambar 10 pada garis berwarna biru untuk pelatihan dan oranye untuk pengujian.

3.3. Hasil Prediksi Menggunakan Metode Linear Regression

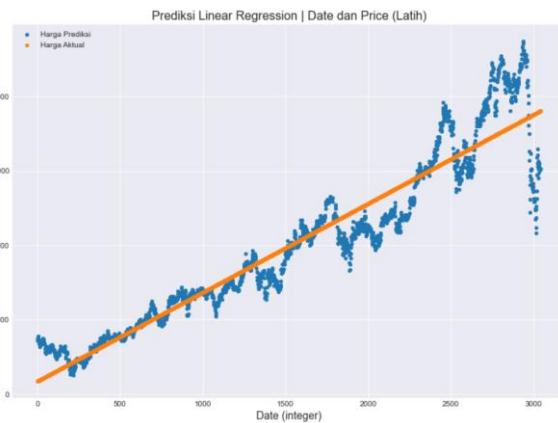
Pada hasil pengujian alur dimulai dari input data harga saham pada Bank BRI pada tahun 2008 sampai 2020 dengan periode harian, kemudian menentukan variabel X dan Y. variabel X dan Y diisi oleh kolom Date dan Price, setelah itu menentukan rasio pembagian data menggunakan rasio 80:20. Selanjutnya, yaitu menghitung nilai intercept (a) dan slope (b) menggunakan persamaan rumus (2) dan (3), Adapun hasil dari nilai a dan b secara berturut-turut yaitu 170.69560 dan 1.19587338, setelah menentukan nilai a dan b kemudian menghitung nilai Y menggunakan persamaan rumus (1) sebagai prediksi. Hasil prediksi tersebut dibagi menjadi tiga jenis sesuai dengan jenis datanya, Adapun hasil prediksi sebagai berikut.

Pertama yaitu melihat prediksi pada data latih, pada data latih sendiri proses pelatihan menggunakan 2432 set data. Hasil dari proses pelatihan ditunjukkan pada Tabel 2 dan Gambar 11.

Tabel 2 Hasil prediksi pada pelatihan

Date	Date (X)	Price (Y)	Prediksi
Mar 23 2011	781	1040	1104,673
Aug 13 2012	1128	1430	1519,641
Dec 10 2008	227	405	442,1589
Jul 07 2009	365	650	607,1894
Aug 15 2013	1370	1660	1809,042
.	.	.	.

.	.	.	.
.	.	.	.
Sep 23 2014	1638	2130	2129,536
Jun 27 2012	1095	1240	1480,177
Aug 15 2012	1130	1420	1522,033
Apr 22 2013	1294	1740	1718,156
Jul 18 2011	860	1340	1199,147



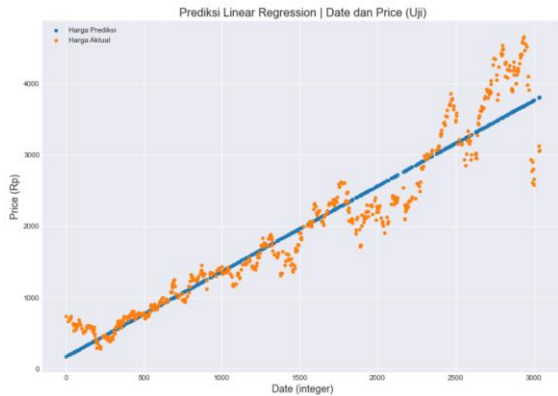
Gambar 11 Grafik hasil prediksi pelatihan

Kedua yaitu melihat hasil prediksi pada data uji, pada data uji sendiri proses pengujian menggunakan 609 set data. Hasil dari proses pengujian ditunjukkan pada Tabel 3 dan Gambar 12.

Tabel 3 Hasil prediksi pada pengujian

Date	Date (X)	Price (Y)	Prediksi
Mar 31 2009	298	420	527,0659
Jul 01 2019	2791	4420	3508,378
Jul 28 2008	139	605	336,922
Mar 23 2010	535	800	810,4879
Nov 22 2011	945	1340	1300,796
.	.	.	.
.	.	.	.
.	.	.	.
Mar 19 2019	2727	3990	3431,842

Nov 09 2017	2401	3290	3041,988
Jan 27 2010	498	755	766,2405
Aug 19 2016	2104	2400	2686,813
Nov 14 2008	210	328	421,829



Gambar 12 Grafik hasil prediksi pengujian

Terakhir yaitu melihat hasil prediksi menggunakan keseluruhan data, proses pengujian menggunakan 3041 set data. Hasil dari proses pengujian ditunjukkan pada Tabel 4.

Tabel 4 Hasil prediksi data keseluruhan

Date	Date (X)	Price (Y)	Prediksi
Jan 02 2008	1	735	170,6956
Jan 03 2008	2	720	171,8915
Jan 04 2008	3	750	173,0873
Jan 07 2008	4	745	174,2832
Jan 08 2008	5	750	175,4791
.	.	.	.
.	.	.	.
.	.	.	.
Jun 24 2020	3037	3120	3801,367199
Jun 25 2020	3038	3060	3802,563072
Jun 26 2020	3039	3030	3803,758946
Jun 29 2020	3040	3040	3804,954819
Jun 30 2020	3041	3030	3806,150692

Adapun hasil akurasi yang dihasilkan dari model yang sudah diuji dengan data prediksi dalam bentuk R^2 yaitu sebesar 0.9127. Hasil tersebut mengindikasikan bahwa hubungan linear bersifat positif yang kuat antara atribut Date dan Price. Semakin mendekati angka satu, model prediksi yang dikeluarkan oleh metode ini akan semakin baik.

3.4. Hasil MAPE

Tabel 5 Hasil MAPE

Data	MAPE
Data Latih	13.75140304407601
Data Uji	13.77304136319174
Data Keseluruhan	13.755736400321164

Pada Tabel 5 merupakan hasil MAPE dari masing-masing data yaitu sebesar 13.7514030440 untuk data pelatihan, 13.7730413 untuk data pengujian, dan 13.7557364 untuk data keseluruhan. Ketiga hasil tersebut mengindikasikan bahwa persentase error berada pada rentang 10% - 20%, bisa disimpulkan bahwa prediksi dari model yang digunakan penulis memiliki error yang kecil dan bisa dikatakan masuk kategori akurat.

4. Kesimpulan

Berdasarkan hasil dan analisis pada BAB IV, terkait prediksi harga saham pada Bank BRI menggunakan metode Linear Regression berdasarkan harga penutupan saham, maka didapatkan kesimpulan sebagai berikut:

1. Metode linear regression dapat digunakan untuk memprediksi harga saham Bank BRI dari periode 2008 sampai 2020 dengan menggunakan proses dimulai dari, preprocessing data yaitu pemilihan atribut yang tepat lalu pemilihan rasio pembagian data antara data latih dan data uji terbaik. Adapun yang digunakan yaitu atribut Date dan Price, serta menggunakan rasio pembagian data dengan akurasi tertinggi yaitu 80:20 didapatkan akurasi sebesar 0.8992 untuk pelatihan dan 0.9125 untuk pengujian.
2. Metode linear regression menghasilkan akurasi R^2 sebesar 0.9125, Ini menunjukkan hubungan linear positif yang kuat antara atribut Date dan Price, karena berdasarkan model linear regression ini menunjukkan bahwa variabilitas harga saham dapat diprediksi menggunakan waktu berdasarkan hari sebesar 91,25%.
3. Nilai error pada prediksi yang diukur dengan

menggunakan MAPE dari model linear regression pada data latih, data uji, dan data keseluruhan secara berturut-turut sebesar 13.751%, 13.773% dan 13.755%. Ketiga hasil tersebut mengindikasikan bahwa persentase error berada pada rentang 10% - 20%, bisa disimpulkan bahwa prediksi dari model yang digunakan penulis memiliki error yang kecil dan bisa dikatakan masuk kategori akurat.

Ucapan Terimakasih

Berikut ini merupakan ucapan terima kasih atas bantuan dana dan fasilitas yang telah diberikan kepada penulis, ucapan terima kasih ditujukan kepada:

1. Bapak Syahwal Sebagai Bapak dari peneliti.
2. Ibu NurLely Sebagai Ibu dari peneliti.

Daftar Rujukan

- [1] F. B. Alrhafynza and E. Siswanto, "Reaksi Pasar Modal Indonesia Terhadap Peristiwa Politik Nasional (Studi Pada Saham LQ45 atas Putusan Sidang Kasus Penistaan Agama oleh Gubernur DKI Jakarta, Basuki Tjahaja Purnama Tahun 2017)," *Magister Manaj. UGM*, pp. 47–56, 2017, [Online]. Available: <https://jurnal.unej.ac.id/index.php/prosiding/article/view/9136>.
- [2] Permata, Citra Puspa and M. A. Ghoni, "Peranan Pasar Modal Dalam Perekonomian Negara Indonesia," *J. Akun Stie*, vol. 5, no. 2, pp. 50–61, 2019.
- [3] Y. S. J. Nasution, "Peranan Pasar Modal Dalam Perekonomian Negara," *Res. Gate*, vol. 2, no. 1, pp. 95–112, 2015.
- [4] W. P. Sari, "Pengaruh Rasio Keuangan terhadap Harga Saham pada Perusahaan Manufaktur Go Public yang Terdaftar di Bursa Efek Indonesia," *J. Ilm. Syaklandsea*, vol. 2, no. 1, pp. 43–52, 2018.
- [5] H. Romli, M. F. Wulandari, and T. S. Pratiwi, "Faktor-Faktor Yang Mempengaruhi Volatilitas Harga Saham Pada PT Waskita Karya Tbk," *Ilm. Ekon. Glob. Masa Kini*, vol. 8, no. 01, pp. 1–5, 2017, [Online]. Available: <http://ejournal.uigm.ac.id/index.php/EGMK/article/view/281>.
- [6] I. Sulistiana, "Pengaruh Pengungkapan Corporate Social Responsibility Dan Rasio Profitabilitas Terhadap Harga Saham Pada Perusahaan Manufaktur Yang Terdaftar Di Bursa Efek Indonesia," *J. Akunt.*, vol. 4, no. 2, pp. 65–74, 2017.
- [7] T. S. Nuryanto, A. Prahutama, and A. Hoyyi, "HISTORICAL SIMULATION UNTUK MENGHITUNG VALUE AT RISK PADA PORTOFOLIO OPTIMAL BERDASARKAN SINGLE INDEX MODEL MENGGUNAKAN GUI MATLAB (Studi Kasus: Kelompok Saham JII Periode Juni - November 2017)," *J. Gaussian*, vol. 7, no. 4, pp. 408–418, 2018, doi: 10.14710/j.gauss.v7i4.28869.
- [8] S. Ayem and S. Wahyuni, "Pengaruh Loan To Deposit Ratio, Capital Adequacy Ratio, Return on Asset Dannon Perfoming Loan Terhadap Return Saham," *J. Akunt.*, vol. 5, no. 1, p. 71, 2017, doi: 10.24964/ja.v5i1.258.
- [9] A. N. Firdausi, I. Fahmi, and I. T. Saptono, "Pengaruh Indeks Harga Saham Regional ASEAN dan Variabel Makroekonomi terhadap Indeks Harga Saham Syariah Indonesia (ISSI)," *Al-Muzara'ah*, vol. 4, no. 2, pp. 76–96, 2016, doi: 10.29244/jam.4.2.76-96.
- [10] Y. Rachmawati, "Pengaruh Inflasi dan Suku Bunga Terhadap Harga Saham Pada Perusahaan Perbankan Yang Terdaftar Di LQ45 Bursa Efek Indonesia," *Media Akunt.*, vol. 1, no. 1, p. 69, 2018, [Online]. Available: [https://jurnal.univpgri-](https://jurnal.univpgri-pt.bursaefekindonesia.com)
- [11] PT. Bursa Efek Indonesia, "Indonesia Stock Exchange," *Stock Summ.*, no. July 1992, p. 2019, 2020, [Online]. Available: <https://www.idx.co.id/>.
- [12] W. Hastomo, R. Dalam, K. Baru, T. Selatan, and U. Gunadarma, "KEMAMPUAN LONG SHORT TERM MEMORY MACHINE," vol. 4, no. September, pp. 229–236, 2020
- [13] A. Izzah and R. Widyastuti, "Prediksi Harga Saham Menggunakan Improved Multiple Linear Regression untuk Pencegahan Data Outlier," *Kinet. Game Technol. Inf. Syst. Comput. Network, Comput. Electron. Control*, vol. 2, no. 3, pp. 141–150, 2017, doi: 10.22219/kinetik.v2i3.268.
- [14] A. Rahmi, W. F. Mahmudy, and B. D. Setiawan, "Prediksi Harga Saham Berdasarkan Data Historis Menggunakan Model Regresi," *DORO Repos. J. Mhs. PTIIK Univ. Brawijaya*, no. 12, pp. 1–9, 2015.
- [15] A. Fikri, "Penerapan Data Mining Untuk Mengetahui Tingkat Kekuatan Beton Yang Dihasilkan Dengan Metode Estimasi Menggunakan Linear Regression," *Fak. Ilmu Komput. UDINUS*, pp. 1–12, 2013.
- [16] A. H. Hutasuhut, "Pembuatan Aplikasi Pendukung Keputusan untuk Peramalan Persediaan Bahan Baku Produksi Plastik Blowing dna Inject Menggunakan Metode ARIMA (Autoregressive Integrated Moving Average) di CV. Asia," *J. Tek. Pomits*, vol. 3, no. 2, pp. 70–171, 2014, [Online]. Available: [http://ejournal.its.ac.id/index.php/teknik/article/view](http://ejournal.its.ac.id/index.php/teknik/article/view/File/8114/1846)
- [17] Ramadhan, Nur Ghaniaviyanto, and Imelda Atastina. "Neural Network on Stock Prediction using the Stock Prices Feature and Indonesian Financial News Titles." *International Journal on Information and Communication Technology*

- (IJoICT) 7.1. 2021. 54-63.
- [18]G. S. Lilipaly, D. Hatidja, and J. S. Kekenusa, “PREDIKSI HARGA SAHAM PT . BRI , Tbk . MENGGUNAKAN METODE ARIMA,” J. Ilm. Sains, vol. 14, pp. 61–66, 2014.
- [19]H. Harliana and A. Syafrianto, “PREDIKSI JUMLAH PENDAFTARAN CALON MAHASISWA BARU DENGAN METODE REGRESI LINIER.”