



Perbandingan metode linear regresi dan polynomial regresi untuk memprediksi harga saham studi kasus Bank BCA

Comparison of linear regression and polynomial regression methods to predict stock prices BCA Bank case study

Satriyo Aji Laksono^{1*}, Adi Rizky Pratama¹, Rahmat¹

^{1*} Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Buana Perjuangan, Karawang, Jawa Barat, Jl. Hs. Ronggo Waluyo, Teluk Jambe Karawang, Jawa Barat, Indonesia

INFORMASI ARTIKEL

Article History:

Submission: 26-05-2023

Revised: 03-06-2023

Accepted: 06-06-2023

Kata Kunci:

Investasi; prediksi; linear regresi; polynomial regresi.

Keywords:

Investment; prediction; MAPE, linear regression; polynomial regression.

* Korespondensi:

Satriyo Aji Laksono

If18.satriyolaksono@mhs.ubpkarawang.a
c.id

ABSTRAK

Era ekonomi modern seperti saat ini perkembangan investasi di Indonesia semakin pesat. Saham adalah jenis investasi yang memiliki dampak kuat dalam perekonomian masyarakat. Salah satu saham yang diminati masyarakat yaitu saham Bank BCA. Saham Bank BCA dalam lima tahun terakhir mengalami naik turun relatif cepat. Maka dari itu, dibutuhkan model suatu data yang dapat mengelola data dalam waktu yang lama untuk memprediksi harga suatu saham. Pada penelitian ini metode Linear Regresi dan Polynomial Regresi digunakan sebagai metode untuk memprediksi harga saham. Proses penelitian ini berfokus pada penerapan dan membandingkan metode mana yang lebih baik, serta untuk mengetahui nilai error pada metode. Data yang digunakan adalah data saham Bank BCA selama lima tahun sebanyak 1252 data. Perhitungan nilai MAPE digunakan untuk pengujian prediksi. Hasil penelitian ini diketahui bahwa kedua metode masuk kriteria sangat baik karena nilai MAPE kurang dari 10%, metode Polynomial Regresi mendapatkan nilai yang lebih baik dibandingkan dengan metode Linear Regresi. Hasil pengujian pada metode Linear Regresi mendapatkan nilai MAPE = 6,55%, sedangkan pada metode Polynomial Regresi mendapatkan nilai MAPE = 6,54%.

ABSTRACT

In this modern economic era, the development of investment in Indonesia is growing rapidly. Shares are a type of investment that has a strong impact on the community's economy. One of the stocks that the public is interested in is BCA Bank shares. BCA Bank shares in the last five years have fluctuated relatively quickly. Therefore, a data model is needed that can manage data over a long time to predict the price of a stock. In this research, Linear Regression and Polynomial Regression methods are used as methods to predict stock prices. This research process focuses on applying and comparing which method is better, and to determine the error value of the method. The data used is Bank BCA stock data for five years totaling 1252 data. MAPE value calculation is used for prediction testing. The results of



this study show that both methods are included in the excellent criteria because the MAPE value is less than 10%, the Polynomial Regression method gets a better value than the Linear Regression method. The test results on the Linear Regression method get a MAPE value = 6.55%, while the Polynomial Regression method gets a MAPE value = 6.54%

1. PENDAHULUAN

Pada era modern seperti saat ini perkembangan Indonesia di berbagai bidang terutama sektor ekonomi semakin pesat. Salah satu kegiatan yang mendorong perekonomian Indonesia adalah investasi. Hal ini didukung dengan mudahnya akses dalam menggunakan *platform* penyedia layanan investasi. Investasi dapat diartikan sebagai kegiatan penanaman dana atau sumberdaya saat ini dengan berharap mendapat keuntungan di masa depan [1]. Aktivitas investasi biasanya berkaitan dengan penanaman modal yang diklasifikasikan menjadi dua yaitu aset riil dan aset finansial. Aset riil meliputi tanah, emas dan rumah, sedangkan aset finansial meliputi saham, obligasi dan reksadana. Saham adalah salah satu jenis investasi yang memiliki dampak kuat dalam perekonomian masyarakat. Arti saham adalah suatu sertifikat sebagai alat bukti atas kepemilikan dari sebuah perusahaan [2]. Naik atau turunnya saham sangat menentukan keuntungan pemilik saham. Ada beberapa faktor yang mempengaruhi naik turunnya saham seperti kondisi ekonomi, laju inflasi serta penawaran dan permintaan. Bursa saham di Indonesia dapat dilihat di Bursa Efek Indonesia (BEI). Saham Bank BCA merupakan salah satu saham di dalamnya. Bank BCA adalah perusahaan yang beroperasi di bidang perbankan sejak tahun 1956. Menurut CNBC Indonesia pada tahun 2021, saham Bank BCA merupakan salah satu saham yang paling menghasilkan dalam lima tahun terakhir. Dari **Gambar 1** data dalam lima tahun terakhir saham Bank BCA mengalami naik turun relatif cepat [3]. Karena pergerakan saham saat ini relatif cepat, para pemilik saham diharapkan dapat mengambil keputusan kapan saham akan dijual atau dipertahankan. Untuk itu prediksi harga saham sangat penting dalam pengambilan keputusan tersebut. Prediksi adalah proses meramalkan suatu keadaan pada masa depan dengan melakukan analisa data masa lampau [4]. Dalam memprediksi harga saham diperlukan suatu metode pengolahan data yang biasa disebut *data mining*.



Gambar 1. Grafik saham Bank BCA lima tahun terakhir

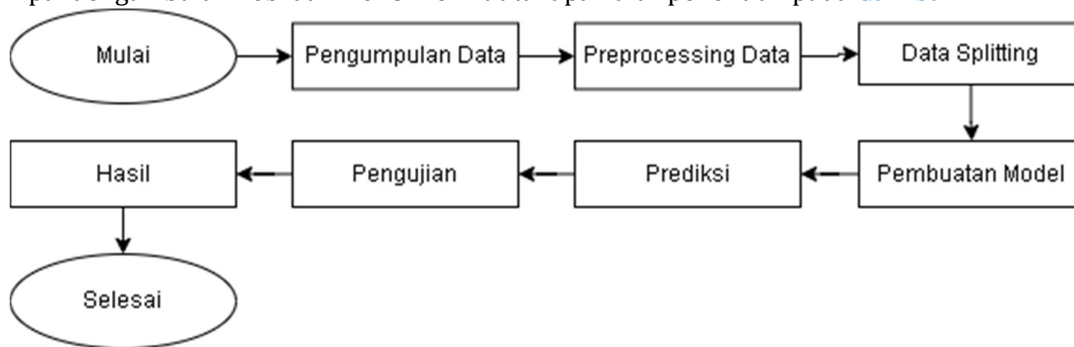
Data mining merupakan aktivitas yang berhubungan dengan pengumpulan data, penggunaan data histori untuk mendapatkan pengetahuan, informasi dan data yang besar dalam suatu pola [5]. *Data mining* membutuhkan suatu metode untuk melakukan ekstraksi data yang berisi informasi, pola dan pengetahuan. Dalam penerapannya *data mining* menggunakan berbagai metode, diantaranya metode Linear Regresi dan Polynomial Regresi. Linear Regresi adalah analisis statistika yang memodelkan beberapa hubungan variabel menurut bentuk hubungan persamaan linier eksplisit [6], sedangkan Polynomial Regresi adalah sebuah model linear regresi berganda yang dibentuk dengan menjumlahkan pengaruh masing-masing variabel prediktor (x) yang dipangkatkan meningkat sampai orde ke-n [7].

Berdasarkan penelitian sebelumnya yang telah dilakukan, penggunaan metode linear mendapatkan hasil dengan kriteria sangat baik dalam memprediksi harga emas [6]. Pada penelitian lain dengan metode yang sama juga mendapatkan hasil dengan kategori sangat baik dalam memprediksi penjualan properti pada PT XYZ [8]. Sementara, penelitian menggunakan metode Polynomial Regresi dapat berjalan dengan baik dalam memprediksi nilai jual objek pajak tanah di Kabupaten Gresik [9]. Penggunaan metode Polynomial Regresi dalam memprediksi produksi daging sapi nasional juga mendapatkan nilai yang baik berdasarkan dataset produksi daging sapi nasional tahun 1984 s.d. 2019 [10].

Berdasarkan penelitian-penelitian sebelumnya yang telah diuraikan, dapat diketahui bahwa metode Linear Regresi dan Polynomial Regresi mendapatkan hasil dengan kategori yang baik dalam melakukan sebuah peramalan atau prediksi. Maka dari itu metode Linear Regresi dan Polynomial Regresi digunakan dalam penelitian ini. Pengolahan model pada kedua metode tersebut menggunakan python dengan *tools* Anaconda. Tujuan penelitian ini adalah menerapkan metode Linear Regresi dan Polynomial Regresi untuk memprediksi pergerakan harga saham Bank BCA dan mengukur nilai error dari setiap metode yang digunakan serta membandingkan mana metode terbaik untuk melakukan prediksi sehingga dapat menjadi acuan untuk para pemilik saham dalam mengambil keputusan.

2. METODE

Penelitian ini menggunakan jenis data kuantitatif, yaitu data yang bersifat numerik. Data yang digunakan untuk penelitian ini adalah data harga saham Bank BCA dari bulan Maret 2018 sampai dengan bulan Februari 2023. Berikut tahapan alur penelitian pada **Gambar 2**.



Gambar 2. Tahapan penelitian

2.1 Pengumpulan data

Pengumpulan data dilakukan pada 01 Maret 2023. Data diperoleh pada *website yahoofinance.com*. Data yang diambil adalah harga saham dari tanggal 01 Maret 2018 sampai dengan 28 Februari 2023. Jumlah data yang diperoleh adalah 1.252 dataset. Data yang telah diperoleh akan dikumpulkan dan dianalisis agar menjadi informasi yang baru sehingga dapat digunakan untuk memprediksi masalah yang diteliti.

2.2 Preprocessing data

Preprocessing data merupakan tahapan yang harus dikerjakan dalam proses *data mining*. *Preprocessing data* sangat penting karena tidak semua data atau atribut data yang ada di dalam data dibutuhkan dalam proses *data mining* [11]. Pada penelitian ini tahapan *preprocessing data* menggunakan dua tahapan yaitu *data selection* dan *data cleaning*. *Data selection* atau seleksi data adalah tahapan untuk memilih data atau atribut yang dibutuhkan dalam proses *data mining*, karena tidak semua atribut dipakai untuk proses data mining [12]. *Data cleaning* atau pembersihan data merupakan tahapan untuk memeriksa duplikasi data, data yang inkonsisten dan memperbaiki kesalahan pada data [12].

2.3 Data Splitting

Data Splitting atau pemisahan data adalah proses membagi data menjadi dua bagian atau lebih data yang memiliki banyak nilai [13]. Pada tahap data *splitting* dilakukan pemisahan data yang digunakan untuk memisahkan dari *independent* variabel dan *dependent* variabel. Selain itu, pada tahap data *splitting* dilakukan pemisahan *training model* dan *testing model*. Pembagian data antara *training model* dan *testing model* digunakan untuk proses melatih model dan untuk mengetahui kualitas dari metode yang digunakan [14]. *Training model* digunakan untuk melatih metode sebelum metode tersebut membentuk model pelatihan, kemudian untuk menghitung akurasi dari model yang sudah dilatih dilakukan *testing model*.

2.4 Pembuatan model

Dalam tahap ini dilakukan pembuatan model untuk prediksi dengan menggunakan data *training*. Metode yang digunakan adalah Linear Regresi dan Polynomial Regresi. Perbedaan antara kedua metode ini terletak pada cara variabel (x) diproses sebelum dimasukkan ke dalam model [10]. Dalam Linear Regresi, variabel (x) tidak mengalami perubahan, dan model mencoba untuk menemukan hubungan linier antara variabel (x) dan variabel (y). Sementara itu, dalam Polynomial Regresi variabel (x) diubah atau dinaikkan ke pangkat tertentu (misalnya X^2 , X^3 , dst.) sebelum dimasukkan ke dalam model.

2.4.1 Linear regresi

Dalam tahapan pembuatan model ini dilakukan perhitungan untuk mencari nilai *intercept* dan *slope* menggunakan python dan perhitungan manual. *Intercept* adalah suatu titik persinggungan suatu garis antara sumbu y pada grafik saat nilai $x = 0$, sedangkan *slope* adalah ukuran kemiringan dari suatu garis yang menunjukkan besarnya pengaruh variabel x terhadap variabel y. Rumus untuk Linear Regresi dengan metode kuadrat terkecil atau sederhana [8]. Dijelaskan pada persamaan 1, 2 dan 3.

$$a = \frac{(\sum y)(\sum x^2) - (\sum x)(\sum xy)}{n(\sum x^2) - (\sum x)^2} \quad (1)$$

$$b = \frac{n(\sum xy) - (\sum x)(\sum y)}{n(\sum x^2) - (\sum x)^2} \quad (2)$$

$$y = a + bx \quad (3)$$

Dengan y adalah jumlah kasus, x adalah periode waktu, a adalah *intersep*, konstanta untuk menunjukkan besarnya nilai y ketika $x = 0$, dan b adalah *slope*, besaran perubahan nilai y.

2.4.2 Polynomial regresi

Dalam tahapan pembuatan model Polynomial Regresi perhitungan dilakukan dengan cara menjumlahkan pengaruh tiap-tiap variabel prediktor (x) yang dipangkatkan meningkat sampai orde ke-n. Perhitungan dilakukan dengan pemrograman python dan perhitungan manual. Perhitungan manual akan menggunakan metode eliminasi gauss untuk menghitung nilai koefisien yang optimal sehingga dapat meningkatkan akurasi prediksi. Rumus untuk Polynomial Regresi secara umum ditulis dalam persamaan 4 dan 5 [10].

$$Y = b_0 + b_1X + b_2X^2 + \dots + b_nX^n + \epsilon \quad (4)$$

Apabila Y merupakan nilai yang diprediksi:

$$Y = b_0 + b_1X + b_2X^2 + \dots + b_nX^n \quad (5)$$

Dimana Y adalah variabel yang diprediksi, b_0 adalah *intersep*, b_1 , b_2 , ..., b_n adalah *slope* atau koefisien-koefisien regresi, X adalah variabel yang bebas, n adalah orde atau derajat polynomial dan ϵ faktor galat.

2.5 Prediksi

Prediksi dilakukan setelah pembuatan model selesai. Perhitungan prediksi dilakukan dengan mencari nilai y sesuai rumus pada metode yang digunakan. Perhitungan prediksi menggunakan data test hasil dari *splitting* data sebanyak 251 data. Prediksi dapat dilakukan dengan menentukan atribut prediktor (x) untuk dimasukkan ke dalam model yang telah dibuat pada pembuatan model.

2.6 Pengujian

Akurasi suatu peramalan atau prediksi ditentukan oleh seberapa besar penyimpangan antara data yang diramalkan dengan data aktual [15]. Penyimpangan atau kesalahan sebuah prediksi ada beberapa cara untuk menghitung besarnya. Salah satunya adalah *mean absolute percentage error* (MAPE). MAPE merupakan rata-rata diferensiasi absolut antara nilai yang diprediksi dan aktual. Hasil dari suatu prediksi dinyatakan sangat baik jika nilai MAPE kurang dari 10%. Kriteria MAPE dapat dilihat pada **Tabel 1**.

Tabel 1. Kriteria MAPE

MAPE	Pengertian
<10%	Sangat Baik
10% - 20%	Baik
20% - 50%	Cukup
50%	Buruk

Rumus cara pengujian ini ditampilkan pada rumus persamaan 6:

$$MAPE = \frac{1}{n} \sum_{t=1}^n \frac{|Y_t - Y'_t|}{Y_t} \times 100 \quad (6)$$

dengan Y_t adalah nilai aktual periode t , Y'_t adalah nilai prediksi periode t , dan n adalah banyaknya periode.

2.7 Hasil

Pada tahap ini hasil pengujian dari metode Linear Regresi dan Polynomial Regresi akan dibandingkan metode manakah yang lebih baik untuk memprediksi. Acuan perbandingan dilihat dari nilai MAPE.

3. HASIL DAN PEMBAHASAN

3.1 Pengumpulan data

Pengumpulan data dilakukan pada 01 Maret 2023. Data diperoleh pada *website yahoofinance.com*. Data yang diambil adalah harga saham dari tanggal 01 Maret 2018 sampai dengan 28 Februari 2023. Data ini berisi enam atribut dengan jumlah 1.252 data. Atribut-atribut tersebut diantaranya yaitu: *date*, *open*, *high*, *low*, *close* dan *volume*. Isi dataset dapat dilihat pada **Tabel 2**.

Tabel 2. Isi Dataset

No	Date	Open	High	Low	Close	Adj Close	Volume
1	01/03/2018	4620	4645	4620	4635	4280	125474500
2	02/03/2018	4615	4625	4550	4575	4225	151511000
3	05/03/2018	4630	4660	4575	4575	4225	87744000
4	06/03/2018	4635	4640	4545	4555	4206	75907000
5	07/03/2018	4620	4630	4505	4520	4174	106152000
...
1248	22/02/2023	8700	8750	8625	8675	8675	38671900
1249	23/02/2023	8625	8750	8625	8725	8725	53412300

No	Date	Open	High	Low	Close	Adj Close	Volume
1250	24/02/2023	8775	8850	8675	8675	8675	44123400
1251	27/02/2023	8675	8775	8650	8775	8775	59983900
1252	28/02/2023	8775	8775	8675	8750	8750	106349400

3.2 Preprocessing data

Pada *preprocessing* data akan dilakukan dua tahap proses yaitu *data selection* dan *data cleaning*. Data yang telah diperoleh mempunyai enam atribut yaitu *date*, *open*, *high*, *low*, *close* dan *volume*. Pada tahap *data selection* akan dipilih atribut yang dibutuhkan yaitu atribut *date* dan *close*. Penyeleksian data dilakukan dengan pemrograman python dapat dilihat pada **Gambar 3**.

	Date	Close
0	2018-03-01	4635.0
1	2018-03-02	4575.0
2	2018-03-05	4575.0
3	2018-03-06	4555.0
4	2018-03-07	4520.0
...
1247	2023-02-22	8675.0
1248	2023-02-23	8725.0
1249	2023-02-24	8675.0
1250	2023-02-27	8775.0
1251	2023-02-28	8750.0

1252 rows x 2 columns

Gambar 3. Hasil data selection

Gambar 3 merupakan atribut yang akan digunakan untuk prediksi yaitu kolom Date dan Close. Atribut Date dan Close berkaitan karena, pada Atribut Close dapat bertambah dan berkurang nilainya berdasarkan Atribut Date, maka dari itu kedua atribut tersebut yang digunakan. Setelah itu dilakukan proses *data cleaning* yaitu pengecekan *missing value* pada data. Hasilnya ditampilkan dalam **Gambar 4**.

```
In [10]: df.isnull().sum()
Out[10]: Date          0
         Open          0
         High          0
         Low           0
         Close         0
         Adj Close     0
         Volume        0
         dtype: int64
```

Gambar 4. Hasil data cleaning

Berdasarkan **Gambar 4** data yang digunakan tidak terdapat data yang kosong atau *missing value* sehingga data siap digunakan pada tahap selanjutnya.

3.3 Data splitting

Pada tahap ini dilakukan pemisahan data untuk proses pembuatan model yang terdiri dari *independent* variabel (x) yaitu atribut *date* dan *dependent* variabel (y) yaitu atribut *close*. *Source code* untuk pemisahan variabel x dan y ditampilkan dalam **Gambar 5**.

```
In [16]: x = data_selection['Open'].values.reshape(-1,1)
         y = data_selection['Close']
```

Gambar 5. Source code pemisahan variabel x dan y

Kemudian dilakukan pembagian *training model* dan *testing model*. Untuk menentukan rasio pembagian antara data *training* dan data *testing* dibutuhkan suatu perhitungan akurasi dari kedua metode dengan mengambil beberapa contoh rasio yang umum digunakan. Adapun beberapa rasio pembagian data ditunjukkan pada **Tabel 3**.

Tabel 3. Rasio akurasi metode

No	Rasio	Akurasi Linear Regresi	Akurasi Polynomial Regresi
1	80:20	0,8229	0,8350
2	85:15	0,8177	0,8331
3	90:10	0,8062	0,8187
Rata-rata		0,8156	0,8289

Berdasarkan Tabel 3 akurasi tertinggi dari kedua metode yang digunakan pada pembagian data ada pada rasio 80:20 dengan akurasi Linear Regresi dan *Polynomial Regresi* sebesar 0,8229 dan 0,8350. Maka dari itu, rasio tersebut yang akan digunakan pada proses pembuatan model. Kemudian dilakukan proses split data menggunakan pemrograman phyton. *Source code* untuk pembagian *training model* dan *testing model* ditampilkan dalam **Gambar 6**.

```
In [22]: x_train, x_test, y_train, y_test = train_test_split(x, y, test_size = 0.2, random_state = 0)

In [23]: print('Shape of x_train :', x_train.shape) #80% dari 1252
         print('Shape of y_train :', y_train.shape)
         print('Shape of x_test :', x_test.shape)
         print('Shape of y_test :', y_test.shape)

Shape of x_train : (1001, 1)
Shape of y_train : (1001,)
Shape of x_test : (251, 1)
Shape of y_test : (251,)
```

Gambar 6. Pembagian training model dan testing model

Berdasarkan **Gambar 6** pembagian data menggunakan rasio 80:20. Jumlah *data train* 1001 data dan jumlah *data test* 251 data. *Data train* akan digunakan untuk pemodelan metode dan *data test* digunakan untuk mengetahui akurasi metode.

3.4 Pembuatan model

Pada tahap ini dilakukan pembuatan model dengan menggunakan metode *Linear Regresi* dan *Polynomial Regresi*.

3.4.1 Linear regresi

Pada tahap ini dilakukan pemanggilan model Linear Regresi dan dimasukkan kedalam variabel linear. Kemudian dilakukan proses *training* pada data. Pembuatan model Linear Regresi dapat dilihat pada **Gambar 7**.

```
In [26]: linear = LinearRegression()
In [27]: model_linear = linear.fit(x_train, y_train)
```

Gambar 7. Pembuatan model linear regresi

Setelah model terbentuk, maka kita dapat melihat nilai *intercept* (a) dan *slope* (b) yang digunakan sebagai parameter untuk melakukan prediksi. *Source Code* untuk menampilkan nilai *intercept* dan *slope* ditampilkan dalam **Gambar 8**.

```
In [18]: print('Nilai intercept (a): ', linear.intercept_)
print('Nilai slope (b): ', linear.coef_)
Nilai intercept (a): 4445.719031550297
Nilai slope (b): [3.0744244]
```

Gambar 8. Nilai *intercept* dan *slope* linear regresi

Berdasarkan **Gambar 8** pada model Linear Regresi yang telah dibuat, nilai *intercept* (a) yaitu 4445,719031550297 dan nilai *slope* (b) yaitu 3,0744244. Sehingga untuk mencari nilai prediksi (y) pada model Linear Regresi yang telah dibuat yaitu sebagai berikut.

$$Y = 4445,719031550297 + 3,0744244x$$

3.4.2 Polynomial regresi

Pada metode *Polynomial Regresi variable independent* (x) dilakukan proses transformasi dengan memanggil fungsi *PolynomialFeatures*. Setelah itu, melakukan *training* model pada data dengan menggunakan *variable* x yang telah ditransformasi. *Source code* pembuatan model *Polynomial Regresi* dapat pada **Gambar 9**.

```
In [33]: poly = PolynomialFeatures()
x_polynomial = poly.fit_transform(x_train)

polynomial = LinearRegression()
polynomial.fit(x_polynomial, y_train)

Out[33]: LinearRegression()
```

Gambar 9. Pembuatan model *polynomial* regresi

Pada **Gambar 9** pada variabel x dilakukan transformasi sebelum proses pembuatan model. Setelah model terbentuk, maka kita dapat melihat nilai *intercept* (b_0) dan *slope* atau koefisien-koefisien regresi (b_1 , b_2) yang digunakan sebagai parameter untuk melakukan prediksi. Nilai *intercept* dan *slope* pada *Polynomial Regresi* ditampilkan dalam **Gambar 10**.

```
In [36]: print('Nilai intercept: ', polynomial.intercept_)
print('Nilai slope: ', polynomial.coef_)
Nilai intercept: 4704.200617204707
Nilai slope: [0.00000000e+00 1.82481842e+00 9.99392400e-04]
```

Gambar 10. Nilai *intercept* dan *slope polynomial* regresi

Berdasarkan **Gambar 10** pada model *Polynomial Regresi*, nilai *intercept* (b_0) yaitu 4704,200617204707, nilai $b_1 = 1,82481842$, nilai $b_2 = 0,0009993924$. Sehingga untuk mencari nilai prediksi (y) pada model *Polynomial Regresi* yang telah dibuat yaitu sebagai berikut.

$$Y = 4704,200617204707 + 1,82481842 x + 0,0009993924 x^2$$

3.5 Prediksi

Pada tahap ini dilakukan prediksi dengan metode Linear Regresi dan *Polynomial Regresi* dengan menggunakan model yang sudah dibuat sebelumnya.

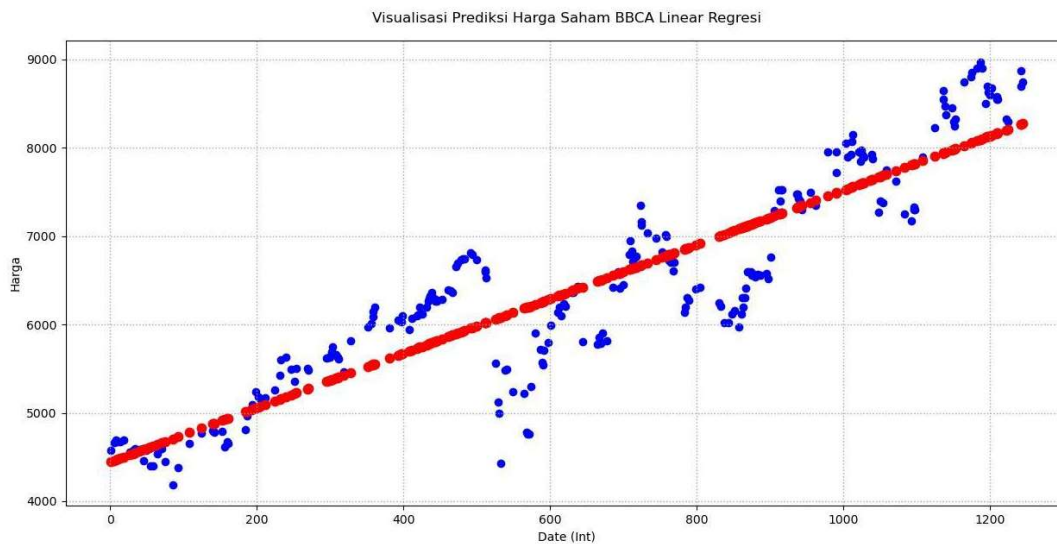
3.5.1 Prediksi linear regresi

Prediksi dilakukan dengan menghitung nilai Y dengan persamaan rumus (2) setelah nilai *intercept* (a) dan *slope* (b) telah didapatkan pada tahap pemodelan. Data yang digunakan untuk menghitung prediksi adalah data *testing*. Data *testing* adalah data hasil split data yang berjumlah 251 data. Untuk melihat hasil prediksi menggunakan metode Linear Regresi dapat dilihat pada **Tabel 4**.

Tabel 4. Prediksi linear regresi

Date	Date(x)	Harga Close(y)	Prediksi
08 March 2018	5	4660	4461,091
24 April 2019	299	5630	5364,972
24 August 2022	1108	7900	7852,181
23 December 2019	472	6660	5896,847
14 May 2018	52	4595	4605,589
...
...
13 April 2018	31	4580	4541,026
11 September 2019	399	6100	5672,414
30 May 2018	64	4540	4642,482
19 June 2020	589	5575	6256,555
10 October 2018	159	4675	4934,553

Kemudian untuk melihat gambaran dari model *Linear Regresi* dengan mudah dapat dilihat pada **Gambar 11**.



Gambar 11. Visualisasi model linear regresi

Pada **Gambar 11** menampilkan hasil visualisasi prediksi metode *Linear Regresi*, titik biru sebagai harga aktual dan garis lurus merah sebagai harga prediksi dengan metode Linear Regresi.

3.5.2 Prediksi *polynomial* regresi

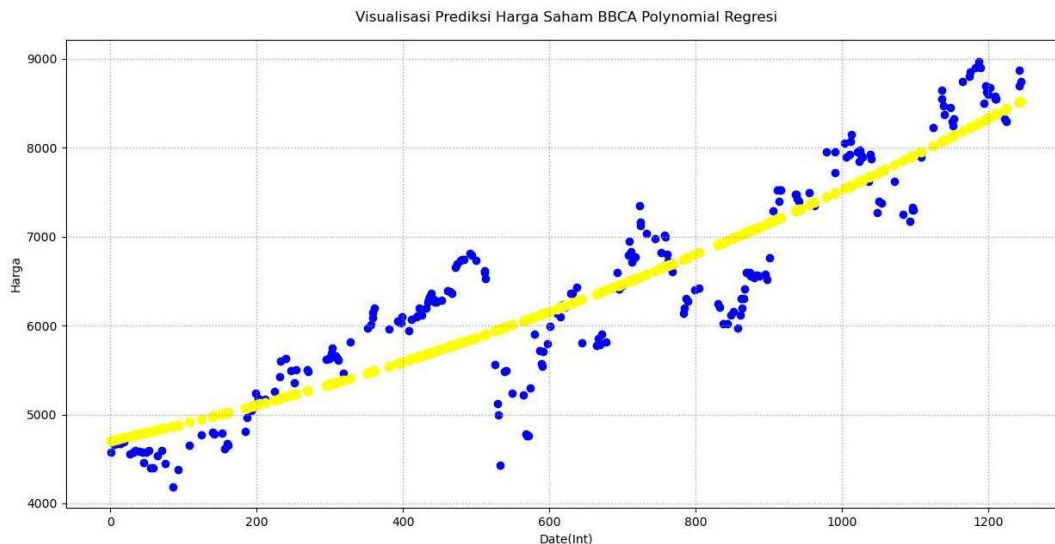
Prediksi dilakukan dengan menghitung nilai Y dengan persamaan rumus (5) setelah nilai *intercept* serta koefisien b_1 dan b_2 didapatkan pada proses pemodelan. Data yang digunakan untuk menghitung prediksi adalah data *testing*. Data *testing* adalah data hasil split data yang

berjumlah 251 data. Untuk melihat hasil prediksi menggunakan metode Linear Regresi dapat dilihat pada **Tabel 5**.

Tabel 5. Prediksi *polynomial* regresi

Date	Date(x)	Harga Close(y)	Prediksi
08 March 2018	5	4660	4713,35
24 April 2019	299	5630	5339,168
24 August 2022	1108	7900	7953,017
23 December 2019	472	6660	5788,164
14 May 2018	52	4595	4801,794
...
...
13 April 2018	31	4580	4761,73
11 September 2019	399	6100	5591,407
30 May 2018	64	4540	4825,083
19 June 2020	589	5575	6125,729
10 October 2018	159	4675	5019,612

Kemudian untuk melihat gambaran dari model *Polynomial* Regresi dengan mudah dapat pada **Gambar 12**.



Gambar 12. Visualisasi model *polynomial* regresi

Gambar 12 menampilkan hasil visualisasi prediksi metode *Polynomial Regresi*, titik biru sebagai harga aktual dan garis lengkung kuning sebagai harga prediksi dengan metode *Polynomial Regresi*.

3.6 Pengujian

Pada tahap pengujian didasarkan pada nilai penyimpangan antara nilai aktual dengan nilai yang diprediksi sesuai dengan model yang dibuat. Metode Perhitungan yang digunakan yaitu *mean absolute percentage error* (MAPE). Nilai MAPE dari Linear regresi sebesar 6,55%, sedangkan nilai MAPE dari *Polynomial* Regresi sebesar 6,54%. Dari hasil ini, maka dapat dikatakan kedua metode yang digunakan untuk memprediksi harga saham Bank BCA dalam kriteria sangat baik karena hasilnya dibawah 10%.

3.7 Hasil

Prediksi harga saham Bank BCA menggunakan metode Linear Regresi dan *Polynomial* Regresi. Hasil dari perhitungan diuji menggunakan *mean absolute percentage error* (MAPE). Untuk melihat perbandingan nilainya dapat pada **Tabel 6**.

Tabel 6. Perbandingan nilai pengujian

Metode	Mean Absolut Percentage Error
Linear Regresi	6,55%
Polynomial Regresi	6,54%

Dari metode yang digunakan dalam menguji hasil prediksi, dapat dilihat bahwa model dengan menggunakan metode *Polynomial* Regresi dapat memprediksi lebih baik dibandingkan dengan metode Linear Regresi. Karena model dengan metode *Polynomial* Regresi mendapatkan nilai *MAPE* yang lebih kecil yaitu 6,54%.

4. SIMPULAN

Berdasarkan hasil dan pembahasan dari penelitian, dapat disimpulkan bahwa Linear Regresi dan *Polynomial* Regresi dapat diterapkan untuk memprediksi saham Bank BCA periode 2018 sampai dengan 2023. Hasil pengujian kedua metode mendapatkan nilai dalam kriteria sangat baik, yaitu nilai *MAPE* dibawah 10%. Metode *Polynomial* Regresi mendapatkan nilai yang lebih baik dibandingkan dengan metode Linear Regresi. Hasil pengujian pada metode Linear Regresi mendapatkan nilai *MAPE* = 6,55%, sedangkan pada metode *Polynomial* Regresi mendapatkan nilai *MAPE* = 6,54%.

REFERENSI

- [1] A. Darmawan, K. Kurnia, and S. Rejeki, "PENGETAHUAN INVESTASI, MOTIVASI INVESTASI, LITERASI KEUANGAN DAN LINGKUNGAN KELUARGA PENGARUHNYA TERHADAP MINAT INVESTASI DI PASAR MODAL," *J. Ilm. Akunt. dan Keuang.*, vol. 8, no. 2, 2019, doi: 10.32639/jiak.v8i2.297.
- [2] S. D. Nabella, A. Munandar, and R. Tanjung, "LIKUIDITAS, SOLVABILITAS, AKTIVITAS DAN PROFITABILITAS TERHADAP HARGA SAHAM PADA PERUSAHAAN SEKTOR TAMBANGAN BATU BARA YANG TERDAFTAR DI BURSA EFEK INDONESIA," *Meas. J. Akunt.*, vol. 16, no. 1, 2022, doi: 10.33373/mja.v16i1.4264.
- [3] "PT Bank Central Asia Tbk (BBCA.JK)." <https://finance.yahoo.com/chart/bbca.jk#eyJpbmRlcnZhbCI6ImRheSI6InBlcmVZGljaXR5IjoxLCJ0aW1lVW5pdCI6bnVsbCwiY2FuZGxlV2lkdGgiOjgsImZsaXBwZWQiOmZhbHNILCJ2b2x1bWVVbmlcmxheSI6dHJ1ZSwiYWwRqjlp0cnVILCJjcm9zc2hhaXIiOnRydWUImNoYXJ0VHlwZSI6ImxpbmUiLCJleHRlbnRlZCI.>
- [4] S. Adiguno, Y. Syahra, and M. Yetri, "Prediksi Peningkatan Omset Penjualan Menggunakan Metode Regresi Linier Berganda," *J. Sist. Inf. Triguna Dharma (JURSI TGD)*, vol. 1, no. 4, 2022, doi: 10.53513/jursi.v1i4.5331.
- [5] Z. Nabila, A. Rahman Isnain, and Z. Abidin, "ANALISIS DATA MINING UNTUK CLUSTERING KASUS COVID-19 DI PROVINSI LAMPUNG DENGAN ALGORITMA K-MEANS," *J. Teknol. dan Sist. Inf.*, vol. 2, no. 2, 2021.
- [6] N. Nafi'iyah, "Perbandingan Regresi Linear , Backpropagation Dan Fuzzy Mamdani Dalam Prediksi Harga Emas," *Semin. Nas. Inov. dan Apl. Teknol. di Ind.*, 2016.
- [7] N. Y. I. Ginting, A. Novianty, and M. T. A. L. Prasasti, "Estimasi Magnitudo Gempa Bumi Dari Sinyal Seismik Gelombang P Menggunakan Metode Regresi Polinomial," *eProceedings Eng.*, vol. 7, no. 2, pp. 4624-4632, 2020.
- [8] G. N. Ayuni and D. Fitrianah, "Penerapan metode Regresi Linear untuk prediksi penjualan properti pada PT XYZ," *J. Telemat.*, vol. 14, no. 2, 2019.
- [9] M. Firdaus, M. Hafiyusholeh, and S. Widodo, "Prediksi Nilai Jual Objek Pajak (NJOP) Tanah di Kabupaten Gresik Menggunakan Regresi Polinomial," *J. Mhs. Mat. Algebr.*, vol. 1, no. 1,

2020.

- [10] "Prediksi Produksi Daging Sapi Nasional dengan Metode Regresi Linier dan Regresi Polinomial," *J. Ilm. Komputasi*, vol. 20, no. 2, 2021, doi: 10.32409/jikstik.20.2.2722.
- [11] R. R. Rerung, "Penerapan Data Mining dengan Memanfaatkan Metode Association Rule untuk Promosi Produk," *J. Teknol. Rekayasa*, vol. 3, no. 1, 2018, doi: 10.31544/jtera.v3.i1.2018.89-98.
- [12] J. Eska, "Penerapan Data Mining Untuk Prediksi Penjualan Wallpaper Menggunakan Algoritma C4.5 STMIK Royal Ksieran," *JURTEKSI (Jurnal Teknol. dan Sist. Informasi)*, vol. 2, 2016.
- [13] R. A. Zuama, S. Rahmatullah, and Y. Yuliani, "Analisis Performa Algoritma Machine Learning pada Prediksi Penyakit Cerebrovascular Accidents," *J. MEDIA Inform. BUDIDARMA*, vol. 6, no. 1, 2022, doi: 10.30865/mib.v6i1.3488.
- [14] J. S. Putra, R. D. Ramadhani, and A. Burhanuddin, "Prediksi Harga Saham Bank Bri Menggunakan Algoritma Linear Regression Sebagai Strategi Jual Beli Saham," *J. Dinda Data Sci. Inf. Technol. Data Anal.*, vol. 2, no. 1, 2022, doi: 10.20895/dinda.v2i1.273.
- [15] A. N. Maharadja, I. Maulana, and B. A. Dermawan, "Penerapan Metode Regresi Linear Berganda untuk Prediksi Kerugian Negara Berdasarkan Kasus Tindak Pidana Korupsi," *J. Appl. Informatics Comput.*, vol. 5, no. 1, 2021, doi: 10.30871/jaic.v5i1.3184.