

PENERAPAN MODEL REGRESI LINIER UNTUK PREDIKSI HARGA SAHAM BANK BCA PADA BURSA EFEK INDONESIA

Rizky Zapar¹, Denni Pratama², Kaslani³, Cep Lukaman Rohmat⁴, Faturrohman⁵

¹ Teknik Informatika, STMIK IKMI Cirebon

^{2,3} Komputerisasi Akuntansi, STMIK IKMI Cirebon

^{4,5} Rekayasa Perangkat Lunak, STMIK IKMI Cirebon

Jalan Perjuangan No.10B Karyamulya Kec. Kesambi Kota. Cirebon Jawa Barat 45153, Indonesia

Rizky.zapar@gmail.com

ABSTRAK

Dalam era globalisasi dan kompleksitas pasar modal, analisis prediksi harga saham menjadi elemen krusial bagi keberhasilan investor dan perusahaan. Fluktuasi harga saham yang tidak menentu menciptakan tantangan dalam meramalkan pergerakan pasar. Metodologi KDD (Knowledge Discovery in Databases) digunakan untuk mengekstraksi pengetahuan berharga dari data historis harga saham Bank BCA. Dengan fokus pada metode regresi linier, penelitian ini bertujuan untuk meningkatkan akurasi prediksi harga saham dan memvalidasi model menggunakan K-Fold Cross-Validation. Hasil evaluasi menunjukkan nilai RMSE sebesar 0.032, menandakan tingkat kesalahan yang rendah dan konsistensi kinerja model. Absolute Error sebesar 0.024 dengan rentang 0.007 mengindikasikan kemampuan model memberikan perkiraan yang akurat terhadap pergerakan harga saham. Root Relative Squared Error sebesar 0.138 dengan rentang 0.036 mencerminkan tingkat kesalahan relatif terhadap variasi data yang dapat dipertahankan oleh model. Dengan Squared Error sebesar 0.001 dan rentang 0.001, sebaran kesalahan prediksi model menunjukkan bahwa sebagian besar prediksi model berada dalam kisaran nilai sebenarnya. Model regresi linier dengan pendekatan KDD mampu memberikan prediksi harga saham Bank BCA dengan tingkat akurasi yang tinggi dan konsistensi yang baik, memberikan landasan yang kuat untuk pengambilan keputusan investasi di masa depan.

Kata kunci: *Prediksi Harga Saham, Metodologi Knowledge Discovery in Databases (KDD), Regresi Linier, K-Fold Cross-Validation, Bank BCA.*

1. PENDAHULUAN

Dalam era globalisasi, pasar modal memiliki peran penting dalam ekonomi global, terutama di negara-negara dengan ekonomi pasar yang kuat. Saham, sebagai instrumen utama di pasar modal, mewakili kepemilikan dan memiliki nilai yang berfluktuasi. Pergerakan harga saham dipengaruhi oleh faktor kompleks, dan pemahaman tentang hal ini penting bagi investor dan perusahaan di pasar modal [1]. Menurut [2] Bahwa investasi dalam saham memiliki dampak yang signifikan terhadap perusahaan-perusahaan di seluruh dunia. Keberhasilan atau kegagalan suatu perusahaan dapat tercermin dalam pergerakan harga sahamnya di pasar modal. Oleh karena itu, pemahaman yang akurat dan analisis mendalam terhadap faktor-faktor yang mempengaruhi pergerakan harga saham sangat penting bagi perusahaan dan investor.

Pergerakan harga saham cenderung tidak dapat diprediksi dengan pasti. Saat ini, kondisi pasar global sangat kompleks dan dipengaruhi oleh berbagai aspek ekonomi, politik, sosial, dan teknologis. Fluktuasi harga saham yang tidak menentu menciptakan tantangan dalam meramalkan pergerakan pasar. Untuk melakukan prediksi harga saham menjadi suatu hal yang relevan dan penting. Prediksi harga saham memberi panduan untuk evaluasi investasi di masa depan. Dalam ranah perdagangan saham, terdapat dua jenis analisis yang umum dikenal, yaitu analisis fundamental dan analisis teknikal. Jika data yang

digunakan dalam penelitian ini mengandung noise atau data yang tidak valid, seperti data yang hilang atau data yang tidak konsisten, maka hasil prediksi yang dihasilkan juga akan tidak akurat. Oleh karena itu, penting untuk melakukan preprocessing data yang baik untuk mengatasi masalah ini.

Dalam penelitian [3], digunakan dua metode, yaitu regresi linier berganda dan ARIMA. Metode regresi linier berganda digunakan untuk menganalisis hubungan antara variabel independen (harga pembukaan, harga tertinggi, harga terendah, dan volume perdagangan) dengan variabel dependen (harga penutupan) saham PT BSI, Tbk. Sedangkan metode ARIMA digunakan untuk menganalisis data deret waktu dan membuat prediksi berdasarkan pola historis. Dalam penelitian [4] Hasil penelitian menunjukkan bahwa model Regresi Linier (LR) memiliki nilai kesalahan terendah, dengan RMSE sebesar 0,010 untuk saham TLKM, 0,011 untuk saham KAEF, dan 0,012 untuk saham INDF. Model Regresi Random Forest (RFR) memiliki RMSE sebesar 0,013 untuk saham KAEF, 0,014 untuk saham TLKM, dan 0,018 untuk saham INDF. Hasil prediksi ini dianggap baik karena memiliki nilai kesalahan yang relatif rendah dibandingkan dengan metode MLP. Dan dalam penelitian [5] memiliki nilai MAPE yang baik dalam memprediksi harga saham Bank BCA. Metode regresi linear menghasilkan nilai MAPE sebesar 6,55%, sedangkan metode regresi polinomial menghasilkan nilai MAPE sebesar 6,54%

Tujuan dari penelitian ini untuk menganalisa pola prediksi harga saham dan memvalidasi model prediksi harga saham Bank BCA dengan menerapkan pendekatan regresi linier dan kerangka kerja KDD (Knowledge Discovery in Databases). Peneliti ingin meningkatkan akurasi dan keandalan prediksi harga saham dengan memanfaatkan metode regresi linier dalam mengatasi kompleksitas data historis harga saham, dan untuk memvalidasi kinerja model dengan menggunakan K-Fold Cross-Validation, memastikan bahwa model dapat menggeneralisasi dengan baik ke data baru dan memberikan hasil prediksi yang dapat diandalkan.

Linear Regression dipilih sebagai metode analisis karena dapat digunakan pada dataset dengan nilai yang tidak stabil. Kelebihan lainnya adalah kemampuan metode ini untuk diaplikasikan pada data tren, seperti pergerakan harga saham [2]. Dalam upaya meningkatkan akurasi prediksi harga saham Bank BCA, peneliti berencana untuk menerapkan model KDD (Knowledge Discovery in Databases) sebagai kerangka kerja analisis. Dan untuk memvalidasi kinerja model, Peneliti juga akan menerapkan K-Fold Cross-Validation. Dengan membagi data ke dalam subset yang saling non-overlapping, untuk dapat mengukur sejauh mana model dapat menggeneralisasi dan memberikan prediksi yang akurat pada data baru.

Hasil dari penelitian ini memberikan wawasan tentang prediksi harga saham Bank BCA dengan menerapkan model regresi linier. Dengan memanfaatkan data historis harga saham dan variabel yang relevan. Peneliti berharap perlu dilakukan penelitian lebih lanjut untuk mengatasi potensi keterbatasan dalam data yang digunakan, termasuk keberlanjutan ketersediaan data historis harga saham dan peningkatan kualitas data. Implementasi strategi pre-processing yang lebih akurat atau penerapan teknik imputasi data yang efektif dapat menjadi fokus penelitian untuk meminimalkan dampak data yang tidak lengkap atau tidak konsisten.

2. TINJAUAN PUSTAKA

2.1. Metode Literatur Review

Metode yang digunakan dalam penelitian [4] Regresi Linier (LR), Regresi Random Forest (RRF), dan Multilayer Perceptron (MLP). Penelitian ini menggunakan model Cross-Standard Industry for Data Mining (CRISP-DM) yang terdiri dari tahapan Business Understanding, Data Understanding, Data Preparation, Modelling, Evaluation, dan Deployment. Proses preprocessing meliputi pembacaan file dataset, penyaringan atribut, pembebasan data yang tidak lengkap, normalisasi, dan partitioning data. Dalam penelitian [5] metode yang digunakan adalah metode regresi linear dan regresi polinomial. Metode regresi linear digunakan untuk menemukan hubungan linier antara variabel (x) dan variabel (y), sedangkan metode regresi polinomial mengubah atau meningkatkan variabel (x) ke pangkat tertentu sebelum dimasukkan ke dalam model. Penelitian ini menggunakan metode

regresi linear dan regresi polinomial untuk memprediksi harga saham Bank BCA.

Metode yang digunakan dalam penelitian [6] adalah metode regresi linear. Metode ini digunakan untuk memprediksi harga saham Bank BRI berdasarkan atribut Date dan Price. Penelitian ini menggunakan dataset historis harga saham Bank BRI sebagai input untuk model regresi linear. Dalam penelitian [7], metode yang digunakan untuk analisis data adalah Meta-Analisis. Meta-analisis merupakan sebuah analisis statistik yang memadukan hasil berbagai kajian ilmiah untuk pertanyaan yang sama. Penelitian ini menggunakan metode Linear Regression, ARIMA, dan Neural Network dalam prediksi harga saham. Metode ini dipilih karena mampu memberikan tingkat akurasi yang tinggi dalam prediksi harga saham. Selain itu, penelitian ini juga melakukan analisis variable moderator berdasarkan tahun terbit menggunakan JASP.

Dalam penelitian [8], Metode yang diterapkan dalam penelitian ini melibatkan Regresi Linear Berganda dan Moving Average. Regresi linear berganda digunakan untuk melakukan prediksi harga saham Bank BCA pada rentang waktu tahun 2015-2017. Keputusan untuk menggunakan metode ini didasarkan pada kemampuannya untuk menghasilkan tren harga saham baru, terutama jika garis kurva Linear Regression dapat menembus harga. Selain itu, Moving Average digunakan untuk mengukur momentum dan mengidentifikasi area support yang mungkin relevan. Penting untuk dicatat bahwa metode Regresi Linear Berganda memiliki asumsi klasik yang harus dipenuhi, termasuk normalitas, multikolinearitas, autokorelasi, heteroskedastisitas, dan asumsi linearitas. Sementara itu, dalam mengukur performa prediksi, digunakan Moving Average Percentage Error. Rumus ini menggambarkan persentase kesalahan hasil peramalan terhadap permintaan aktual selama periode tertentu. Dengan mengintegrasikan kedua metode ini, penelitian ini berusaha untuk memberikan pemahaman yang lebih holistik terhadap pergerakan harga saham Bank BCA dan meningkatkan ketepatan prediksi.

2.2. Hasil Literatur Review

Dalam Penelitian [4] “Analisa Perbandingan Model Prediction Dalam Prediksi Harga Saham Menggunakan Metode Linear Regression, Random Forest Regression dan Multilayer Perceptron” Hasil penelitian menunjukkan bahwa model Regresi Linier (LR) memiliki nilai kesalahan terendah, dengan RMSE sebesar 0,010 untuk saham TLKM, 0,011 untuk saham KAEF, dan 0,012 untuk saham INDF. Model Regresi Random Forest (RRF) memiliki RMSE sebesar 0,013 untuk saham KAEF, 0,014 untuk saham TLKM, dan 0,018 untuk saham INDF. Hasil prediksi ini dianggap baik karena memiliki nilai kesalahan yang relatif rendah dibandingkan dengan metode MLP. Hasil dan pembahasan dari penelitian [5], Berdasarkan hasil pengujian yang telah dilakukan

untuk memprediksi saham Bank BCA pada periode 2018 hingga 2023 menggunakan metode Linear Regresi dan Polynomial Regresi, dapat disimpulkan bahwa kedua metode tersebut memberikan hasil yang sangat baik. Kriteria keberhasilan diukur dengan nilai Mean Absolute Percentage Error (MAPE) yang berada di bawah 10%. Metode Polynomial Regresi menunjukkan kinerja yang lebih unggul dibandingkan dengan metode Linear Regresi. Dengan nilai MAPE sebesar 6,54%, Polynomial Regresi menghasilkan prediksi yang lebih akurat jika dibandingkan dengan Linear Regresi yang memiliki nilai MAPE sebesar 6,55%.

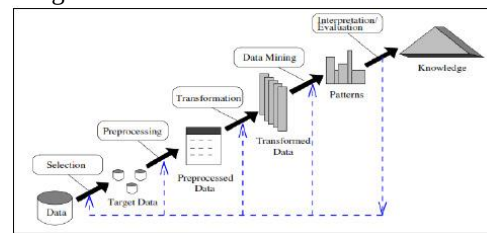
Berdasarkan penelitian yang dilakukan oleh [6] “Prediksi Harga Saham Bank Bri Menggunakan Algoritma Linear Regresion Sebagai Strategi Jual Beli Saham” Hasil penelitian menunjukkan bahwa metode regresi linear efektif digunakan untuk memprediksi harga saham Bank BRI. Dalam penelitian ini, variabel input yang digunakan adalah tanggal (Date) dan harga (Price), dengan pembagian dataset menjadi data latih dan data uji dalam rasio 80:20. Model regresi linear menghasilkan tingkat akurasi sebesar 0,89 untuk data latih dan 0,91 untuk data uji. Evaluasi menggunakan Mean Absolute Percentage Error (MAPE) menunjukkan hasil yang cukup baik. Untuk data latih, MAPE sebesar 13,751%, untuk data uji sebesar 13,773%, dan untuk keseluruhan data sebesar 13,755%. Hasil MAPE yang relatif rendah ini menandakan bahwa model regresi linear mampu memberikan prediksi dengan tingkat kesalahan yang kecil dan dapat dianggap akurat.

Hasil penelitian dalam [7] menunjukkan bahwa metode Meta-Analisis digunakan untuk menggabungkan hasil berbagai kajian ilmiah yang ditujukan untuk pertanyaan yang sama. Penelitian ini juga mencatat bahwa terdapat berbagai metode yang digunakan dalam prediksi harga saham, seperti Linear Regression, ARIMA, dan Neural Network. Selain itu, analisis variable moderator berdasarkan tahun terbit dilakukan menggunakan JASP. Data yang dikumpulkan dalam penelitian sebanyak 300 publikasi kemudian masuk kedalam kriteria inklusi sebanyak 36 data yang memenuhi dengan diketahui jumlah data (N) dan nilai Koefisien Korelasi sebanyak 36 dengan rincian 17 data Metode Linear Regression, 9 data Metode ARIMA, dan 10 data Metode Neural Network. Hasil dari penelitian [8] menunjukkan bahwa model regresi yang dibentuk oleh interaksi variabel bebas dan variabel terikat memiliki koefisien determinasi sebesar 65.6%, yang menunjukkan seberapa bagus model tersebut digunakan untuk memprediksi harga saham.

3. METODE PENELITIAN

Metodologi yang digunakan dalam analisis data dan pengembangan model ini adalah kerangka kerja KDD (Knowledge Discovery in Databases) [9]. Metodologi ini serangkaian langkah atau proses yang digunakan untuk mengekstraksi pengetahuan atau

informasi berharga dari data yang besar, kompleks, dan sering kali tidak terstruktur.



Gambar 1. Alur penelitian kdd

- Pemilihan dan Pemahaman Data (Data Selection)**
Melakukan identifikasi dan pilih data yang relevan dan sesuai dengan tujuan analisis. Menentukan sumber data yang akan digunakan.
- Pre-processing Data (Data Preprocessing)**
Mengidentifikasi dan mengisi nilai-nilai yang hilang dalam dataset, mendeteksi dan menangani nilai-nilai ekstrem yang mungkin mempengaruhi kinerja model dan menormalkan skala data untuk memastikan konsistensi bobot variabel [10].
- Transformasi Data (Data Transformation)**
Melakukan transformasi pada variabel input atau target menjadi label.
- Data Mining**
Melakukan implementasi model regresi linier pada data yang telah diolah sebelumnya, termasuk penyesuaian parameter agar sesuai dengan pola data dan melakukan validasi silang (Cross-Validation) menerapkan metode validasi silang, seperti K-Fold Cross Validation, untuk membagi data menjadi subset pelatihan dan pengujian secara iteratif.
- Interpretasi dan Evaluasi Hasil (Interpretation and Evaluation of Results)**
Melakukan evaluasi Kinerja Model Mengukur kinerja model menggunakan metrik evaluasi Root Mean Squared Error (RMSE), Absolute Error, Root Relatif squared Error dan Squared Error dan melakukan visualisasi hasil seperti scatter plot antara nilai sebenarnya dan prediksi model untuk memberikan gambaran yang jelas dan intuitif tentang performa model [11].

3.1. Sumber Data

Dalam penelitian ini, sumber data utama yang digunakan adalah website resmi Yahoo Finance melalui website <https://www.yahoofinance.com>, peneliti dapat mengakses dan mengumpulkan data historis harga saham perusahaan PT Bank Central Asia (BCA) yang terdaftar di BEI, seperti data harga saham yaitu Date, Price, Open, High, Low, Close, Adj Close dan Volume.

3.2. Teknik Pengumpulan Data

Teknik pengumpulan yang dilakukan dalam penelitian ini adalah lebih bersifat primer, pengumpulan data historis harga saham Bank BCA dilakukan dari sumber resmi pasar modal Indonesia

yaitu yahoofinance dapat diakses dengan link <https://finance.yahoo.com/quote/BBCA.JK/history?p=BBCA.JK>. Data ini merupakan sumber informasi utama yang diperoleh langsung dari pasar modal sebagai hasil dari aktivitas perdagangan saham.

3.3. Teknik Analisis Data

Proses analisis data yang cermat diperlukan untuk mengungkap pola prediksi dan hubungan yang signifikan dalam dataset harga saham Bank BCA yaitu regresi linier metode statistik digunakan untuk menganalisis hubungan antara variabel independen dan dependen, Regresi terdiri dari dua model dasar regresi linear sederhana dan regresi linear berganda. Model regresi linear sederhana digunakan untuk meramalkan hubungan antara dua variabel, sedangkan regresi linear berganda melibatkan dua atau lebih variabel independen. [12]. Dan variabel yang digunakan variabel independen (X) tanggal (Date), harga pembukaan (Open), harga tertinggi (High), harga terendah (Low), penutupan (Close) dan volume perdagangan dengan variabel dependen (Y) harga penutupan (Close) saham Bank BCA.

$$Y = a + bX \quad (1)$$

Y adalah variabel dependen

X adalah variabel independen

a adalah koefisien intercept

b adalah koefisien regresi

untuk mengukur seberapa besar perubahan dalam Y yang diharapkan ketika X mengalami perubahan sebesar satu unit

Mencari nilai a (intercept) dengan menggunakan persamaan rumus :

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

Mencari nilai b (slope) dengan menggunakan persamaan rumus :

$$b = \frac{n(\sum XY) - (\sum X)(\sum Y)}{n(\sum X^2) - (\sum X)^2} \quad (3)$$

Dan melakukan penerapan K-Fold Cross-Validation memungkinkan evaluasi kinerja model regresi linier secara lebih teliti. Dengan membagi dataset menjadi subset non-overlapping, peneliti dapat mengukur sejauh mana model dapat menggeneralisasi ke data baru.

4. HASIL DAN PEMBAHASAN

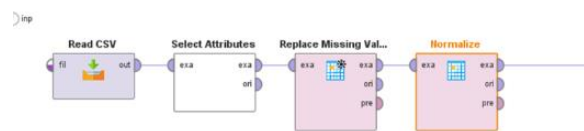
4.1. Data Selection

Data saham historis PT. Bank BCA (BBCA.JK) Tbk, diperoleh dari data public web <https://www.yahoofinance.com> dengan rentang waktu mulai dari 1 januari 2018 sampai dengan 31 agustus 2023, terkumpul jumlah data sebanyak 1410 data sudah sesuai dengan format data dalam penelitian ini.

4.2. Data Preprocessing

Setelah data harga saham diperoleh, langkah pertama dalam analisis adalah melakukan pengecekan

data awal melalui proses preprocessing. Proses ini memiliki tujuan utama untuk melakukan pembersihan terhadap data harga saham dengan mengidentifikasi dan menangani potensi ketidaksempurnaan, seperti data yang hilang atau tidak valid. [13]. Selain itu, tahap preprocessing juga mencakup normalisasi data historis harga saham. Adapun tahapan yang dilakukan dalam proses preprocessing pada penelitian yang dilakukan ini yaitu terlihat pada gambar berikut:



Gambar 2. Tahapan preprocessing

- File Reader merupakan langkah awal dalam membaca dataset yang berformat CSV dengan menggunakan node CSV Reader.
- Pada tahap pemilihan atribut, digunakan dalam proses pra-pemrosesan sebagai langkah penyaringan atau pemilihan atribut yang akan digunakan dalam penelitian. Parameter yang digunakan dalam metode seleksi atribut ini adalah "attribute filter type subset," yang memungkinkan pemilihan beberapa atribut melalui daftar yang ditampilkan.
- Replace missing values, digunakan untuk membersihkan data dari catatan yang tidak lengkap atau atribut yang tidak relevan, sehingga meningkatkan integritas dan kualitas dataset.
- Normalize, proses yang bertujuan untuk menyeimbangkan nilai antara fitur-fitur yang ada dalam dataset. Teknik normalisasi yang digunakan dalam konteks ini adalah Min-Max Normalization.

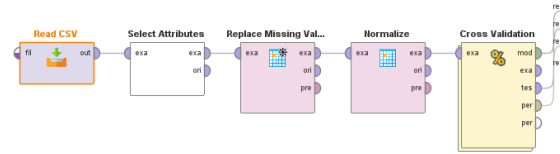
4.3. Transformasi Data

Transformasi data pada saham Bank BCA dapat dilakukan dengan merubah atribut "Close" menjadi variabel dependen atau variabel target. Tujuan dari transformasi ini adalah mengubah data historis harga penutupan (Close) menjadi label yang akan dijadikan fokus prediksi oleh model. Model akan dilatih untuk mengidentifikasi pola dan tren dalam perubahan harga penutupan saham, memungkinkan untuk memberikan prediksi yang lebih akurat terkait pergerakan harga di masa depan [14].

4.4. Data Mining

Proses data mining tahap kunci dalam pengembangan model prediksi untuk harga saham Bank BCA, dalam pengembangan model digunakan untuk memastikan bahwa model dapat belajar dari data yang ada dan dapat diuji pada data yang belum pernah dilihat sebelumnya. atribut "Close" diidentifikasi sebagai variabel dependen yang akan diprediksi. Pemilihan model regresi linier dipertimbangkan

karena kemampuannya dalam menangkap hubungan linier antara atribut open, high, low, close, dan volume dan close.



Teknik yang dilakukan validasi cross validation pada model regresi linear menggunakan sampling type linear sampling. Dalam teknik K fold 10 cross-validation diterapkan pada data untuk meningkatkan reliabilitas model. Ini melibatkan pembagian data model menjadi subset yang saling bergantian, memastikan bahwa model dievaluasi dengan berbagai kombinasi data.

4.5. Interpretasi dan Evaluasi Hasil

Setelah dilakukan proses modeling pada data model menggunakan K fold 10 cross validation regresi linier menunjukkan hasil koefisien dari beberapa attribute untuk mengetahui hasil perhitungan sistem.

Attribute	Coefficient	Std. Error	Std. Coefficient	Tolerance	t-Stat	p-Value
Open	-0.445	0.023	-0.445	0.001	-19.156	0
High	0.649	0.020	0.649	0.003	31.657	0
Low	0.795	0.024	0.795	0.002	33.734	0
Volume	-0.000	0.000	-0.001	0.995	0	1
(Intercept)	0.002	0.002	?	?	0.914	0.361

Gambar 4. Hasil nilai koefisien

Dari hasil regresi linier dengan menggunakan cross-validation pada data model, dapat diambil beberapa rangkum: Open (Koefisien: -0.4447): Koefisien negatif menandakan adanya hubungan negatif antara harga pembukaan (Open) dan harga penutupan saham. Artinya, penurunan harga pembukaan berkorelasi dengan kenaikan harga penutupan, dan sebaliknya. High (Koefisien: 0.6489): Koefisien positif untuk variabel "High" (harga tertinggi) menunjukkan bahwa kenaikan harga tertinggi saham berkorelasi dengan kenaikan harga penutupan. Low (Koefisien: 0.7955): Korelasi positif dari koefisien "Low" (harga terendah) menandakan bahwa kenaikan harga terendah saham berkorelasi dengan kenaikan harga penutupan. Volume (Koefisien: -2.1137E-11): Koefisien yang mendekati nol untuk variabel "Volume" menandakan bahwa volume perdagangan tidak memiliki dampak yang signifikan terhadap perubahan harga penutupan saham dalam model regresi linier. Intercept (Koefisien: 0.0017): Intercept adalah nilai prediksi dasar harga penutupan saham ketika semua variabel independen diatur pada nol.

Dengan mempertimbangkan koefisien hasil regresi linier yang telah dihitung, hasil prediksi yang didapat sebagai berikut:

Row No.	Close	prediction(Close)	Open	High	Low	Date	Volume
1	1.274	1.273	1.256	1.242	1.288	Jan 20, 2023	96859500
2	1.220	1.263	1.274	1.270	1.270	Jan 24, 2023	42474500
3	1.202	1.225	1.274	1.225	1.252	Jan 25, 2023	66124500
4	1.400	1.343	1.220	1.350	1.270	Jan 26, 2023	77582300
5	1.562	1.550	1.562	1.618	1.505	Jan 27, 2023	196804700
6	1.562	1.524	1.580	1.565	1.523	Jan 30, 2023	68544300
7	1.400	1.451	1.508	1.493	1.451	Jan 31, 2023	200599000
8	1.418	1.408	1.472	1.421	1.433	Feb 1, 2023	77396200
9	1.382	1.404	1.364	1.386	1.396	Feb 2, 2023	99109500
10	1.562	1.502	1.454	1.511	1.469	Feb 3, 2023	71716700
11	1.580	1.548	1.599	1.547	1.577	Feb 6, 2023	117757700
12	1.670	1.644	1.617	1.618	1.650	Feb 7, 2023	116014700
13	1.652	1.658	1.653	1.618	1.686	Feb 8, 2023	54495500
14	1.705	1.681	1.653	1.654	1.686	Feb 9, 2023	72381700
15	1.652	1.607	1.689	1.654	1.613	Feb 10, 2023	62060900

Gambar 5. Hasil nilai prediksi

Pada tanggal 20 Januari 2023, data saham Bank BCA menunjukkan beberapa parameter penting. Harga penutupan aktual pada hari tersebut adalah 1.274, sedangkan model regresi linier memprediksi harga penutupan sebesar 1.273. Harga pembukaan saham (Open) pada hari itu tercatat sebesar 1.256, sementara harga tertinggi (High) dan harga terendah (Low) berada dalam kisaran 1.242 hingga 1.288. Volume perdagangan pada hari tersebut mencapai 96.859.500, mencerminkan tingkat aktivitas perdagangan yang signifikan.

Berdasarkan perhitungan koefisien, hasil prediksi menunjukkan bahwa penurunan harga pembukaan saham dapat berhubungan dengan peningkatan harga penutupan. Selain itu, kenaikan harga tertinggi dan harga terendah saham dapat dihubungkan dengan kenaikan harga penutupan. Volume perdagangan, dalam model ini, tidak memberikan dampak signifikan terhadap perubahan harga penutupan saham. Hasil prediksi ini menggambarkan bagaimana setiap variabel independen dapat memengaruhi pergerakan harga penutupan saham Bank BCA.

Terdapat informasi mengenai nilai minimum (min) dan maksimum (max) dari beberapa variabel yang relevan, yaitu "Close," "Prediction Close," "Open," "High," "Low," dan "Volume".

Close	Real	0	Min: -1.734	Max: 2.065	Average: 0.000
Prediction(Close)	Real	0	Min: -1.695	Max: 2.075	Average: 0.001
Open	Real	0	Min: -1.686	Max: 2.085	Average: -0.000
High	Real	0	Min: -1.793	Max: 2.048	Average: 0.000
Low	Real	0	Min: -1.699	Max: 2.102	Average: 0.000
Date	Date	0	Start date: Jan 1, 2018	End date: Aug 30, 2023	Duration: 2067 days
Volume	Real	0	Min: 0	Max: 722827900	Average: 62052162.385

Gambar 6. Nilai statistics

Close (Min: -1.734, Max: 2.065): Variabel "Close" merupakan harga penutupan saham. Nilai minimum dan maksimum menunjukkan rentang variasi harga penutupan pada dataset. Nilai terendah (-1.734) dan tertinggi (2.065) memberikan gambaran tentang sejauh mana variasi harga saham Bank BCA. Prediction Close (Min: -1.695, Max: 2.075): "Prediction Close" adalah nilai yang diprediksi oleh model untuk harga penutupan saham. Rentang nilai ini

mencerminkan seberapa jauh prediksi model dapat bervariasi dari nilai sebenarnya. Nilai minimum dan maksimum (-1.695 hingga 2.075) memberikan gambaran tentang seberapa akurat model dalam memprediksi harga penutupan. Open (Min: -1.686, Max: 2.085): Variabel "Open" adalah harga pembukaan saham. Nilai minimum dan maksimum mencerminkan variasi harga pembukaan pada dataset. Rentang nilai ini memberikan gambaran tentang variasi awal harga saham Bank BCA. High (Min: -1.703, Max: 2.048): "High" mencerminkan harga tertinggi saham pada suatu periode. Nilai minimum dan maksimum menunjukkan rentang variasi harga tertinggi. Low (Min: -1.699, Max: 2.102): Variabel "Low" adalah harga terendah saham pada suatu periode. Nilai minimum dan maksimum menunjukkan variasi harga terendah dalam dataset. Volume (Min: 0, Max: 722827900): "Volume" mencerminkan jumlah perdagangan saham. Nilai minimum 0 menunjukkan bahwa ada hari di mana tidak ada perdagangan saham. Nilai maksimum 722827900 menunjukkan jumlah perdagangan tertinggi pada periode tertentu. Rentang yang lebih kecil antara nilai minimum dan maksimum menunjukkan bahwa model memiliki tingkat konsistensi yang baik dalam memprediksi harga penutupan (close). Jika nilai prediksi berada dalam rentang yang lebih terbatas, hal ini dapat dianggap sebagai indikasi bahwa model cenderung memberikan prediksi yang stabil. Sebaliknya, jika rentang nilai prediksi lebih luas, model mungkin menghadapi kesulitan dalam memberikan prediksi yang konsisten.



Gambar 7. Hasil visualisasi

Membuat visualisasi seperti scatter plot antara nilai sebenarnya dan prediksi model. Visualisasi ini membantu memberikan gambaran yang jelas tentang seberapa baik model dapat merepresentasikan pergerakan harga saham.

Hasil visualisasi pergerakan dari data variabel close dari tahun 2018 sampai dengan tahun 2023 yang ditandai dengan garis biru dan pergerakan prediksi close ditandai dengan garis hijau pada grafik.

Evaluasi kinerja model mengukur kinerja model menggunakan metrik evaluasi Root Mean Squared Error (RMSE), Absolute Error, Root Relative Squared Error dan Squared Error.

Hasil evaluasi performa model regresi linier pada data model, yang direpresentasikan dalam PerformanceVector, dapat dijelaskan sebagai berikut: RMSE mengukur rata-rata besarnya kesalahan prediksi model. Nilai yang lebih rendah menunjukkan

PerformanceVector

```
PerformanceVector:
root_mean_squared_error: 0.032 +/- 0.010 (micro average: 0.033 +/- 0.000)
absolute_error: 0.024 +/- 0.007 (micro average: 0.024 +/- 0.023)
root_relative_squared_error: 0.138 +/- 0.036 (micro average: 0.128)
squared_error: 0.001 +/- 0.001 (micro average: 0.001 +/- 0.004)
```

Gambar 8. Hasil performancevektor

tingkat akurasi yang lebih baik. Root Mean Squared Error (RMSE) menjadi indikator utama seberapa dekat prediksi model dengan nilai sebenarnya. Nilai RMSE sebesar 0.032 dengan rentang 0.010 mengindikasikan bahwa model mampu memberikan perkiraan harga penutupan dengan tingkat kesalahan yang rendah. Rentang 0.010 menunjukkan konsistensi kinerja model dalam mengukur kesalahan. Absolute Error sebesar 0.024 dengan rentang 0.007 memberikan gambaran tentang seberapa dekat nilai prediksi model dengan nilai sebenarnya. Nilai yang relatif rendah menandakan bahwa model memiliki kemampuan untuk memberikan perkiraan yang akurat terhadap pergerakan harga saham. Root Relative Squared Error sebesar 0.138 dengan rentang 0.036 mengukur tingkat kesalahan relatif terhadap variasi data. Nilai yang relatif rendah menunjukkan kemampuan model untuk mempertahankan tingkat ketepatan prediksi dalam berbagai kondisi. Squared Error sebesar 0.001 dengan rentang 0.001 memberikan gambaran tentang sebaran kesalahan prediksi model. Nilai yang rendah menunjukkan bahwa sebagian besar prediksi model berada dalam kisaran yang dekat dengan nilai sebenarnya.

5. KESIMPULAN DAN SARAN

Dalam kesimpulan, hasil evaluasi kinerja model regresi linier pada prediksi harga penutupan saham Bank BCA menunjukkan kinerja yang baik. RMSE yang rendah (0.032) menunjukkan bahwa model cenderung memberikan prediksi harga penutupan dengan tingkat kesalahan yang minim. Absolute Error yang relatif kecil (0.024) menandakan bahwa prediksi model secara konsisten mendekati nilai sebenarnya. Root Relative Squared Error yang rendah (0.138) menggambarkan kemampuan model untuk menjaga tingkat kesalahan relatif terhadap variasi data. Keberhasilan model dalam menggambarkan variasi harga saham dapat menjadi sumber informasi yang berharga bagi para pemangku kepentingan di pasar keuangan. Tingkat akurasi yang tinggi ini dapat memberikan kepercayaan dalam pengambilan keputusan investasi, membantu para pelaku pasar

untuk membuat keputusan yang lebih informasional dan dapat diandalkan.

DAFTAR PUSTAKA

- [1] V. Yuningsih, "Pengaruh net profit margin (NPM) dan earning per share (EPS) terhadap return saham dengan struktur modal sebagai variabel intervening," *Entrep. bisnis Manaj.* ..., 2020, [Online]. Available: <http://ejournal.widyamataram.ac.id/index.php/j-mae/article/view/215>
- [2] I. Himawan, O. Nurdian, G. Dwilestari, and ..., "Irvan Himawan PREDIKSI HARGA SAHAM DENGAN ALGORITMA REGRESI LINIER DENGAN RAPIDMINER," *JURSIMA (Jurnal ...)*, vol. 10, no. 3, 2022, [Online]. Available: <https://ejournal.indobarunasional.ac.id/index.php/jursima/article/view/475>
- [3] H. Winnos, R. Septima, and ..., "Perbandingan Metode Regresi Linier Berganda dan Autoregressive Integrated Moving Average (ARIMA) Untuk Prediksi Saham PT. BSI, Tbk.," ... *J. Ilmu Tek. dan ...*, 2022, [Online]. Available: <https://journal.unimar-amni.ac.id/index.php/ocean/article/view/350>
- [4] E. Fitri and D. Riana, "Analisa Perbandingan Model Prediction Dalam Prediksi Harga Saham Menggunakan Metode Linear Regression, Random Forest Regression Dan Multilayer ...," *METHOMIKA J. Manaj.* ..., 2022, [Online]. Available: <https://ejournal.methodist.ac.id/index.php/methomika/article/view/926>
- [5] S. A. L. Satriyo and A. R. Pratama, "Perbandingan metode linear regresi dan polynomial regresi untuk memprediksi harga saham studi kasus Bank BCA," ... *J. Inform.* ..., 2023, [Online]. Available: <http://jurnal.sttmcileungsi.ac.id/index.php/infotech/article/view/602>
- [6] J. S. Putra, R. D. Ramadhani, and ..., "Prediksi Harga Saham Bank Bri Menggunakan Algoritma Linear Regresion Sebagai Strategi Jual Beli Saham," *J. Dinda Data* ..., 2022, [Online]. Available: <https://journal.ittelkom-pwt.ac.id/index.php/dinda/article/view/273>
- [7] E. S. Ashri, S. Syaharuddin, and ..., "Metode Linear Regression, ARIMA, dan Neural Network Dalam Prediksi Harga Saham: Sebuah Meta-Analisis," *Indones. J. ...*, 2022, [Online]. Available: <https://unu-ntb.ejournal.id/ije/article/view/152>
- [8] M. Keren and R. S. Oetama, "Prediksi prospek harga saham perusahaan perbankan Menggunakan Regresi Linear (Study Kasus Bank BCA Tahun 2015-2017)," *JSI J. Sist. Inf.*, 2019, [Online]. Available: <https://ejournal.unsri.ac.id/index.php/jsi/article/view/8105>
- [9] O. Nurdian, A. Irma Purnamasari, and I. Ali, "Analisa Penjualan Mobil Dengan Menggunakan Algoritma K-Means Di PT. Mulya Putra Kencana," *J. Data Sci. dan Inform.*, vol. 1, no. 2, pp. 32–35, 2021.
- [10] Syafi'i, O. Nurdian, and G. Dwilestari, "Penerapan Machine Learning Untuk Menentukan Kelayakan Kredit Menggunakan Metode Support Vektor Machine," *J. Sist. Inf. dan Manaj.*, vol. 10, no. 2, pp. 1–6, 2022.
- [11] Melawati, A. Nur Hakim, A. Irma Purnamasari, O. Nurdian, and S. Anwar, "Information Management For Educators And Professionals Berbasis Game 2D Menggunakan Metode Systems Development Life Cycle," *Inf. Manag. Educ. Prof.*, vol. 6, no. 1, pp. 91–100, 2021, [Online]. Available: <http://ejournal-binainsani.ac.id/index.php/IMBI/article/view/1681/1368>
- [12] A. S. Nuriyah, R. Saladini, and ..., "Prediksi Harga Saham Bank BCA Menggunakan Metode Regresi Linier Berganda," *Prosiding ... prosidingsnsa.statistics.unpad.ac.id*, 2023, [Online]. Available: <https://prosidingsnsa.statistics.unpad.ac.id/?journal=prosidingsnsa&page=article&op=view&path%5B%5D=323>
- [13] M. Walid and F. Halimiyah, "Klasifikasi Kemandirian Siswa SMA/MA Double Track Menggunakan Metode Naive Bayes," *J. ICT Inf. Commun. Technol.*, vol. 22, pp. 190–197, 2022.
- [14] O. Nurdian, A. Faqih, and ..., "Pemanfaatan Sistem Informasi Layanan Berbasis Digital Desa Sinarancang," *AMMA J.* ..., vol. 2, no. 8, pp. 922–928, 2023, [Online]. Available: <https://journal.mediapublikasi.id/index.php/amma/article/view/3561%0Ahttps://journal.mediapublikasi.id/index.php/amma/article/download/3561/1875>