

Pengembangan Model Prediksi Harga Saham Dengan Menggunakan Regresi Linear Berganda Pada Saham BRI

Yuan Anisa¹, Muhammad Hafiz², Nanda Novita³

^{1,3} Universitas Medan Area, Medan, Indonesia,

² Universitas Pembangunan Pancabudi, Medan, Indonesia

ABSTRAK

Penelitian ini bertujuan untuk mengembangkan model prediksi harga saham menggunakan regresi linear berganda, dengan studi kasus pada saham PT Bank Rakyat Indonesia (Persero) Tbk. Data yang digunakan berasal dari Yahoo Finance dan PT Stockbit Sekuritas Digital, yang meliputi harga saham historis. Model regresi linear berganda digunakan untuk memprediksi harga saham di masa depan, dengan menggunakan variabel independen mencakup harga pembukaan, harga tertinggi, harga terendah, dan net buy/sell asing, sedangkan untuk harga penutupan berperan sebagai variabel dependen. Hasil analisis menghasilkan persamaan regresi: $Y = 1.725 - 0.529x_1 + 0.689x_2 + 0.840x_3 + 3.221E-8x_4$, yang menunjukkan hubungan signifikan antara variabel independen terhadap harga penutupan saham. Penelitian ini memberikan wawasan bagi investor dan pelaku pasar untuk memahami faktor-faktor utama yang memengaruhi pergerakan harga saham, sehingga dapat digunakan sebagai alat bantu dalam pengambilan keputusan investasi. Model yang dihasilkan memiliki potensi untuk dioptimalkan lebih lanjut dengan melibatkan variabel makroekonomi dan teknik machine learning untuk meningkatkan akurasi prediksi.

ABSTRACT

This study aims to develop a stock price prediction model using multiple linear regression, with a case study on the shares of PT Bank Rakyat Indonesia (Persero) Tbk. The data used was sourced from Yahoo Finance and PT Stockbit Sekuritas Digital, consisting of historical stock prices. Multiple linear regression is applied to predict future stock prices, where independent variables include opening price, highest price, lowest price, and net foreign buy/sell, while the closing price serves as the dependent variable. The analysis resulted in the regression equation: $Y = 1.725 - 0.529x_1 + 0.689x_2 + 0.840x_3 + 3.221E-8x_4$, indicating a significant relationship between independent variables and the closing price. This study provides insights for investors and market participants to understand key factors influencing stock price movements, serving as a decision-support tool for investments. The resulting model has the potential to be further optimized by incorporating macroeconomic variables and machine learning techniques to enhance prediction accuracy.

Kata Kunci: Saham, Pemodelan, Regresi Linear, Prediksi

Email Address: ¹ yuan@staff.uma.ac.id

DOI: <http://dx.doi.org/10.30829/jistech.v9i2.22213>

Received 15 October 2024; Received in revised form 2 December 2024; Accepted 18 December 2024



This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

Pendahuluan

Regresi linier adalah metode statistik yang digunakan untuk memodelkan hubungan antara variabel dependen dan satu atau lebih variabel independen dengan menyesuaikan persamaan linier dengan titik data yang diamati. Metode ini biasanya digunakan untuk tujuan prediksi dan peramalan, karena memungkinkan kita untuk memahami hubungan antara variabel dan membuat prediksi berdasarkan hubungan tersebut. Pada paragraf ini, kita akan membahas asumsi-asumsi regresi linier dan bagaimana menginterpretasikan hasil analisis regresi linier. Salah satu asumsi utama dari regresi linier adalah adanya hubungan linier antara variabel independen dan dependen. Ini berarti bahwa ketika variabel independen berubah, variabel dependen berubah dengan cara yang konstan dan dapat diprediksi. Selain itu, regresi linier mengasumsikan bahwa residual (perbedaan antara titik data yang diamati dan nilai yang diprediksi oleh model) terdistribusi secara normal. Ketika menginterpretasikan hasil analisis regresi linier, penting untuk melihat koefisien determinasi (R-squared) untuk menentukan seberapa baik model sesuai dengan data dan signifikansi koefisien untuk memahami dampak dari setiap variabel independen

terhadap variabel dependen.

Selain itu, sangat penting untuk menilai asumsi-asumsi regresi linier, seperti linieritas, independensi kesalahan, homoskedastisitas, dan normalitas residual, untuk memastikan validitas model. Pelanggaran terhadap asumsi-asumsi ini dapat menyebabkan estimasi yang bias dan kesimpulan yang salah. Selain itu, penting untuk mempertimbangkan potensi multikolinearitas di antara variabel-variabel independen, karena hal ini dapat mempengaruhi stabilitas dan interpretabilitas koefisien regresi. Secara keseluruhan, pemahaman yang menyeluruh mengenai asumsi dan keterbatasan regresi linier sangat penting untuk menginterpretasikan dan menerapkan hasil analisis secara akurat. Lebih lanjut, sangat penting untuk menilai keberadaan outlier dalam data, karena dapat sangat memengaruhi hasil analisis regresi. Penculan dapat mengubah estimasi koefisien dan mempengaruhi kecocokan model secara keseluruhan. Selain itu, peneliti juga harus menyadari titik data yang berpengaruh yang memiliki dampak kuat pada hasil regresi. Dengan mengatasi masalah ini dan memastikan bahwa asumsi regresi linier terpenuhi, peneliti dapat memiliki kepercayaan diri dalam keandalan dan validitas temuan mereka.

Memprediksi harga saham adalah tugas yang kompleks dan menantang yang memerlukan pertimbangan cermat atas berbagai faktor. Untuk membuat prediksi yang akurat, peneliti harus menganalisis data historis, tren pasar, dan indikator ekonomi dengan cermat. Selain itu, penting juga untuk mempertimbangkan faktor eksternal seperti peristiwa politik, bencana alam, dan kemajuan teknologi yang dapat memengaruhi harga saham. Dengan menggunakan teknik statistik canggih dan algoritme pembelajaran mesin, para peneliti dapat mengembangkan model yang dapat membantu memprediksi harga saham di masa depan dengan tingkat akurasi tinggi. Prediksi ini dapat bermanfaat bagi investor, analis keuangan, dan pembuat kebijakan yang mengandalkan prakiraan akurat untuk membuat keputusan yang tepat. Selain itu, para peneliti juga harus mempertimbangkan aspek psikologis perilaku pasar, seperti sentimen investor dan sentimen pasar, yang dapat sangat memengaruhi harga saham. Dengan memahami faktor-faktor ini dan memasukkannya ke dalam model mereka, para peneliti dapat membuat prediksi yang lebih akurat tentang pergerakan saham di masa depan. Pada akhirnya, kemampuan untuk memprediksi harga saham secara akurat dapat memberikan keunggulan kompetitif yang signifikan di pasar keuangan dan mengarah pada strategi investasi dan pengambilan keputusan yang lebih baik.

Selain memahami perilaku pasar dan memasukkan faktor psikologis, peneliti juga harus selalu mengikuti tren ekonomi global, peristiwa politik, dan berita spesifik industri yang dapat memengaruhi harga saham. Dengan menganalisis beragam data dan informasi, peneliti dapat memperoleh pemahaman yang lebih komprehensif mengenai faktor-faktor yang mendorong pergerakan pasar saham. Pendekatan holistik ini memungkinkan prediksi yang lebih akurat dan keputusan investasi yang lebih baik. Selain itu, terus mengikuti perkembangan terbaru di pasar dapat membantu peneliti mengidentifikasi peluang dan risiko potensial sebelum diketahui secara luas, sehingga memberikan keunggulan kompetitif di pasar keuangan. Dengan terus memantau dan menginterpretasikan berbagai sumber informasi, para peneliti dapat tetap menjadi yang terdepan dan menyesuaikan strategi investasi mereka. Pendekatan proaktif ini tidak hanya membantu investor memaksimalkan imbal hasil, tetapi juga meminimalkan potensi kerugian yang tidak terduga. Dalam lingkungan pasar yang serba cepat dan bergejolak saat ini, memiliki pengetahuan dan proaktif sangat penting untuk sukses di pasar saham.

Penelitian ini bertujuan untuk mengeksplorasi dampak pembelian/penjualan bersih asing (Net Buy/Sell) pada saham BRI. Dengan mengkaji perilaku investor yang secara aktif mencari informasi baru dan menyesuaikan strategi mereka, kami berharap dapat menjelaskan keefektifan pendekatan ini dalam mencapai tujuan finansial. Melalui analisis mendalam terhadap data pasar dan perilaku investor, kami berupaya memberikan wawasan berharga bagi investor yang ingin meningkatkan kinerja mereka dan menavigasi kompleksitas pasar saham.

Untuk mencapai tujuan ini, kami akan melakukan tinjauan komprehensif terhadap literatur yang ada mengenai strategi investasi proaktif dan dampaknya terhadap kinerja pasar saham. Selain itu, kami akan menganalisis data historis dan studi kasus untuk mengidentifikasi faktor-faktor kunci yang berkontribusi terhadap keberhasilan investor proaktif. Melalui penelitian ini, kami berharap dapat memberikan wawasan berharga yang dapat membantu investor dalam mengambil keputusan yang tepat dan memaksimalkan imbal hasil di pasar saham. Dengan memahami strategi yang digunakan oleh investor proaktif yang sukses dan faktor-faktor yang menyebabkan kesuksesan mereka, investor akan lebih siap untuk menavigasi kompleksitas pasar saham. Penelitian kami bertujuan untuk menjelaskan pentingnya untuk selalu mendapatkan informasi, tetap berada di depan tren pasar, dan mengambil risiko yang diperhitungkan untuk mencapai tujuan keuangan. Pada akhirnya, temuan kami akan menjadi sumber daya yang berharga bagi para investor yang ingin mengambil pendekatan proaktif terhadap strategi investasi mereka dan meningkatkan kinerja mereka secara keseluruhan di pasar saham.

Metodologi Penelitian

a. Proses pengumpulan data

Sumber data utama yang dipakai dalam penelitian ini adalah website resmi Yahoo Finance melalui [website https://www.yahoofinance.com](https://www.yahoofinance.com) dan PT Stockbit Sekuritas Digital sebagai sekuritas yang juga menyajikan data secara lengkap. Peneliti dapat mengumpulkan data dan mengkases data historis harga saham perusahaan PT Bank Rakyat Indonesia (BRI) yang terdaftar di bursa efek Indonesia (BEI). Data yang dikumpulkan dapat berupa Date, Price, Open, High, Low, Close, Adj Close, Volume dan net buy atau sell asing pada suatu saham dalam satu tahun.

b. Variabel yang digunakan dalam model regresi

Variabel yang digunakan dalam penelitian ini adalah :

1. Variabel independen (x) adalah harga pembukaan ($Open/x_1$), harga tertinggi ($High/x_2$), harga terendah (Low/x_3), dan volume pembelian asing ($Net\ buy/sell/x_4$) dalam satu hari perdagangan yang dikumpulkan dalam beberapa bulan ke belakang.
2. Variabel dependen (Y) adalah harga penutupan ($Close$) saham Bank BRI dalam satu hari perdagangan yang dikumpulkan dalam beberapa bulan ke belakang.

c. Teknik Analisis Data

Untuk menemukan pola prediksi dan hubungan yang signifikan dalam data harga saham Bank Rakyat Indonesia (BRI), proses analisis data yang cermat sangat diperlukan. Penelitian ini menggunakan regresi linier berganda, yaitu yang melibatkan beberapa variabel bebas. Model regresi linear berganda digunakan untuk meramalkan hubungan antara dua variabel, sedangkan model berganda melibatkan dua atau lebih variabel independen. Persamaan regresi linear berganda adalah:

$$Y = a + b_1x_1 + b_2x_2 + \dots + b_nx_n$$

Berdasarkan persamaan diatas, variabel dependen atau terikat dinotasikan sebagai huruf Y dan variabel bebas atau independen (x). Notasi a adalah suatu konstanta dan b merupakan suatu koefisien regresi variabel x . Nilai a dan b terhadap variabel x dapat diperoleh dengan persamaan sebagai berikut ini:

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

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

Hasil dan Pembahasan

Berikut ini adalah pembahasan hasil peneitian mengenai perhitungan prediksi harga saham bank BRI periode 1 Januari 2024 samai 18 agustus 2024 sebagai berikut:

a. **Data Selection**

Data saham dari PT. Bank Rakyat Indonesia Tbk (BBRI), diperoleh dari data [website https://www.yahoofinance.com](https://www.yahoofinance.com) dan PT Stockbit Sekuritas Digital dengan interval waktu mulai dari 1 januari 2024 sampai dengan 18 agustus 2024. Data yang terkumpul sebanyak 145 data sudah sesuai dengan format yang ada dalam penelitian ini.

b. **Data Preprocessing**

Setelah mendapatkan data harga saham dari *website* yang dijadikan sumber data maka, langkah pertama yang dilakukan adalah mengecek data awal melalui proses *preprocessing*. Tujuan utama dari hal ini adalah melakukan pembersihan data harga saham dengan mengidentifikasi dan evaluasi data, seperti data yang hilang atau tidak valid. Selain itu, tahap *preprocessing* juga mencakup normalisasi data historis harga saham. Adapun tahapan yang dilakukan dalam proses *preprocessing* pada penelitian yang dilakukan ini yaitu:

1. Node CSV Reader dapat digunakan untuk membaca dataset yang berformat CSV dengan menggunakan *file reader*.
2. Pemilihan atribut digunakan dalam proses pra-pemrosesan sebagai langkah penyaringan yang akan digunakan dalam penelitian. Adapun parameter yang dipakai dalam metode seleksi atribut adalah "attribute filter type subset," yang mampu melakukan pemilihan beberapa atribut melalui daftar yang ditampilkan.
3. Replace missing values, dipergunakan untuk membersihkan data dari hal yang tidak lengkap atau atribut yang tidak relevan, sehingga meningkatkan kualitas dataset.

c. **Transformasi Data**

Transformasi Data dilakukan dengan mengubah atribut "Close" menjadi variabel dependen atau variabel target, pada data saham Bank Rakyat Indonesia (BRI). Transformasi ini bertujuan untuk mengubah data historis harga penutupan (*Close*) menjadi label, yang akan menjadi fokus prediksi model. Model akan dilatih untuk menemukan pola dan tren dalam perubahan harga penutupan saham. Ini akan memungkinkan untuk membuat prediksi yang lebih akurat tentang pergerakan harga di masa depan.

d. **Data Mining**

Pada tahap data mining yang merupakan kunci dalam pengembangan model prediksi untuk harga saham Bank BRI, dalam mengembangkan model perlu kehati-hatian karena untuk memastikan bahwa model dapat belajar dari data yang ada dan dapat diuji pada data yang belum pernah dilihat sebelumnya. Data atribut "Close" diidentifikasi sebagai variabel dependen yang akan diprediksi. Pemilihan model regresi linier dipakai dikarenakan kemampuannya dalam menangkap dan menjelaskan hubungan linier antara harga pembukaan (*Open*), harga tertinggi (*High*), harga terendah (*Low*), volume pembelian asing (*Net buy/sell*), dan harga penutupan (*close*).

e. Interpretasi dan Evaluasi Hasil

Berikut Ini adalah hasil analisa data yang telah dilakukan:

Tabel 1. Model Summary
Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.999 ^a	.997	.997	33.091

a. Predictors: (Constant), Nett buy/sell, Open, High, Low

Dari *output* di atas nilai R Square sebesar 0,997, nilai ini mengandung artinya bahwa pengaruh x_1 sampai x_4 terhadap y adalah sebesar 99,7% sedangkan 0,3 % dipengaruhi oleh variabel yang tidak diteliti. Dengan demikian dapat disimpulkan pengaruh variabel independen sangat besar terhadap pengaruh variabel dependen.

Tabel 2. Anova
ANOVA^a

Model		Sum of Square	df	Mean Square	F	Sig.
1	Regression	63352244.60	4	15838061.15	14463.372	.000 ^b
	Residual	163161.889	149	1095.046		
	Total	63515406.49	153			

a. Dependent Variabel: close

b. Predictors: (Constant), Net buy/sell, Open, High, Low

Nilai prob. *F* hitung (*sig.*) pada tabel di atas nilainya 0,000 lebih kecil dari tingkat signifikansi 0,05 sehingga dapat disimpulkan bahwa model regresi linier yang diestimasi layak digunakan untuk menjelaskan pengaruh variabel independen terhadap dependen.

Tabel 3. Coefficients^a
Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig
		B	Std. Error	beta		
1	(Constant)	1.725	24.097		.072	.943
	Open	-.529	.069	-.521	-7.664	.000
	High	.689	.065	.681	10.550	.000
	Low	.840	.073	.831	11.494	.000
	Net buy/sell	3.221E-8	.000	.019	3.649	.000

a. Dependent Variabel: close

Berdasarkan tabel di atas maka model regresi yang didapatkan adalah:

$$Y = 1.725 - 0.529x_1 + 0.689x_2 + 0.840x_3 + 3.221E - 8x_4$$

Di mana:

x_1 = Harga pembukaan (*Open*)

x_2 = Harga Tertinggi (*High*)

x_3 = Harga Terendah (*Low*)

x_4 = Net buy/sel asing (*Net*)

Setelah data dianalisa dan diperoleh suatu model prediksi maka dilakukan uji coba model dengan menggunakan data yang baru dari tanggal 2 September 2024 sampai tanggal 1 Oktober 2024. Berikut hasil dari uji coba model:

Tabel 4. Hasil uji coba

Open	High	Low	Net	Prediksi	Close
5150	5225	5150	247365557.0	5211.37	5175.00
5225	5250	5150	308351775.0	5190.88	5175.00
5100	5175	5100	-333138695	5142.67	5100.00
5125	5150	5075	39259210.00	5103.21	5150.00
5200	5250	5150	435867420.0	5208.21	5225.00
5275	5300	5150	-99862313.0	5185.73	5175.00
5200	5250	5125	49821223.00	5174.78	5250.00
5225	5250	5125	-451042323	5145.42	5175.00
5175	5225	5150	190174897.0	5196.30	5175.00

5200	5300	5175	804930000.0	5275.55	5300.00
5350	5350	5250	-131281263	5263.50	5300.00
5350	5375	5325	687876667.00	5350.17	5350.00
5425	5500	5400	559704420	5475.43	5400.00
5400	5450	5375	-40018565.0	5413.89	5375.00
5350	5500	5350	669690852.0	5476.65	5500.00
5550	5575	5525	296938548.0	5557.51	5525.00
5400	5425	5200	-1643183470	5198.02	5325.00
5150	5200	5025	-2782272875	4991.65	5050.00
5050	5150	5000	-441152908	5064.42	5100.00
5050	5050	4950	-1219756544	4928.44	4950.00
4970	5075	4970	-42802239.0	5042.69	5075.00

Kesimpulan

Hasil pengujian menggunakan metode regresi linear menunjukkan adanya pengaruh yang signifikan antara variabel bebas (harga pembukaan, harga tertinggi, harga terendah, dan net buy/sell asing) terhadap variabel terikat, yaitu harga penutupan saham. Model regresi yang dihasilkan memiliki potensi sebagai alat prediksi, meskipun masih dapat ditingkatkan. Untuk penelitian selanjutnya, disarankan untuk menambahkan variabel lain seperti indikator makroekonomi (inflasi, suku bunga) atau sentimen pasar yang dapat memberikan informasi tambahan untuk meningkatkan akurasi prediksi. Penggunaan metode lain, seperti machine learning, juga dapat dipertimbangkan untuk membandingkan performa model.

Daftar Pustaka

- [1] B. Boris, "Determinants of Alliance Portfolio Complexity and Its Effect on Innovative Performance of Companies," *Journal of Product Innovation Management*, 2011. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1540-5885.2011.00824.x>.
- [2] C. Chu, "Investor mood and financial markets," *Journal of Economic Behavior & Organization*, 2010. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0167268110001058>.
- [3] D. Chris, "Linear regression and the normality assumption," *Journal of Statistical Planning and Inference*, 2018. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0895435617304857>.
- [4] M. Del and N. N., "Simple linear and multivariate regression models," *Journal of Statistical Planning and Inference*, 2011. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0301054611001054>.
- [5] E. Erik, M. Marko, and E. Erik, "Regression analysis," in *Springer Handbook of Computational Statistics*, Berlin, Germany: Springer, 2019, pp. 345–374. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-662-56707-4_7.
- [6] H. Huina and X. Xiaojun, "Twitter mood predicts the stock market," *Journal of Computational Science*, 2011. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S187775031100007X>.
- [7] K. Kai, "A blockchain research framework: What we (don't) know, where we go from here, and how we will get there," *Business & Information Systems Engineering*, vol. 59, pp. 379–409, 2017. [Online]. Available: <https://link.springer.com/article/10.1007/s12599-017-0506-0>.
- [8] S. Samir and A. Alexander, "Short and long-term forecasting using artificial neural networks for stock prices in Palestine: a comparative study," *Electronic Journal of Applied Statistical Analysis*, vol. 10, no. 1, pp. 58–76, 2017. [Online]. Available: <http://siba-ese.unisalento.it/index.php/ejasa/article/view/15803>.
- [9] T. T., F. Felix, and S. Sai, "Scientometric review and analysis of recent approaches to stock market forecasting: Two decades survey," *Expert Systems with Applications*, vol. 209, p. 117834, 2023. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S095741742202317X>.
- [10] A. Timmermann, "Forecast combinations," *Handbook of Economic Forecasting*, vol. 1, pp. 135–196, 2006. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1574070605010049>.
- [11] M. Walid and F. Halimiyah, "Klasifikasi Kemandirian Siswa SMA/MA Double Track Menggunakan Metode Naive Bayes," *Jurnal ICT: Information and Communication Technology*, vol. 22, pp. 190–197, 2022.