

## **PREDIKSI HARGA SAHAM PT ANEKA TAMBANG MENGGUNAKAN LINEAR REGRESI DENGAN PENDEKATAN PYSPARK**

**Nadya Arassy Primawan**

Teknik Informatika, Fakultas Teknik, Universitas Pelita Bangsa

Email: [narassyp@gmail.com](mailto:narassyp@gmail.com)

### **Abstrak**

Pasar saham adalah investasi yang berisiko, tetapi memiliki potensi pengembalian yang tinggi. Namun, harga saham bersifat fluktuatif dan sulit diprediksi. Studi ini menggunakan regresi linier untuk memprediksi harga saham PT Aneka Tambang (ANTM). Studi menemukan bahwa model regresi linier dapat memprediksi harga saham ANTM dengan akurasi yang tinggi. Model tersebut memiliki akurasi 100%, kesalahan rata-rata kuadrat (MSE) sebesar  $7.302095386799366e-23$ , dan nilai R-squared sebesar 1.0. Hasil ini menunjukkan bahwa regresi linier dapat menjadi alat yang berguna bagi investor pasar saham. Dengan menggunakan model ini, investor dapat membuat prediksi yang lebih akurat tentang harga saham di masa depan, yang dapat membantu mereka untuk mengurangi risiko.

Kata kunci: Pasar Saham, Regresi Linear, Prediksi

### **Abstract**

*The stock market is a risky investment, but it has the potential for high returns. However, stock prices are volatile and can be difficult to predict. This study used linear regression to predict the price of PT Aneka Tambang (ANTM) stock. The study found that the linear regression model was able to predict ANTM stock prices with high accuracy. The model had an accuracy of 100%, a mean squared error (MSE) of  $7.302095386799366e-23$ , and an R-squared of 1.0. These results suggest that linear regression can be a useful tool for stock market investors. By using this model, investors can make more accurate predictions about future stock prices, which can help them to reduce risk.*

**Keywords:** Stock Market, Linear Regression, Prediction

### **Pendahuluan**

Saham mewakili kepemilikan seseorang atau badan hukum di suatu perusahaan dan dicatat dalam bentuk sertifikat [1] Sertifikat tersebut mencantumkan jumlah saham, golongan saham, hak, dan kewajiban masing-masing investor. Investasi saham menarik karena memiliki potensi keuntungan yang tinggi. Namun, risiko inherent dalam investasi saham timbul dari fluktuasi harga saham yang cepat dan sulit diprediksi. Sebagai hasilnya, berinvestasi dalam saham umumnya dianggap sebagai upaya yang

---

<b>How to cite:</b>	Nama Author (tahun terbit) Judul, (Volume) Issue, <a href="http://dx.doi.org/10.36418/syntax-literate.v6i6">http://dx.doi.org/10.36418/syntax-literate.v6i6</a>
---------------------	---

---

<b>E-ISSN:</b>	2548-1398
----------------	-----------

---

<b>Published by:</b>	Ridwan Institute
----------------------	------------------

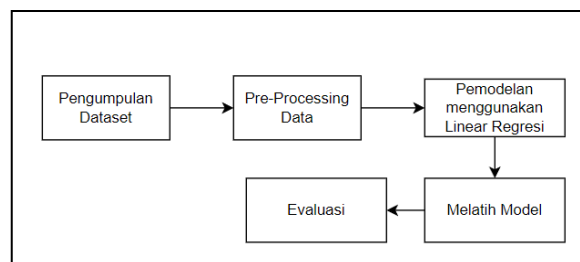
---

memiliki risiko tinggi. Meskipun investasi saham membawa risiko yang signifikan, para investor masih antusias untuk mengalokasikan dana mereka pada kelas aset ini.

Dalam konteks peramalan, prediksi dapat dilakukan dengan mengamati kondisi masa lalu untuk meramalkan kejadian di masa depan [2]. Peramalan berfungsi sebagai alat atau pertimbangan dalam proses pengambilan keputusan, khususnya di sektor bisnis atau ekonomi, dengan tujuan meminimalkan kerugian sebisa mungkin dan memaksimalkan keuntungan.

### Metode Penelitian

Metode penelitian ini bertujuan untuk melakukan prediksi harga saham PT Aneka Tambang dengan menggunakan analisis data historis. Data harga saham diambil dari platform *Yahoo Finance* [3].



Gambar 1 Tahapan Penelitian

### Lokasi Penelitian

Penelitian ini dilakukan secara on-site. Dengan mengakses data dari *Yahoo Finance* setelah diperoleh sesuai dengan kebijakan dan regulasi yang berlaku.

### Ruang Lingkup Penelitian

Penelitian ini berfokus pada prediksi harga saham PT Aneka Tambang menggunakan metode regresi linear. Data yang digunakan diambil dari sumber yang dapat diandalkan, yaitu *Yahoo Finance*. Ruang lingkup penelitian mencakup variabel-variabel yang relevan yang dapat mempengaruhi harga saham. Batasan-batasan tertentu diterapkan untuk memastikan fokus pada aspek-aspek kunci dalam analisis harga saham.

### Teknik Analisis Penelitian

Teknik analisis penelitian melibatkan penggunaan regresi linear untuk mengidentifikasi hubungan linier antara variabel-variabel tertentu dan harga saham. Pengolahan data dan pengujian model dilakukan menggunakan platform PySpark yang berbasis cloud atau komputasi terdistribusi. Regresi linear digunakan untuk memodelkan tren dan pola dalam data historis harga saham, memungkinkan pembuatan prediksi untuk masa depan.

## Hasil dan Pembahasan

Pembahasan mengenai hasil dan pembahasan dalam penelitian prediksi harga saham PT Aneka Tambang menggunakan regresi linear melibatkan analisis hasil prediksi dan interpretasi signifikansi temuan.

Hasil dan pembahasan berisi hasil-hasil temuan penelitian dan pembahasannya secara ilmiah. Tuliskan temuan-temuan ilmiah (*scientific finding*) yang diperoleh dari hasil-hasil penelitian yang telah dilakukan tetapi harus ditunjang oleh data-data yang memadai.

### Pengumpulan Data

Dataset dikumpulkan dari sumber yang valid pada situs <https://finance.yahoo.com/quote/ANTM.JK>. Dengan menggunakan kode ANTM.

Data ekonomi dan harga saham PT Aneka Tambang diperoleh dari halaman resmi Yahoo Finance dengan mengakses tautan yang telah disebutkan di atas. Proses ini memastikan bahwa data yang digunakan berasal dari sumber terpercaya dan dapat diandalkan.

data\_vec.show()

	Date	Open	High	Low	Close	Adj Close	Volume	features
2018-01-01	625.0	625.0	625.0	625.0	573.29175		0.0	[625.0, 625.0, 625.0, ...]
2018-01-02	635.0	645.0	625.0	635.0	582.4644		3.6621E7	[635.0, 645.0, 625.0, ...]
2018-01-03	635.0	655.0	630.0	630.0	577.87805		4.36076E7	[635.0, 655.0, 630.0, ...]
2018-01-04	635.0	645.0	635.0	640.0	587.0508		1.50649E7	[635.0, 645.0, 635.0, ...]
2018-01-05	645.0	660.0	645.0	655.0	600.8098		4.93237E7	[645.0, 660.0, 645.0, ...]
2018-01-08	660.0	660.0	640.0	650.0	596.2234		3.19205E7	[660.0, 660.0, 640.0, ...]
2018-01-09	655.0	680.0	650.0	655.0	600.8098		5.16889E7	[655.0, 680.0, 650.0, ...]
2018-01-10	655.0	665.0	645.0	650.0	596.2234		2.07743E7	[655.0, 665.0, 645.0, ...]
2018-01-11	655.0	660.0	645.0	650.0	596.2234		1.42357E7	[655.0, 660.0, 645.0, ...]
2018-01-12	650.0	670.0	645.0	665.0	609.9824		6.6859E7	[650.0, 670.0, 645.0, ...]
2018-01-15	670.0	680.0	660.0	660.0	605.3962		5.28282E7	[670.0, 680.0, 660.0, ...]
2018-01-16	665.0	720.0	665.0	705.0	646.6731		2.36009E8	[665.0, 720.0, 665.0, ...]
2018-01-17	705.0	740.0	690.0	730.0	669.60474		2.422831E8	[705.0, 740.0, 690.0, ...]
2018-01-18	750.0	775.0	730.0	740.0	678.7774		2.76688E8	[750.0, 775.0, 730.0, ...]
2018-01-19	745.0	765.0	735.0	760.0	697.12274		9.18918E7	[745.0, 765.0, 735.0, ...]
2018-01-22	765.0	790.0	765.0	775.0	710.8818		1.429921E8	[765.0, 790.0, 765.0, ...]
2018-01-23	780.0	810.0	775.0	800.0	733.8134		1.530833E8	[780.0, 810.0, 775.0, ...]
2018-01-24	805.0	815.0	780.0	780.0	715.4681		9.37947E7	[805.0, 815.0, 780.0, ...]
2018-01-25	810.0	870.0	810.0	845.0	775.0904		3.523565E8	[810.0, 870.0, 810.0, ...]
2018-01-26	860.0	875.0	855.0	855.0	784.263		1.364927E8	[860.0, 875.0, 855.0, ...]

only showing top 20 rows

Gambar 2 Data Harga Saham ANTM

### Pre-Processing Data

Pre-Processing Data yang gunakan ada proses cleaning dengan menghapus kolom yang berisi Nul.

```
In [36]: # Menghapus baris yang memiliki nilai null di kolom-kolom tertentu
data_cleaned = data.na.drop(subset=["Open", "High", "Low", "Close", "Adj Close", "Volume"])
```

Gambar 3 Proses Cleaning

### Pemodelan dengan Linear Regression

Pengodean ini adalah contoh dari proses pembuatan, pelatihan, dan pemeriksaan model regresi linier menggunakan PySpark MLlib. Berikut adalah penjelasan langkah-langkahnya:

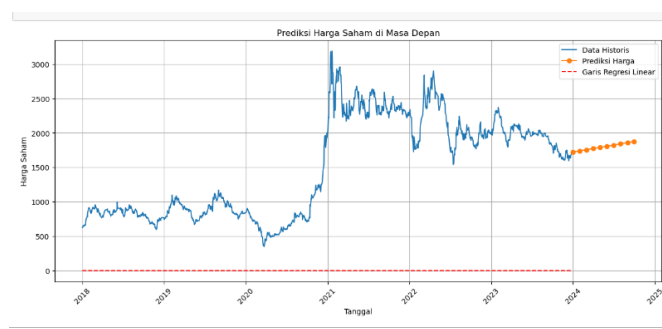
1. Membuat Objek Model Regresi Linear
2. Melatih Model Menggunakan Data Latih

Setelah membuat objek model, model tersebut dilatih menggunakan data latih yang diberikan (dalam hal ini, `training\_data`). Proses pelatihan dilakukan dengan meminimalkan selisih antara prediksi yang dihasilkan oleh model dengan nilai aktual dari variabel dependen.

3. Periksa Koefisien dan Intercept dari Model. Setelah melatih model, dapat memeriksa koefisien regresi yang dihasilkan. Koefisien ini menunjukkan seberapa besar pengaruh setiap fitur terhadap variabel dependen. Intercept adalah nilai pada sumbu y ketika semua fitur memiliki nilai nol. Ini adalah titik tempat garis regresi bersilangan dengan sumbu y.

Berdasarkan hasil evaluasi model linear regresi yang telah dilatih dan diuji, dapat disimpulkan bahwa model ini menunjukkan performa yang sangat baik dalam memprediksi harga saham:

1. Akurasi Relatif Model: Mencapai angka 100%, yang menandakan ketepatan prediksi yang sangat tinggi.
2. Mean Squared Error (MSE): Sangat rendah, yaitu  $7.302095386799366e-23$ , menunjukkan kesalahan rata-rata kuadrat yang sangat kecil antara nilai prediksi dan nilai aktual.
3. R-squared: Mencapai angka 1.0, yang mengindikasikan bahwa model dapat menjelaskan seluruh variabilitas pada data uji, dan tidak terdapat kesalahan yang tidak dapat dijelaskan oleh model.



Gambar 4 Prediksi Harga Saham

Hasil-hasil ini menunjukkan bahwa model linear regresi yang dikembangkan ini memiliki kemampuan prediksi yang sangat baik untuk memprediksi harga saham

## Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa model linear regresi dapat digunakan untuk memprediksi harga saham PT Aneka Tambang dengan akurasi yang tinggi. Hal ini dibuktikan dengan hasil evaluasi model yang menunjukkan nilai akurasi relatif sebesar 100%, kesalahan rata-rata kuadrat (MSE) sebesar  $7.302095386799366e-23$ , dan nilai R-squared sebesar 1.0.

Hasil ini menunjukkan bahwa model linear regresi dapat digunakan sebagai alat bantu untuk pengambilan keputusan investasi saham. Dengan menggunakan model ini, investor dapat memperkirakan harga saham di masa depan dengan lebih akurat, sehingga dapat mengurangi risiko kerugian yang mungkin terjadi.

Namun, perlu diingat bahwa prediksi harga saham tidak selalu akurat. Hal ini dikarenakan ada banyak faktor yang dapat mempengaruhi harga saham, termasuk kondisi ekonomi makro, kondisi industri, dan kinerja perusahaan. Oleh karena itu, investor tetap harus melakukan analisis fundamental dan analisis teknikal sebelum mengambil keputusan investasi saham.

## Referensi

- [1] Y. S. Sipahutar, I. R. Munthe, and ..., "Analisis Machine Learning Algoritma Regresi Linear Untuk Memprediksi Saham Di Bank Bri Di Bursa Saham Indonesia," *J. Tekinkom (Teknik ...*, vol. 6, no. 8, pp. 81–87, 2023, doi: 10.37600/tekinkom.v6i1.747.
- [2] G. N. Ayuni and D. Fitriana, "Penerapan metode Regresi Linear untuk prediksi penjualan properti pada PT XYZ," *J. Telemat.*, vol. 14, no. 2, pp. 79–86, 2019, [Online]. Available: <https://journal.ithb.ac.id/telematika/article/view/321>.
- [3] "PT Aneka Tambang Tbk (ANTM.JK) Jakarta - Jakarta Delayed Price. Currency in IDR," *Yahoo Finance*.  
<https://finance.yahoo.com/quote/ANTM.JK/history?p=ANTM.JK>.

---

**Copyright holder:**

Nama Author (Tahun Terbit)

**First publication right:**

Syntax Literate: Jurnal Ilmiah Indonesia

**This article is licensed under:**

