



## HYBRID SVR-GS UNTUK PREDIKSI SAHAM PT ANEKA TAMBANG TBK

Muhammad Afif Aunur Rohman<sup>1)</sup>, Taghfirul Azhima Yoga Siswa<sup>2)</sup>, Rofilde Hasudungan<sup>3)</sup>

<sup>1,2,3</sup> Teknik Informatika, Sains dan Teknologi, Universitas Muhammadiyah Kalimantan Timur, Indonesia

email: <sup>1</sup> [2211102441185@umkt.ac.id](mailto:2211102441185@umkt.ac.id), <sup>2</sup> [tay758@umkt.ac.id](mailto:tay758@umkt.ac.id), <sup>3</sup> [rh219@umkt.ac.id](mailto:rh219@umkt.ac.id)

ARTICLE INFO	ABSTRACT
<p><b>Article History:</b> Recieved : 3 Maret 2026 Accepted : 9 Mei 2026 Published : 3 Juni 2026</p> <hr/> <p><b>Keywords:</b> Machine Learning, Support Vector Regression, Grid Search, Prediksi, Saham</p>	<p>Sektor pertambangan di Indonesia memiliki peran strategis dalam perekonomian nasional, namun pergerakan harga sahamnya bersifat fluktuatif akibat pengaruh faktor eksternal, seperti harga komoditas dan kondisi pasar global. Kondisi tersebut menjadikan prediksi harga saham sebagai permasalahan yang kompleks. Penelitian ini bertujuan untuk memprediksi harga saham PT Aneka Tambang Tbk (ANTM) menggunakan metode <i>Support Vector Regression (SVR)</i> dengan optimasi hyperparameter melalui <i>Grid Search (GS)</i>, sehingga membentuk model <i>hybrid SVR-GS</i>. Data yang digunakan berupa data historis saham ANTM periode 2020–2025 sebanyak 1.202 data yang diperoleh dari Investing.com. Tahapan penelitian meliputi <i>preprocessing data</i>, <i>feature engineering</i> dengan pendekatan lag time, normalisasi <i>Min-Max</i>, pembagian data berbasis <i>time series</i>, serta evaluasi model. Kinerja model diukur menggunakan <i>Mean Absolute Error (MAE)</i>, <i>Root Mean Square Error (RMSE)</i>, dan <i>Mean Absolute Percentage Error (MAPE)</i>. Hasil penelitian menunjukkan bahwa <i>SVR Default</i> menghasilkan MAE 173,78 dan MAPE 9,65%, sedangkan <i>SVR Semi-Tuned</i> menurunkan kesalahan menjadi MAE 71,72 dan MAPE 3,19%. Model <i>SVR-GS</i> memberikan performa terbaik dengan MAE 45,16, RMSE 67,94, dan MAPE 2,24% pada rasio data 70:30. Dengan demikian, optimasi <i>Grid Search</i> terbukti meningkatkan akurasi prediksi harga saham ANTM secara signifikan.</p>
<p><b>IEEE style in citing this article:</b> M. A. A. Rohman, T. A. Y. Siswa, R. Hasudungan, "Hybrid SVR-GS Prediksi Saham PT Aneka Tambang Tbk", <i>jurnal.ilmiah.informatika</i>, vol. 11, no. 1, pp. xx-xx, Juni. 2026.</p>	
<p><b>Corresponding Author:</b> Taghfirul Azhima Yoga Siswa Universitas Muhammadiyah Kalimantan Timur</p>	

© 2026 Jurnal Ilmiah Informatika (Scientific Informatics Journal) with CC BY NC licence



## 1. PENDAHULUAN

Saham adalah sebuah aset finansial yang merepresentasikan sebagian porsi kepemilikan dalam suatu perusahaan. Dengan demikian, ketika investor membeli suatu saham pada dasarnya mereka sedang memperoleh sebagian kecil dari perusahaan tersebut [1]. Sektor tambang di Indonesia memainkan peran penting dalam mendorong perekonomian dengan dorongan dari potensi sumber daya alam yang melimpah, yang menjadikannya menarik bagi investor. Namun untuk memprediksi saham merupakan tantangan yang sulit karena saham bersifat fluktuatif [2].

Salah satu contoh kasus masalah saham di Indonesia muncul selama pandemi COVID-19 tahun 2020. Saat itu, Indeks Harga Saham Gabungan (IHSG) mengalami penurunan signifikan hingga mencapai titik terendah di angka 1,19 pada bulan Maret akibat kepanikan di kalangan investor dan penurunan aktivitas ekonomi nasional. Meskipun Bank Indonesia telah menurunkan BI Rate hingga 3,75% untuk mendorong pemulihan ekonomi, dan inflasi terjaga rendah karena lemahnya permintaan, kinerja IHSG tetap mengalami tekanan. Hal ini menunjukkan bahwa faktor-faktor eksternal, seperti ketidakpastian ekonomi memiliki dampak yang lebih besar terhadap fluktuasi saham dibandingkan variabel makro ekonomi seperti inflasi dan suku bunga [3].

Prediksi harga saham awalnya menggunakan pendekatan tradisional. Namun dengan kemajuan teknologi, prediksi saham mulai menggunakan pendekatan modern seperti *Machine Learning (ML)*, ML mampu mengolah jumlah data besar dengan cepat dan akurat yang membuat prediksi menjadi

lebih efisien [4]. ML adalah salah satu bidang dalam kecerdasan buatan yang memberikan kemampuan kepada sistem untuk belajar dari data dan menerapkannya dalam melaksanakan tugas tanpa perlu pemrograman yang khusus [5]. Terdapat berbagai pendekatan dalam ML seperti klasifikasi, *clustering*, prediksi, peramalan, dan asosiasi [6].

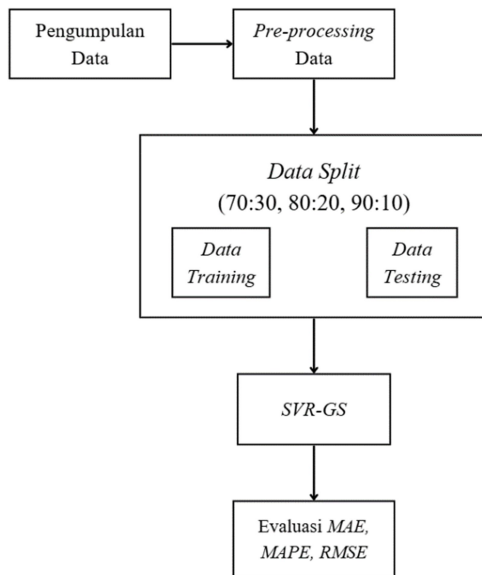
Metode prediksi yang sering digunakan pada saham, diantaranya mencakup *Adaptive Boosting*, *K-Nearest Neighbors*, *Linear Regression*, *Artificial Neural Network*, *Random Forest*, *Stochastic Gradient Descent*, *Support Vector Regression (SVR)*, *Decision Tree*, dan *Long Short-Term Memory* [7]. Dalam sebuah penelitian mengenai prediksi saham indeks MASI mengungkapkan bahwa SVR terbukti dapat memberikan hasil prediksi yang sangat memuaskan dengan nilai *Mean Absolute Error (MAE)* 28.69, *Root Mean Squared Error (RMSE)* 36.316, dan koefisien determinasi ( $R^2$ ) 0.997 pada dataset pengujian. Temuan ini menunjukkan tingkat akurasi yang lebih tinggi jika dibandingkan dengan *MultiLayer Perceptron* dan *XGBoost* [8].

SVR merupakan metode regresi yang didasarkan pada *Support Vector Machine (SVM)* yang dapat menangani data nonlinier dan berdimensi tinggi secara efisien. Dengan regulasi, SVR dapat menghindari *overfitting* dan meningkatkan generalisasi, sehingga menghasilkan prediksi yang akurat [9]. Berbeda dengan sebuah studi tentang prediksi bitcoin yang menunjukkan penerapan SVR memiliki performa yang kurang optimal jika dibandingkan dengan LSTM dengan nilai MAE berkisar antara 6.500-6.900 dan *Mean Squared Error (MSE)* yang melebihi 83 juta. Sebaliknya, LSTM menunjukkan akurasi yang lebih baik dengan nilai MAE antara 1.400-1.700 dan MSE yang berada di

kisaran 3-5 juta, sehingga peneliti merekomendasikan penerapan teknik optimasi hyperparameter untuk meningkatkan kemampuan prediktif [10].

Metode optimasi yang sering diterapkan dalam prediksi meliputi *Grid Search (GS)*, *Random Search*, *Bayesian Optimization (BO)*, *Genetic Algorithms*, dan *Particle Swarm Optimization (PSO)* yang menunjukkan perbedaan efektivitas dalam meningkatkan kinerja model [11]. *Grid Search* digunakan untuk mengatur faktor-faktor seperti kekuatan regulasi, tingkat pembelajaran, jumlah layer tersembunyi, dan aspek lainnya. *GS* mampu menemukan kombinasi hyperparameter paling optimal untuk model, sehingga meningkatkan akurasi prediksi dan kinerja [12]. Dalam sebuah penelitian tentang prediksi kualitas beton *SVR-GS* terbukti memberikan hasil terbaik dengan nilai *RMSE* 0,56, *MAE* 0,42, dan *MAPE* 1,18, yang lebih baik daripada *PSO-SVR* dan *BO-SVR* [13].

**2. METODE PENELITIAN**



Gambar 1. Alur Penelitian

Pada Gambar 1 menunjukkan tahapan yang diambil dalam alur penelitian. Studi

literatur dilakukan untuk membangun pemahaman teori dan metode yang relevan sebelum pengumpulan data, yang akan digunakan dalam penelitian. Setelah data dikumpulkan, tahap *pre-processing* dilakukan untuk memastikan kualitas data siap diolah. Selanjutnya, data dibagi dengan mencoba beberapa rasio untuk mencari rasio terbaik, yaitu 70:30, 80:20, dan 90:10. Dua metode digunakan untuk pemodelan: *SVR* standar tanpa optimasi dan *SVR* yang dioptimasi dengan *Grid Search*. Setelah itu, kedua model dievaluasi menggunakan *MAE*, *MAPE*, dan *RMSE* untuk membandingkan tingkat performa prediksi, sehingga dapat diputuskan apakah model mana yang lebih baik.

**2.1 Pengumpulan Data**

Data yang digunakan dalam penelitian berupa data historis PT Aneka Tambang yang diambil dari situs *Investing.com* dengan jumlah 1202 data selama 5 tahun terakhir (2020-2025), dengan beberapa mencakup informasi seperti Tanggal (*Date*) harga pembukaan (*Open*), harga tertinggi (*High*), harga terendah (*Low*), harga penutupan (*Close*), volume, dan perubahan [14].

**2.2 Preprocessing**

**2.2.1 Data Selection**

Pada tahap ini dilakukan penyaringan untuk memilih fitur-fitur yang paling berpengaruh terhadap hasil analisis, sementara fitur yang dianggap tidak relevan akan diabaikan atau dibuang dari data [15].

Tabel 1. Data Selection

Atribut (Sebelum)	Atribut (Sesudah)
Tanggal, Terakhir, Pembukaan, Tertinggi, Terendah, Volume, Perubahan	Terakhir, Pembukaan, Tertinggi, Terendah

Pada Tabel 1 menampilkan hasil dari proses *data selection* yang di mana fitur yang digunakan dalam analisis telah dipilih. Kolom "Atribut (Sebelum)" menunjukkan bahwa data awal terdiri dari tujuh elemen: Tanggal, Terakhir, Pembukaan, Tertinggi, Terendah, Volume, dan Perubahan. Setelah melalui proses seleksi, kolom "Atribut (Sesudah)" memperlihatkan bahwa hanya empat elemen yang dipertahankan, yaitu Terakhir, Pembukaan, Tertinggi, dan Terendah. Ini menunjukkan bahwa atribut Tanggal, Volume, dan Perubahan dinilai tidak relevan untuk pemodelan dan telah dikeluarkan dari data, agar analisis dapat lebih terfokus pada fitur harga yang paling berpengaruh.

### 2.2.2 Data Transformation

Tahap ini bertujuan untuk mengonversi data dari format non-numerik seperti teks atau label kategori ke dalam format angka agar dapat diproses secara matematis oleh algoritma atau model [15]. Teknik yang digunakan pada tahap ini adalah *type conversion* dan *numeric parsing*, yaitu metode dalam *machine learning* untuk mengubah nilai-nilai *nonnumeric* menjadi nilai numerik dan menjaga format tetap konsisten.

Tabel 2. Data Transformation

Sebelum		Sesudah	
Atribut	Nilai	Atribut	Nilai
Pembukaan	710.0	Pembukaan	710
Tertinggi	780.0	Tertinggi	780
Terendah	3.38	Terendah	3380
Terakhir	3.39	Terakhir	3390

Pada Tabel 2 menampilkan proses *data transformation* yang menunjukkan bahwa data diubah dari bentuk desimal menjadi angka bulat dengan skala harga yang sama. Nilai pembukaan dan tertinggi tetap dalam rentang yang sama, sedangkan nilai terendah dan terakhir

disesuaikan skalanya agar semua variabel harga memiliki format angka yang konsisten. Transformasi ini tidak mengurangi atau menambah jumlah data, hanya menyesuaikan skala dan bentuk angka tanpa menghilangkan atau menambah baris data, sehingga data setelah diubah menjadi lebih konsisten.

### 2.2.3 Data Split

Pada tahap ini data terbagi menjadi dua, yakni data latih dan data uji, di mana data latih digunakan untuk melatih model, sementara data uji digunakan untuk menguji model setelah proses pelatihan [16]. Dalam penelitian ini, akan menguji 3 rasio data dengan proporsi 70:30, 80:20, dan 90:10. Pengujian rasio ini untuk mencari rasio terbaik yang kemudian akan digunakan sebagai rasio utama.

### 2.2.4 Normalisasi dan Denormalisasi

Pada tahap ini normalisasi data dilakukan dengan memanfaatkan fungsi skalar *Min-Max*, sehingga data disesuaikan agar nilai fiturnya berada dalam kisaran 0 hingga 1. Proses pra-pemrosesan adalah tahap pembersihan dan pengolahan data agar model SVR dapat bekerja secara optimal [17]. Di sisi lain, denormalisasi bertujuan untuk mengembalikan nilai data ke skala aslinya setelah prediksi selesai, yang memungkinkan pemahaman hasil prediksi dalam satuan atau nilai yang lebih berkaitan dengan praktik [18], [19], [20]. Perhitungan normalisasi dan denormalisasi menggunakan persamaan 1 dan 2.

$$x_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (1)$$

Pada Persamaan 1,  $x_{norm}$  merupakan nilai baru hasil normalisasi, sedangkan  $x$  adalah nilai asli dari data sebelum

dilakukan normalisasi. Nilai  $x_{min}$  menyatakan nilai minimum dari seluruh data, dan  $x_{max}$  menyatakan nilai maksimum dari keseluruhan kumpulan data. Proses normalisasi ini bertujuan untuk menskalakan data ke dalam rentang 0 hingga 1 agar setiap fitur memiliki skala yang seragam.

$$x^i = x_{new} \times (x_{max} - x_{min}) + x_{min} \quad (2)$$

Pada Persamaan 2,  $x^i$  merupakan nilai data setelah dilakukan proses denormalisasi, sedangkan  $x_{norm}$  adalah nilai data yang telah dinormalisasi pada rentang 0 hingga 1. Nilai  $x_{min}$  dan  $x_{max}$  masing-masing menyatakan nilai minimum dan maksimum dari data asli sebelum normalisasi. Denormalisasi dilakukan untuk mengembalikan hasil prediksi ke skala nilai sebenarnya agar mudah diinterpretasikan.

### 2.2.5 Feature Engineering

Di tahap ini langkah penting untuk mendapatkan informasi dari data temporal adalah feature engineering pada data time series [21]. Salah satu prosedur utama adalah Lag Time, yaitu jeda waktu antara nilai saat ini dan masa lalu. Konsep ini krusial dalam analisis time series karena hubungan temporal antar data memengaruhi prediksi. Sebagai contoh, nilai satu hari sebelumnya disebut lag-1, dua hari sebelumnya lag-2, dan seterusnya. Penentuan orde lag dilakukan secara empiris atau berdasar pengetahuan domain. Dalam penelitian ini, lag time ditentukan dengan data harian satu minggu untuk prediksi nilai hari berikutnya [22]. Perhitungan lag dapat menggunakan persamaan 3.

$$Lagk(xt) = xt - k \quad (3)$$

Pada Persamaan 3,  $xt - k$  merupakan nilai lag atau nilai data pada periode

sebelumnya, sedangkan  $xt$  adalah nilai data pada waktu saat ini. Parameter  $k$  menyatakan jumlah periode waktu yang digeser ke belakang. Penggunaan lag time ini bertujuan untuk menangkap ketergantungan temporal pada data deret waktu.

### 2.3 Support Vector Regression

Di tahap ini SVR yang merupakan cabang dari SVM, telah diterapkan secara luas dalam berbagai disiplin ilmu untuk tujuan prediksi. Salah satu kelebihan dari SVR adalah kemampuannya dalam menangani hubungan *nonlinear* dan data dengan banyak dimensi, sehingga mampu menggambarkan interaksi yang rumit antara variabel lingkungan dengan lebih akurat. [9]. Perhitungan fungsi utama, persamaan linear dapat menggunakan persamaan 4.

$$f(x) = w^T x + b \quad (4)$$

Pada Persamaan 4,  $x$  merupakan vektor input,  $w$  adalah vektor bobot yang menentukan arah dan besar pengaruh setiap fitur, dan  $b$  merupakan *bias* yang berfungsi sebagai konstanta penyesuaian. Persamaan ini merepresentasikan fungsi regresi linear yang digunakan dalam *Support Vector Regression*.

Untuk menangani model mengalami *overfitting*, Vapnik memperkenalkan kesalahan yang dikenal sebagai fungsi  $\epsilon$ -insensitive pada tahun 1998 [23]. Perhitungan  $\epsilon$ -insensitive loss dapat menggunakan pada persamaan 5.

$$L_{\epsilon}(y, f(x)) = \begin{cases} 0 & \\ |y - f(x)| - \epsilon, & \text{jika } |y - f(x)| \leq \epsilon \\ & \text{jika } |y - f(x)| > \epsilon \end{cases} \quad (5)$$

Pada Persamaan 5,  $L_{\epsilon}(y, f(x))$  merupakan fungsi kerugian  $\epsilon$ -insensitive yang digunakan untuk menghitung

kesalahan prediksi. Variabel  $y$  menyatakan nilai aktual,  $f(x)$  adalah nilai prediksi yang dihasilkan model, dan  $\epsilon$  merupakan batas toleransi kesalahan yang diperbolehkan. Nilai kesalahan hanya akan diperhitungkan apabila selisih absolut antara nilai aktual dan prediksi melebihi nilai  $\epsilon$ .

Tujuan utama optimasi SVR berpusat pada meminimalkan fungsi  $\epsilon$ -insensitive serta mengurangi magnitudo vektor  $w$  [23]. Fungsi utama SVR dapat dihitung menggunakan persamaan 6 dan 7.

$$\min \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*) \quad (6)$$

dengan syarat :

$$\begin{cases} y_i - w^T \phi(x_i) - b \leq \epsilon + \xi_i \\ w^T \phi(x_i) + b - y_i \leq \epsilon + \xi_i^* \\ \xi_i + \xi_i^* \geq 0 \end{cases} \quad (7)$$

Pada Persamaan 6 dan 7, parameter  $C$  merupakan parameter regularisasi yang mengatur trade-off antara kompleksitas model dan besarnya kesalahan prediksi. Variabel  $\xi_i, \xi_i^*$  merupakan *slack variables* yang merepresentasikan kesalahan prediksi di luar margin  $\epsilon$ . Fungsi objektif ini bertujuan untuk meminimalkan kesalahan prediksi sekaligus menjaga margin tetap optimal.

Dengan memperkenalkan *multiplier Lagrange*  $\alpha_i^*$  dan  $\alpha_i$ , SVR nonlinier dapat diubah menjadi masalah dual [23]. Perhitungannya dapat menggunakan dalam persamaan 8.

$$f(x) = \sum_{i=1}^n (\alpha_i - \alpha_i^*) K(x_i, x) + b \quad (8)$$

Pada Persamaan 8,  $x_i$  merupakan vektor input dari data pelatihan ke- $i$ ,  $n$  menyatakan jumlah data pelatihan, dan  $\alpha_i$  serta  $\alpha_i^*$  adalah *multiplier Lagrange*. Fungsi

$K(x_i, x)$  merupakan fungsi kernel yang mengukur tingkat kesamaan antara data pelatihan dan data uji, sedangkan  $b$  adalah *bias*. Bentuk dual ini digunakan untuk menyelesaikan permasalahan regresi nonlinier pada SVR.

### 2.3.1 Kernel

Metode *kernel* bekerja dengan baik untuk memindahkan data ke ruang fitur dengan dimensi yang lebih tinggi, memungkinkan pemrosesan secara *linear* tanpa harus melakukan transformasi ( $\phi$ ) yang rumit secara langsung [24]. Perhitungan RBF dapat menggunakan persamaan 9.

$$K(x, x_i) = \exp\left(-\frac{|x, x_i|^2}{2\gamma^2}\right) \quad (9)$$

Pada Persamaan (9),  $K(x, x_i)$  merupakan fungsi kernel *Radial Basis Function* yang mengukur tingkat kesamaan antara dua  $x$  dan  $x_i$ . Jarak *Euclidean* antara kedua vektor tersebut dinyatakan dalam bentuk eksponensial, sedangkan parameter  $\gamma > 0$  berfungsi untuk mengontrol lebar kernel, yang memengaruhi tingkat fleksibilitas model dalam menangkap pola data.

### 2.3.2 Parameter

Penggunaan  $\epsilon$ -SVR untuk regresi tergolong efektif karena kinerjanya bergantung pada pengaturan tiga parameter kunci:  $C$ ,  $\gamma$ , dan  $\epsilon$ . Parameter  $C$  berperan sebagai koefisien penalti yang mengendalikan toleransi terhadap kesalahan. Selanjutnya,  $\gamma$  berfungsi menentukan sejauh mana pengaruh dari setiap titik data yang telah dipetakan ke dalam ruang fitur yang baru dan  $\epsilon$  bertindak sebagai batas dalam fungsi kerugian yang memungkinkan adanya kesalahan dalam prediksi selama masih dalam rentang toleransi tanpa dikenakan penalti [25].

## 2.4 Grid Search

*Grid Search* menetapkan batasan nilai untuk setiap hyperparameter, meliputi nilai terendah, nilai tertinggi, serta jumlah langkah di antaranya. *GS* lalu menyusun sebuah *grid* yang mencakup seluruh kombinasi *hyperparameter* yang mungkin dan menerapkan setiap kombinasi tersebut dengan menggunakan *cross validation* (*CV*) untuk menilai performanya [26]. Setiap kombinasi dinilai berdasarkan nilai performa dan kerugiannya, dan kombinasi dengan hasil terbaik kemudian dipilih untuk membuat model akhir [27].

## 2.5 Evaluasi

Dalam penelitian ini indeks yang digunakan untuk memvalidasi kinerja model prediksi adalah *Mean Absolute Error* (*MAE*), *Root Mean Square Error* (*RMSE*), *Mean Absolute Percentage Error* (*MAPE*) [28].

### 2.5.1 Mean Absolute Error

*MAE* adalah rata-rata perbedaan mutlak antara nilai aktual dan nilai prediksi [29]. Untuk menghitung *MAE* dapat menggunakan persamaan 10.

$$MAE = \frac{1}{N} \sum_{i=1}^N |\gamma_i - \bar{\gamma}_i| \quad (10)$$

Pada Persamaan 10,  $N$  merupakan jumlah total data pengamatan,  $\gamma_i$  adalah nilai aktual pada data ke- $i$ , dan  $\bar{\gamma}_i$  merupakan nilai prediksi pada data ke- $i$ . *MAE* dihitung sebagai rata-rata nilai absolut dari selisih antara nilai aktual dan nilai prediksi.

### 2.5.2 Root Mean Square Error

*Root Mean Square Error* (*RMSE*) menghitung rata-rata ukuran kesalahan

antara nilai yang diperkirakan dan nilai sebenarnya, dengan memberikan penekanan lebih pada kesalahan yang signifikan [29]. Untuk menghitung *RMSE* dapat menggunakan persamaan 11.

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (\gamma_i - \bar{\gamma}_i)^2} \quad (11)$$

Pada Persamaan 11,  $N$  menyatakan jumlah total data,  $\gamma_i$  merupakan nilai aktual, dan  $\bar{\gamma}_i$  adalah nilai prediksi. *RMSE* diperoleh dengan menghitung akar dari rata-rata kuadrat selisih antara nilai aktual dan nilai prediksi, sehingga memberikan penalti lebih besar pada kesalahan yang besar.

### 2.5.3 Mean Absolute Percentage Error

*Mean Absolute Percentage Error* (*MAPE*) digunakan untuk menilai ketepatan ramalan dalam bentuk kesalahan persentase. Nilai ini diperoleh dengan menghitung rata-rata kesalahan persentase mutlak antara nilai yang diprediksi dan nilai yang sebenarnya [29]. Untuk menghitung *MAPE* dapat menggunakan persamaan 12.

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{\gamma_i - y_i}{y_i} \right| \times 100 \quad (12)$$

Pada Persamaan 12,  $N$  merupakan jumlah data pengamatan,  $\gamma_i$  menyatakan nilai aktual, dan  $\bar{\gamma}_i$  adalah nilai prediksi. *MAPE* digunakan untuk mengukur tingkat kesalahan prediksi dalam bentuk persentase dengan menghitung rata-rata selisih absolut antara nilai aktual dan nilai prediksi terhadap nilai aktual.

### 3 HASIL DAN PEMBAHASAN

Dalam penelitian ini kami membandingkan tiga model SVR untuk memprediksi harga saham PT Aneka Tambang Tbk (ANTM), yaitu *SVR Default*, *SVR Semi-Tuned*, dan *SVR-GS*. Evaluasi dilakukan dengan menggunakan tiga metrik utama: *Mean Absolute Error (MAE)*, *Root Mean Squared Error (RMSE)*, dan *Mean Absolute Percentage Error (MAPE)*, yang mengukur kesalahan prediksi dari model yang digunakan.

Tabel 3. Hasil Pengujian Rasio *Data Split*

<i>Split</i>	<i>Train</i>	<i>Test</i>	<i>MAE</i>	<i>RMSE</i>	<i>MAPE%</i>
70:30	840	361	45.16	67.94	2.24
80:20	960	241	55.24	80.86	2.49
90:10	1080	121	80.78	105.16	3.07

Pada Tabel 3 menunjukkan hasil evaluasi dari pengujian 3 rasio *data split*, yaitu 70:30, 80:20, dan 90:10. Untuk rasio 70:30, model dilatih menggunakan 840 data dan diuji pada 361 data, dengan nilai *MAE* sebesar 45.16, *RMSE* 67.94, dan *MAPE* 2.24% yang paling rendah dibandingkan dengan rasio lainnya. Untuk rasio 80:20 dan 90:10, masing-masing menunjukkan nilai error yang lebih tinggi seiring dengan peningkatan proporsi data pelatihan dan penurunan. Berdasarkan perbandingan nilai error pada *data testing*, rasio 70:30 dipilih sebagai rasio terbaik karena memberikan kemampuan generalisasi model yang paling baik.

Tabel 4. Hasil Evaluasi

	<i>Default</i>	<i>Semi Tuned</i>	<i>Grid Search</i>
<i>MAE</i>	173.78	71.72	45.16
<i>RMSE</i>	245.83	149.42	67.94
<i>MAPE</i>	9.65%	3.19%	2.24%

Pada Tabel 4 menyajikan hasil evaluasi dari masing-masing model. Berdasarkan hasil yang diperoleh, *SVR Default* menunjukkan nilai *MAE* sebesar

173.78, *RMSE* sebesar 245.83, dan *MAPE* sebesar 9.65%. *SVR Semi-Tuned* yang menggunakan beberapa penyesuaian parameter menghasilkan *MAE* 71.72, *RMSE* 149.42, dan *MAPE* 3.19% yang menunjukkan peningkatan kinerja dibandingkan dengan model *default*. Dan *SVR-GS* yang menggunakan teknik optimasi *Grid Search* untuk menemukan kombinasi parameter terbaik menunjukkan hasil yang paling baik dengan *MAE* 45.16, *RMSE* 67.94, dan *MAPE* 2.24%.

Tabel 5. Kombinasi Parameter Terbaik

<i>Rank</i>	<i>C</i>	<i>Epsilon n</i>	<i>Gamma ma</i>	<i>mean_M ae</i>	<i>Std_MA E</i>
1	500	0.001	0.001	46.07372	32.59313
2	500	0.010	0.001	46.36383	31.94360
3	50	0.001	0.01	46.96579	34.81073
4	50	0.010	0.001	47.60907	30.51357
5	100	0.010	0.001	47.78134	33.21508

Selanjutnya pada Tabel 5 menunjukkan kombinasi parameter terbaik yang ditemukan menggunakan *Grid Search*. Berdasarkan pencarian ditemukan bahwa parameter  $C = 500$ ,  $\epsilon = 0.001$ , dan  $\gamma = 0.001$  memberikan kinerja terbaik.

Tabel 6. Pengujian *Fold best Parameter*

<i>Fold</i>	<i>MAE</i>	<i>RMSE</i>	<i>MAPE (%)</i>
1	136.92	167.40	5.39
2	48.46	62.87	2.00
3	33.17	45.20	1.39
4	56.21	84.47	2.51
5	52.23	67.49	2.41
6	36.77	48.00	1.84
7	28.63	37.18	1.38
8	26.51	35.73	1.33
9	17.97	25.67	0.97
10	23.86	30.62	1.48
<b>Rata-Rata</b>	46.07	60.46	2.07

Untuk pengujian lebih lanjut, dilakukan *cross-validation* dengan 10 *fold*, yang dapat dilihat pada Tabel 6. Hasilnya menunjukkan rata-rata menunjukkan  $MAE = 46.07$ ,  $RMSE = 60.46$ , dan  $MAPE = 2.07\%$ , ini menggambarkan stabilitas dan akurasi model yang sangat baik selama pelatihan dan pengujian.

Tabel 7. Aktual Vs Prediksi

Tanggal	Aktual	Default	SemiTuned	GS
2024-02-26	1.515	1687.998 596	1579.899 373	1546.449 761
2024-02-27	1.480	1687.603 116	1523.996 436	1517.105 939
2024-02-28	1.490	1659.790 928	1493.312 808	1480.707 365
2024-02-29	1.460	1655.518 359	1515.424 989	1489.515 511
2024-03-01	1.480	1648.425 408	1473.156 883	1463.058 730

Terakhir pada Tabel 7 memperlihatkan hasil perbandingan antara nilai aktual dan prediksi dari berbagai model *Default*, *Semi Tuned*, dan *Grid Search* menunjukkan bahwa model *Grid Search* memberikan prediksi yang sangat mendekati nilai aktual, dengan deviasi yang relatif kecil, yang mengindikasikan bahwa pengoptimalan parameter memberikan hasil yang paling akurat di antara semua model yang diuji.

Berdasarkan hasil evaluasi yang telah dilakukan, dapat disimpulkan bahwa model *hybrid SVR-GS* menghasilkan  $MAE$  45.16,  $RMSE$  67.94, dan  $MAPE$  2.24%, yang menunjukkan akurasi tinggi dan deviasi prediksi rendah. Hasil ini konsisten dengan penelitian lainnya [14] yang juga menunjukkan bahwa *Grid Search* mampu meminimalkan error dengan  $MAPE$  0,87%. Pengujian dengan *TimeSeriesSplit* menunjukkan kinerja stabil dengan rata-rata  $MAE$  46.07 dan  $MAPE$  2.07%, membuktikan kemampuan generalisasi model *SVR-GS* pada berbagai periode waktu.

Model *SVR* tanpa optimasi parameter menunjukkan kesalahan tinggi, dengan  $MAE$  173.78,  $RMSE$  245.83, dan  $MAPE$  9.65%. Penyesuaian manual pada model *SVR Semi-Tuned* mengurangi kesalahan menjadi  $MAE$  71.72,  $RMSE$  149.42, dan  $MAPE$  3.19%. Optimasi parameter dengan *Grid Search* menghasilkan peningkatan signifikan, dengan *SVR-GS* secara konsisten mengurangi kesalahan dibandingkan *SVR Default* dan *Semi-Tuned*. Hal ini sejalan dengan sebuah penelitian [24] yang menunjukkan peningkatan akurasi sebesar 10.74% dengan *SVR-GS*. Oleh karena itu, optimasi parameter meningkatkan kinerja prediksi model *SVR* secara signifikan.

#### 4 KESIMPULAN

Berdasarkan penelitian ini, model *SVR-GS* yang dioptimalkan menggunakan *Grid Search* menunjukkan kinerja terbaik dalam memprediksi harga saham PT Aneka Tambang Tbk. Optimasi parameter melalui *Grid Search* secara signifikan meningkatkan akurasi, terbukti dengan nilai  $MAE$ ,  $RMSE$ , dan  $MAPE$  yang lebih rendah dibandingkan model lainnya. Model *SVR Default* menunjukkan kesalahan tinggi, sementara *SVR Semi-Tuned* memberikan perbaikan, namun masih di bawah *SVR-GS*. Penelitian ini menegaskan pentingnya optimasi parameter untuk meningkatkan performa prediksi model *SVR*, menghasilkan prediksi yang lebih akurat dan sesuai dengan kondisi pasar saham yang dinamis.

Untuk penelitian selanjutnya, disarankan untuk menambahkan fitur teknikal lain seperti *MA*, *RSI*, atau *MACD*, serta menggunakan metode optimasi parameter lain untuk perbandingan. Penelitian bisa juga mengembangkan prediksi dengan horizon waktu lebih panjang dan menguji model pada saham

atau sektor yang berbeda untuk menguji generalisasi model.

## 5 REFERENSI

- [1] Furizal, A. Ritonga, A. Ma'arif, and I. Suwarno, "Stock Price Forecasting with Multivariate Time Series Long Short-Term Memory: A Deep Learning Approach," *J. Robot. Control*, vol. 5, no. 5, pp. 1322–1335, 2024, doi: 10.18196/jrc.v5i5.22460.
- [2] S. I. M. Rajagukguk and N. Hasanuh, "Pengaruh Return on Asset Dan Return on Equity Terhadap Harga Saham Pada Entitas Pertambangan Periode 2018-2022," *J. Econ. Bus. Eng.*, vol. 6, no. 1, pp. 157–164, 2024, doi: 10.32500/jebe.v6i1.7301.
- [3] A. R. Siringoringo *et al.*, "Analisis Pengaruh BI Rate dan Inflasi Terhadap Indeks Harga Saham Gabungan Di Indonesia Periode 2019-2023: Studi Dengan Model VAR," *EKOMA J. Ekon. Manajemen, Akunt.*, vol. 4, no. 3, pp. 5483–5490, 2025, doi: 10.56799/ekoma.v4i3.7227.
- [4] L. N. Mintarya, J. N. M. Halim, C. Angie, S. Achmad, and A. Kurniawan, "Machine learning approaches in stock market prediction: A systematic literature review," *Procedia Comput. Sci.*, vol. 216, pp. 96–102, 2022, doi: 10.1016/j.procs.2022.12.115.
- [5] M. Wazid, A. K. Das, V. Chamola, and Y. Park, "Uniting cyber security and machine learning: Advantages, challenges and future research," *ICT Express*, vol. 8, no. 3, pp. 313–321, 2022, doi: 10.1016/j.ict.2022.04.007.
- [6] W. Lestari and S. Sumarlinda, "Studi Komparatif Model Klasifikasi Kerentanan Penyakit Jantung Menggunakan Algoritma Machine Learning," *SATIN - Sains dan Teknol. Inf.*, vol. 9, no. 1, pp. 107–115, 2023, doi: 10.33372/stn.v9i1.918.
- [7] D. G. Singh, "Machine Learning Models in Stock Market Prediction," *Int. J. Innov. Technol. Explor. Eng.*, vol. 11, no. 3, pp. 18–28, 2022, doi: 10.35940/ijitee.c9733.0111322.
- [8] H. Kadiri, H. Oukhouya, and K. Belkhoutout, "A comparative study of hybrid and individual models for predicting the Moroccan MASI index: Integrating machine learning and deep learning approaches," *Sci. African*, vol. 28, p. e02671, 2025, doi: 10.1016/j.sciaf.2025.e02671.
- [9] Z. Liu, X. Huang, and X. Wang, "PM2.5 prediction based on modified whale optimization algorithm and support vector regression," *Sci. Rep.*, vol. 14, no. 1, pp. 1–15, 2024, doi: 10.1038/s41598-024-74122-z.
- [10] I. F. Riziq and A. R. Dzikrillah, "Implementasi Algoritma LSTM Dan SVR Untuk Prediksi Harga Bitcoin Menggunakan Data Yahoo Finance," *Metik J.*, vol. 9, no. 2, p. 2025, 2025, doi: 10.47002/metik.v9i2.1077.
- [11] L. Vancsura, T. Tatay, and T. Bareith, "Navigating AI-Driven Financial Forecasting: A Systematic Review of Current Status and Critical Research Gaps," *Forecast.*, vol. 7, no. 3, p. 36, 2025, doi: 10.3390/forecast7030036.
- [12] H. Oukhouya and K. El Himdi, "Comparing Machine Learning

- Methods—SVR, XGBoost, LSTM, and MLP— For Forecasting the Moroccan Stock Market,” in *Computer Sciences & Mathematics Forum*, 2023, p. 39. doi: 10.3390/iocma2023-14409.
- [13] M. Açıkkar, “Fast grid search: A grid search-inspired algorithm for optimizing hyperparameters of support vector regression,” *Turkish J. Electr. Eng. Comput. Sci.*, vol. 32, no. 1, pp. 68–92, 2024, doi: 10.55730/1300-0632.4056.
- [14] I. Miraltamirus, F. Fitri, D. Vionanda, and D. Permana, “Prediksi Harga Saham PT Bank Syariah Indonesia Tbk Menggunakan Support Vector Regression,” *UNP J. Stat. Data Sci.*, vol. 1, no. 3, pp. 112–119, 2023, doi: 10.24036/ujsds/vol1-iss3/43.
- [15] S. Muawwanah, T. A. Y. Siswa, and W. J. Pranoto, “Model Optimasi SVM-GSBE dalam Menangani High Dimensional Data Stunting Kota Samarinda,” *J. Teknol. Sist. Inf. dan Apl.*, vol. 7, no. 3, pp. 1246–1258, 2024, doi: 10.32493/jtsi.v7i3.41545.
- [16] C. A. Salsabila, F. Yulianto, and T. A. Y. Siswa, “Implementasi Metode Naive Bayes Untuk Klasifikasi Kecelakaan Lalu Lintas Di Kota Samarinda,” *J. Inform. dan Tek. Elektro Terap.*, vol. 13, no. 1, pp. 1268–1277, 2025, doi: 10.23960/jitet.v13i1.5890.
- [17] M. AL-Ghamdi, A. A. M. AL-Ghamdi, and M. Ragab, “A Hybrid DNN Multilayered LSTM Model for Energy Consumption Prediction,” *Appl. Sci.*, vol. 13, no. 20, pp. 1–19, 2023, doi: 10.3390/app132011408.
- [18] P. A. Nugroho, “KOMPUTA®: Jurnal Ilmiah Komputer dan Informatika IMPLEMENTASI JARINGAN SYARAF TIRUAN MULTI-LAYER PERCEPTRON UNTUK PREDIKSI PENYINARAN KOMPUTA®: Jurnal Ilmiah Komputer dan Informatika,” vol. 12, no. 1, pp. 83–90, 2023.
- [19] T. B. Sianturi, I. Cholissodin, and N. Yudistira, “Penerapan Algoritma Long Short-Term Memory ( LSTM ) berbasis Multi Fungsi Aktivasi Terbobot dalam Prediksi Harga Ethereum,” vol. 7, no. 3, pp. 1101–1107, 2023.
- [20] M. R. Patiallo, M. Fathurahman, S. Prangga, and E. Nadhilah, “Prediksi Curah Hujan di Kabupaten Berau Menggunakan Support Vector Regression Rainfall Prediction in Berau Regency Using Support Vector Regression,” vol. 16, pp. 112–121, 2025, doi: 10.30872/eksponensial.v16i2.1508.
- [21] E. Katya, “Exploring Feature Engineering Strategies for Improving Predictive Models in Data Science,” *Res. J. Comput. Syst. Eng.*, vol. 4, no. 2, pp. 201–215, 2023, doi: 10.52710/rjce.88.
- [22] V. Kumar, N. Kedam, K. V. Sharma, D. J. Mehta, and T. Caloiero, “Advanced Machine Learning Techniques to Improve Hydrological Prediction: A Comparative Analysis of Streamflow Prediction Models,” *Water (Switzerland)*, vol. 15, no. 14, pp. 1–24, 2023, doi: 10.3390/w15142572.
- [23] Z. Qi, Y. Feng, S. Wang, and C. Li, “Enhancing hydropower generation Predictions: A comprehensive study of XGBoost and Support Vector Regression

- models with advanced optimization techniques," *Ain Shams Eng. J.*, vol. 16, no. 1, p. 103206, 2025, doi: 10.1016/j.asej.2024.103206.
- [24] I. Fadil, M. A. Helmiawan, and Y. Sofiyana, "Optimization Parameters Support Vector Regression using Grid Search Method," in *2021 9th International Conference on Cyber and IT Service Management, CITSM 2021*, IEEE, 2021, pp. 21–25. doi: 10.1109/CITSM52892.2021.9589028.
- [25] B. Xiao, H. Wu, X. Zhang, R. Wu, and Y. Liu, "A Novel Approach for the Open-circuit Voltage Estimation of Lithium-ion Batteries by epsilon SVR," *Int. J. Electrochem. Sci.*, vol. 17, no. 5, p. 22059, 2022, doi: 10.20964/2022.05.14.
- [26] Z. Arifin, D. F. Rahman, B. S. Rintyarna, and D. Daryanto, "Penerapan Algoritma Support Vector Machine Berbasis Kernel Radial Basis Function dalam Klasifikasi Sel Kanker," *BIOS J. Teknol. Inf. dan Rekayasa Komput.*, vol. 4, no. 2, pp. 100–106, 2023, doi: 10.37148/bios.v4i2.165.
- [27] K. Alemerien, S. Alsarayreh, and E. Altarawneh, "Diagnosing Cardiovascular Diseases using Optimized Machine Learning Algorithms with GridSearchCV," *J. Appl. Data Sci.*, vol. 5, no. 4, pp. 1539–1552, 2024, doi: 10.47738/jads.v5i4.280.
- [28] M. N. Elizabeth, M. Mishra, S. Hasan, and A. Al-Durra, "Short-Term Solar Power Predicting Model Based on Multi-Step CNN Stacked LSTM Technique," *Energies*, vol. 15, no. 6, pp. 1–20, 2022, doi: 10.3390/en15062150.
- [29] H. H. Nuha, A. Balghonaim, R. R. Pahlevi, S. Rehman, and M. Mohandes, "Vertical Wind Speed Extrapolation Using Statistical Approaches," *FME Trans.*, vol. 52, no. 1, pp. 78–89, 2024, doi: 10.5937/fme2401078N.