

PENERAPAN METODE *SUPPORT VECTOR REGRESSION* (SVR) DENGAN ALGORITMA *GRID SEARCH* DALAM PERAMALAN HARGA SAHAM

Ni Putu Sri Yuli Artini^{1§}, I Wayan Sumarjaya², Desak Putu Eka Nilakusmawati³

¹Program Studi Matematika, Fakultas MIPA – Universitas Udayana [Email: niputusriyuliartini@gmail.com]

²Program Studi Matematika, Fakultas MIPA – Universitas Udayana [Email: sumarjaya@unud.ac.id]

³Program Studi Matematika, Fakultas MIPA – Universitas Udayana [Email: nilakusmawati@unud.ac.id]

[§]Corresponding Author

ABSTRACT

Stocks are the investment that is much in demand by investors because they are able to provide a high level of profit with a certain risk. Therefore, stock price forecasting is very important to maximize investment returns. The purpose of this study was to forecast stock prices using the support vector regression (SVR) method by utilizing linear, RBF, sigmoid, and polynomial kernel functions. Parameter optimization is carried out using a grid search algorithm that applies the concept of cross validation. After training and testing the model, the best SVR model is obtained using a polynomial kernel with parameters $C = 5$, $\epsilon = 0,01$, $\gamma = 1$, and $d = 1$, which produces an accuracy of 0,99211, RMSE of 0,01027, and MAE of 0,00723 on the training data and produces an accuracy value of 0,99389, RMSE of 0,01988, MAE of 0,01323, and MAPE of 0,02709 on data testing. Forecasting results for the next 85 periods using the best SVR model have a MAPE of 6,45%, this means that the SVR model obtained is able to predict closing stock prices much better than the ARIMA model which has a MAPE of 20,68%.

Keywords: *Stock Price, SVR, Grid Search, Kernel Function*

1. PENDAHULUAN

Investasi secara umum bisa diartikan sebagai proses penempatan dana yang seharusnya bisa digunakan pada masa sekarang di suatu perusahaan atau proyek dalam jangka waktu tertentu dengan harapan bisa memperoleh manfaat berupa balas jasa atau keuntungan pada masa yang akan datang (Hidayati, 2017). Salah satu jenis instrumen atau alat investasi yang paling diminati saat ini oleh para investor dan cukup banyak diperdagangkan oleh perusahaan adalah investasi saham. Hal ini terjadi karena tingkat keuntungan yang diperoleh melalui investasi saham cukup tinggi walaupun memiliki tingkat resiko tertentu (Indiani & Dewi, 2016).

Peramalan harga saham sangat penting dilakukan untuk dapat memaksimalkan keuntungan investasi yang diperoleh oleh investor. Dengan adanya peramalan, maka investor dapat memahami waktu yang benar-benar tepat untuk membeli atau menjual saham yang dimilikinya. Namun harga saham cenderung mengalami fluktuasi, adanya

fluktuasi ini dapat berpengaruh pada kinerja harga saham dan volume penjualan saham atau *trading volume*, sehingga dapat menyebabkan kesulitan dalam melakukan peramalan. Selain itu, untuk memperoleh hasil peramalan yang lebih detail juga diperlukan data dengan frekuensi yang tinggi (Gharehchopogh et al., 2013). Data dengan frekuensi yang tinggi ini bersifat sangat nonlinear (Jobson & Korkie, 1980).

Salah satu metode analisis yang dapat digunakan untuk meramalkan data deret waktu adalah metode *support vector regression* (SVR) (Muller et al., 1997). SVR dapat digunakan untuk memodelkan data deret waktu finansial baik linear maupun nonlinear serta mampu mengatasi masalah *over-fitting*.

SVR dapat mengatasi masalah data nonlinear dengan menggunakan trik kernel. Trik kernel adalah metode sederhana yang dipakai untuk memetakan data nonlinear berdimensi rendah dan mengubahnya ke dalam ruang dimensi yang lebih tinggi, dengan menjadikan

data nonlinear terpisah secara linear (Awad & Khanna, 2015). Selain itu, SVR juga dapat digunakan untuk mengatasi masalah *overfitting*, yaitu suatu keadaan ketika model pembelajaran mesin memberikan prediksi yang akurat pada data *training* tetapi tidak untuk data *testing*. Metode SVR menerapkan prinsip *structural risk minimization* yang meminimalkan batas atas *generalization error* daripada meminimalkan kesalahan *training*. Hal inilah yang membuat metode SVR memiliki generalisasi yang baik sehingga mampu mengatasi masalah *over-fitting* (Tay & Cao, 2001).

Terlepas dari keunggulan SVR tersebut, penentuan parameter dari model SVR masih sangat sulit untuk dilakukan (Guo et al., 2018). Salah satu metode optimasi parameter yang dapat digunakan adalah metode *grid search*. Metode ini memungkinkan untuk melakukan pemilihan model dan *hyperparameter tuning* dengan lebih cepat. *Grid search* adalah suatu algoritma yang berisi kombinasi parameter yang akan diuji pada suatu model SVR untuk mencari nilai *error* dalam fungsi regresi (Hsu et al., 2003).

Penelitian ini bertujuan untuk meramalkan harga saham menggunakan metode *Support Vector Regression* (SVR) dengan bantuan algoritma *Grid Search* untuk menentukan parameter model yang optimal. Berdasarkan dari buku yang berjudul "*Efficient Learning Machines*" oleh Awad & Khanna (2015), ada beberapa jenis fungsi kernel yang dapat digunakan untuk meramalkan data dengan metode SVR. Namun untuk mempersempit pembahasan hasil penelitian, peramalan dalam penelitian ini hanya menggunakan empat jenis fungsi kernel yang umum digunakan, yaitu kernel linear, RBF, sigmoid, dan polinomial.

Penelitian sebelumnya yang telah dilakukan dengan metode SVR adalah "*Application of Support Vector Machines in Financial Time Series Forecasting*" oleh Tay & Cao (2001). Berdasarkan penelitian tersebut, diperoleh kesimpulan bahwa metode *Support Vector Machine* yang digunakan untuk melakukan peramalan atau yang sering disebut dengan SVR dapat digunakan untuk meramalkan data deret waktu finansial dengan sangat baik. Hal ini juga didukung dengan penelitian yang berjudul "*An Adaptive SVR for High-Frequency Stock Price Forecasting*" oleh Guo, dkk (2018). Penelitian lain tentang metode SVR adalah "*Analisis Support Vector Regression (SVR) dalam Memprediksi Kurs Rupiah terhadap Dollar*

Amerika Serikat" oleh Amanda, dkk (2014). Pada penelitian ini disimpulkan bahwa peramalan menggunakan metode SVR dengan kernel linear memberikan akurasi yang sedikit lebih baik dibandingkan dengan kernel polinomial. Oleh karena itu peneliti tertarik untuk melakukan peramalan data deret waktu finansial dengan metode SVR dan membandingkan empat jenis fungsi kernel yang umum digunakan, yaitu kernel linear, RBF, sigmoid, dan polinomial.

2. METODE PENELITIAN

2.1 Jenis dan Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder berupa data *closing price* dari saham harian PT Adaro Energy Tbk (ADRO). Data dari BP *Statistical Review of World Energy* tahun 2022 menunjukkan bahwa Indonesia telah menjadi negara penghasil batu bara terbesar nomor dua di dunia. Menurut Bursa Efek Indonesia (2022), PT Adaro Energy Tbk adalah salah satu perusahaan yang bergerak di sektor pertambangan serta telah menjadi produsen unggul dan terbesar dalam produksi batu bara di Indonesia. Meskipun telah menjadi salah satu perusahaan terbesar, namun harga saham ADRO masih mengalami fluktuasi sehingga perlu adanya peramalan harga saham untuk dapat memaksimalkan keuntungan investasi.

Data saham ADRO diperoleh melalui *website* resmi *Yahoo Finance*, dengan periode analisis mulai dari 28 Oktober 2014 sampai 27 Oktober 2022 selama hari aktif. Berdasarkan dari percobaan yang telah dilakukan oleh peneliti, dapat disimpulkan bahwa penggunaan data saham ADRO yang meliputi masa pandemi COVID-19 tidak mengganggu hasil peramalan. Pada penelitian ini, diasumsikan bahwa harga penutupan saham besok hanya dipengaruhi oleh harga saham hari ini.

2.2 Metode Analisis Data

Pada penelitian ini, data diolah menggunakan bantuan *software* Jupyter Notebook dengan bahasa pemrograman Python. Adapun langkah-langkah dalam analisis data, yaitu:

a. Analisis Deskriptif

Melakukan analisis deskriptif untuk mengetahui gambaran umum mengenai data.

b. Menyiapkan Data

Dalam penelitian ini, variable bebas (X) yang

digunakan adalah data *closing price* pada waktu t dan variable terikat (Y) adalah data *closing price* pada waktu $t + 1$.

c. Melakukan Normalisasi Data

Proses normalisasi dilakukan menggunakan normalisasi *min-max* pada persamaan (1) dengan tujuan untuk mengubah data asli ke dalam rentang 0 sampai 1, sehingga dapat membantu membuat perhitungan numerik yang lebih tepat dan meningkatkan akurasi hasil peramalan.

$$x'_i = \frac{x_i - x_{min}}{x_{max} - x_{min}} \quad (1)$$

d. Data Splitting

Melakukan pembagian data menjadi data *training* dan data *testing* dengan perbandingan 80% untuk data *training* dan 20% untuk data *testing*. Perbandingan ini ditentukan sesuai dengan keinginan peneliti, yang mana dalam penerapannya persentase data *training* dan data *testing* tidak selalu bernilai 80% dan 20%.

e. Inisialisasi Parameter

Pada tahapan ini akan dilakukan inisialisasi parameter SVR, dengan nilai dari masing-masing parameter harus lebih besar dari nol sehingga rentangnya berada pada interval $(0, \infty)$ (Cholissodin & Soebroto, 2019). Pada penelitian ini, inisialisasi parameter SVR untuk percobaan model SVR menggunakan parameter acak yaitu dengan $C = \{0.01, 0.05, 0.1, 0.5, 1, 2, 5\}$, dan $\varepsilon = \{0.01, 0.1, 0.2, 0.5, 1\}$ untuk kernel linear, $C = \{0.01, 0.05, 0.1, 0.5, 1, 2, 5\}$, $\varepsilon = \{0.01, 0.1, 0.2, 0.5, 1\}$, dan $\gamma = \{0.1, 0.2, 0.3, 0.4, 0.5\}$ untuk kernel RBF dan kernel sigmoid, serta $C = \{0.01, 0.05, 0.1, 0.5, 1, 2, 5\}$, $\varepsilon = \{0.01, 0.1, 0.2, 0.5, 1\}$, dan $d = \{1, 2\}$ untuk kernel polinomial.

f. Optimasi Parameter dengan Grid Search

Melakukan *tuning* parameter atau pelatihan pada data *training* untuk setiap kombinasi parameter dari kernel linear pada model SVR menggunakan algoritma *grid search* dengan *cross validation*, tahapannya adalah sebagai berikut.

1. Membagi data *training* ke dalam tiga bagian dengan perbandingan yang sama karena dalam penelitian ini digunakan tiga *fold cross validation*. Pada proses *splitting*, urutan datanya harus diperhatikan karena akan dilakukan peramalan data *time series*.
2. Masing-masing *fold* ini akan dibagi lagi menjadi data *training* dan data *testing*.

3. Kemudian melakukan uji coba model SVR dengan kernel linear pada *fold* yang menjadi data *training*.
4. Selanjutnya akan dilakukan uji validasi model SVR pada *fold* yang menjadi data *testing*.
5. Selanjutnya akan dihitung *error rate* pada data *testing* dari masing-masing iterasi dengan menggunakan formula koefisien determinasi pada (2).

$$R^2 = 1 - \frac{\sum_{i=1}^M (y_i - \hat{y}_i)^2}{\sum_{i=1}^M (y_i - \bar{y})^2} \quad (2)$$

dengan

$$\bar{y} = \frac{1}{M} \sum_{i=1}^M y_i.$$

6. Menghitung rata-rata *error rate* untuk mendapatkan *error rate* secara keseluruhan menggunakan (3).

$$Error\ rate = \frac{1}{K} \sum_{i=1}^K Error_i \quad (3)$$

7. Ulangi langkah (3) sampai semua kombinasi parameter dari kernel linear telah diuji.
8. Karena dalam fitur *cross validation* menggunakan koefisien determinasi untuk *scoring* model, maka kombinasi parameter yang dipilih adalah kombinasi parameter yang memiliki rata-rata eror tertinggi untuk setiap kombinasi parameter yang diuji.
9. Ulangi langkah (1-9) hingga semua kombinasi parameter dari masing-masing kernel telah diuji menggunakan algoritma *grid search* dengan *time series cross validation*, sehingga akan diperoleh pasangan parameter yang optimal dari masing-masing kernel.

g. Retrained Model

Selanjutnya akan dilakukan pelatihan ulang untuk model SVR pada data *training* dengan estimasi pasangan parameter terbaik dari masing-masing kernel yang telah diperoleh menggunakan metode *grid search cross validation*. Dalam pelatihan model SVR, langkah pertama yang perlu dilakukan adalah menghitung matriks kernel yang digunakan. Adapun formulasi dari masing-masing fungsi kernel adalah sebagai berikut.

1. Fungsi untuk kernel linear

$$K(x_i, x_j) = x_i^T x_j$$

2. Fungsi untuk kernel RBF

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2)$$

3. Fungsi untuk kernel sigmoid

$$K(x_i, x_j) = \tanh(\gamma x_i^T x_j + c)$$

4. Fungsi untuk kernel polinomial

$$K(x_i, x_j) = (\gamma x_i^T x_j + c)^d$$

Setelah diperoleh matriks kernel, maka selanjutnya akan dicari Lagrange *multiplier* α dan α^* yang memaksimalkan formula $Q(\alpha, \alpha^*)$ pada persamaan (4).

$$\text{maksimalkan } Q(\alpha, \alpha^*) = -\frac{1}{2} \sum_{i,j=1}^M (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*) \phi^T(x_i) \phi(x_j) - \varepsilon \sum_{i=1}^M (\alpha_i + \alpha_i^*) + \sum_{i=1}^M y_i (\alpha_i - \alpha_i^*) \quad (4)$$

$$\text{dengan batasan } \sum_{i=1}^M (\alpha_i - \alpha_i^*) = 0$$

$$0 \leq \alpha_i \leq C, \quad 0 \leq \alpha_i^* \leq C, \quad i = 1, \dots, M.$$

Penyelesaian untuk mendapatkan Lagrange *multiplier* dapat dilakukan dengan menggunakan bantuan program Python pada *software* Jupyter Notebook. Data *training* yang dipilih sebagai *support vector* adalah x_i dengan $0 < \alpha_i \leq C$ atau $0 < \alpha_i^* \leq C$.

Dimisalkan $(\alpha_i - \alpha_i^*) = \beta_i$, maka setelah didapatkan nilai β_i dan *support vector* x_i , selanjutnya akan dihitung nilai b dengan menggunakan persamaan (5) untuk $0 < \alpha_i < C$ dan dengan persamaan (6) untuk $0 < \alpha_i^* < C$.

$$b = y_i - w^T \phi(x_i) - \varepsilon \quad (5)$$

$$b = y_i - w^T \phi(x_i) + \varepsilon \quad (6)$$

Untuk menghindari kesalahan perhitungan, akan dihitung nilai rata-rata dari b yang memenuhi persamaan (5) dan (6).

Dari nilai β_i , *support vector* x_i , dan b yang diperoleh, maka fungsi untuk model SVR dapat dinyatakan dengan

$$f(x) = \sum_{i=1}^M \beta_i K(x_i, x) + b, \quad (7)$$

dimana $K(x_i, x)$ didefinisikan sebagai fungsi kernel dan $K(x_i, x) = \phi^T(x_i) \phi(x)$.

h. Uji Validasi

Proses uji validasi dilakukan pada data *testing* untuk dapat memperoleh model SVR terbaik. Tahap pengujian ini dilakukan untuk menguji model yang telah dilatih dengan memasukkan data baru (data *testing*) yang belum pernah dilatih sebelumnya. Hal ini dilakukan untuk mengetahui model yang memberikan hasil prediksi paling baik dibandingkan model lainnya. Proses uji validasi pada masing-masing kernel dilakukan dengan menyubstitusikan data *testing* secara satu persatu ke dalam persamaan (7).

i. Pemilihan Model Terbaik

Menganalisis kinerja prediksi dari masing-masing model SVR pada data *training* dan data *testing* menggunakan nilai akurasi R^2 pada persamaan (2) dan nilai *root mean squared error*

(RMSE) pada persamaan (8).

$$RMSE = \sqrt{\frac{\sum_{t=1}^M (y_i - \hat{y}_i)^2}{M}} \quad (8)$$

dengan y_i adalah data aktual pada waktu ke- i , \hat{y}_i adalah data hasil prediksi pada waktu ke- i , dan M adalah jumlah data.

Model SVR terbaik adalah model dengan nilai R^2 terbesar dan nilai RMSE terkecil.

j. Peramalan dengan Model SVR terbaik

Pada tahapan ini dilakukan peramalan harga penutupan saham PT Adaro Energy Tbk untuk 85 periode ke depan menggunakan model SVR terbaik yang telah diperoleh. Proses ini dilakukan dengan menyubstitusikan variabel *input* x yang telah dinormalisasi ke dalam persamaan (7), sehingga diperoleh hasil peramalan pada periode ke-1. Kemudian substitusikan *output* peramalan periode ke-1 ke persamaan (7) sehingga akan diperoleh hasil peramalan pada periode ke-2. Begitu seterusnya hingga diperoleh hasil peramalan untuk 85 periode ke depan.

k. Mengukur Kinerja Hasil Peramalan

Melakukan analisis kinerja prediksi harga saham mendatang dengan menggunakan *mean absolute percentage error* (MAPE) untuk mengetahui akurasi dari hasil peramalan.

$$MAPE = \left(\frac{1}{M} \sum_{t=1}^M \left| \frac{y_i - \hat{y}_i}{y_i} \right| \right) \times 100 \quad (9)$$

Nilai MAPE dapat ditafsirkan atau diinterpretasikan ke dalam empat kategori, yaitu: $< 10\%$ artinya sangat akurat, $10\% - 20\%$ artinya baik, $20\% - 50\%$ masih dapat dikatakan wajar, namun apabila sudah $> 50\%$ maka hasil peramalan dikatakan tidak akurat (Zhang et al., 2015).

l. Denormalisasi

Proses denormalisasi dilakukan untuk mengembalikan nilai dari data hasil peramalan ke awal data sebelum dilakukannya normalisasi.

$$x_d = (x'_i (x_{max} - x_{min}) + x_{min}) \quad (10)$$

3. HASIL DAN PEMBAHASAN

3.1 Analisis Deskriptif

Analisis deskriptif dilakukan untuk mengetahui gambaran umum mengenai data. Grafik harga penutupan saham ADRO untuk periode analisis mulai dari 28 Oktober 2014 sampai dengan 27 Oktober 2022 dapat dilihat pada Gambar 3.1.



Gambar 3. 1 Grafik Harga Penutupan Saham ADRO

Berdasarkan Gambar 3.1, dapat diamati bahwa pola data aktual dari harga penutupan saham ADRO pada periode Oktober 2014 sampai Oktober 2022 mengalami fluktuasi yang cukup tinggi. Rendahnya harga saham ADRO dari tahun 2014 sampai pertengahan tahun 2016 diakibatkan oleh penurunan harga batu bara acuan. Berdasarkan penelitian yang dilakukan oleh Gunarto & Wulansari (2020), harga batu bara acuan berdampak positif dan signifikan terhadap harga saham batu bara, baik dalam jangka pendek maupun jangka panjang. Penurunan harga batu bara memberi dampak pada laba perusahaan sehingga investor cenderung melepas sahamnya karena muncul pesimisme atas dividen perusahaan. Hal ini mengakibatkan terjadinya penurunan harga saham perusahaan batu bara, salah satunya adalah PT Adaro Energy Tbk.

Pada tahun 2018, harga saham ADRO juga kembali mengalami penurunan karena melemahnya harga batu bara acuan.

Berdasarkan data yang diperoleh dari *website* Yahoo Finance, pada tahun 2022 harga saham ADRO meningkat sekitar 150% dalam setahun. Hal ini juga dipengaruhi oleh peningkatan harga batu bara acuan yang terjadi dari awal tahun 2022. Peningkatan harga batu bara ini terjadi karena adanya krisis energi di benua Eropa akibat perang Rusia dan Ukraina. Hal ini menyebabkan harga gas alam di Eropa terus mengalami peningkatan, seiring dengan adanya ketidakpastian pasokan gas di Eropa. Untuk mengatasi masalah ini, banyak perusahaan mulai beralih ke bahan bakar batu bara sebagai cadangan dalam memenuhi

kebutuhan energi untuk produksi. Selain itu, beberapa negara di Eropa juga telah mengaktifkan kembali pembangkit listrik dengan bahan bakar batu bara untuk mengantisipasi adanya krisis listrik menjelang musim dingin (Deputi Bidang Ekonomi, 2022). Di samping itu, China juga masih menjadi pengimpor batu bara terbesar di RI karena harga batu bara dari Indonesia jauh lebih murah dari China sendiri. Hal inilah yang menyebabkan terjadinya peningkatan harga batu bara Indonesia pada tahun 2022, sehingga berimbas pada meningkatnya harga saham batu bara, salah satunya adalah saham ADRO.

3.2 Penentuan Variabel

Pada penelitian ini hanya menggunakan satu fitur dengan asumsi bahwa harga penutupan saham besok hanya dipengaruhi oleh harga saham hari ini. Sehingga variabel bebas yang digunakan adalah data *closing price* pada waktu t dan variabel terikatnya adalah data *closing price* pada waktu $t + 1$.

3.3 Melakukan Normalisasi Data

Tahapan normalisasi dalam penelitian ini dilakukan dengan *min-max normalization* agar data pada variabel (X) dan variabel (Y) berada pada rentang yang sama, yaitu pada interval tertutup 0 sampai 1 sehingga dapat mengurangi kompleksitas perhitungan dan tingkatan eror dalam komputasi. Data harga penutupan saham tertinggi atau maksimum terjadi pada tanggal 7 Oktober 2022 dengan nilai sebesar Rp 4.140, sedangkan data dengan nilai terkecil atau minimum terjadi pada tanggal 20 Januari 2016

yaitu sebesar Rp 437. Dengan menggunakan persamaan (1) maka diperoleh normalisasi data ke-1 pada fitur X adalah sebagai berikut.

$$x'_i = \frac{x_i - x_{min}}{x_{max} - x_{min}}$$

$$x'_1 = \frac{(1030 - 437)}{(4140 - 437)} = 0,16014$$

3.4 Data Splitting

Pada tahap ini, data dibagi menjadi dua jenis yaitu data *training* dan data *testing*. Penentuan data *training* dan data *testing* pada penelitian ini didapatkan setelah melakukan normalisasi data. Pada penelitian ini dilakukan pembagian data dengan perbandingan 80% untuk data *training* dan 20% untuk data *testing*. Dengan jumlah data sebanyak 1995 data, maka diperoleh data *training* sebanyak 1596 data dan data *testing* sebanyak 399 data.

3.5 Pemilihan Kombinasi Parameter Kernel dengan Algoritma *Grid Search*

Pemilihan parameter terbaik pada kernel linear, RBF, sigmoid, dan polinomial bertujuan untuk mendapatkan kombinasi parameter terbaik dari masing-masing kernel. Proses ini dilakukan dengan melatih model SVR pada data *training* dengan menggunakan tiga *fold cross validation*, sehingga data *training* akan dibagi ke dalam tiga bagian dengan perbandingan yang sama dengan memperhatikan urutan data. *Cross*

validation bertujuan untuk mengukur kekonsistenan model dalam melakukan prediksi.

Kombinasi parameter optimal dari masing-masing kernel yang dipilih adalah kombinasi parameter yang memiliki nilai rata-rata akurasi tertinggi untuk setiap kombinasi parameter yang diuji. Hal ini karena dalam fitur *cross validation* menggunakan koefisien determinasi untuk *scoring* model. Hasil luaran dari penggunaan *grid search cross validation* untuk estimasi pasangan parameter terbaik dari masing-masing kernel tidak jauh berbeda, yaitu $C = 3,0$ dan $\varepsilon = 0,01$ untuk kernel linear dengan nilai akurasi sebesar 0,9720, $C = 5$, $\varepsilon = 0,01$ dan $\gamma = 0,5$ untuk kernel RBF dengan nilai akurasi sebesar 0,9699, $C = 5$, $\varepsilon = 0,01$ dan $\gamma = 0,4$ untuk kernel sigmoid dengan nilai akurasi sebesar 0,9681, serta $C = 5$, $\varepsilon = 0,01$, $\gamma = 1$, dan $degree = 1$, untuk kernel polinomial dengan nilai akurasi sebesar 0,9722.

3.6 Pemilihan Model Terbaik

Pemilihan model SVR terbaik dilakukan berdasarkan nilai akurasi R^2 , RMSE, MAE, dan MAPE yang diperoleh dari pengujian data *testing*. Model SVR terbaik adalah model dengan nilai akurasi R^2 terbesar dan *error* terkecil. Perbandingan nilai akurasi dan *error* dari masing-masing model dapat dilihat pada Tabel 3.1

Tabel 3.1 Perbandingan Akurasi dan Error dari Masing-masing Kernel

Kernel	Data <i>testing</i>			
	R^2	RMSE	MAE	MAPE
Linear ($C = 3, \varepsilon = 0,01$)	0,99388	0,01989	0,01325	0,02713
RBF ($C = 5, \varepsilon = 0,01, \gamma = 0,5$)	0,96492	0,04762	0,02908	0,04503
Sigmoid ($C = 5, \varepsilon = 0,01, \gamma = 0,4$)	0,43542	0,19106	0,11466	0,15528
Polinomial ($C = 5, \varepsilon = 0,01, \gamma = 1, d = 1$)	0,99389	0,01988	0,01323	0,02709

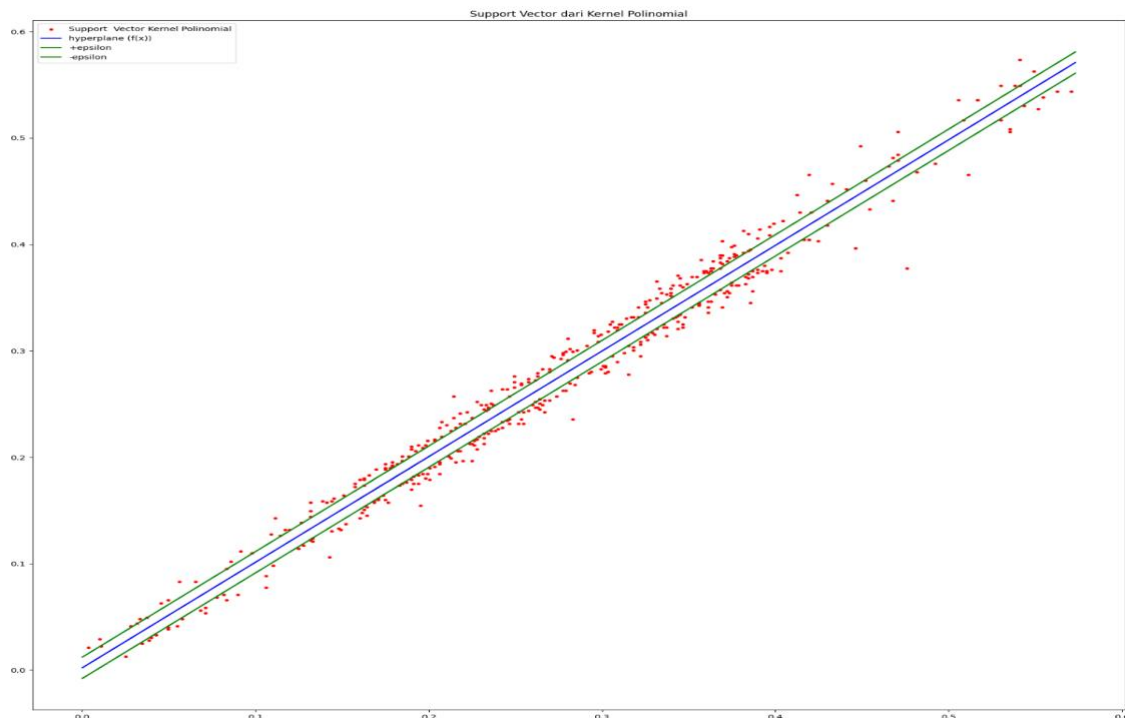
Berdasarkan penjabaran pada sub bab sebelumnya, hasil prediksi *closing price* saham ADRO menggunakan model dari kernel linear, RBF, sigmoid, dan polinomial memberikan hasil yang tidak jauh berbeda. Keempatnya memberikan akurasi yang baik dan *error* yang kecil pada data *training* maupun data *testing*. Namun jika dilihat pada Tabel 3.1, hasil prediksi yang memberikan nilai akurasi R^2 terbesar dan nilai eror terkecil pada data *testing* adalah prediksi dengan menggunakan model dari kernel polinomial. Oleh karena itu, dapat di simpulkan bahwa model SVR terbaik untuk memprediksi harga penutupan saham harian PT Adaro Energy Tbk adalah menggunakan kernel polinomial dengan $C = 5, \varepsilon = 0,01, \gamma = 1$, dan $d = 1$.

3.7 Peramalan

Dari tahap sebelumnya diperoleh model SVR terbaik yaitu menggunakan kernel polinomial dengan $C = 5, \varepsilon = 0,01, \gamma = 1$, dan $d = 1$. Persamaan dari model tersebut adalah sebagai berikut:

$$f(x) = \sum_{i=1}^{415} \beta_i ((1,0)x_i^T x) + 0,00208787 \quad (11)$$

dengan β_i adalah Lagrange *multiplier* dan *support vector* x_i dapat dilihat pada Gambar 3.2.



Gambar 3. 2 Support Vector dari Kernel Polinomial

Pada penelitian ini dilakukan peramalan untuk 85 langkah ke depan mulai dari 28 Oktober 2022 sampai dengan 24 Februari 2023. Variabel *input* yang digunakan yaitu data *closing price* saham ADRO pada tanggal 27 Oktober 2022 yang bernilai Rp 3.980. Sebelum melakukan proses peramalan, variabel *input* tersebut harus dinormalisasi terlebih dahulu untuk menyamakan rentang data yang digunakan. Dengan menggunakan persamaan (1) maka diperoleh hasil normalisasi sebagai berikut:

$$x_{\text{norm}} = \frac{(3980-437)}{(4140-437)} = 0,956792$$

Selanjutnya substitusikan nilai variabel *input* x yang telah dinormalisasi ke dalam persamaan (11), sehingga diperoleh hasil peramalan pada periode ke-1 sebagai berikut:

$$f(1) = \sum_{i=1}^{415} \beta_i ((1,0)x_i^T [0,956792]) + 0,002087 = 0,95149294$$

Kemudian substitusikan hasil peramalan periode ke-1 ke persamaan (11), sehingga diperoleh:

$$f(2) = \sum_{i=1}^{415} \beta_i ((1,0)x_i^T [0,951493]) + 0,002087 = 0,94623478$$

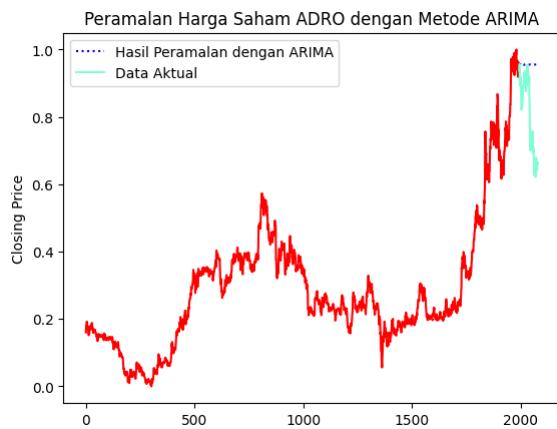
Begitu seterusnya hingga diperoleh hasil peramalan untuk 85 periode ke depan.

Setelah hasil peramalan didapatkan, maka langkah selanjutnya adalah melakukan evaluasi dengan cara membandingkan antara nilai aktual *closing price* saham ADRO dengan data hasil peramalan. Evaluasi ini bertujuan untuk membandingkan tingkat kesalahan hasil

ramalan dengan menggunakan MAPE pada persamaan (9).

Peramalan harga penutupan saham PT Adaro Energy Tbk untuk 85 periode ke depan dilakukan menggunakan model SVR terbaik yang telah diperoleh sebelumnya, yaitu menggunakan fungsi kernel polinomial dengan $C = 5, \varepsilon = 0,01, \gamma = 1, \text{ dan } d = 1$. Menurut Zhang et al. (2015), nilai MAPE yang kurang dari 10% menunjukkan bahwa hasil peramalan sangat akurat. Sedangkan nilai MAPE antara 10% sampai 20% menunjukkan bahwa peramalan baik. Berdasarkan persamaan (9), diperoleh nilai evaluasi hasil peramalan saham ADRO memiliki tingkat kesalahan MAPE sebesar 0,0645 atau 6,45%. Dengan demikian, dapat disimpulkan bahwa model SVR tersebut dapat memprediksi harga penutupan saham ADRO dengan sangat baik karena nilai MAPE di bawah 10%. Selain itu, peneliti juga mencoba membandingkan hasil peramalan saham ADRO dengan metode SVR dan metode ARIMA. ARIMA adalah salah satu metode statistik yang paling umum digunakan untuk meramalkan data *times series*. Dengan menggunakan ARIMA diperoleh nilai MAPE sebesar 20,68%, hal ini menunjukkan bahwa metode ARIMA memiliki kinerja yang kurang baik apabila digunakan untuk meramalkan data harga saham yang memiliki tren jangka panjang.

Visualisasi dari hasil peramalan untuk 85 periode ke depan dengan menggunakan metode ARIMA dan metode SVR secara berturut-turut dapat dilihat pada Gambar 3.2 dan Gambar 3.3.



Gambar 3. 3 Hasil Peramalan Saham ADRO dengan Metode ARIMA



Gambar 3. 4 Hasil Peramalan dengan Metode SVR untuk 85 Periode ke Depan

Pada Gambar 3.4, hasil peramalan yang berupa garis lurus merupakan akibat dari penggunaan kernel polinomial dengan $d = 1$, sehingga akan membentuk pola data yang linear. Selain itu, berdasarkan dari beberapa percobaan yang telah dilakukan oleh peneliti, dapat diambil kesimpulan bahwa hasil peramalan dengan metode SVR dapat mendekati nilai aktual secara general. Meskipun model SVR tidak dapat membaca pola fluktuasi secara spesifik, namun model ini mampu membaca kecenderungan naik atau turun (*tren*) yang akan terjadi pada pola data finansial ke depannya. Hal ini terjadi karena metode SVR menerapkan prinsip *structural risk minimization* yang meminimalkan batas atas *generalization error* daripada meminimalkan kesalahan *training*, sehingga metode SVR memiliki generalisasi yang baik dan mampu mengatasi masalah *over-fitting* (Tay & Cao, 2001).

3.8 Denormalisasi & Interpretasi Hasil Peramalan

Pada tahapan ini akan dilakukan proses denormalisasi data untuk mengembalikan nilai dari data hasil akhir ke awal data sebelum dilakukannya normalisasi. Dengan menggunakan persamaan (10) maka diperoleh denormalisasi hasil peramalan data ke-1 sebagai berikut.

$$x_d = (x'_i(x_{max} - x_{min}) + x_{min})$$

$$x_1 = (0,951493(4140 - 437) + 437)$$

$$x_1 = 3960$$

Tabel 3. 2 Denormalisasi Hasil Peramalan Saham ADRO

No.	Tanggal	Aktual	Hasil Ramalan	No.	Tanggal	Aktual	Hasil Ramalan
1	28/10/2022	3910	3960	5	03/11/2022	3750	3883
2	31/10/2022	3980	3941	6	04/11/2022	3750	3865
3	01/11/2022	3750	3922	7	07/11/2022	3830	3846
4	02/11/2022	3740	3902	8	08/11/2022	3760	3827
9	09/11/2022	3640	3809	48	03/01/2023	3520	3190
10	10/11/2022	3480	3790	49	04/01/2023	3300	3177
11	11/11/2022	3490	3772	50	05/01/2023	3100	3163
12	14/11/2022	3500	3754	51	06/01/2023	3140	3150
13	15/11/2022	3510	3736	52	09/01/2023	3030	3137
14	16/11/2022	3570	3719	53	10/01/2023	3140	3124
15	17/11/2022	3590	3701	54	11/01/2023	3180	3111
16	18/11/2022	3590	3684	55	12/01/2023	3090	3098
17	21/11/2022	3690	3666	56	13/01/2023	3140	3085
18	22/11/2022	3700	3649	57	16/01/2023	3100	3072
19	23/11/2022	3750	3632	58	17/01/2023	3140	3060

Tabel 3. 2 Denormalisasi Hasil Peramalan Saham ADRO (Lanjutan)

No.	Tanggal	Aktual	Hasil Ramalan	No.	Tanggal	Aktual	Hasil Ramalan
20	24/11/2022	3750	3615	59	18/01/2023	3170	3047
21	25/11/2022	3730	3598	60	19/01/2023	3230	3035
22	28/11/2022	3740	3582	61	20/01/2023	3240	3022
23	29/11/2022	3880	3565	62	24/01/2023	3200	3010
24	30/11/2022	3870	3549	63	25/01/2023	3130	2998
25	01/12/2022	3900	3532	64	26/01/2023	3010	2986
26	02/12/2022	3850	3516	65	27/01/2023	3060	2974
27	05/12/2022	3780	3500	66	30/01/2023	3010	2962
28	06/12/2022	3810	3484	67	31/01/2023	2960	2951
29	07/12/2022	3830	3468	68	01/02/2023	2950	2939
30	08/12/2022	3740	3453	69	02/02/2023	2870	2927
31	09/12/2022	3680	3437	70	03/02/2023	2760	2916
32	12/12/2022	3800	3422	71	06/02/2023	2780	2904
33	13/12/2022	3740	3406	72	07/02/2023	2890	2893
34	14/12/2022	3880	3391	73	08/02/2023	2920	2882
35	15/12/2022	3900	3376	74	09/02/2023	2840	2871
36	16/12/2022	3890	3361	75	10/02/2023	2740	2860
37	19/12/2022	3970	3346	76	13/02/2023	2780	2849
38	20/12/2022	3920	3332	77	14/02/2023	2830	2838
39	21/12/2022	3920	3317	78	15/02/2023	2950	2827
40	22/12/2022	3930	3302	79	16/02/2023	2920	2816
41	23/12/2022	3810	3288	80	17/02/2023	2880	2806
42	26/12/2022	3800	3274	81	20/02/2023	2900	2795
43	27/12/2022	3750	3260	82	21/02/2023	2820	2785
44	28/12/2022	3750	3246	83	22/02/2023	2850	2774
45	29/12/2022	3800	3232	84	23/02/2023	2900	2764
46	30/12/2022	3850	3218	85	24/02/2023	2890	2754
47	02/01/2023	3590	3204				

Berdasarkan Tabel 3.2 dan Gambar 3.3, baik data aktual maupun data hasil peramalan cenderung mengalami penurunan harga saham. Untuk memaksimalkan keuntungan investasi, para investor disarankan untuk tidak membeli saham ADRO pada tanggal 27 Oktober 2022 karena dapat menyebabkan adanya kerugian atau *capital loss*.

Penurunan harga saham ADRO pada bulan November ke Februari kemungkinan juga diakibatkan oleh menurunnya harga batu bara acuan. Pada bulan-bulan sebelumnya, Eropa telah membeli sejumlah besar batu bara untuk dijadikan cadangan energi listrik selama musim dingin. Sehingga pada bulan selanjutnya, beberapa negara di Eropa diperkirakan akan

mengurangi permintaan batu bara, karena cadangan batu bara yang telah diimpor sebelumnya dirasa sudah mencukupi. Pada saat yang bersamaan, perusahaan batu bara telah memaksimalkan produksi karena naiknya harga batu bara dari awal tahun 2022. Sehingga dengan melemahnya permintaan terhadap batu bara di Indonesia dan meningkatnya produksi maka perusahaan akan cenderung berfokus pada penjualan batu bara guna menghindari persediaan berlebih. Hal ini dapat mengakibatkan adanya penurunan harga batu bara, sehingga berdampak pada penurunan harga saham ADRO sebagai salah satu perusahaan yang bergerak di sektor pertambangan batu bara.

4. KESIMPULAN DAN SARAN

4.1 Kesimpulan

Berdasarkan pemaparan dari hasil penelitian ini, maka dapat disimpulkan bahwa:

1. Estimasi pasangan parameter terbaik yang diperoleh menggunakan algoritma *grid search* CV dari masing-masing kernel yaitu $C = 3,0$ dan $\varepsilon = 0,01$ untuk kernel linear, $C = 5$, $\varepsilon = 0,01$ dan $\text{gamma} = 0,5$ untuk kernel RBF, $C = 5$, $\varepsilon = 0,01$ dan $\text{gamma} = 0,4$ untuk kernel sigmoid, serta $C = 5$, $\varepsilon = 0,01$, $\text{gamma} = 1$, dan $\text{degree} = 1$, untuk kernel polinomial.
2. Model SVR terbaik yang terbentuk dari data harga penutupan saham ADRO adalah menggunakan kernel polinomial dengan parameter $C = 5$, $\varepsilon = 0,01$, $\text{gamma} = 1$, dan $\text{degree} = 1$. Pada data *training* kernel polinomial menghasilkan akurasi sebesar 0,99211, nilai RMSE sebesar 0,01027, dan MAE sebesar 0,0073. Sedangkan untuk data *testing* kernel polinomial menghasilkan akurasi sebesar 0,99389, nilai RMSE sebesar 0,01988, MAE sebesar 0,01323, dan MAPE sebesar 0,02709. Dengan demikian dapat disimpulkan bahwa model ini dapat memberikan hasil peramalan yang sangat baik tanpa menyebabkan *overfitting*.
3. Hasil peramalan menggunakan model SVR dengan kombinasi parameter terbaik dari kernel polinomial memiliki MAPE sebesar 6,45%. Hal ini membuktikan bahwa model SVR tersebut dapat memprediksi harga penutupan saham ADRO jauh lebih baik dibandingkan dengan peramalan dengan metode ARIMA yang memiliki nilai MAPE sebesar 20,68%.

4.2 Saran

Adapun saran dari penelitian yaitu:

1. Pada penelitian ini diketahui bahwa metode *grid search* memiliki kelemahan dalam proses *tuning* yang memakan waktu ketika *hyperparameter* ditambahkan, karena jumlah kombinasi parameter akan meningkat secara eksponensial. Oleh karena itu, untuk penelitian selanjutnya diharapkan dapat menggunakan metode optimasi parameter lain seperti *fruit fly optimization* (FOA).
2. Pembagian data *training* dan *testing* pada penelitian ini hanya menggunakan perbandingan 80% dan 20%. Sehingga penelitian selanjutnya diharapkan dapat

mencoba beberapa perbandingan persentase lainnya untuk memperoleh kombinasi persentase terbaik dalam peramalan harga saham dengan metode SVR.

3. Hasil peramalan harga saham yang diperoleh dengan metode SVR juga dapat digunakan untuk menghitung harga kontrak opsi pada penelitian selanjutnya.

DAFTAR PUSTAKA

- Amanda, R., Yasin, H., & Prahutama, A. (2014). Analisis Support Vector Regression (SVR) dalam Memprediksi Kurs Rupiah terhadap Dollar Amerika Serikat. *Jurnal Gaussian*, 3(4), 849–857.
- Awad, M., & Khanna, R. (2015). *Efficient Learning Machines*. Apress Open.
- Cholissodin, I., & Soebroto, A. A. (2019). *Buku Ajar AI, Machine Learning & Deep Learning*. Malang: Fakultas Ilmu Komputer, Universitas Brawijaya.
- Dale, S. (2022). *BP Statistical Review of World Energy*. British Porelum.
- Deputi Bidang Ekonomi. (2022). *Perkembangan Ekonomi Indonesia dan Dunia Triwulan I Tahun 2022*.
- Gharehchopogh, F. S., Bonab, T. H., & Khaze, S. R. (2013). A Linear Regression Approach to Prediction of Stock Market Trading Volume: A Case Study. *International Journal of Managing Value and Supply Chains*, 4(3), 25–31.
- Gunarto, M., & Wulansari, R. (2020). Analisis Pergerakan Harga Saham berdasarkan Harga Acuan dan Volume Penjualan: Studi pada PT Bukit Asam Tbk. *Jurnal Manajemen Dan Bisnis Sriwijaya*, 18(4), 1412–4521.
- Guo, Y., Han, S., Shen, C., Li, Y., Yin, X., & Bai, Y. (2018). An Adaptive SVR for High-Frequency Stock Price Forecasting. *IEEE Access*, 6, 11397–11404.
- Hidayati, A. N. (2017). Investasi: Analisis dan Relevansinya dengan Ekonomi Islam. *Jurnal Ekonomi Islam*, 8(2), 227–242.
- Hsu, C.-W., Chang, C.-C., & Lin, C.-J. (2003). *A Practical Guide to Support Vector Classification*. 1–16.

- Indiani, N. P. L., & Dewi, S. K. S. (2016). Pengaruh Variabel Tingkat Kesehatan Bank terhadap Harga Saham Perbankan di Bursa Efek Indonesia. *Jurnal Manajemen Unud*, 5(5), 2756–2785.
- Jobson, J. D., & Korkie, B. (1980). Estimation for Markowitz Efficient Portfolios. *Journal of the American Statistical Association*, 75(371), 544–554.
- Muller, K. R., Smola, A. J., Ratsch, G., Scholkopf, B., Kohlmorgen, J., & Vapnik, V. (1997). Predicting Time Series with Support Vector Machines. *International Conference on Artificial Neural Networks*, 999–1004.
- Tay, F. E. H., & Cao, L. (2001). Application of Support Vector Machines in Financial Time Series Forecasting. *Omega*, 29, 309–317.
- Zhang, T., Wang, K., & Zhang, X. (2015). Modeling and analyzing the transmission dynamics of HBV epidemic in Xinjiang, China. *PLoS ONE*, 10(9), 1–14.