

## PREDIKSI PERGERAKAN HARGA SAHAM MENGUNAKAN SUPPORT VECTOR REGRESSION

Lisa Yuli Kurniawati<sup>1\*)</sup>, Handayani Tjandrasa<sup>2)</sup>, Isye Arieshanti<sup>3)</sup>

<sup>1,2,3)</sup>Teknik Informatika, Fakultas Teknologi Informasi  
Institut Teknologi Sepuluh Nopember  
Surabaya, Indonesia

<sup>\*)</sup>[yuli.kurniawati11@mhs.if.its.ac.id](mailto:yuli.kurniawati11@mhs.if.its.ac.id)

### ABSTRAK

Dalam pasar saham, harga suatu saham dapat berubah secara cepat dari waktu ke waktu. Para pemilik saham diharapkan dapat segera memutuskan kapan saham sebaiknya dijual atau tetap dipertahankan. Karenanya prediksi pergerakan harga saham sampai saat ini masih menjadi topik hangat untuk diperbincangkan dalam dunia jual beli saham. Model prediksi pergerakan harga saham yang akurat dapat membantu para investor dalam pertimbangan pengambilan keputusan transaksi saham. Di dalam praktiknya, harga suatu saham dapat diprediksi dengan menggunakan konsep analisa teknikal. Analisa teknikal didasarkan pada prinsip penggunaan data histori harga saham untuk memprediksi pergerakan saham dimasa mendatang. Tujuan penelitian ini adalah mengimplementasikan metode *Support Vector Regression* dalam analisa teknikal untuk memprediksi pergerakan harga saham di masa mendatang. *Support Vector Regression* (SVR) merupakan pengembangan dari metode support vector machine untuk kasus regresi. Metode ini mampu mengatasi overfitting serta mampu menunjukkan performa yang bagus. Dari serangkaian uji coba yang dilakukan, dapat disimpulkan bahwa metode SVR dapat memprediksi pergerakan harga saham dengan cukup baik. Hal ini terlihat dari nilai NRMSE terbaik yang didapatkan sebesar 0.14.

*Kata kunci:* Analisa Teknikal, Prediksi harga saham, Regresi, SVR.

### ABSTRACT

*In stock markets, the price of a stock can change rapidly over time. The investor may decide when the stock should be sold or retained. For this reason, the stock price movement prediction is still a hot topic to be discussed in the world of buying and selling stocks. The model which was accurate for stock price movement prediction may help investors in consideration of the decision-making of stock transactions. In practice, stock price may be predicted with the technical analysis approach. Technical analysis is based on the principle of using historical data to predict stock price movement in the future. The purpose of this study is to implement Support Vector Regression in technical analysis to predict the movement of stock prices. Support Vector Regression (SVR) was the development of support vector machine for regression case. This method was able to overcome the over fitting and be able to show good performance. From a series of experiments, it can be concluded that the SVR method may predict the movement of stock prices pretty well. It can be seen from the best values of NRMSE is 0.14.*

*Keywords:* Technical Analysis, Stock price forecasting, Regression, SVR,.

## PENDAHULUAN

Pasar modal atau pasar saham adalah tempat bertemunya penjual dan pembeli yang melakukan transaksi jual beli sertifikat kepemilikan status perusahaan, atau disebut juga dengan saham. Harga suatu saham didalam pasar modal dapat berubah dengan cepat dari waktu ke waktu. Bagi investor, harga saham dan pergerakannya merupakan faktor penting dalam investasi di pasar modal. Para investor dituntut agar dapat memutuskan membeli, menjual atau mempertahankan saham tertentu dengan cepat. Kesalahan dalam pengambilan keputusan dapat berdampak besar bagi keberlangsungan hidup para pemilik saham.

Berbagai macam penelitian telah dilakukan untuk memprediksi pergerakan harga saham. Penelitian awal mengenai prediksi harga saham melahirkan sebuah model statistika *General Autoregressive Conditional Heteroskedasticity (GARCH)*[1] dan *Autoregressive Integrated Moving Average (ARIMA)* [2]. Sayangnya model tersebut hanya didasarkan pada asumsi bahwa data yang diolah adalah linier. Pada kenyataannya, data pasar saham biasanya memiliki dimensi yang kompleks dan penuh dengan *noise*. Karenanya, seringkali terjadi ketidakkonsistenan antara hasil prediksi dengan kenyataan yang ada. Untuk meminimalisir hal tersebut, penelitian kemudian dilanjutkan kearah pendekatan nonlinier menggunakan konsep kecerdasan komputasi diantaranya Fuzzy Neural Networks (FNN), Artificial Neural Networks (ANN) dan Genetic Algorithm (GA) [3-5].

Penelitian serupa juga dilakukan oleh Yakup Kara dan rekannya dengan menggunakan data pasar saham Istanbul untuk prediksi arah pergerakan harga saham menggunakan model *Artificial Neural Network(ANN)* dan *Support Vector Machine(SVM)* [6]. Penelitian ini mampu memprediksi pergerakan harga

saham dengan baik apakah harga selanjutnya akan naik atau turun. Meski demikian penelitian ini belum dapat mengeluarkan nilai keluaran untuk prediksi harga saham selanjutnya. Melihat hal ini, penelitian ini dapat dikembangkan kearah pendekatan regresi untuk memprediksi pergerakan harga saham pada data pasar saham Indonesia.

Penelitian-penelitian yang sudah berkembang sampai saat ini pada umumnya menggunakan *Artificial Neural Network* sebagai model prediksi. Penelitian yang ada menunjukkan bahwa ANN memiliki performa akurasi yang lebih baik dibandingkan model statistika terdahulu. Namun, kelemahan dari ANN adalah seringkali terjebak dalam perangkap solusi yang *local minimum*.

Di sisi lain, penelitian terkait SVM sejauh ini memperlihatkan bahwa performa akurasi dari metode ini lebih baik dibandingkan dengan pendekatan non-linier lainnya termasuk *case based reasoning*, *linear discriminant analysis*, *quadratic discriminant analysis* dan *Elman back-propagation neural networks*[7-10]. *Support vector machine* terdiri atas dua jenis yaitu model klasifikasi dan juga regresi yang biasa disebut *support vector regression*. *Support Vector Regression (SVR)* merupakan metode regresi yang mampu mengatasi overfitting serta mampu menunjukkan performa yang bagus. Hal inilah yang menjadi dasar pemikiran dalam penelitian ini untuk mengajukan implementasi algoritma *Support Vector Regression* dalam kasus prediksi harga pasar saham Indonesia.

## PREDIKSI HARGA SAHAM

Harga Saham adalah harga dari suatu saham yang ditentukan pada saat pasar saham sedang berlangsung berdasarkan permintaan dan penawaran pada saham yang dimaksud. Harga saham yang berlaku di pasar modal biasanya ditentukan oleh para pelaku

pasar yang sedang melangsungkan perdagangan sahamnya.

Dalam transaksi saham di Bursa Efek Indonesia (BEI) para investor setidaknya mengenal beberapa macam harga pasar yakni harga pembukaan (*preopening*) dan harga penutupan (*closing*). Perubahan harga saham dapat terjadi dikarenakan adanya tawar-menawar antara penjual saham dan pembeli saham. Proses tawar-menawar ini terjadi terus menerus hingga berakhirnya jam perdagangan saham. Selanjutnya harga yang terbentuk pada akhir jam perdagangan itulah yang disebut sebagai harga penutupan.

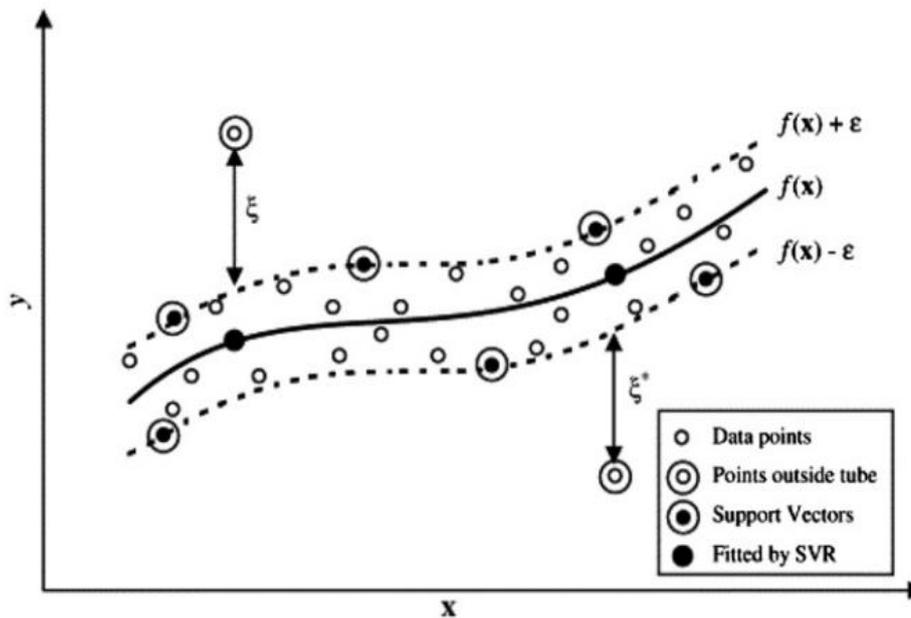
Biasanya harga pembukaan akan sama dengan harga penutupan hari sebelumnya, namun tidak selalu demikian bisa juga terjadi ketidaksamaan. Mengapa ketidaksamaan ini dapat terjadi? Berikut penjelasannya, setelah jam perdagangan ditutup banyak sekali informasi yang dianggap bisa mengubah keputusan investasi investor pada keesokan harinya. Karenanya adalah tidak fair kalau berbagai informasi itu tidak diakomodir oleh pelaku pasar, utamanya faktor-faktor yang datang di luar jam perdagangan. Sebagaimana kita ketahui bisa saja informasi itu tidak seluruhnya diserap oleh investor. Untuk menghindari hal tersebut maka perlu diadakan proses *adjustment* untuk menetapkan harga pembukaan sebuah saham sebelum memulai perdagangan. Proses *adjustment* ini diadakan pada saat *pre-opening* yakni waktu dimana pasar menentukan harga yang paling pantas bagi saham tertentu setelah penutupan sehari sebelumnya itu. Faktor-faktor yang menjadi pertimbangan dalam menentukan harga pada saat *pre-opening* ini antara lain, adalah informasi dalam 12 jam terakhir antara harga penutupan hingga menjelang pasar saham dibuka, lalu jumlah saham (*volume*) pada posisi penawaran jual dan penawaran beli pada saat terakhir saham di perdagangan, serta berbagai kondisi ekonomi dan finansial baik

secara lokal maupun regional yang terjadi pada bursa-bursa di luar negeri.

Waktu perdagangan saham di Bursa Efek Indonesia (BEI) dibagi dalam dua sesi perdagangan yaitu pagi dan siang hari. Perdagangan sesi pagi dimulai jam 09.30 - 12.00 WIB dan sesi kedua pada pukul 13.30 - 16.00 WIB tiap hari Senin hingga Kamis. Sedangkan Jumat pukul 09.30 hingga 11.30 untuk sesi pagi dan pukul 14.00-16.00 WIB. Untuk proses *Pre-opening* dibuka pada pukul 09.10-09.25. Di dalam *pre-opening* ini, sebelum dibukanya pasar saham investor melalui perusahaan efek melakukan penawaran beli maupun jual atas saham yang diinginkan. Lalu melalui mekanisme Jakarta Automatic Trading System (JATS), penawaran jual dan beli saham itu diolah dan akan muncul harga pembukaan yang menjadi patokan bagi order investor pada awal perdagangan.

Prediksi harga saham sangat dibutuhkan oleh para pemilik saham dalam menentukan keputusan yang akan diambil pada transaksi pasar saham. Harga yang biasanya diprediksi adalah harga penutupan di hari berikutnya. Hal ini dimaksudkan agar para pemilik saham atau para investor dapat mengetahui perkiraan harga penutupan selanjutnya sebagai bahan pertimbangan dalam mengambil keputusan apakah akan tetap dipertahankan sahamnya ataukah dijual.

Terdapat dua pendekatan untuk menganalisis pergerakan harga saham, yaitu analisis fundamental dan teknikal. Analisis fundamental memanfaatkan informasi dari sisi faktor ekonomi untuk memprediksi nilai intrinsik saham. Sedangkan analisis teknikal didasarkan pada prinsip Teori Dow menggunakan data historis harga untuk memprediksi pergerakan saham dimasa depan [11]. Menurut Murphy, analisa teknikal adalah studi dari aksi-aksi pasar menggunakan grafik untuk meramalkan tren harga saham masa depan. Aksi pasar diamati melalui dua variabel utama yaitu harga dan volume perdagangan.



Gambar 1. Ilustrasi SVR

## SVR SEBAGAI MODEL PREDIKSI SAHAM

*Support Vector Regression (SVR)* merupakan penerapan *Support Vector Machine (SVM)* dalam kasus regresi. Berbeda dengan klasifikasi, di dalam kasus regresi outputnya berupa sebuah bilangan riil atau kontinu. SVR pertama kali diperkenalkan oleh Drucker [12]. SVR merupakan metode regresi yang mampu mengatasi *overfitting* serta mampu menunjukkan performansi yang bagus [13].

Semisal diberikan sejumlah data latih (training)  $\{(x_1, y_1), \dots, (x_\ell, y_\ell)\} \subset X$  dimana  $X$  menyatakan ruang input maka tujuan utama dari SVR adalah menemukan sebuah fungsi regresi  $y = f(x)$  yang memiliki deviasi yang besar pada rentang sedemikian hingga mampu memprediksi nilai aktual dari  $y_i$  untuk semua data latih. Selain daripada deviasi yang besar pada rentang  $\epsilon$ , secara bersamaan SVR juga mencari persamaan regresi yang sedatar mungkin. Maksudnya adalah fungsi regresi yang mendekati dengan nilai sebenarnya. Algoritma SVR mencoba untuk

menempatkan sebuah tabung disekitar data seperti yang terlihat pada **Error! Reference source not found.**  $\epsilon$  adalah sebuah parameter yang mewakili radius tabung disekitar fungsi regresi. Wilayah region yang dikelilingi oleh tabung disebut *zone- $\epsilon$ -insensitive*. Pendek kata, SVR akan mentolerir adanya kesalahan (error) selama kesalahan tersebut kurang dari  $\epsilon$ , sebaliknya jika kesalahan terjadi melebihi nilai  $\epsilon$  maka akan dikenakan penalti. Di dalam SVR, fungsi regresi dinyatakan dalam bentuk persamaan umum sebagai berikut:

$$f(x) = (w \cdot x) + b, \quad w \in X, b \in R \quad (1)$$

dimana  $(\cdot)$  adalah operator dot product dalam  $X$ . Yang dimaksud sebagai kedataran fungsi dalam persamaan (1) adalah mencari nilai  $w$  seminimal mungkin. Salah satu caranya adalah dengan meminimalkan bentuk *Euclidian* seperti  $\|w\|^2$ . Secara *metematis* dapat dituliskan kedalam permasalahan *convex optimization* sebagai berikut:

$$\text{Minimize } \frac{1}{2} \|w^2\| \quad (2)$$

$$\text{Subject to } \begin{cases} y_i - (w \cdot x_i) - b \leq \varepsilon \\ (w \cdot x_i) - y_i + b \leq \varepsilon \end{cases} \quad (3)$$

Persamaan tersebut dapat digunakan dengan asumsi bahwa fungsi regresi  $f(x)$  dapat mengaproksimasi semua titik  $(x_i, y_i)$  dengan nilai presisi berada dalam zona  $\varepsilon$ -insensitive. Dalam kasus ini, diasumsikan bahwa semua titik ada dalam rentang  $(x) \pm \varepsilon$ , kondisi yang demikian disebut sebagai permasalahan convex optimization yang feasible. Sedangkan pada kenyataannya tidak selalu demikian, ada kemungkinan beberapa titik yang mungkin berada diluar area  $f(x) \pm \varepsilon$ . Untuk mengatasi permasalahan tersebut, dapat ditambahkan sebuah variabel slack yaitu  $\xi, \xi^*$  sebagai batasan baru (*infeasible constraint*) dalam problem optimasi.

$$\begin{aligned} &\text{Min } \frac{1}{2} \|w^2\| + C \sum_{i=1}^{\ell} (\xi_i + \xi_i^*) \\ &\text{Subject to} \\ &\begin{cases} y_i - \langle w \cdot x_i \rangle - b \leq \varepsilon + \xi_i \\ \langle w \cdot x_i \rangle + b - y_i \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0 \end{cases} \end{aligned} \quad (4)$$

Konstanta  $C > 0$  menyatakan harga tawar menawar (*trade off*) antara kedataran fungsi dan batas toleransi kesalahan (*error*). Setiap kesalahan yang nilainya lebih besar dari  $\varepsilon$  akan dikenakan penalti sebesar  $C$ . Jika nilai  $C$  terlalu besar, itu artinya model SVR terlalu ketat akan toleransi error sehingga setiap kesalahan yang nilainya melebihi  $\varepsilon$  akan dikenakan penalti yang besar. Sebaliknya, jika nilai  $C$  terlalu kecil, maka rentang toleransi kesalahan ( $\varepsilon$ ) akan bernilai besar sehingga terlalu beresiko pada hasil prediksi model. Dikatakan terlalu beresiko karena hasil prediksi yang menyimpang dengan harga yang sebenarnya selama masih berada pada zona  $\varepsilon$ -insensitive tidak dianggap sebagai error (diasumsikan zona  $\varepsilon$ -insensitive memiliki rentang yang sangat lebar). Hal ini berkaitan dengan yang disebut sebagai fungsi kerugian  $\varepsilon$ -insensitive yang diusulkan oleh Vapnik [14] sebagai berikut:

$$|\xi|_{\varepsilon} := \begin{cases} 0 & , \text{if } |\xi| \leq \varepsilon \\ |\xi| - \varepsilon & , \text{otherwise} \end{cases} \quad (5)$$

Untuk itulah diperlukan suatu optimasi untuk memilih nilai  $C$  yang tepat sehingga hasil prediksi SVR menjadi lebih akurat.

Permasalahan convex optimization pada persamaan (4) dapat diselesaikan dengan fungsi Lagrange menjadi bentuk berikut:

$$\begin{aligned} L = &\frac{1}{2} \|w\|^2 + C \sum_{i=1}^{\ell} (\xi_i + \xi_i^*) \\ &- \sum_{i=1}^{\ell} \alpha_i (\varepsilon + \xi_i - y_i + \langle w, x_i \rangle + b) \\ &- \sum_{i=1}^{\ell} \alpha_i^* (\varepsilon + \xi_i^* + y_i - \langle w, x_i \rangle - b) \\ &- \sum_{i=1}^{\ell} (\eta_i \xi_i + \eta_i^* \xi_i^*) \end{aligned} \quad (6)$$

dimana  $L$  merupakan Lagrangian dan  $\eta_i, \eta_i^*, \alpha_i$  dan  $\alpha_i^*$  adalah *Lagrange multipliers*. Oleh karena itu, variabel-variabel dual pada persamaan (6) harus memenuhi batasan berikut:

$$\eta_i, \eta_i^*, \alpha_i, \alpha_i^* \geq 0 \quad (7)$$

Selain itu pula, juga memenuhi kondisi turunan partial  $L$  terhadap variabel-variabel primal  $(w, b, \xi_i, \xi_i^*)$  sebagai berikut:

$$\partial_b L = \sum_{i=1}^{\ell} (\alpha_i^* - \alpha_i) = 0 \quad (8)$$

$$\partial_w L = w - \sum_{i=1}^{\ell} (\alpha_i - \alpha_i^*) x_i = 0 \quad (9)$$

$$\partial_{\xi_i^{(*)}} L = C - \alpha_i^{(*)} - \eta_i^{(*)} = 0 \quad (10)$$

Dengan melakukan substitusi persamaan (8), (9), dan (10) kedalam persamaan (6) membangun bentuk permasalahan optimasi dual sebagai berikut:

$$\text{Max} \begin{cases} -\frac{1}{2} \sum_{i,j=1}^{\ell} (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*) \langle x_i, x_j \rangle \\ -\varepsilon \sum_{i=1}^{\ell} (\alpha_i + \alpha_i^*) + \sum_{i=1}^{\ell} y_i (\alpha_i - \alpha_i^*) \end{cases}$$

Subject to  
 $\sum_{i=1}^{\ell} (\alpha_i - \alpha_i^*) = 0, \alpha_i, \alpha_i^* [0, C]$  (11)

Dengan menurunkan persamaan (11) telah mengeliminasi variabel dual  $\eta_i, \eta_i^*$  melalui formula  $\eta_i^{(*)} = C - \alpha_i^{(*)}$  yang didapatkan dari persamaan (12). Dari persamaan (11) didapatkan bahwa:

$$w = \sum_{i=1}^{\ell} (\alpha_i - \alpha_i^*) x_i$$

sehingga:

$$\begin{aligned} f(x) &= (w \cdot x) + b \\ f(x) &= \sum_{i=1}^{\ell} (\alpha_i - \alpha_i^*) \langle x_i, x \rangle + b \end{aligned} \quad (12)$$

dimana  $x_i$  adalah Support Vector (SV) yakni titik-titik yang tepat berada pada  $f(x) \pm \varepsilon$ . SV ini merupakan poin data yang paling informatif yang memampatkan isi informasi dari sekumpulan data training, sehingga SV inilah yang dapat mewakili fungsi SVR secara keseluruhan. Nilai bias  $b$  dihitung dengan menggunakan prinsip Karush-Lhun Tucker (KKT) [15] sebagai berikut:

$$\begin{aligned} b &= y_i - \langle w^* - x_i \rangle - \varepsilon \\ b^* &= y_i - \langle w^* - x_i \rangle + \varepsilon, \leq C \end{aligned} \quad (13)$$

dimana  $0 \leq \alpha_i, \alpha_i^* \leq C$

Jika dimensi data berukuran besar, permasalahan kompleksitas komputasional dapat diatasi dengan menerapkan fungsi kernel pada persamaan (12) menggantikan dot product dari vektor input sebagai berikut:

$$f(x, \alpha_i, \alpha_i^*) = \sum_{i=1}^N (\alpha_i - \alpha_i^*) K(x_i, x) + b \quad (14)$$

dimana  $K(x_i, x) = (\phi(x_i)\phi(x))$  adalah fungsi kernel. Ada berbagai pilihan kernel yang dapat digunakan seperti fungsi linier, polinomial, gaussian radial basis (RBF), spline dan B spline.

## HASIL UJI COBA

Data yang digunakan dalam penelitian ini adalah data sepuluh saham Indonesia yang terdaftar dalam LQ45. Kesepuluh saham ini dipilih karena saham ini merupakan saham yang sehat (tidak terindikasi adanya praktik permainan saham). Adapun saham-saham tersebut seperti yang terdapat pada Tabel 1. Dari kesepuluh perusahaan tersebut kemudian diambil data histori harga sahamnya selama 3 tahun (Januari 2010 – Juli 2012). Data histori harga saham adalah data pergerakan dari hari ke hari dari suatu saham. Setelah mendapatkan data histori harga saham, selanjutnya adalah melakukan perhitungan indikator analisa teknikal. Hasil perhitungan inilah yang kemudian menjadi data masukan pada penelitian ini. Indikator analisa teknikal yang digunakan sebagai fitur untuk masing-masing saham dalam penelitian ini didapatkan dari [6]. Ada 7 indikator analisis teknikal seperti yang tertera pada Tabel 2.

Uji coba yang dilakukan bertujuan untuk melihat kinerja dari algoritma Support Vector Regression dalam memprediksi harga saham di masa mendatang. Untuk mengevaluasi performa dari model prediksi SVR pada penelitian ini digunakan nilai *root mean squared error* (RMSE) dengan menggunakan persamaan berikut:

$$RMSE = \sqrt{\frac{\sum_{i=1}^l (a_i - p_i)^2}{l}} \quad (15)$$

Selain menggunakan nilai RMSE juga digunakan perhitungan NRMSE (*Normalized Root Mean Square Error*)

untuk mengetahui RMSE yang didapatkan termasuk dalam kesalahan tingkat tinggi atau rendah. Semakin kecil nilai RMSE dan NRMSE maka semakin kecil pula kesalahan prediksi (model prediksi semakin akurat). Adapun Tabel 2. Daftar fitur analisa teknikal saham

formula dari NRMSE adalah sebagai berikut:

$$NRMSE = \frac{RMSE}{Max_{Ak} - Min_{Ak}} \quad (16)$$

No	Nama Indikator	Formula
1	Simple 10-day Moving Average (MA10)	$\frac{C_t + C_{t-1} + \dots + C_{t-10}}{10}$
2	Weighted 10-day Moving Average (WMA10)	$\frac{(n) \times C_t + (n-1)C_{t-1} + \dots + C_{t-10}}{(n + (n-1) + \dots + 1)}$
3	Momentum	$C_t - C_{t-n}$
4	Relative Strength Index (RSI)	$100 - \frac{100}{1 + (\sum_{i=0}^{n-1} Up_{t-i}/n) / (\sum_{i=0}^{n-1} Dw_{t-i}/n)}$
5	Moving average convergence Divergence (MACD)	$EMA(12) - EMA(26)$
6	BIAS10	$C_t [(C_t - MA10)/MA10] \times 100$
7	Psychological line for 10 days (PSY10)	$(Up_{10}/10) \times 100$

Keterangan:

$C_t$  adalah harga penutupan saham (closing price) pada hari  $t$ ,  $EMA$  adalah nilai exponential moving average,  $EMA(k)_t = EMA(k)_{t-1} + \alpha \times (C_t - EMA(k)_{t-1})$ ,  $\alpha$  adalah nilai smoothing factor yaitu  $\frac{2}{k+1}$ ,  $k$  adalah periode waktu dari  $k$ -hari,  $Up_t$  adalah upward price change,  $Dw_t$  adalah downward price change pada waktu  $t$ .

dimana RMSE adalah Root Mean Square Error yang didapatkan,  $Max_{Ak}$  adalah nilai tertinggi yang ada pada data aktual, dan  $Min_{Ak}$  adalah nilai terendah yang ada pada data aktual.

Data saham yang digunakan dalam penelitian ini berjumlah 638 hari terhitung dari 4 Januari 2010-30 Juli 2012. Data dari hari pertama hingga hari ke 616 digunakan sebagai data latih. Sedangkan 20 hari sisanya (hari ke 617 hingga 638) tidak digunakan dalam proses, namun digunakan sebagai data aktual untuk membandingkan hasil prediksi yang didapat dari model SVR-ABC dengan data aktual yang ada.

Prediksi pergerakan harga saham dilakukan secara bertahap dari hari ke hari mulai dari hari ke 617 hingga hari ke 638. Untuk dapat memprediksi hari ke 617, maka dilakukan pembentukan data menggunakan perhitungan fitur sesuai yang tertera pada Tabel 2. Kemudian data yang terbentuk ini dimasukkan ke dalam model prediksi untuk diprediksi harga penutupan selanjutnya. Setelah mendapatkan prediksi harga penutupan selanjutnya di hari 617, maka harga tersebut yang digunakan untuk

menghitung fitur untuk memprediksi harga penutupan di hari ke 618. Proses ini berlanjut terus menerus hingga hari terakhir (hari ke 638). Setelah didapatkan hasil prediksi selama 20 hari, maka selanjutnya adalah melakukan evaluasi terhadap hasil prediksi yang telah didapatkan.

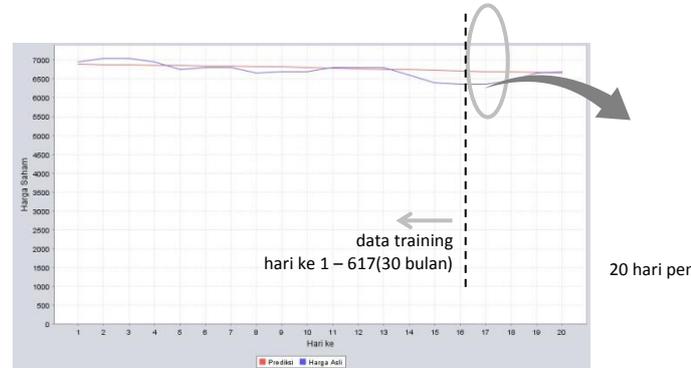
Uji coba pada penelitian ini dilakukan dengan menggunakan kernel linear untuk parameter SVRnya. Adapun parameter dari support vector regression yang di uji coba dalam penelitian ini meliputi nilai  $\epsilon$  dan  $C$ . Untuk nilai  $\epsilon$  adalah  $2^{-3}$ ,  $2^{-5}$ ,  $2^{-7}$ ,  $2^{-9}$  dan nilai  $C$  sebesar  $2^{-1}$ ,  $2^1$ ,  $2^3$ ,  $2^5$ ,  $2^7$ . Kombinasi parameter-parameter inilah yang akan memperlihatkan performa yang berbeda satu sama lain. Uji coba dilakukan terhadap sepuluh perusahaan yang terdaapt pada Tabel 1 dan satu saham index gabungan yaitu saham LQ45.

Dari serangkaian uji coba yang dilakukan, didapatkan bahwa performa terbaik yang diperoleh tiap-tiap perusahaan dapat dilihat pada Gambar 2, Gambar 3 dan Gambar 4. Grafik bagian kiri adalah plot dari data training selama 30 bulan dan hasil prediksi 20 hari setelahnya sedangkan grafik bagian

kanan adalah plot dari data aktual dibandingkan dengan data hasil prediksi. Garis berwarna merah merupakan hasil prediksi dan garis berwarna biru adalah data aktualnya. Jika garis berwarna merah (hasil prediksi) dapat mengikuti pergerakan dari garis berwarna biru

dengan jarak yang relatif kecil maka menunjukkan bahwa hasil prediksi semakin akurat.

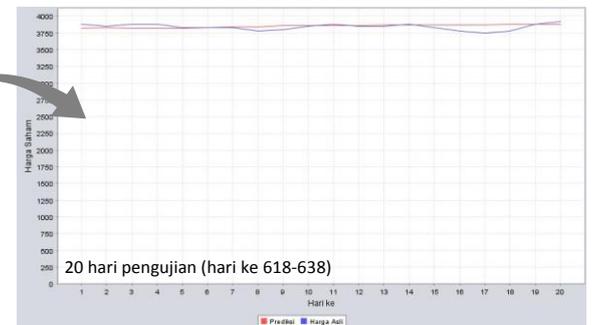
Pada Gambar 2,3 dan 4 terlihat bahwa mayoritas hasil prediksi dapat



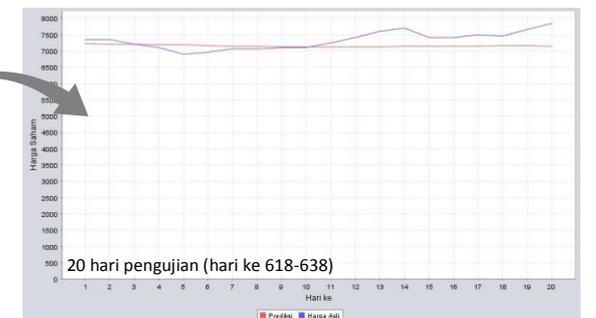
(a) ASII



(b) BBCA

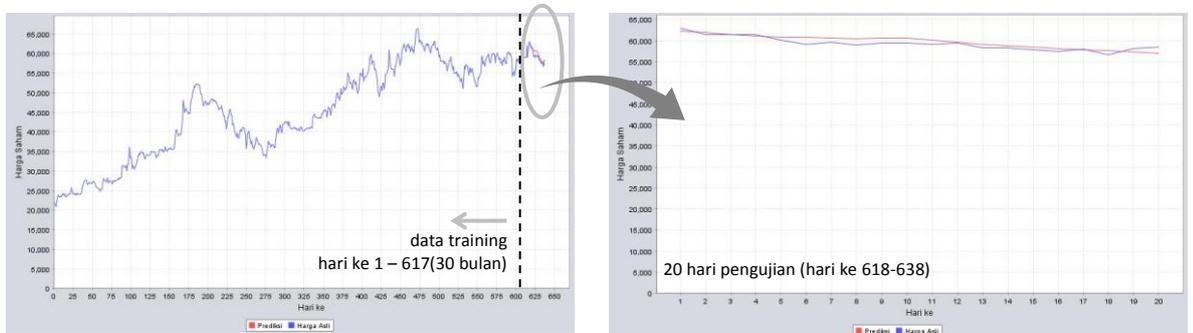


(c) BBNI



(d) BMRI

Gambar 2. Hasil uji coba prediksi pergerakan harga saham perusahaan (a) Astra International (b) Bank BCA (c) Bank BNI (d) Bank Mandiri,Tbk. Grafik bagian kiri adalah plot data latih selama 30 bulan dan hasil prediksi 20 hari kedepan. Garis biru menunjukkan harga aktual, Garis merah menunjukkan harga prediksi.



(a) GGRM



(b) JSMR



(c) KLBf



(d) PGAS

Gambar 3. Hasil uji coba prediksi pergerakan harga saham perusahaan (a) Gudang Garam (b) Jasa Marga (c) Kalbe Farma (d) Perusahaan Gas Negara. Grafik bagian kiri adalah plot data latih selama 30 bulan dan hasil prediksi 20 hari kedepan. Garis biru menunjukkan harga aktual, Garis merah menunjukkan harga prediksi.

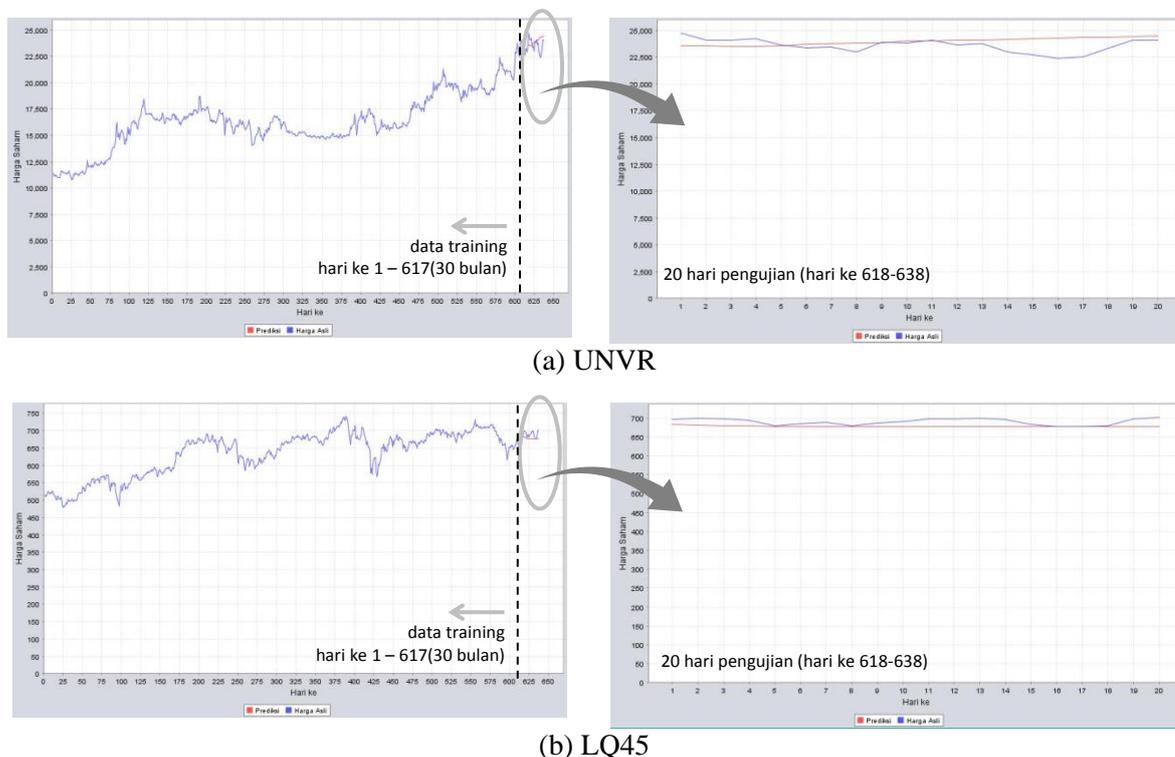
Tabel 3. Hasil uji coba Support Vector Regression menggunakan data saham ASII, BBCA, BBNI, BMRI, GGRM & JSMR

$\epsilon$	C	ASII		BBCA		BBNI		BMRI		GGRM		JSMR	
		RMSE	NRSME	RMSE	NRMSE	RMSE	NRMSE	RMSE	NRMSE	RMSE	NRMSE	RMSE	NRMSE
$2^{-5}$	$2^{-1}$	171.17	0.25	631.82	0.90	125.03	0.72	428.01	0.45	2241.66	0.35	452.89	1.13
	$2^1$	196.71	0.28	624.90	0.89	134.78	0.77	375.82	0.39	2056.06	0.32	380.47	0.95
	$2^3$	193.17	0.28	644.54	0.92	85.76	0.49	373.12	0.39	927.06	0.15	158.65	0.39
	$2^5$	183.51	0.26	680.20	0.97	100.97	0.58	330.02	0.35	1169.08	0.18	142.74	0.36
	$2^7$	190.42	0.27	625.23	0.89	68.32	0.39	324.77	0.34	<b>892.01</b>	<b>0.14</b>	<b>132.66</b>	<b>0.33</b>
$2^{-7}$	$2^{-1}$	202.96	0.29	547.67	0.78	59.63	0.34	355.04	0.37	3684.05	0.58	233.78	0.58
	$2^1$	181.98	0.26	542.00	0.77	82.99	0.47	350.12	0.37	3714.81	0.59	220.55	0.55
	$2^3$	<b>170.59</b>	<b>0.24</b>	531.77	0.76	59.76	0.34	320.72	0.34	3457.11	0.54	177.77	0.44
	$2^5$	211.82	0.30	478.65	0.68	<b>53.23</b>	<b>0.30</b>	374.96	0.39	3291.56	0.52	184.85	0.46
	$2^7$	212.72	0.30	441.87	0.63	69.33	0.39	378.91	0.39	3616.81	0.57	161.92	0.41
$2^{-9}$	$2^{-1}$	252.31	0.36	555.80	0.79	63.70	0.36	320.83	0.34	3473.21	0.55	236.16	0.59
	$2^1$	274.63	0.39	<b>438.26</b>	<b>0.62</b>	54.38	0.31	309.17	0.33	3911.97	0.62	206.60	0.52
	$2^3$	251.31	0.36	560.23	0.80	55.47	0.32	<b>308.16</b>	<b>0.32</b>	4021.88	0.63	205.26	0.51
	$2^5$	261.90	0.37	517.62	0.74	86.81	0.49	358.55	0.38	4299.39	0.68	164.86	0.41
	$2^7$	267.88	0.38	512.09	0.73	55.00	0.31	316.72	0.33	4036.33	0.64	156.18	0.39
<b>MIN</b>		<b>170.59</b>	<b>0.24</b>	<b>438.26</b>	<b>0.62</b>	<b>53.23</b>	<b>0.30</b>	<b>308.16</b>	<b>0.32</b>	<b>892.01</b>	<b>0.14</b>	<b>132.66</b>	<b>0.33</b>

Tabel 4. Hasil uji coba Support Vector Regression menggunakan data saham ASII, BBCA, BBNI, BMRI, GGRM & JSMR

$\epsilon$	C	KLBF		PGAS		SMGR		UNVR		LQ45	
		RMSE	NRMSE	RMSE	NRMSE	RMSE	NRMSE	RMSE	NRMSE	RMSE	NRMSE
$2^{-5}$	$2^{-1}$	21.54	0.54	663.87	2.21	566.03	0.31	1,998.44	0.85	40.80	1.70
	$2^1$	20.64	0.52	681.20	2.27	584.93	0.33	2,087.86	0.89	36.14	1.51
	$2^3$	23.06	0.58	680.30	2.27	588.58	0.33	1,734.33	0.74	32.30	1.35
	$2^5$	16.96	0.42	689.49	2.30	535.74	0.30	1,681.61	0.72	27.06	1.13
	$2^7$	16.99	0.42	<b>663.07</b>	<b>2.21</b>	539.63	0.30	<b>1,660.27</b>	<b>0.71</b>	20.12	0.84
$2^{-7}$	$2^{-1}$	12.81	0.32	816.52	2.72	547.00	0.30	3,038.91	1.29	20.21	0.84
	$2^1$	<b>11.92</b>	<b>0.30</b>	817.10	2.72	507.33	0.28	3,487.37	1.48	15.92	0.66
	$2^3$	13.19	0.33	834.31	2.78	504.03	0.28	3,067.16	1.31	17.47	0.73
	$2^5$	13.65	0.34	825.90	2.75	466.25	0.26	3,144.07	1.34	20.41	0.85
	$2^7$	13.40	0.33	833.13	2.78	469.06	0.26	3,465.71	1.47	20.41	0.85
$2^{-9}$	$2^{-1}$	13.09	0.33	769.66	2.57	533.64	0.30	2,786.31	1.19	16.81	0.70
	$2^1$	13.13	0.33	776.06	2.59	515.54	0.29	2,828.60	1.20	<b>14.69</b>	<b>0.61</b>

$2^3$	13.31	0.33	803.24	2.68	557.21	0.31	2,727.08	1.16	14.89	0.62
$2^5$	13.65	0.34	808.65	2.70	584.51	0.32	2,655.65	1.13	15.57	0.65
$2^7$	13.07	0.33	801.02	2.67	537.55	0.30	2,778.75	1.18	16.08	0.67
<b>MIN</b>	<b>11.92</b>	<b>0.30</b>	<b>663.07</b>	<b>2.21</b>	<b>466.25</b>	<b>0.26</b>	<b>1,660.27</b>	<b>0.71</b>	<b>14.69</b>	<b>0.61</b>



Gambar 4. Hasil uji coba prediksi pergerakan harga saham perusahaan (a) Unilever (b) indeks LQ45. Grafik bagian kiri adalah plot data latih selama 30 bulan dan hasil prediksi 20 hari kedepan. Garis biru menunjukkan harga aktual, Garis merah menunjukkan harga prediksi.

mengikuti pergerakan dari plot data aktual. Performa hasil prediksi dari model SVR ditentukan oleh nilai  $\epsilon$  dan C. Adapun hasil prediksi terbaik dari masing-masing saham perusahaan dan indeks LQ45 ditunjukkan pada Tabel 3 dan 4. Terlihat bahwa saham ASII, BBNI, BMRI, GGRM, JSMR, KLBF dan SMGR menunjukkan nilai NRMSE yang relatif kecil yakni berkisar antara 0.14-0.32. Hal ini mengindikasikan bahwa prediksi 20 hari kedepan makin mendekati dengan data aktual. Berbeda halnya dengan saham BBKA, PGAS, UNVR dan indeks LQ45 nilai NRMSE

terbaik yang didapatkan cukup besar berkisar antara 0.61-2.21. Hal ini dapat disebabkan karena parameter kernel SVR yang digunakan adalah linear. Untuk itu perlu dilakukan penelitian lebih lanjut dengan menggunakan parameter kernel lainnya, karena ada kemungkinan distribusi data histori saham adalah fungsi nonlinear. Tabel 3 dan 4 memperlihatkan bahwa NRMSE terbaik diperoleh saat memprediksi saham Gudang Garam, Tbk (GGRM) sebesar 0.14 dengan perolehan RMSE sebesar 892.01 menggunakan parameter kernel linear,  $\epsilon$

=  $2^{-5}$  dan  $C=2^7$  sedangkan NRMSE terendah didapatkan ketika memprediksi data saham Perusahaan Gas Negara, Tbk (PGAS) sebesar 2.21 dengan nilai RMSE sebesar 663.07.

## SIMPULAN DAN SARAN

Prediksi pergerakan harga saham merupakan hal yang penting untuk investor karena dapat digunakan sebagai bahan pertimbangan pengambilan keputusan transaksi dalam pasar saham. Pada penelitian ini diusulkan prediksi pergerakan harga saham secara komputasional menggunakan metode Support Vector Regression terhadap beberapa data saham perusahaan yang tergabung dalam indeks LQ45. Dari serangkaian uji coba yang dilakukan, dapat disimpulkan bahwa metode SVR dapat memprediksi pergerakan harga saham dengan cukup baik. Hal ini terlihat dari nilai NRMSE yang didapatkan terbaik adalah 0.14. Akan tetapi ada beberapa saham yang diprediksi dengan model ini tidak menunjukkan performa yang cukup baik. Untuk itu model prediksi ini dapat dikembangkan lagi dengan menggunakan parameter kernel lainnya yakni pendekatan kernel nonlinear seperti misalnya RBF. Selain daripada itu, dari hasil uji coba yang dilakukan didapatkan bahwa pemilihan parameter SVR berpengaruh besar terhadap hasil prediksi. Oleh karena itu, penelitian ini dapat dikembangkan lagi dengan ikut sertakan optimasi pemilihan parameter dari Support Vector Regression.

## DAFTAR PUSTAKA

- [1] Franses, P., & Ghijssels, H. (1999). Additive outliers, GARCH and forecasting. *International Journal of Forecasting*, 15(1), 1–9.
- [2] Box, G., & Jenkins, G. (1994). *Time Series Analysis: Forecasting and Control* (Vol. III). Englewood Cliffs: Prentice Hall.
- [3] Kim, K., & Han, I. (2000). Genetic algorithms approach to feature discretization in artificial neural networks for the prediction of stock price index. *Expert Systems with Applications*, 19(2), 125–132.
- [4] Armano, G., Marchesi, M., & Murru, A. (2005). A hybrid genetic-neural architecture for stock indexes forecasting. *Information Sciences*, 170(1):3–33.
- [5] Erkam, G., Gulgun, K., & Tugrul, U. (2011). Using artificial neural network models in stock market index prediction. *Expert Systems with Applications Elsevier*, 10389–10397.
- [6] Yakup, K., Melek, A., & Ömer, K. (2011). Predicting direction of stock price index movement using artificial neural networks and support vector machines: The sample of the Istanbul Stock Exchange. *Expert Systems with Applications ELSEVIER*, 5311–5319.
- [7] Sapankevych, N., & Sankar, R. (2009). Time series prediction using support vector machines: a survey. *IEEE Computational Intelligence Magazine*, 24–38.
- [8] Min, J., & Lee, Y. (2005). Bankruptcy prediction using support vector machine with optimal choice of kernel function parameters. *Expert Systems with Application*, 603–614.
- [9] Kim, K. (2003). Financial time series forecasting using support vector machines. *Neuro computing*, 307–309.
- [10] Huang, W., Nakamori, Y., & Wang, S. (2005). Forecasting stockmarket movement direction with support vector machine. *Computers and Operations Research*, 32(2513–2522), 2513–2522.

- [11] Murphy, J. (1999). Technical analysis of the financial markets. New York Institute of Finance.
- [12] Drucker, H., Burges, C., Smola, A., Kaufmann, L., & Vapnik, V. (1996). Support Vector Regression Machines. *Advances in Neural Information Processing Systems*, 155-161.
- [13] Smola, A., & Scholkopf, B. (2004). A tutorial on support vector regression\*. *Statistics and Computing*, 199–222.
- [14] Vapnik, V. (1995). *The Nature of Statistical Learning Theory*. New York: Springer-Verlag.
- [15] Karush, W. (1939). Minima of functions of several variables with inequalities as side constraints.