

PREDIKSI INDEKS HARGA SAHAM DENGAN METODE SUPPORT VECTOR REGRESSION DUA TAHAP

PREDICTION STOCK INDEX USING SUPPORT VECTOR REGRESSION TWO STAGE

¹Arfian Nurdiansyah, ²Rian Febrian Umbara, S.Si, M.Si, ³Annisa Aditsania M.Si
^{1,2,3}Ilmu Komputasi Fakultas Informatika Universitas Telkom, Bandung
¹arfian.nurdiansyah@gmail.com, ²rianum123@gmail.com, ³annisaaditsania@gmail.com

ABSTRACT

In this research, the stock price index prediction with two-stages Support Vector Regression (SVR) method will be discussed. This study uses Compositon Stock Price Index (CSPI) index of Indonesia's stock market. These experiments are based on historical daily data in 5 years. Prediction done using one-stage SVR and two-stages SVR to predict closing price of the period ahead (t+1), two periods ahead (t+2), three periods ahead (t+3), four period ahead (t+4), and five periods ahead (t+5). MAPE for (t+1) scenario is 16.4214%, MAPE for (t+2) scenario is 15.026%, MAPE for (t+3) scenario is 15.453%, MAPE for (t+4) scenario is 15.605%, and MAPE for (t+5) scenario is 8.8449%. At the same time using the one-stage Support Vector Regression (SVR) it is obtained testing MAPE of five scenarios, there are MAPE for (t+1) scenario is 15.316%, MAPE for (t+2) scenario is 15.026%, MAPE for (t+3) scenario is 15.453%, MAPE for (t+4) scenario is 15.605%, and MAPE for (t+5) scenario is 15.7007%. So it can be concluded that the two-stages Support Vector Regression (SVR) method has better accuracy in every scenario except in the (t+2) scenario.

Keywords: Prediction of stock, Compositon Stock Price Index (CSPI), Support Vector Regression (SVR)

ABSTRAK

Dalam penelitian ini akan dibahas prediksi indeks harga saham dengan metode *Support Vector Regression* (SVR) dua tahap. Penelitian ini menggunakan indeks saham dari pasar saham Indonesia yaitu Indeks Harga Saham Gabungan (IHSG). Percobaan ini didasarkan pada data historis harian selama 5 tahun. Prediksi dilakukan dengan menggunakan SVR satu tahap dan SVR dua tahap untuk meramal harga penutupan saham pada periode kedepan (t+1) (dua periode kedepan (t+2) dan tiga periode kedepan (t+3), empat periode kedepan (t+4) dan lima periode kedepan (t+5)). Metode *Support Vector Regression* (SVR) dua tahap didapatkan *Mean Absolute Percentage Error* (MAPE) testing dari 5 skenario yang diuji yaitu (t+1) adalah 16,4214%, (t+2) adalah 15,026%, (t+3) adalah 15,453%, (t+4) adalah 15,605% dan (t+5) adalah 8,8449%. Sedangkan dengan menggunakan *Support Vector Regression* (SVR) satu tahap didapatkan MAPE testing dari 5 skenario yang diuji yaitu MAPE untuk skenario (t+1) adalah 15,316%, MAPE untuk skenario (t+2) adalah 15,026%, MAPE untuk skenario (t+3) adalah 15,453%, MAPE untuk skenario (t+4) adalah 15,605%, dan MAPE untuk skenario (t+5) adalah 15,7007%. Jadi dapat disimpulkan bahwa metode *Support Vector Regression* (SVR) dua tahap memiliki akurasi yang lebih baik disetiap skenario kecuali di skenario (t+2).

Kata kunci : prediksi saham, Indeks Harga Saham Gabungan (IHSG), Support Vector Regression (SVR)

1. Pendahuluan

1.1 Latar Belakang

Saham adalah surat berharga dengan satuan nilai sebagai tanda kepemilikan suatu perusahaan. Harga saham di pasaran berjalan secara acak. Oleh karena itu untuk mengetahui harga saham di masa yang akan datang perlu adanya prediksi. Prediksi harga saham akan sangat bermanfaat bagi investor untuk dapat melihat bagaimana prospek investasi saham sebuah perusahaan di masa yang akan datang

Penelitian yang luas telah menghasilkan berbagai aplikasi prediksi yang menggunakan Jaringan Syaraf Tiruan (JST), logika fuzzy, *Genetic Algorithm* (GA) dan teknik lainnya. Jaringan Syaraf Tiruan (JST) dan *Support Vector Regression* (SVR) adalah dua algoritma *machine learning* yang paling banyak digunakan untuk memprediksi harga saham dan nilai-nilai indeks pasar saham.

Ou dan Wang pada tahun 2009 [1] menggunakan total sepuluh teknik data mining untuk memprediksi pergerakan harga indeks Hang Seng pasar saham Hong Kong. Pendekatan tersebut termasuk *Linear Discriminant Analysis* (LDA), *Quadratic Discriminant Analysis* (QDA), *K Nearest neighbor classification*, *Naive Bayes based on*

kernel estimation, Logit model, Tree based classification, jaringan saraf, klasifikasi Bayesian dengan proses Gaussian, Support Vector Machine (SVM) dan Least Squares Support Vector Machine (LS-SVM). Hasil penelitian menunjukkan bahwa SVM dan LS-SVM menghasilkan kinerja prediktif unggul di antara model-model lain.

Support Vector Machine (SVM) (Vapnik, 1999) telah mendapatkan popularitas dan dianggap sebagai teknik state-of-the-art untuk regresi dan aplikasi klasifikasi. Baru kemudian pada tahun 2013 Kazem, Sharifi, Hussain, Saberi, dan Hussain mengusulkan sebuah model peramalan berdasarkan chaotic mapping, firefly algorithm, dan Support Vector Regression (SVR) untuk memprediksi harga pasar saham. Model SVR-CFA yang baru diperkenalkan dalam studi mereka, dibandingkan dengan SVR-GA (Genetic Algorithm), SVR-CGA (Genetic Algorithm Chaotic), SVR-FA (Algoritma Firefly), ANN dan ANFIS model dan hasilnya menunjukkan bahwa model SVR-CFA tampil lebih baik daripada model lainnya.

Dalam tugas akhir ini, akan dianalisis aplikasi metode Support Vector Regression (SVR) dua tahap agar dapat menghasilkan nilai prediksi harga saham yang mendekati nilai sebenarnya atau dengan kata lain meminimalisasi galat dalam prediksi tersebut.

1.2 Perumusan Masalah

Permasalahan yang akan dibahas dalam penelitian ini adalah sebagai berikut:

1. Bagaimana memprediksi indeks harga saham gabungan menggunakan metode Support Vector Regression (SVR) dua tahap?
2. Bagaimana perbandingan akurasi metode Support Vector Regression (SVR) dua tahap dengan Support Vector Regression (SVR) satu tahap?

1.3 Tujuan dari penelitian adalah :

1. Data yang digunakan adalah data Indeks Harga Saham Gabungan (IHSG) pada tahun 2010-2015.
2. Faktor-faktor makro ekonomi seperti situasi politik, kebijakan perusahaan, tingkat suku bunga dan nilai tukar bank, yang mempengaruhi harga saham tidak diperhitungkan dalam tugas akhir ini.

2. Landasan Teori

2.1 Support Vector Regression (SVR)

Support Vector Regression (SVR) merupakan bagian dari Support Vector Machine (SVM). Support Vector Machine (SVM) dapat dijelaskan secara sederhana sebagai usaha mencari hyperplane (garis pemisah) terbaik yang berfungsi sebagai pemisah dua buah kelas -1 dan +1 pada input space. Margin adalah jarak antara hyperplane tersebut dengan titik terdekat dari masing-masing kelas. Titik yang paling dekat ini disebut sebagai support vector. Usaha untuk

mencari lokasi hyperplane ini merupakan inti dari proses pembelajaran pada SVM.

Data yang tersedia dinotasikan sebagai $x_i \in \mathbb{R}^n$ sedangkan label class masing masing dinotasikan $y_i \in \{-1, +1\}$ untuk $i = 1, 2, \dots, n$, yang mana n adalah banyaknya data. Diasumsikan kedua kelas -1 dan +1 dapat terpisahkan secara sempurna oleh hyperplane yang didefinisikan :

$$w^T x_i + b = 0$$

SVR merupakan penerapan support vector machine (SVM) untuk kasus regresi yang bisa menghasilkan performansi yang lebih baik dan bisa mengatasi masalah overfitting. Overfitting adalah keadaan dimana model yang didapat dari hasil training set terlalu pintar atau dengan kata lain rasio kesalahan sangat kecil saat dibandingkan dengan training set tersebut. Dalam SVR hyperplane yang dicari adalah sebuah fungsi regresi yang output dari fungsi tersebut adalah sebuah bilangan riil. Ide dasar dari SVR dengan menentukan set data yang dibagi menjadi set training dan set testing. Kemudian dari set training tersebut ditentukan suatu fungsi regresi dengan batasan deviasi tertentu sehingga dapat menghasilkan prediksi yang mendekati target aktual

Fungsi regresi akan sempurna apabila batas deviasinya sama dengan 0. Dengan asumsi data masukan $x_i \in \mathbb{R}^n, i = 1, 2, \dots, n$ dan output $y_i \in \mathbb{R}, i = 1, 2, \dots, n$ sehingga dapat dituliskan sebagai berikut :

$$f(x_i) = w^T x_i + b$$

Dengan x_i menunjukkan suatu titik di dalam feature space yang merupakan hasil pemetaan x di dalam input space. Koefisien w dan b berfungsi untuk meminimalkan fungsi resiko.

Pada persamaan di bawah menunjukkan Insensitivity loss function. Loss function adalah fungsi yang menunjukkan hubungan antara galat dengan galat yang dikenai pinalti.

$$|y_i - f(x_i)| = \begin{cases} 0, & |y_i - f(x_i)| \leq \epsilon \\ |y_i - f(x_i)| - \epsilon, & |y_i - f(x_i)| > \epsilon \end{cases}$$

Dengan ini regresi linier diestimasi secara bersamaan dengan meminimalkan $\|w\|$ dan jumlah besarnya Insensitivity loss. Kostanta C menentukan tawar-menawar (trade-off) antar batas deviasi yang masih bisa ditoleransi.

$$Q = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n |y_i - f(x_i)|$$

Suatu fungsi dianggap layak dan baik apabila semua titik ada dalam rentang yang seharusnya, jika terjadi sebaliknya maka fungsi tersebut tidak layak digunakan untuk sebuah prediksi. Untuk mengatasi masalah tersebut

diperlukan variable tambahan yang disebut variable slack. Sehingga fungsi sebelumnya dioptimasi lagi menjadi :

$$\begin{aligned} & 1 \\ & \phi = \frac{1}{2} \|w\|^2 + c \sum_{i=1}^n \xi_i \\ & \xi_i \geq 0 \\ & \phi_i - (w_i^T x_i + b) \leq \xi_i + \epsilon \\ & \xi_i^* \geq 0, \phi_i = 1, 2, \dots \end{aligned}$$

Di sini, ξ_i dan ϕ_i adalah variabel yang saling bebas atau tidak terikat, satu untuk melebihi nilai dari target ϵ dan satu

lagi untuk membuat ϵ berada di bawah target. Konstanta c berfungsi untuk mengontrol aproksimasi error dan nilai

vector $\|w\|$. R berfungsi untuk meminilakan resiko.

Mirip dengan SVM, masalah optimasi diselesaikan

dengan menggunakan teori Lagrangian dan kondisi Karush-Kuhn-Tucker untuk mendapatkan bobot fungsi regresi yang diinginkan. SVR dapat memprediksi nilai indeks harga saham saat penutupan dengan menggunakan bantuan kernel yang di masukkan ke dalam sepuluh indikator sebagai input. Fungsi kernel yang dimaksud adalah

Dimana :

1. ϕ_t = harga penutupan pada saat t
2. ϕ_{t-1} = harga saham terendah pada saat t
3. ϕ_{t-2} = harga saham tertinggi pada saat t
4. ϕ_{t-k} = harga terendah dalam rentang waktu t
5. ϕ_{t+k} = harga tertinggi dalam rentang waktu t
6. ϕ_{t-12} = ϕ_{t-12} - ϕ_{t-26}
7. $\phi_{t-1} = \frac{\phi_t + \phi_{t-1} + \dots + \phi_{t-k} - (k-1)}{k}$
8. $\alpha = \frac{2}{\phi_{t+1}}$
9. ϕ_{t+k} = periode waktu k-hari untuk EMA
10. $\phi_{t+k} = \frac{\phi_t + \phi_{t+1} + \dots + \phi_{t+k}}{k+1}$
11. $\phi_{t+k} = \frac{(\sum_{i=1}^n \phi_{t+i})}{n+1}$
12. $\phi_{t+k} = \frac{(\sum_{i=1}^n \phi_{t+i} - \phi_{t-1})}{n}$
13. UP_t = kenaikan harga saat perubahan waktu t
14. DN_t = penurunan saat perubahan waktu t

2.2 Algoritma PSO (particle swarm optimization)

Algoritma PSO ini awalnya diusulkan oleh James Kennedy(1995) dan Russel Eberhart(1995). PSO (particle swarm optimization) merupakan suatu metode optimasi yang algoritmanya meniru perilaku sosial organisme ini. Perilaku sosial terdiri dari tindakan individu dan pengaruh dari individu-individu lain dalam suatu kelompok. Kata partikel menunjukkan, misalnya, seekor burung dalam kawanan burung. Setiap individu atau partikel berperilaku secara terdistribusi dengan cara menggunakan kecerdasannya (intelligence) sendiri dan juga dipengaruhi perilaku kelompok kolektifnya. Dengan demikian, jika satu partikel atau seekor burung menemukan jalan yang tepat atau pendek menuju ke sumber makanan, sisa kelompok yang lain juga akan dapat segera mengikuti jalan tersebut meskipun lokasi mereka jauh di kelompok tersebut. Metode optimasi yang didasarkan pada swarm intelligence ini disebut algoritma behaviorally inspired. Dalam konteks optimasi multivariabel, kawanan diasumsikan mempunyai ukuran tertentu atau tetap dengan setiap partikel posisi awalnya terletak di suatu lokasi yang acak dalam ruang multidimensi. Setiap partikel diasumsikan memiliki dua karakteristik: posisi dan kecepatan. Setiap partikel bergerak dalam ruang tertentu dan mengingat posisi terbaik yang pernah dilalui atau ditemukan terhadap sumber makanan atau nilai fungsi objektif. Setiap partikel menyampaikan informasi atau posisi bagusnya kepada partikel yang lain dan menyesuaikan posisi dan kecepatan masing-masing berdasarkan informasi yang diterima mengenai posisi yang bagus tersebut. Sebagai contoh, misalnya perilaku burung-burung dalam dalam kawanan burung. Meskipun setiap

No	Indikator	Formula
1	Dimana n-day moving average	$\phi_t = \phi_{t-1} + \dots + \phi_{t-n}$ Konstanta dari fungsi basis radial
2	Exponential moving average	Ada sembilan indikator sebagai masukan untuk SVR yang memungkinkan untuk memprediksi harga penutupan saham. (Jigar Patel, Sahil Shah, Priyank Thakkar, K. Kotecha 2015). Indikatornya yaitu :
4	Stochastic K%	$\frac{\phi_t - \phi_{t-1}}{\phi_t - \phi_{t-1}} \times 100$
5	Stochastic D%	$\frac{\sum_{i=0}^{n-1} \phi_{t-i}}{10} \%$
6	Relative Strength Index	$100 - \frac{100}{1 + (\sum_{i=0}^{n-1} \frac{UP_{t-i}}{\phi_{t-i}}) / (\sum_{i=0}^{n-1} \frac{DN_{t-i}}{\phi_{t-i}})}$
7	Moving average convergence divergence	$\frac{\phi_{t+1} + \frac{2}{\phi_{t+1}} \times (\phi_{t-1} - \phi_{t-2})}{\phi_{t+1} - \phi_{t-1}}$
8	Larry William's R%	$\frac{\phi_t - \phi_{t-1}}{\phi_t - \phi_{t-2}} \times 100$
9	Commodity channel index	$\frac{\phi_t - \phi_{t-1}}{0.015 \phi_t}$

burung mempunyai keterbatasan dalam hal kecerdasan, biasanya ia akan mengikuti kebiasaan (rule) seperti berikut

:

1. Seekor burung tidak berada terlalu dekat dengan burung yang lain.
2. Burung tersebut akan mengarahkan terbangnya ke arah rata-rata keseluruhan burung

3. Akan memposisikan diri dengan rata-rata posisi burung yang lain dengan menjaga sehingga jarak antar burung dalam kawanan itu tidak terlalu jauh

Dengan demikian perilaku kawanan burung akan didasarkan pada kombinasi dari 3 faktor simpel berikut:

1. Kohesi - terbang bersama
2. Separasi - jangan terlalu dekat
3. Penyesuaian(alignment) - mengikuti arah bersama

Jadi PSO dikembangkan dengan berdasarkan pada model berikut:

1. Ketika seekor burung mendekati target atau makanan (atau bisa minimum atau maximum suatu fungsi tujuan) secara cepat mengirim informasi kepada burung-burung yang lain dalam kawanan tertentu.
2. Burung yang lain akan mengikuti arah menuju ke makanan tetapi tidak secara langsung.
3. Ada komponen yang tergantung pada pikiran setiap burung, yaitu memorinya tentang apa yang sudah dilewati pada waktu sebelumnya.

Model ini akan disimulasikan dalam ruang dengan dimensi tertentu dengan sejumlah iterasi sehingga di setiap iterasi, posisi partikel akan semakin mengarah ke target yang dituju (minimasi atau maksimasi fungsi). Ini dilakukan hingga maksimum iterasi dicapai atau bisa juga digunakan kriteria penghentian yang lain.

2.2.1 Implementasi PSO

Misalkan kita mempunyai fungsi berikut

$$\min f(x)$$

dimana $X(B) \leq X \leq X(A)$

dimana $X(B)$ adalah batas bawah dan $X(A)$ adalah batas atas dari X . Prosedur PSO dapat dijabarkan dengan langkah-langkah sebagai berikut [2]:

Asumsikan bahwa ukuran kelompok atau kawanan (jumlah partikel) adalah N . Untuk mengurangi jumlah evaluasi fungsi yang diperlukan untuk menemukan solusi, sebaiknya ukuran N tidak terlalu besar, tetapi juga tidak terlalu kecil, agar ada banyak kemungkinan posisi menuju solusi terbaik atau optimal. Jika terlalu kecil, sedikit kemungkinan menemukan posisi partikel yang baik. Terlalu besar juga akan membuat perhitungan jadi panjang. Biasanya digunakan ukuran kawanan adalah 20 sampai 30 partikel.

Bangkitkan populasi awal X dengan rentang $X(B)$ dan $X(A)$ secara random sehingga didapat X_1, X_2, \dots, X_N . Setelah itu, untuk mudahnya, partikel j dan kecepatannya pada iterasi i dinotasikan sebagai $X(i)_j$ dan $V(i)_j$ Sehingga partikel-partikel awal ini akan menjadi $X_1(0), X_2(0), \dots, X_N(0)$. Vektor $X_j(0)$, ($j = 1, 2, \dots, N$) disebut partikel atau vektor koordinat dari partikel. (seperti kromosom dalam algoritma genetika). Evaluasi nilai fungsi tujuan untuk setiap partikel dan nyatakan dengan

$$f[X_1(0)], f[X_2(0)], \dots, f[X_N(0)]$$

Hitung kecepatan dari semua partikel. Semua partikel bergerak menuju titik optimal dengan suatu kecepatan.

Awalnya semua kecepatan dari partikel diasumsikan sama dengan nol. Set iterasi $i = 1$.

Pada iterasi ke- i , temukan 2 parameter penting untuk setiap partikel j yaitu:

Nilai terbaik sejauh ini dari $X_j(i)$ (koordinat partikel j pada iterasi i) dan nyatakan sebagai $P_{best,j}$, dengan nilai fungsi obyektif paling rendah (kasus minimasi), $f[X_j(i)]$, yang ditemui sebuah partikel j pada semua iterasi sebelumnya. Nilai terbaik untuk semua partikel $X_j(i)$ yang ditemukan sampai iterasi ke- i , G_{best} , dengan nilai fungsi tujuan paling kecil/minimum diantara semua partikel untuk semua iterasi sebelumnya, $f[X_j(i)]$.

Hitung kecepatan partikel j pada iterasi ke i dengan rumus sebagai berikut:

$$V_j(i) = V_j(i-1) + c_1 r_1 [P_{best,j} - x_j(i-1)] + c_2 r_2 [G_{best} - x_j(i-1)], j = 1, 2, \dots, N$$

dimana c_1 dan c_2 masing-masing adalah learning rates untuk kemampuan individu (cognitive) dan pengaruh sosial (group), dan r_1 dan r_2 bilangan random yang berdistribusi uniform dalam interval 0 dan 1. Jadi parameters c_1 dan c_2 menunjukkan bobot dari memory (position) sebuah partikel terhadap memory (posisi) dari kelompok (swarm). Nilai dari c_1 dan c_2 biasanya adalah 2 sehingga perkalian $c_1 r_1$ dan $c_2 r_2$.

2.3 Pasar Modal

Pasar modal (capital market) adalah suatu sarana pendanaan suatu perusahaan maupun institusi pemerintahan atau lainnya. Yang diperjualbelikan dalam pasar modal adalah instrumen jangka panjang (lebih dari 1 tahun) seperti saham, obligas (surat hutang), reksa dana dan yang lainnya

2.4 Saham

Saham adalah surat berharga dengan satuan nilai sebagai tanda kepemilikan suatu perusahaan. Dengan menerbitkan saham, perusahaan perusahaan yang membutuhkan suntikan modal bisa menjual saham kepada investor (penanam modal) dengan kata lain saham bisa diperjualbelikan dalam pasar modal untuk kepentingan pendanaan perusahaan. Saham bisa diperjualbelikan dalam pasar primer (primary market) dan pasar sekunder (secondary market).

2.5 Indeks Harga Saham

Indeks harga saham adalah sebuah indikator yang mencerminkan suatu pergerakan harga saham dalam suatu periode tertentu. Dengan demikian para investor bisa mengamati pergerakan harga saham, tahu tren yang berlangsung di pasar dan tahu kapan mereka harus membeli saham atau menjualnya di saat yang tepat.

2.6 Indeks Harga Saham Gabungan (IHSG)

Indeks harga saham gabungan (IHSG) merupakan indeks gabungan dari semua saham yang ada di Bursa Efek Indonesia (BEI). Untuk menggambarkan keadaan pasar yang wajar BEI berwenang untuk mengatur perusahaan mana yang di keluarkan atau dimasukan dalam perhitungan IHSG. Pertimbangan BEI dalam keputusan tersebut adalah jika jumlah saham perusahaan yang tercatat tersebut dimiliki oleh publik yang relatif kecil padahal kapitalisasi pasarnya cukup besar, sehingga perubahan harga sahamnya berpotensi memengaruhi kewajaran pergerakan IHSG.

2.7 Mean Absolute Percentage Error (MAPE)

Mean Absolute Percentage Error (MAPE) adalah total persentase rata-rata nilai galat antara data sebenarnya dengan data hasil prediksi. Ukuran akurasi dicocokkan dengan data time series, dan ditunjukkan dalam persentase. Dengan rumus :

$$MAPE = \frac{1}{n} \sum_{t=1}^n \frac{|A_t - F_t|}{|A_t|} \times 100$$

Dimana :

- At : Nilai sebenarnya
- Ft : Nilai Prediksi

3.1 Deskripsi Umum

Dalam tugas akhir ini akan dirancang sebuah sistem untuk memprediksi indeks harga saham menggunakan metode *Support Vector Regression* (SVR) dua tahap yang di optimasi oleh algoritma *Particle Swarm Optimization* (PSO). Pada *Support Vector Regression* (SVR) tahap pertama untuk input sistem adalah sepuluh indikator teknis yang ada pada Tabel 2.1. kemudian akan digunakan sebagai masukan sistem di *Support Vector Regression* (SVR) tahap kedua. *Output* dari sistem adalah nilai prediksi indeks harga saham dimasa yang akan datang. Dalam setiap proses SVR untuk menentukan parameter SVR digunakan algoritma

PSO untuk mendapatkan nilai parameter yang optimal untuk fungsi SVR.

3.2 Pengumpulan Data

Untuk menyelesaikan permasalahan yang dibahas sebelumnya dikumpulkan data-data yang dibutuhkan antara lain data Saham IHSG historis per hari selama 5 tahun yang di ambil dari www.duniainvestasi.com/bei.

3.3 Perancangan sistem

Sistem prediksi yang akan dibangun dibagi menjadi tiga tahap proses utama. Tahap pertama adalah penghitungan sepuluh indikator yang menggambarkan indeks saham pada hari ke-t sebagai input pada *Support*

Support Vector Regression (SVR) tahap kedua. Output pada *Support Vector Regression* (SVR) tahap kedua adalah nilai prediksi indeks harga saham dimasa yang akan datang. Kemudian untuk menghitung galat hasil prediksi dari metode *Support Vector Regression* (SVR) dua tahap digunakan *Mean Absolute Percentage Error* (MAPE).

Berikut adalah diagram alur (*flowchart*) perancangan sistem :

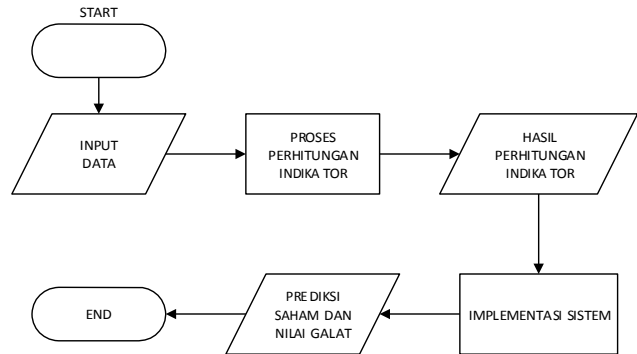
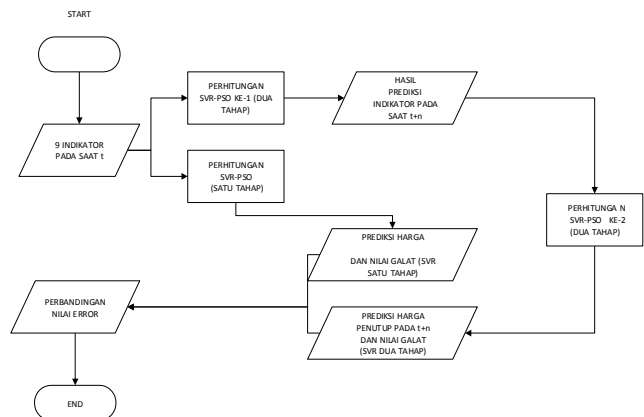


Diagram Alur Perancangan Sistem Prediksi Indeks Harga Saham.

- a.) Input Data : Data historis saham IHSG
Data yang digunakan adalah data pembukan, tertinggi, terendah dan penutup saham selama 5 tahun terhitung tanggal 4 Januari 2010 – 31 Agustus 2015.
- b.) Proses Perhitungan indikator
Proses ini menghitung masing-masing indikator dalam pada subbab (2.1).
- c.) Implementasi sistem subproses akan dijabarkan pada subbab selanjutnya (3.4)
- d.) Prediksi Saham dan nilai galat
Ini akan didapatkan dari hasil implementasi sistem

3.4 Implementasi Sistem

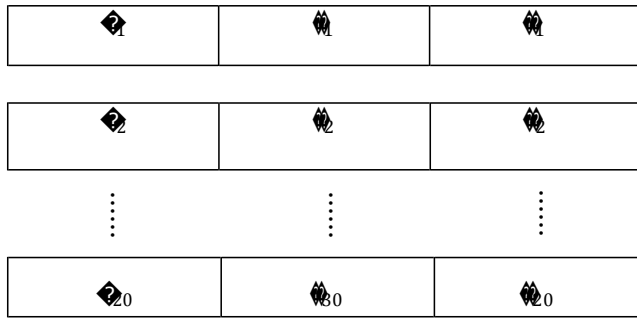


Vector Regression (SVR) tahap pertama . Kemudian output dari SVR tahap pertama tersebut menghasilkan indikator teknis pada hari ke (t+n) yang akan digunakan sebagai input

*Diagram alur Implementasi Sistem***3.4.1 Perhitungan *Particle Swarm Optimization* (PSO).**

Algoritma PSO ini digunakan untuk mencari nilai parameter dari SVR yaitu nilai (C, epsilon, dan gamma). Pertama inialisasi partikel, partikel disini adalah calon

nilai parameter C, epsilon dan gamma dengan jumlah partikel (N) adalah 20 yang dibangkitkan secara acak.



Nilai parameter C direpresentasikan oleh satu nilai acak bilangan riil dari 0,001 – 100, untuk nilai parameter epsilon direpresentasikan oleh satu nilai acak bilangan riil dari 0,001 – 1, dan untuk nilai parameter gamma direpresentasikan oleh satu nilai acak bilangan riil dari 0,001 – 1000. Kemudian dihitung nilai fitnessnya untuk mendapatkan P_{best} , setelah didapatkan P_{best} akan dibandingkan dengan P_{best} yang sebelumnya dan dipilih yang lebih baik yang nantinya partikel tersebut akan dijadikan G_{best} . Setelah G_{best} didapatkan dihitung nilai kecepatan tiap partikelnya dan kemudian posisi tiap partikel di update, jika sudah optimum maka nilai partikel itu akan dipilih untuk nilai C, epsilon dan gamma, sebaliknya jika tidak akan dilakukan iterasi lagi sebanyak 30 kali sampai didapatkan nilai yang paling optimal.

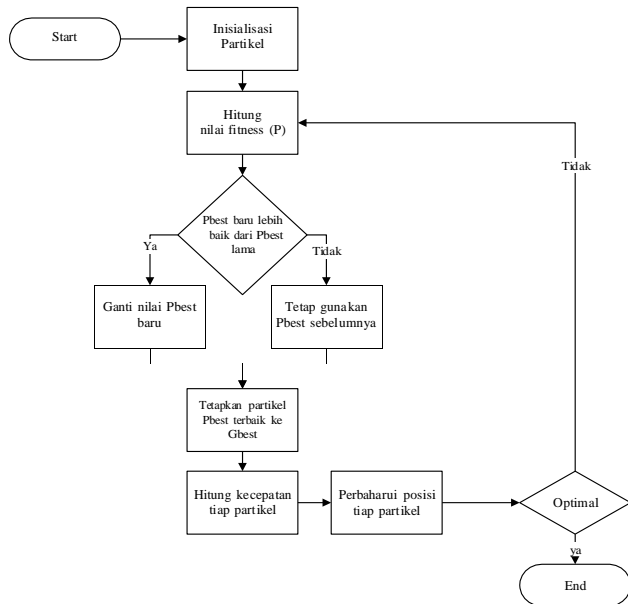


Diagram alur PSO

akan langsung diprediksi oleh satu SVR yang outputnya adalah prediksi harga penutup hari ke $(t+1)$. Dalam penelitian ini prediksi harga penutup adalah pada $(t+2)$, $(t+3)$, $(t+4)$, dan $(t+5)$. Dengan jumlah n hari untuk perhitungan indikator SMA adalah 10 hari, untuk EMA adalah 10 hari, untuk MOM adalah 10 hari, untuk STCK adalah 10 hari, untuk STCD adalah 20 hari, untuk MACD adalah 35, untuk RSI adalah 10hari, untuk WR adalah 10 hari, dan untuk CCI adalah 10 hari.

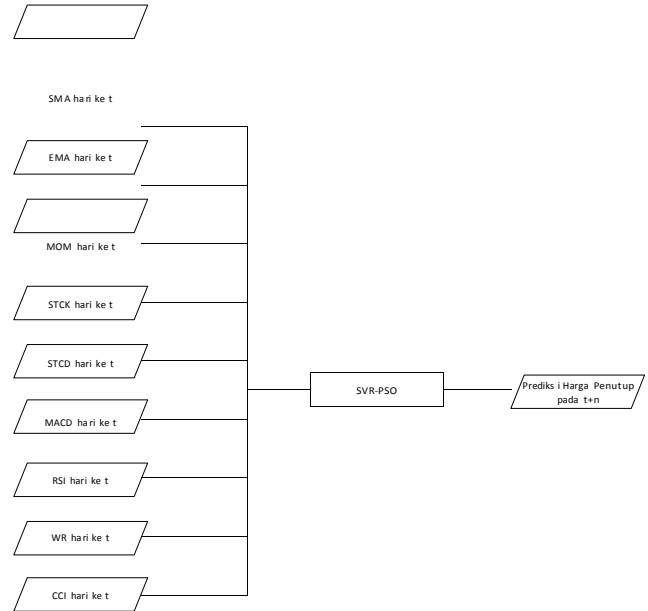


Diagram SVR satu tahap.

3.4.3 Prediksi Support Vector Regression (SVR) dua tahap di tahap pertama.

Pada tahap ini menggunakan Support Vector Regression (SVR) tahap pertama untuk mempersiapkan masukan untuk model prediksi Support Vector Regression (SVR) tahap kedua. Input dalam Support Vector Regression (SVR) tahap pertama ini adalah hasil perhitungan tiap indikator teknis pada hari ke- t dengan output pada proses ini adalah indikator teknis hari ke $(t+1)$. Dalam penelitian ini prediksi harga penutup adalah pada $(t+3)$, $(t+4)$, dan $(t+5)$. Sembilan output indikator ini adalah inputan untuk Support Vector Regression (SVR) tahap kedua. Dengan jumlah n hari untuk perhitungan indikator SMA adalah 10 hari, untuk EMA adalah 10 hari, untuk MOM adalah 10 hari, untuk STCK adalah 10 hari, untuk STCD adalah 20 hari, untuk MACD adalah 35, untuk RSI adalah 10hari, untuk WR adalah 10 hari, dan untuk CCI adalah 10 hari.

3.4.2 Prediksi *Support Vector Regression* (SVR) satu tahap.

Pada prediksi SVR satu tahap, input untuk SVR adalah hasil perhitungan sembilan indikator teknis yang

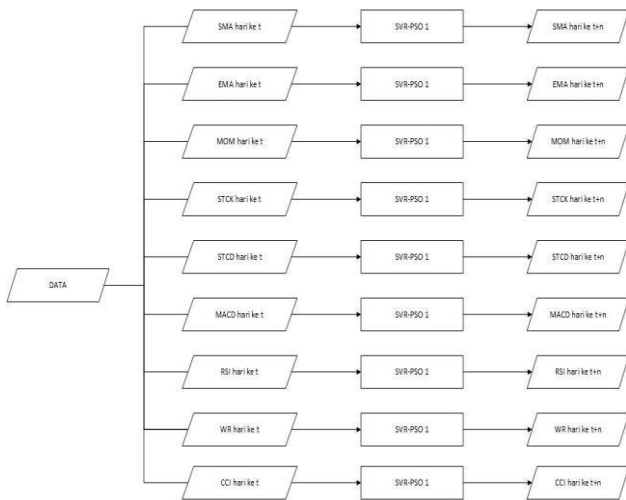


Diagram Model SVR dua tahap di tahap pertama

3.4.4 Prediksi Support Vector Regression (SVR) dua tahap di tahap kedua.

Selanjutnya pada tahap ini setelah prediksi tiap indikator teknis pada hari ke $(t + 1)$ yang didapatkan akan menjadi input di *Support Vector Regression* (SVR) tahap kedua untuk memprediksi harga penutup pada $(t + 2)$, $(t + 3)$, $(t + 4)$, dan $(t + 5)$.

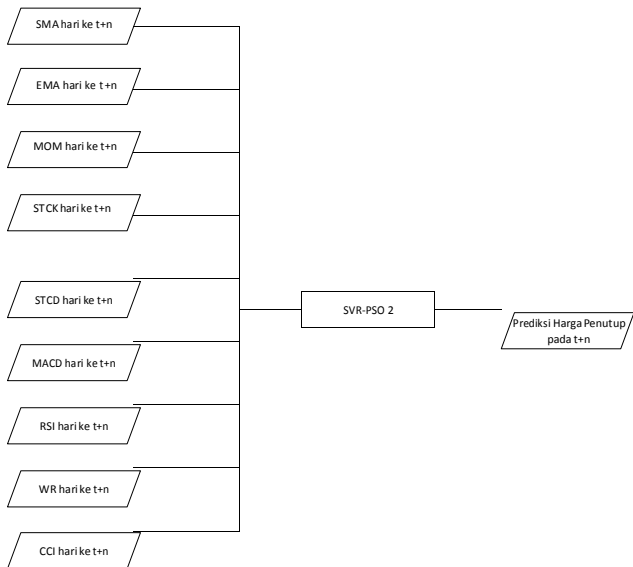


Diagram Model SVR dua tahap di tahap kedua

4 Hasil Dan Pengujian

4.1 Pengujian Sistem

Seperti yang telah dijelaskan pada bab sebelumnya, Tugas Akhir ini bertujuan untuk menerapkan metode SVR dua tahap dioptimasi dengan metode PSO pada prediksi

saham IHSG. Pada tahap pertama yaitu memprediksi nilai masing masing indikator, kemudian pada tahap kedua hasil prediksi tahap pertama digunakan untuk memprediksi nilai harga penutup.

4.2 Tujuan Pengujian Sistem

Tujuan dilakukannya pengujian pada tugas akhir ini adalah sebagai berikut:

1. Menganalisis nilai pediksi saham yang didapat dari sistem dengan menguji nilai harga penutup IHSG.
2. Mendapatkan nilai prediksi yang mendekati nilai aslinya.

4.3 Skenario Pengujian

Pengujian pada Tugas Akhir ini menggunakan data historis IHSG. Selanjutnya akan digunakan beberapa skenario yang akan menguji sistem yang dibentuk.

Skenario yang akan diuji yaitu membandingkan SVR satu tahap dan SVR dua tahap yang dioptimasi PSO dengan skenario $(t + 1)$, $(t + 2)$, $(t + 3)$, $(t + 4)$ dan $(t + 5)$.

4.3.1 Skenario 1 (SVR satu tahap $(t + 1)$ vs SVR dua tahap $(t + 1)$)

Pada skenario ini akan membandingkan hasil prediksi untuk waktu $(t + 1)$ antara SVR satu tahap dan SVR dua tahap. Jika nilai galat SVR satu tahap lebih kecil berarti hasil prediksi dari SVR satu tahap lebih baik daripada SVR dua tahap untuk waktu $(t + 1)$ jika hasilnya sebaliknya 1) SVR dua tahap memprediksi lebih baik.

4.3.2 Skenario 2 (SVR satu tahap $(t + 2)$ vs SVR dua tahap $(t + 2)$)

Pada skenario ini akan membandingkan hasil prediksi untuk waktu $(t + 2)$ antara SVR satu tahap dan SVR dua tahap. Jika nilai galat SVR satu tahap lebih kecil berarti hasil prediksi dari SVR satu tahap lebih baik daripada SVR dua tahap untuk waktu $(t + 2)$ jika hasilnya sebaliknya 2) SVR dua tahap memprediksi lebih baik.

4.3.3 Skenario 3 (SVR satu tahap $(t + 3)$ vs SVR dua tahap $(t + 3)$)

Pada skenario ini akan membandingkan hasil prediksi untuk waktu $(t + 3)$ antara SVR satu tahap dan SVR dua tahap. Jika nilai galat SVR satu tahap lebih kecil berarti hasil prediksi dari SVR satu tahap lebih baik daripada SVR dua tahap untuk waktu $(t + 3)$ jika hasilnya sebaliknya 3) SVR dua tahap memprediksi lebih baik.

4.3.4 Skenario 4 (SVR satu tahap (t + 4) vs SVR dua tahap (t + 4))

Pada skenario ini akan membandingkan hasil prediksi untuk waktu (t + 4) antara SVR satu tahap dan SVR dua tahap. Jika nilai galat SVR satu tahap lebih kecil berarti hasil prediksi dari SVR satu tahap lebih baik daripada SVR dua tahap untuk waktu (t + 4), jika hasilnya sebaliknya SVR dua tahap memprediksi lebih baik.

4.3.5 Skenario 3 (SVR satu tahap (t + 5) vs SVR dua tahap (t + 5))

Pada skenario ini akan membandingkan hasil prediksi untuk waktu (t + 5) antara SVR satu tahap dan SVR dua tahap. Jika nilai galat SVR satu tahap lebih kecil berarti hasil prediksi dari SVR satu tahap lebih baik daripada SVR dua tahap untuk waktu (t + 5), jika hasilnya sebaliknya yaitu SVR dua tahap lebih kecil maka untuk waktu (t + 5) SVR dua tahap memprediksi lebih baik.

4.4 Hasil dan Analisis

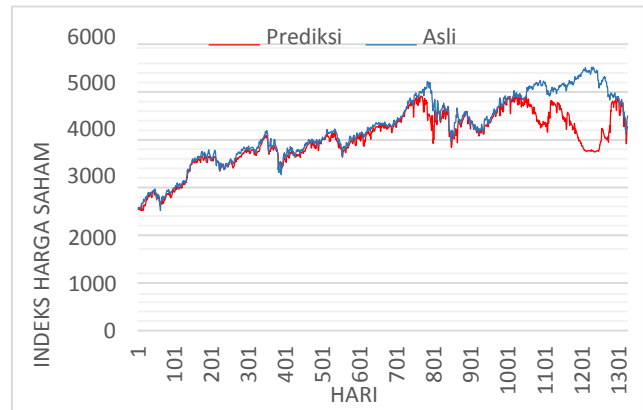
4.4.1 Parameter Indikator SVR Tahap Pertama

Pramater indikator di prediksi SVR yang dioptimasi PSO pada tahap pertama.

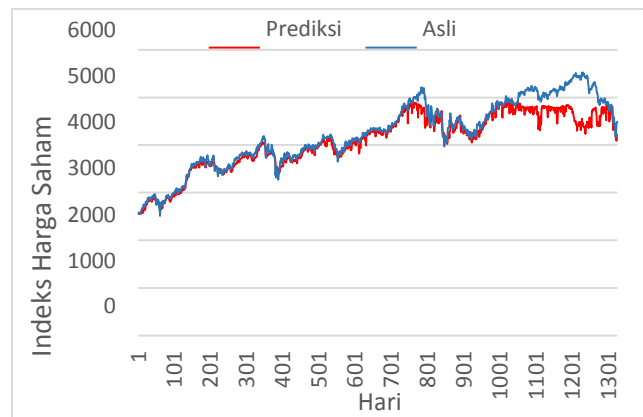
No.	Indikator	Parameter	
1	SMA	C	80,764
		Epsilon	5,6911
		Gamma	516,82
2	EMA	C	66,729
		Epsilon	6,0727
		Gamma	351,92
3	MOM	C	992,36
		Epsilon	38,908
		Gamma	591,91
4	STCK	C	49,638
		Epsilon	1,3491
		Gamma	95,189
5	STCD	C	90,413
		Epsilon	3,3049
		Gamma	67,607
6	MACD	C	91,993
		Epsilon	0,0626
		Gamma	2,5526
7	RSI	C	38,22
		Epsilon	0,4063
		Gamma	33,412
8	WR	C	96,812
		Epsilon	0,0469
		Gamma	95,618
9	CCI	C	77,862
		Epsilon	0,3134
		Gamma	74,352

4.4.2 Hasil Prediksi Untuk Skenario (t + 1).

Proses perhitungan nilai prediksi (t + 1) dilakukan dengan SVR dua tahap dan SVR satu tahap untuk pembandingnya. Parameter untuk setiap SVR dioptimasi dengan PSO termasuk untuk SVR yang memprediksi 9 indikator di SVR dua tahap. Pada SVR dua tahap untuk masing masing SVR yang memprediksi indikator mempunyai parameter yang berbeda. Dan untuk hasil akhir prediksi SVR satu tahap dan dua tahap bisa dilihat dalam tabel.



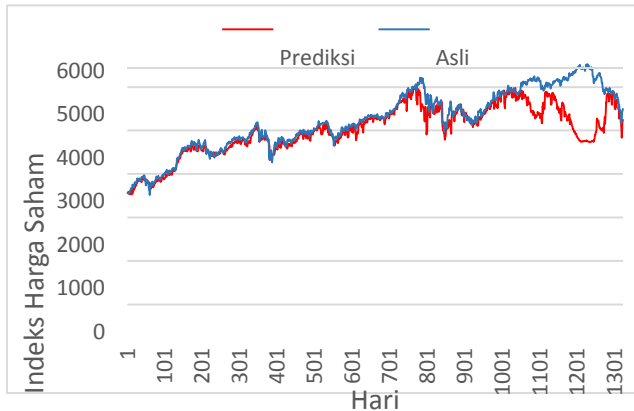
Grafik Hasil (t + 1) Satu Tahap.



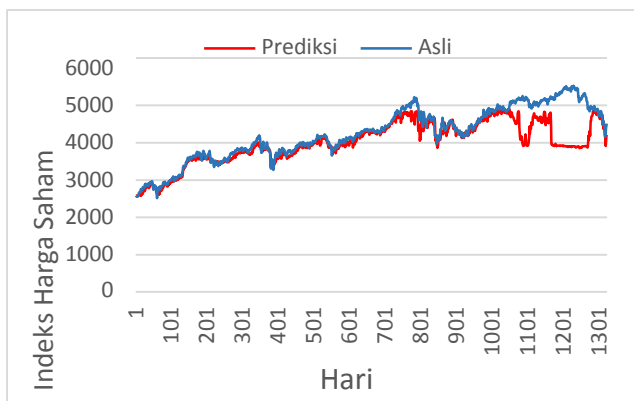
Grafik Hasil (t + 1) Dua Tahap

Skenario	SVR Satu Tahap			
		MAPE	Parameter	
t+1	Training	2,1155 %	C	32,078
	Testing	15,3160 %	Epsilon	0,041
	Total	4,8143 %	Gamma	224,665
	SVR Dua Tahap			
		MAPE	Parameter	
	Training	1,9858 %	C	32,078
	Testing	9,2088 %	Epsilon	0,041
	Total	3,4169 %	Gamma	224,665

4.4.3 Hasil Prediksi Untuk Skenario (t + 2).



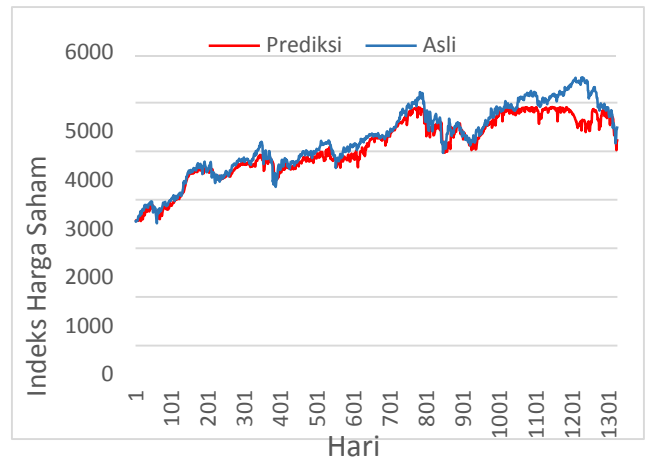
Grafik Hasil (t+2) Satu Tahap.



Grafik Hasil (t+2) Dua Tahap.

Skenario	SVR Satu Tahap			
	MAPE		Parameter	
t+2	Training	2,3517 %	C	32,078
	Testing	15,0260 %	Epsilon	0,041
	Total	4,9252 %	Gamma	224,665
	SVR Dua Tahap			
	MAPE		Parameter	
	Training	2,4441 %	C	32,078
Testing	16,4214 %	Epsilon	0,041	
Total	5,2055 %	Gamma	224,665	

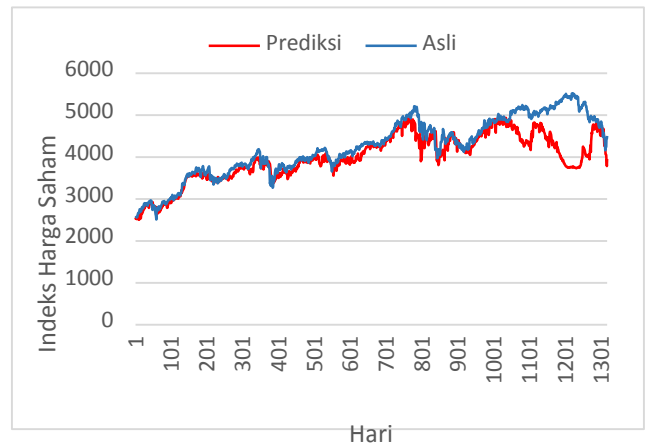
4.4.4 Hasil Prediksi Untuk Skenario (t + 3).



Grafik Hasil (t+3) Dua Tahap

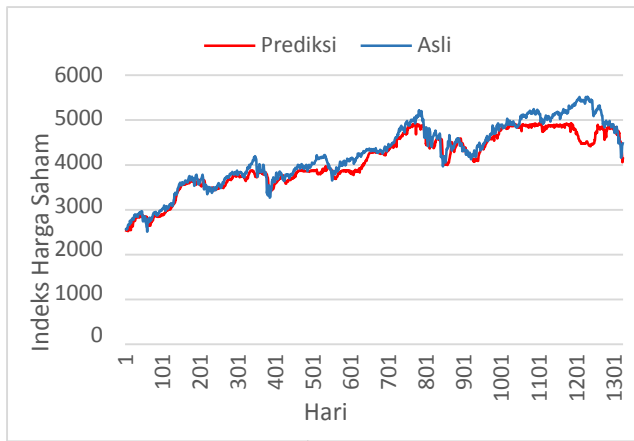
Skenario	SVR Satu Tahap			
	MAPE		Parameter	
t+3	Training	2,7570 %	C	32,078
	Testing	15,4537 %	Epsilon	0,041
	Total	5,3228 %	Gamma	224,665
	SVR Dua Tahap			
	MAPE		Parameter	
	Training	2,9713 %	C	32,078
Testing	7,3080 %	Epsilon	0,041	
Total	3,8260 %	Gamma	224,665	

4.4.5 Hasil Prediksi Untuk Skenario (t + 4).



Grafik Hasil (t+4) Satu Tahap.

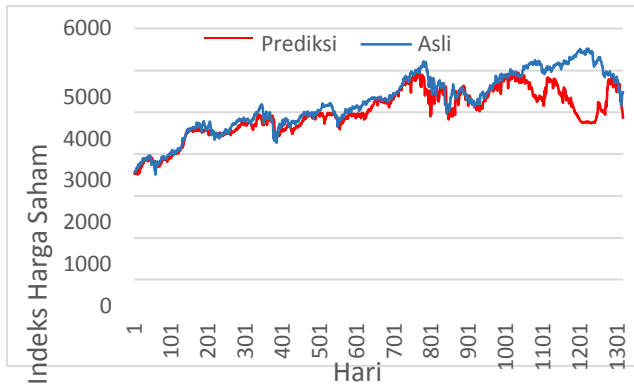
Grafik Hasil (◆+ 3) Satu Tahap.



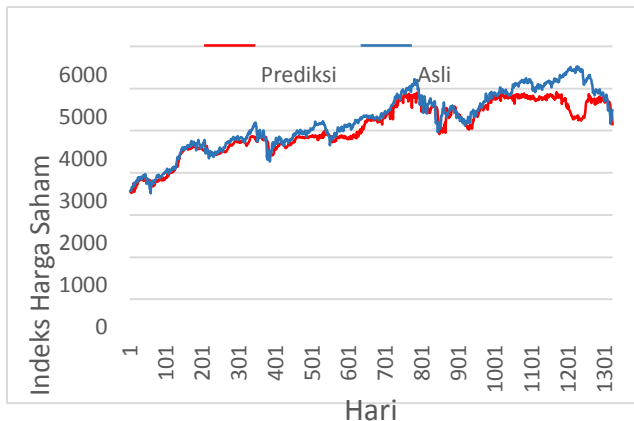
Grafik Hasil (♦+ 4) Dua Tahap.

Skenario	SVR Satu Tahap			
	MAPE		Parameter	
t+4	Training	3,0202 %	C	32,078
	Testing	15,6058 %	Epsilon	0,041
	Total	5,5485 %	Gamma	224,665
	SVR Dua Tahap			
	MAPE		Parameter	
	Training	2,9664 %	C	32,078
Testing	7,1856 %	Epsilon	0,041	
Total	3,7954 %	Gamma	224,665	

4.4.6 Hasil Prediksi Untuk Skenario (t + ♦)



Gambar 4.9 Grafik Hasil (♦+ 5) Satu Tahap.



Skenario	SVR Satu Tahap			
	MAPE		Parameter	
t+5	Training	3,1650 %	C	32,078
	Testing	15,7007 %	Epsilon	0,041
	Total	5,6684 %	Gamma	224,665
	SVR Dua Tahap			
	MAPE		Parameter	
	Training	3,2282 %	C	32,078
Testing	8,8449 %	Epsilon	0,041	
Total	4,3280 %	Gamma	224,665	

5. Kesimpulan dan Saran

5.1 Kesimpulan

Berdasarkan analisis terhadap implementasi sistem dan pengujian sistem prediksi indeks harga saham IHSG dengan metode Support Vector Regression (SVR), maka dapat ditarik kesimpulan sebagai berikut :

1. Metode SVR dua tahap menunjukkan akurasi yang lebih baik dibandingkan dengan akurasi SVR satu tahap kecuali untuk skenario (♦+ 2).
2. Berdasarkan prediksi indeks harga saham dengan metode SVR dua tahap didapatkan nilai MAPE Testing dari lima skenario, yaitu MAPE untuk skenario (♦+ 1) : 9,2099 %, MAPE untuk skenario (♦+ 2) : 16,4214 %, MAPE untuk skenario (♦+ 3) : 7,308%, MAPE untuk skenario (♦+ 4) : 7,1856 % , dan MAPE untuk skenario (♦+ 5) : 8,8449 %.

5.2 Saran

Setelah dilakukan beberapa percobaan, saran untuk penelitian tugas akhir ini adalah metode *Support Vector Regresson* (SVR) dua tahap digunakan untuk prediksi 1 periode kedepan saja.

Daftar Pustaka

[1] Ou, P., & Wang, H. (2009). *Prediction of stock market index movement by ten data mining techniques*. Modern Applied Science, 3, P28.
 [2] Jigar Patel, Sahil Shah, Priyank Thakkar, K Kotecha.

Grafik Hasil (♦+ 5) Dua Tahap.

(20
14).
*Pre
dict
ing
stoc
k
mar
ket
ind
ex
usin
g
fusi
on
of
mac
hin
e
lear
nin
g
tech
niq
ues.*
Co
mp
uter
Sci
enc
e &
Eng
inee
ring
Dep
art
me
nt,
Inst
itut
e of
Tec
hno
log
y,
Nir
ma
Uni
vers
ity,
Ah
me
dab
ad,
Guj
arat
,
Indi
a.

[3] Kara, Y., Acar Boyacioglu, M., & Baykan, Ö. 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*, 38, 5311–5319. mining techniques. *Modern Applied Science*, 3(2), 8.

[4] Budi Susanto. (2007). *Teknik Pemanfaatan Data untung Keperluan Bisnis. : Data Mining*

- [5] Cristianini, N., & Shawe-Taylor, J. (2000). *An introduction to support vector machines*.
- [6] Huang, C.-L., & Tsai, C.-Y. (2009). *A hybrid SOFM–SVR with a filter-based feature selection for stock market forecasting*. *Expert Systems with Applications*, 36, 1529–1539.
- [7] Lufuno Ronald Marwala .(2010). *Forecasting the Stock Market Index Using Artificial Intelligence Techniques*. University of the Witwatersrand, Johannesburg, 30-36.
- [8] Kazem, A., Sharifi, E., Hussain, F. K., Saberi, M., & Hussain, O. K. (2013). *Support vector regression with chaos-based firefly algorithm for stock market price forecasting*. *Applied Soft Computing*, 13, 947–958.
- [9] Francesco Parella. (2000). *Online Support Vector Regression*. Department of Information Science University of Genoa Italy
- [10] J. Kennedy, R. C. Eberhart.(1995). Particle swarm optimization. In *Proceedings of the 1995 IEEE International Conference on Neural Networks*. IEEE Service Center, Piscataway.