

Prediksi Indeks Harga Saham Menggunakan Metode Gabungan *Support Vector Regression* dan *Random Forest*

Prediction Of Stock Market Index Using Hybrid of Support Vector Regression and Random Forest

Barini Harahap¹, Rian Febrain Umbara, M.Si, Danang Triantoro, M.T.²

^{1,2}Prodi S1 Ilmu Komputasi, Fakultas Informatika, Universitas Telkom

¹bariniharahap@gmail.com, ²rianum123@gmail.com, ³danangtri@telkomuniversitv.ac.id

Abstrak

Indeks harga saham adalah indikator atau cerminan pergerakan harga saham. Indeks merupakan salah satu pedoman bagi investor untuk menanamkan modal investasi di pasar modal khususnya saham. Penelitian ini memprediksi harga penutupan indeks harga saham pada hari ke (t+1), (t+5), (t+10), (t+20), dan (t+30) menggunakan metode gabungan Support Vector Regression (SVR) dan Random Forest (RF) lalu membandingkannya dengan menggunakan metode Support Vector Regression dan Random Forest. Data yang digunakan pada penelitian ini adalah data Indeks Harga Saham Gabungan Indonesia selama lima tahun (2011-2015). Hasil yang didapat pada penelitian ini adalah nilai MAPE metode SVR lebih kecil untuk prediksi hari ke (t+1) dan (t+5) dengan nilai MAPE sebesar 1.9119% dan 4.5691%, sedangkan untuk prediksi hari ke (t+10), (t+20), dan (t+30) metode yang memiliki nilai MAPE terkecil adalah metode SVR-RF dengan nilai MAPE sebesar 4.1173%, 8.6391%, dan 9.1708%. Dari hasil penelitian ini dapat disimpulkan bahwa metode SVR baik untuk prediksi jangka pendek sedangkan metode SVR-RF baik untuk prediksi jangka panjang.

Kata kunci : Prediksi, Indeks Harga Saham, IHSG, *Support Vector Regression*, *Random Forest*.

Abstract

Stock price index is an indicator or reflection stock price movements. Index is a guideline for investors to invest capital investment in the capital markets in particular stocks. This study predicts the closing price of the stock price index on day (t + 1), (t + 5), (t + 10), (t + 20), and (t + 30) using Support Vector Regression (SVR) - Random Forest (RF) and then compare it with using Support Vector Regression and Random Forest. The data used in this study is Indonesia Composite Index (IDX Composite) data for five years (2011-2015). The results obtained in this study is the value of MAPE for SVR is smaller for the prediction of day (t + 1) and (t + 5) with a value of 1.9119% and 4.5691%, while for the prediction of day (t + 10), (t + 20), and (t + 30) the method that has the smallest MAPE value is SVR-RF with MAPE value of 4.1173%, 8.6391% and 9.1708%. From these results it can be concluded that the SVR method is good for short-term prediction while SVR-RF good method for long-term predictions.

Keyword : Prediction Indices, IDX Composite, *Support Vector Regression*, *Random Forest*

Utama dan Papan Pengembangan, Indeks Kompas100, Bisnis-27.

1. Pendahuluan

Saham adalah tanda penyertaan modal seseorang atau pihak (badan usaha) dalam satu perusahaan atau perseroan terbatas. Menerbitkan saham menjadi salah satu pilihan perusahaan ketika perusahaan memutuskan untuk berinvestasi. Indeks harga saham adalah indikator atau cerminan pergerakan harga saham. Indeks merupakan salah satu pedoman bagi investor untuk menanamkan modal investasi di pasar modal khususnya saham. Di Indonesia sendiri ada beberapa indeks harga saham seperti Indeks Harga Saham Gabungan (IHSG), Indeks Jakarta Islamic Index (JII), Indeks Papan

Untuk mengetahui naik atau turunnya kegiatan ekonomi, banyak investor akan melihat dari sisi indeks yang dicapai pada saat itu. Naiknya nilai indeks saham dapat menjadi indikator kebangkitan ekonomi di suatu negara. Nilai indeks saham dihitung berdasarkan saham dengan kapitalisasi pasar yang tinggi. Berbagai parameter teknis digunakan untuk mendapatkan informasi statistic dari nilai harga saham. Indeks saham yang berasal dari harga saham dengan kapitalisasi pasar yang tinggi dapat memberikan gambaran

keseluruhan ekonomi dan tergantung dari beberapa factor [1].

Memprediksi indeks harga saham dan pergerakannya telah dianggap sebagai salah satu aplikasi yang paling menantang dari prediksi *time series*. Meskipun ada banyak penelitian empiris yang berhubungan dengan isu-isu memprediksi indeks harga saham, tetapi kebanyakan penelitian dilakukan di negara dengan pasar keuangan yang maju. Prediksi yang akurat dari pergerakan indeks harga saham sangat penting untuk mengembangkan strategi pasar yang efektif [3]. Dengan demikian investor dapat meminimalkan resiko pasar dan berkesempatan untuk membuat keuntungan dengan

perdagangan indeks saham [4]. Prediksi harga saham sendiri sudah banyak dilakukan oleh cendekiawan di seluruh dunia salah satunya oleh Patel, Shah, Thakkar, & Kotecha. 2015 dalam jurnal *Predicting*

stock market index using fusion of machine learning techniques yang memprediksi indeks harga saham CNX Nifty dan S&P Bombay Stock

Exchange (BSE) menggunakan algoritma SVR-ANN, SVR-RF, SVR-SVR [1].

Dalam tugas akhir ini akan digunakan data dari Indeks Harga Saham Gabungan (IHSG) atau *Composite Stock Price Index*. IHSG mempunyai pergerakan yang selalu berubah-ubah maka dari itu akan digunakan dua algoritma *machine learning* yakni *Support Vector Regression (SVR)*-*Random*

Forest (RF). *Support Vector Regression* adalah algoritma yang paling banyak digunakan untuk memprediksi harga saham dan indeks harga saham yang dikemukakan oleh Vladimir N Vapnik pada tahun 1996, sedangkan *Random Forest* adalah algoritma pohon keputusan untuk memperkuat regresi dari SVR. *Random Forest* sendiri pertama kali diperkenalkan oleh Leo Breimen pada tahun 2001 dalam jurnalnya "*Random Forest, Machine Learning*".

2. Dasar Teori dan Perancangan

A. Support Vector Regression

SVR adalah sebuah metode yang pertama kali diperkenalkan oleh Vladimir N. Vapnik, Harris Drucker, Christopher J. C. Burges, Linda Kaufman dan Alexander J. Smola pada tahun 1996. Dimana model ini memperkenalkan Support Vector Classification yang hanya bergantung pada subset dari data pelatihan, karena fungsi biaya untuk membangun model tidak peduli poin pelatihan yang berada di luar margin. SVR merupakan penerapan SVM yang digunakan untuk kasus regresi, yang keluarannya berupa bilangan riil atau kontinu.

Vapnik membuat suatu rumus untuk meminimalkan fungsi insentivitas-ε (loss function) yakni

$$|w \cdot x_i - b| - \epsilon = \begin{cases} 0, & |w \cdot x_i - b| \leq \epsilon \\ |w \cdot x_i - b| - \epsilon, & |w \cdot x_i - b| > \epsilon \end{cases} \quad (2.2)$$

Berdasarkan persamaan (2.2) regresi linier dari $w \cdot x_i - b$ diestimasi sebagai simultan untuk meminimalkan $\|w\|^2$ dan penjumlahan dari linier $\epsilon - |w \cdot x_i - b|$. Dimana factor $\|w\|^2$ dinamakan regularisasi, meminimalkan $\|w\|^2$ akan membuat suatu fungsi setipis mungkin, sehingga dapat mengontrol fungsi kapasitas [7].

$$w = \frac{1}{2} \left(\|w\|^2 + \sum_{i=1}^n |w \cdot x_i - b| \right) \quad (2.3)$$

$$w = \frac{1}{2} \left(\|w\|^2 + \sum_{i=1}^n (\xi_i + \xi_i^*) \right)$$

$$(w \cdot x_i + b) - y_i \leq \epsilon + \xi_i \quad (2.5)$$

$$y_i - (w \cdot x_i + b) \leq \epsilon + \xi_i \quad (2.6)$$

$$\xi_i, \xi_i^* \geq 0, i = 1, 2, \dots, n \quad (2.7)$$

SVR membangun persamaan fungsi regresi linier sebagai berikut

$$y_i = w \cdot x_i + b \quad (2.1)$$

Dimana ξ_0 dan ξ adalah variable yang tidak negatif untuk membuat \hat{y} berada di bawah target dan satu R berfungsi untuk meminilakan resiko. Konstanta c berfungsi untuk mengontrol aproksimasi error dan nilai vector normal $\|w\|$.

Untuk permasalahan optimasi harus menggunakan teori langrang 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 indicator sebagai *input*. Fungsi kernel yang dimaksud adalah

PolynomialFunction :

$$K(x, x') = (\gamma \cdot x \cdot x' + 1)^d \quad (2.8)$$

RadialBasisFunction :

$$K(x, x') = \exp(-\gamma \|x - x'\|^2) \quad (2.9)$$

Dimana,

- d = pangkat dari fungsi polynomial
- γ = konstanta dari fungsi basis radial

B. Indikator Analisis Teknikal

Analisis teknikal adalah beberapa indikator atau dasar untuk menghitung pergerakan indeks saham [9]. Dalam tugas akhir ini akan digunakan 10 indikator yang berfungsi untuk menghitung pergerakan saham. Indikator-indikator tersebut adalah [9]:

1. *Simple Moving Average (SMA)*
Simple Moving Average adalah indikator analisis teknikal yang paling umum digunakan. *SMA* adalah jumlah seluruh data dibagi banyaknya data.
2. *Exponential Moving Average (EMA)*
Exponential Moving Average hampir sama dengan *SMA* bedanya *EMA* menggunakan data saham terbaru sehingga perhitungan saham lebih cepat.
3. *Momentum (MOM)*
Momentum adalah sebuah indikator analisis teknikal yang banyak digunakan. *Momentum* digunakan untuk menghitung kecepatan dari pergerakan saham.
4. *Stochastic K% (STCK)*
Stochastic K% adalah sebuah indikator analisis teknikal yang berfungsi membandingkan harga penutupan saat ini dengan kisaran harga untuk jangka waktu tertentu dimana *K%* menandakan baris utama. Perhitungan *Stochastic K%* menggunakan nilai stokastik.
5. *Stochastic D% (STCD)*
Stochastic D% adalah sebuah indikator analisis teknikal yang berfungsi membandingkan harga penutupan saat ini dengan kisaran harga untuk jangka waktu tertentu dimana untuk *D%* menandakan baris kedua yang berfungsi sebagai garis pengidentifikasi arah. Perhitungan *D%* menggunakan nilai *average movement*.
6. *Moving Average Convergence and Divergence (MACD)*
Moving Average Convergence and Divergence adalah sebuah indikator analisis teknikal yang dinamis, yang dapat menentukan korelasi antara dua harga rata-rata yang bergerak. *MACD* menggunakan perbedaan antara periode *EMA* (12 atau 26). *MACD* baik untuk menghitung nilai garis signal dan pelanggaran arah garis.
7. *Relative Strength Index (RSI)*
Relative Strength Index adalah sebuah indikator analisis teknikal berupa grafik harga yang berkisar antara 0-100. *RSI* menggambarkan kekuatan indeks harga saham.
8. *Larry William's R% (WR)*

Larry William's R% adalah sebuah indikator analisis teknikal yang dinamis, yang dapat menentukan apakah pasar *overbought* atau *oversold*. *WR* memperlihatkan hubungan antar titik penutupan relative terhadap titik-rendah sepanjang periode waktu tertentu.

9. *Accumulation/Distribution Oscillators (ADO)*
Accumulation/Distribution Oscillators adalah indikator analisis teknikal yang berfungsi membandingkan harga penutupan saat ini dengan kisaran harga untuk jangka waktu tertentu. *ADO* menggunakan akumulasi atau distribusi untuk membandingkannya.

10. *Commodity Channel Index (CCI)*
Commodity Channel Index adalah sebuah indikator analisis teknikal yang mengukur penyimpanan harga instrumen dari harga rata-rata. Harga turun jika nilai barang tinggi, jika nilai barang rendah maka harga akan naik.

Adapun formula untuk ke-sepuluh indikator teknis yang mempengaruhi indeks harga saham adalah sebagai berikut^[1].

Tabel 2.1. Indikator teknis yang mempengaruhi indeks harga saham

No	Indikator	Formula
1	Simple n-day moving average	$\frac{C_1 + C_2 + \dots + C_n}{n}$
2	Weighted moving average	$\frac{C_1 + (2-1)C_2 + \dots + (1)C_n}{2 + (2-1) + \dots + 1}$
3	Momentum	$C_t - C_{t(n-1)}$
4	Stochastic K%	$\frac{C_t - C_{(n-1)}}{C_{(n-1)} - C_{(n-1)}} \times 100$
5	Stochastik D%	$\frac{\sum_{i=0}^{n-1} C_i}{10} \%$
6	Relative Strength Index	$100 - \frac{100}{1 + (\sum_{i=0}^{n-1} \frac{C_i - C_{(i-1)}}{C_{(i-1)}}) / (\sum_{i=0}^{n-1} \frac{C_i - C_{(i-1)}}{C_{(i-1)}})}$
7	Moving average convergence divergence	$\frac{C_{t+1} + C_{t-1}}{2} \times (C_t - C_{t-1})$
8	Larry William's R%	$\frac{C_t - C_{(n-1)}}{C_t - C_{(n-1)}} \times 100$
9	A/D oscillator	$\frac{C_t - C_{t-1}}{C_t - C_{t-1}}$
10	Commodity channel index	$\frac{C_t - C_{(n-1)}}{0.015 C_t}$

dimana :

1. \diamond = harga saham pada saat penutupan pada saat t

- 2. P_t = harga saham terendah pada saat t
- 3. P_t^* = harga saham tertinggi pada saat t
- 4. P_{t-1} = harga saham terendah dalam rentang waktu t
- 5. P_{t-1}^* = harga saham tertinggi dalam rentang waktu t
- 6. $P_{t-1}^* = (P_{t-1}^* - P_{t-1}) \times 2$
- 7. $P_{t-1} = (P_{t-1} - P_{t-1}) \times 2$
- 8. $\alpha = \frac{2}{1 + \sqrt{2}}$
- 9. k = periode waktu k-hari untuk EMA
- 10. σ = $\frac{1}{\sqrt{k}}$
- 11. $\sigma^2 = \frac{1}{k}$
- 12. $\sigma^2 = \frac{1}{k}$
- 13. U_t = kenaikan harga saham pada saat perubahan waktu t
- 14. D_t = penurunan harga saham pada saat perubahan waktu t
- 15. N = jumlah saham

node pohon. m harus lebih kecil daripada M.

- Sebagian pohon dibangun secara random yang didapat dari berbagai subset data pelatihan
- Untuk setiap node, split data ke dua anak pohon di bawahnya menggunakan *residual sum of squares*
- Node paling bawah adalah node terminal

2. Untuk regresi nilai prediksinya adalah nilai rata-rata dari semua node.

Untuk regresi gunakan *residual sum of squares*

$$RSS = \sum_{i=1}^n (y_i - \hat{y}_i)^2 + \sum_{j=1}^m (y_j - \hat{y}_j)^2 \tag{2.12}$$

Dimana: \bar{y}_i^* = mean dari nilai y untuk node kiri

kiri

kanan

E. Menghitung Error

Performansi suatu prediksi dari suatu metode pada diukur dari error hasil prediksi tersebut. Nilai error dapat dihitung dengan menggunakan :

Mean Absolute Percentage Error (MAPE)
MAPE adalah rata-rata dari keseluruhan persentase selisih antara data aktual dan data hasil prediksi. Rumus MAPE adalah :

$$MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{|y_i|} \times 100 \tag{2.14}$$

Dimana:

- m sebagai jumlah variabel, m berfungsi untuk mendeterminasi keputusan pada

C. 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. PSO dapat diimplementasikan dalam persamaan berikut

$$\min f(x) \tag{2.9}$$

dimana $X^{(B)} \leq X \leq X^{(A)}$

D. Random Forest

Random Forest pertama kali diperkenalkan pada tahun 2001 oleh Leo Breiman dalam jurnalnya yang berjudul "*Random Forest*". Random forest merupakan pengembangan dari Decision Tree adalah kombinasi dari prediktor pohon sehingga masing-masing pohon tergantung pada nilai-nilai acak vektor sampel independen dan dengan sama distribusi untuk semua pohon di hutan [10].

Algoritma yang harus diikuti ketika membangun sebuah pohon menggunakan *Random Forest* adalah:

1. Bangun pohon binari

- Tempatkan data latih sebagai N, banyak variabel dalam klasifikasi sebagai M

◆= nilai saham sesungguhnya

◆= nilai prediksi

◆= banyaknya data

◆= waktu

3. Rancangan Sistem

Pada tahap ini akan dibangun sebuah rancangan sistem yang berisi alur kinerja sistem untuk mempermudah proses pengerjaan tugas akhir dari awal sampai akhir.



Gambar 3.1. 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 Vector Regression* (SVR) yang dioptimasi algoritma *Particle Swam*

Optimization (PSO) tahap pertama. Kemudian output dari SVR tahap pertama tersebut menghasilkan indikator teknis pada hari ke $(t+n)$ yang akan digunakan sebagai input pada *Random Forest* (RF). Output pada *Random Forest* adalah nilai prediksi indeks harga saham dimasa yang akan datang. Kemudian untuk menghitung akurasi hasil prediksi dari metode *Support Vector Regression* (SVR) dua tahap digunakan *Mean Absolute Percentage Error* (MAPE).

a. Input Data

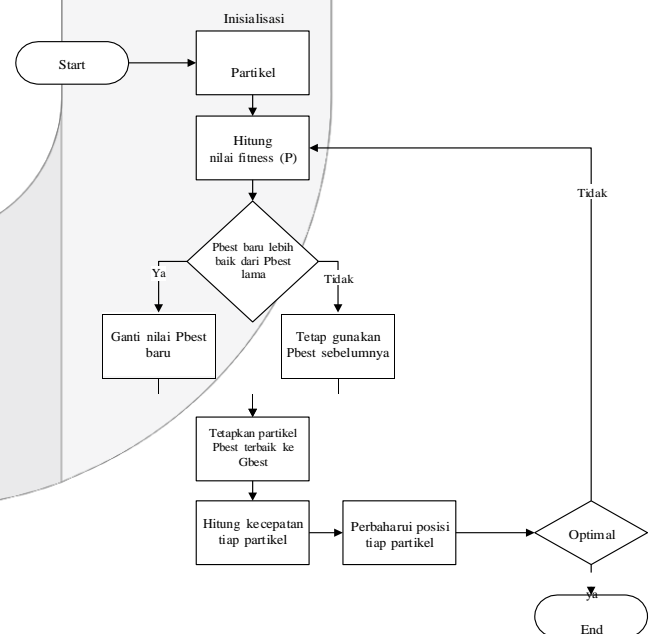
Data yang digunakan adalah data yang diambil dari www.finance.yahoo.com yaitu data indeks harga saham IHSG dan selama lima tahun.

b. Hitung Indikator Analisis Teknikal

Pada tahap ini akan dilakukan proses perhitungan sepuluh indikator analisis teknikal yakni SMA, EMA, MOM, STCK, STCD, RSI, MACD, WR, ADO, CCI menggunakan rumus yang telah dijelaskan sebelumnya.

c. Algoritma SVR-PSO

Pada tahap ini akan dilakukan perhitungan menggunakan algoritma *SVR-PSO*. Algoritma PSO ini digunakan untuk mencari nilai parameter dari SVR yaitu nilai (C , epsilon, dan gamma).



Gambar 3.2. Algoritma PSO

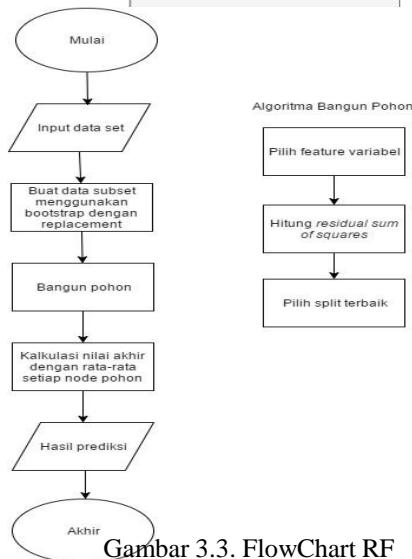
Pertama inisialisasi partikel, partikel disini adalah calon nilai parameter C , epsilon dan

gamma dengan jumlah partikel (N) adalah 20 yang dibangkitkan secara acak.

Kemudian dihitung nilai fitnessnya untuk mendapatkan P_{best} , setelah didapatkan P_{best} akan di bandingkan 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.

d. Algoritma RF

Data yang telah diolah oleh SVR berupa sepuluh SVR akan menjadi data masukan metode *Random Forest* lalu akan dibuat subset data baru dari data set yang telah ada menggunakan *bootstrap* dengan *replacement*. Setelah itu akan dibangun pohon sebanyak n pohon, dimana untuk membangun pohon dibutuhkan feature variable lalu untuk split node pohon dilakukan perhitungan *Residual sum of squares*. Setelah itu dilakukan kalkulasi nilai akhir dengan cara merata-ratakan setiap node pohon.



Gambar 3.3. FlowChart RF

e. Analisis Akurasi

Pada tahap ini, akurasi yang telah didapat dari metode *Support Vector Regression*, *Random Forest*, *Support Vector Regression-Random Forest* pada hari ke (t+n) dibandingkan.

4. Pembahasan

A. Pengujian

Seperti yang telah dijelaskan pada bab sebelumnya, Tugas Akhir ini bertujuan untuk menerapkan metode SVR, RF dan SVR-RF 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, sedangkan untuk SVR dan RF sendiri hanya memprediksi nilai harga penutup.

B. Skenario Pengujian

Pada penelitian ini menggunakan 5 skenario yang memprediksi harga penutup IHSG untuk (t+n) menggunakan metode SVR, RF menggunakan 50 pohon, 100 pohon dan 150 pohon dan SVR-RF dimana pada tahap kedua, RF, menggunakan 50 pohon, 100 pohon dan 150 pohon. Metode dengan nilai MAPE terkecil dari SVR, RF dan SVR-RF adalah metode terbaik untuk prediksi (t+30). Ada pun skenario penilitan ini adalah sebagai berikut

1. SVR vs RF vs SVR-RF untuk prediksi hari ke (t+1)
2. SVR vs RF vs SVR RF untuk prediksi hari ke (t+5)
3. SVR vs RF vs SVR RF untuk prediksi hari ke (t+10)
4. SVR vs RF vs SVR RF untuk prediksi hari ke (t+20)
5. SVR vs RF vs SVR RF untuk prediksi hari ke (t+30)

C. Hasil dan Analisis

Hasil yang didapat dalam penelitian ini adalah sebagai berikut

1. Parameter indikator

Berikut parameter indikator yang digunakan sebagai parameter masukan SVR.

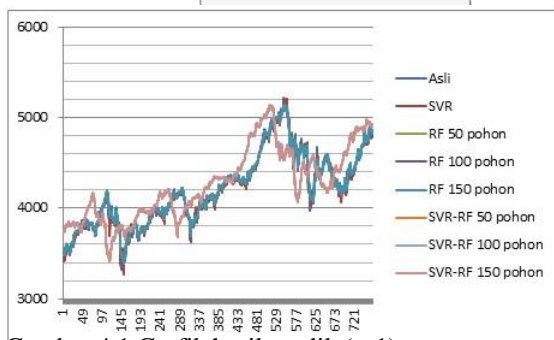
1	SMA	C	91556.4154
		Epsilon	6.0026
		Gamma	740.1296
2	EMA	C	91887.4231
		Epsilon	0.70707
		Gamma	47886.4169
3	MOM	C	7.6843

		Epsilon	511.1023
		Gamma	0.70707
4	STK	C	18527.6314
		Epsilon	0.94299
		Gamma	873.0132
		C	30092.3277
5	STD	Epsilon	1.0919
		Gamma	40.7086
6	RSI	C	91556.4145
		Epsilon	6.0026
		Gamma	740.1296
		C	91887.4231
7	A/D	Epsilon	0.70707
		Gamma	224.2873
8	LW	C	77595.9969
		Epsilon	3.3223
		Gamma	155.6078
		C	68192.5925
9	CCI	Epsilon	8.808
		Gamma	60.1705
10	MACDt	C	24492.2901
		Epsilon	6.2264
		Gamma	783.9811

Tabel 4.1. Tabel parameter indikator

2. Grafik hasil prediksi SVR vs RF vs SVR-RF untuk prediksi hari ke (t+n)

a. Grafik hasil prediksi SVR vs RF vs SVR-RF untuk prediksi hari ke (t+1)



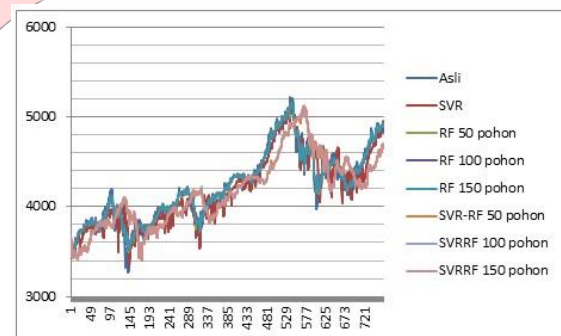
Gambar 4.1 Grafik hasil predik (t+1)

b. Grafik hasil prediksi SVR vs RF vs SVR-RF untuk prediksi hari ke (t+5)



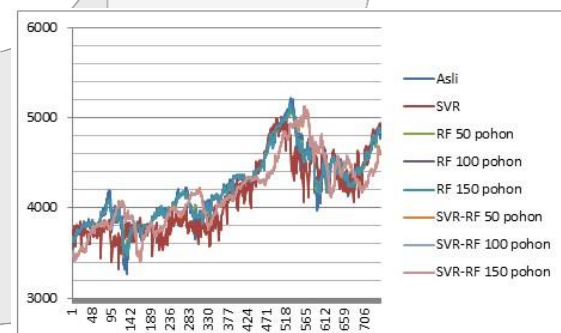
Gambar 4.2 Grafik hasil predik (t+5)

c. Grafik hasil prediksi SVR vs RF vs SVR-RF untuk prediksi hari ke (t+10)



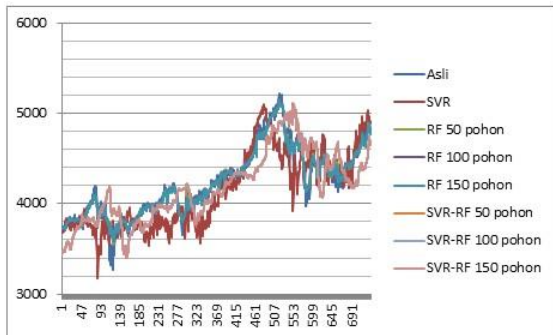
Gambar 4.3 Grafik hasil predik (t+10)

d. Grafik hasil prediksi SVR vs RF vs SVR-RF untuk prediksi hari ke (t+20)



Gambar 4.4 Grafik hasil predik (t+20)

e. Grafik hasil prediksi SVR vs RF vs SVR-RF untuk prediksi hari ke (t+30)



Gambar 4.5 Grafik hasil predik (t+30)

D. Hasil prediksi keseluruhan

Hasil keseluruhan nilai MAPE pada metode SVR vs RF vs SVR-RF dapat dilihat pada tabel 4.2.

Dari tabel dapat di lihat bahwasannya pada prediksi hari ke (t+1) metode SVR mendapatkan nilai MAPE sebesar 1.9119%, jauh lebih kecil dibanding metode lainnya, pada metode RF dan SVR-RF jumlah pohon mempengaruhi hasil nilai MAPE. Untuk prediksi hari ke (t+5) metode SVR mendapatkan hasil 4.5691%, dan jumlah pohon pada metode RF dan SVR-RF mempengaruhi hasil nilai MAPE tetapi tidak linier. Untuk prediksi hari ke (t+10) perbandingan nilai MAPE ketiga metode signifikan dan metode SVR-RF lebih kecil dengan nilai MAPE 4.1173% untuk jumlah pohon 150 pohon.. Pada prediksi hari ke (t+20) perbandingan nilai MAPE signifikan, tetapi metode SVR-RF lebih kecil dengan nilai 8.6391%. Terakhir, pada prediksi hari ke (t+30) hasil nilai MAPE untuk

		Hari ke	t+1	t+5	t+10	t+20	t+30
SVR		training	0.9225%	1.8621%	2.3186%	4.1512%	4.2022%
		testing	1.9119%	4.5691%	6.5795%	19.249%	21.5306%
RF	50 pohon	training	0.6298%	0.8187%	0.9104%	1.043%	1.1628%
		testing	4.4927%	5.8148%	7.6525%	9.6026%	9.3617%
	100 pohon	training	0.6306%	0.8192%	0.9037%	0.9991%	1.125%
		testing	4.2849%	6.1357%	7.6275%	9.6702%	9.4255%
	150 pohon	training	0.6149%	0.809%	0.9073%	0.9863%	1.1108%
		testing	4.2723%	6.5787%	7.1848%	9.9783%	9.3771%
SVR-RF	50 pohon	training	0.7483%	0.8319%	0.8475%	0.7867%	0.6615%
		testing	3.9876%	4.7314%	4.3164%	8.9719%	9.3378%
	100 pohon	training	0.7454%	0.7963%	0.845%	0.7625%	0.651%
		testing	3.9693%	4.7484%	4.1302%	8.6391%	9.3082%
	150 pohon	training	0.7402%	0.8104%	0.8208%	0.7554%	0.6414%
		testing	3.9004%	4.6504%	4.1173%	8.7069%	9.1708%
Training Terkecil			0.6149%	0.7963%	0.8208%	0.7554%	0.6414%
Testing Terkecil			1.9119%	4.5691%	4.1173%	8.6391%	9.1708%
Algoritma Terbaik			SVR	SVR	SVR-RF	SVR-RF	SVR-RF

Tabel 4.2. Hasil prediksi keseluruhan SVR vs RF vs SVR-RF pada hari ke (t+n)

metode SVR dan SVR-RF sangat jauh, 21.5306% untuk SVR dan 9.1708% untuk SVR-RF dan jumlah pohon pada metode RF dan SVR-RF mempengaruhi nilai MAPE tetapi tidak terlalu signifikan.

Gabungan SVR-RF didapat kesimpulan sebagai berikut

5. Kesimpulan

A. Kesimpulan

Berdasarkan analisis terhadap implementasi sistem dan analisis system Prediksi Indeks Harga Saham Menggunakan Metode

1. Jika dibandingkan hasil terbaik diantara metode SVR, RF dan metode gabungan SVR-RF, hasil SVR lebih baik untuk prediksi jangka pendek untuk hari ke (t+1) dengan nilai MAPE sebesar 1.9119% dan hari ke (t+5) dengan nilai MAPE sebesar 4.5691%, sedangkan SVR-RF lebih baik untuk prediksi jangka panjang untuk hari ke (t+10) dengan nilai MAPE sebesar

- 4.1173%, hari ke (t+20) dengan nilai MAPE sebesar 8.6391% dan hari ke (t+30) dengan nilai MAPE sebesar 9.1708%.
- Untuk prediksi hari ke (t+20) dan hari ke (t+30) *RF* masih lebih baik daripada *SVR* dengan nilai MAPE terbaik 9.6026% untuk hari ke (t+20) pada 50 pohon dan 9.3617% pada hari ke (t+30) pada 50 pohon.
 - Jumlah pohon yang digunakan pada metode *RF* dan *SVR-RF* mempengaruhi hasil nilai MAPE tetapi tidak terlalu signifikan.

B. Saran

Setelah dilakukan beberapa percobaan saran untuk penelitian ini adalah untuk memprediksi Indeks Harga Saham gunakan metode *SVR* satu tahap untuk prediksi jangka pendek, sedangkan untuk prediksi jangka panjang gunakan metode *SVR-RF*.

Indonesia, Jakarta: PT Bursa Efek Indonesia.

- [7] Santosa, Budi (2007). *Data Mining, Teori dan Aplikasi*. Graha Ilmu, Yogyakarta.
- [8] Parrella, F. (2007). *Online support vector regression (Master's thesis)*. Italy: Department of Information Science, University of Genoa.
- [9] Pring, Martin.(2001), *Introduction to Technical Analysis*. McGraw-Hill. Singapore.
- [10] Breiman, L. (2001), *Random forests, Machine Learning*, Kluwer Academic Publishers, 45, 5-32.
- [11] J. Kennedy, R. C. Eberhart.(1995). *Particle swarm optimization*. In *Proceedings of the 1995 IEEE International Conference on Neural Networks*. IEEE Service Center, Piscataway.

Daftar Pustaka

- [1] Patel, Jigar., Shah, Sahil., Thakkar, Priyank., & Kotecha, K. (2015). Predicting stock market index using fusion of machine learning techniques. *Expert Systems with Applications* 42 (2015) 2162-2172.
- [2] Miao, K., Chen, F., & Zhao, Z.-g. (2007). Stock price forecast based on bacterial colony RBF neural network [j]. *Journal of Qingdao University (Natural Science Edition)*, 2, 011.
- [3] Leung, M. T., Daouk, H., & Chen, A. S. (2000). Forecasting stock indices: A comparison of classification and level estimation models. *International Journal of Forecasting*, 16, 173–190.
- [4] Manish, K., & Thenmozhi, M. (2005). Forecasting stock index movement: A comparison of support vector machines and random forest. In *Proceedings of ninth Indian institute of capital markets conference*, Mumbai, India. <<http://ssrn.com/abstract=876544>>.
- [5] Bursa Efek Indonesia, "IDX," [Online]. Available: <http://www.idx.co.id/id-id/beranda/informasi/bagiinvestor/saham.aspx>. [Diakses 16 03 2015].
- [6] Indonesia Stock Exchange (2008), *Buku Panduan: Indeks Harga Saham Bursa Efek*