

## MODEL PREDIKSI HARGA SAHAM MEDIA SOSIAL BERDASARKAN ALGORITMA SVM YANG DIOPTIMASIKAN DENGAN PSO

Eka Puspita Sari

Program Studi Manajemen Informatika

AMIK BSI Jakarta

Jl. Kramat Raya No.18, Jakarta Pusat

eka.eps@bsi.ac.id

**Abstrak** - Stock is a unit of value or books in a variety of financial instruments that spur on the ownership of a company. By issuing shares allow companies who require long-term financing to 'sell' business interests in shares in exchange (equity securities) for cash. This is the primary method for businesses to raise capital other than issuing bonds. Shares sold through the primary market or secondary market. This study aims to demonstrate the application of optimization Particle Swarm Optimization (PSO) on Support Vector Machine (SVM) can provide better accuracy than just using Support Vector Machine (SVM) algorithm in predicting stock prices. After the test is done by using two model of SVM algorithm dan previously discovered SVMPSO after optimization parameters on SVM, then from experiments shown that the kernel gives the result is a DOT to give an accuracy of 94.8% for training, 94.6% for testing compared with the results obtained from the RBF kernel for 94.4% for training and 90.5% for testing and Polynomial kernel for 58.6% for training and 47.3% for testing.

**Intisari** - Saham adalah satuan nilai atau pembukuan dalam berbagai instrumen financial yang mengacu pada bagian kepemilikan sebuah perusahaan. Dengan menerbitkan saham, memungkinkan perusahaan-perusahaan yang membutuhkan pendanaan jangka panjang untuk 'menjual' kepentingan dalam bisnis saham (efek ekuitas) – dengan imbalan uang tunai. Ini adalah metode utama untuk meningkatkan modal bisnis selain menerbitkan obligasi. Saham dijual melalui pasar primer (*primary market*) atau pasar sekunder (*secondary market*). Penelitian ini bertujuan untuk membuktikan penerapan optimasi Particle Swarm Optimization (PSO) pada Support Vector Machine (SVM) dapat memberikan keakurasian yang lebih baik dari pada hanya menggunakan algoritma SVM dalam memprediksi harga saham. Setelah dilakukan pengujian dengan menggunakan dua model yaitu algoritma SVM dan SVMPSO setelah sebelumnya ditemukan parameter-parameter optimasi pada SVM, maka dari eksperimen yang dilakukan terlihat bahwa kernel yang memberikan hasil terbaik adalah DOT dengan memberikan nilai

akurasi sebesar 94.8% untuk training, 94.6% untuk testing dibandingkan dengan hasil yang diperoleh dari kernel RBF sebesar 94.4% untuk training dan 90.5% untuk testing dan kernel Polynomial sebesar 58.6% untuk training dan 47.3% untuk testing.

**Kata kunci:** Harga Saham, Optimasi Parameter, Particle Swarm Optimization, Support Vector Machines.

### PENDAHULUAN

Saham adalah satuan nilai atau pembukuan dalam berbagai instrumen financial yang mengacu pada bagian kepemilikan sebuah perusahaan. Dengan menerbitkan saham, memungkinkan perusahaan-perusahaan yang membutuhkan pendanaan jangka panjang untuk 'menjual' kepentingan dalam bisnis saham (efek ekuitas) – dengan imbalan uang tunai. Ini adalah metode utama untuk meningkatkan modal bisnis selain menerbitkan obligasi. Saham dijual melalui pasar primer (*primary market*) atau pasar sekunder (*secondary market*). Investasi saham menjadi pilihan investasi yang diminati dewasa ini dikarenakan dapat memberikan keuntungan besar dan cepat. Tetapi banyak faktor ketidakpastian dalam investasi saham yang menyebabkan kerugian dalam berinvestasi. Data harga saham perusahaan merupakan hal yang paling menarik perhatian bagi investor. Dengan mengetahui harga saham, investor dapat mengambil keputusan untuk membeli saham suatu perusahaan atau menjual saham miliknya. Akan tetapi, harga saham bersifat fluktuatif atau berubah-ubah dikarenakan faktor-faktor tertentu.

SVM telah digunakan secara luas untuk peramalan harga saham serta menunjukkan performa yang lebih baik dari pada algoritma lainnya termasuk ANN, dimana ANN sudah terlebih dahulu banyak digunakan untuk proses peramalan termasuk sebagai alternatif yang menjanjikan untuk memprediksi harga saham, dimana ANN menemukan solusi berupa lokal optimal sedangkan SVM menemukan solusi yang global optimal (Santosa, 2007), serta hasil yang

diberikan oleh ANN akan selalu berbeda-beda setiap kali kita melakukan *training*, dikarenakan untuk proses *training* pada ANN berjalan dengan mengoreksi nilai *weights* secara berulang (Santosa, 2007) berbeda dengan SVM yang akan memberikan hasil yang selalu sama, dan juga SVM tidak tergantung pada dimensi ruang input yang membuat SVM fleksibel dalam pemilihan parameter (Premanode, Toumazou, 2012), yaitu parameter kernel dan juga parameter C, berbeda dengan ANN, banyak hal yang perlu di atur nilai parameternya seperti jumlah *hidden layer*, *neuron* untuk *hidden layer*, metode untuk *training*, dan lain lain (Santosa, 2007) dan juga SVM cocok untuk menangani data set yang besar dan kompleks sepertinya halnya prediksi harga saham. Di dalam SVM, pemilihan parameter C,  $\sigma$ ,  $\epsilon$  memberikan dampak yang sangat besar bagi performace SVM (Wei Fei, Bin Miao, Liang Liu, 2009).

Berdasarkan latar belakang yang telah diuraikan terdapat permasalahan yang terjadi, prediksi harga saham adalah hal yang penting dan sangat menarik karena keberhasilan prediksi harga saham dapat menjanjikan keuntungan bagi para investor. Pergerakan harga saham yang bersifat fluktuatif atau berubah-ubah, sehingga diperlukan metode peramalan yang bisa membantu investor dalam mengambil keputusan membeli atau menjual saham.

## BAHAN DAN METODE

Data mining sering juga disebut *Knowledge Discovery in Database* (KDD), adalah kegiatan yang meliputi pengumpulan, pemakaian data historis untuk menemukan keteraturan, pola atau hubungan dalam set data berukuran besar (Santosa 2007), yang disimpan di dalam repositori, menggunakan teknologi pengenalan pola dan juga teknik statistik dan matematika (Larose, 2005). Fungsi-fungsi yang terdapat dalam data mining adalah (Larose, 2005) : Fungsi deskripsi (*Description*), Fungsi estimasi (*Estimation*), Fungsi prediksi (*Prediction*), Fungsi klasifikasi (*Classification*), Fungsi pengelompokan (*Clustering*), Fungsi asosiasi (*Associaton*).

Pada *Conference on Learning Theory* (COLT), Boser, Bernhard, Guyon, dan Vapnik tahun 1992, memperkenalkan SVM (Premanode, Tzoumazou 2012) yaitu sebuah teknik *supervised learning* dari bidang *machine learning* yang dapat di aplikasikan kedalam klasifikasi dan regresi (Sewell, Taylor 2012). SVM merupakan salah satu teknik yang relatif baru untuk melakukan prediksi, SVM berada dalam satu kelas dengan ANN dalam hal fungsi dan kondisi permasalahan yang bisa di selesaikan (Santosa, 2007).

Yang menjadi karakteristik dari *Support Vector Machine* (SVM) adalah sebagai berikut:

1. Secara prinsip SVM adalah *linear classifier*.
2. *Pattern recognition* dilakukan dengan mentransformasikan data pada *input space* ke ruang yang berdimensi lebih tinggi, dan optimisasi dilakukan pada ruang vector yang baru tersebut. Hal ini membedakan SVM dari solusi *pattern recognition* pada umumnya, yang melakukan optimisasi parameter pada ruang hasil transformasi yang berdimensi lebih rendah daripada dimensi *input space*.
3. Menerapkan strategi *Structural Risk Minimization* (SRM).
4. Prinsip kerja SVM pada dasarnya hanya mampu menangani klasifikasi dua *class*.

Secara sederhana konsep SVM adalah sebagai usaha mencari *hyperlane* terbaik yang berfungsi sebagai pemisah dua buah *class* pada *input space*.

Banyak teknik data mining atau machine learning yang dikembangkan dengan asumsi kelinearan, sehingga algoritma yang di hasilkan terbatas untuk kasus-kasus yang linear, maka untuk mengatasi nya kita bisa menggunakan metode kernel. Fungsi kernel yang biasa digunakan dalam SVM :

- a. *Linear* :  $x^T x$ ,
- b. *Polinomial* :  $(x^T x_i + 1)^p$ ,
- c. *Radial basis function* (RBF) :  $\exp(-\frac{1}{2\sigma} \|x - x_i\|^2)$ ,
- d. *Tangent hyperbolic* (sigmoid) :  $\tanh(\beta x^T x_i + \beta_1)$ , dimana  $\beta, \beta_1 \in R$  ?

Fungsi kernel mana yang harus digunakan untuk substitusi dot product di *feature space* sangat bergantung pada data. Biasanya metode *cross validation* (Santosa 2007) digunakan untuk pemilihan fungsi kernel.

*Algoritma Particle Swarm Optimization* (PSO) merupakan teknik optimasi berbasis populasi yang diperkenalkan oleh DR. Eberhart dan Dr. Kennedy pada tahun 1995 (Ratna Wati, 2011) yang terinspirasi oleh perilaku sosial kawanan burung yang mencoba mencapai tujuan yang tidak di ketahui (Jung Hsieh, Fen Hsiao, Chang Yeh, 2012). *Particle Swarm Optimization* (PSO) adalah alat untuk menangani masalah optimasi (Yuan Hsu et al, 2011). Di dalam PSO, populasi disebut sebagai kawanan dan individu-individu sebagai partikel (Zhou, Fang, Li, Zhang, Peng 2009). Kawanan tersebut terdiri atas partikel-partikel volume-less dengan kecepatan stokastik, dimana masing-masing partikel tersebut merepresentasikan sebuah solusi yang layak, dimana algoritma PSO tersebut menemukan solusi optimal melalui pergerakan partikel-partikel di dalam ruang solusi (Yusup,

Zain, Hashim 2012). Untuk mendapatkan solusi yang optimal atau kurang lebih optimal, setiap partikel mengubah arah pencariannya berdasarkan dua faktor, pengalaman terbaiknya dan pengalaman terbaik dari kawanan (Zhou, Fang, Li, Zhang, Peng 2009). Untuk menemukan solusi yang optimal, masing-masing partikel bergerak ke arah posisi sebelumnya terbaik (pbest) dan terbaik posisi global (gbest).

Salah satu pendekatan alternatif untuk "train dan test" yang sering di adopsi dalam beberapa kasus (dan beberapa lainnya terlepas dari ukurannya) yang di sebut dengan k-fold cross validation (Bramer, 2007), dengan cara menguji besarnya error pada data test (santosa, 2007). Kita gunakan k-1 sampel untuk training dan 1 sampel sisanya untuk testing. Misalnya ada 10 subset data, kita menggunakan 9 subset untuk training dan 1 subset sisanya untuk testing. Ada 10 kali training dimana pada masing-masing training ada 9 subset data untuk training dan 1 subset digunakan untuk testing. Dari situ lalu di hitung rata-rata error dan standar deviasi error (Santosa, 2007). Setiap bagian k pada gilirannya digunakan sebagai ujian menetapkan dan k lainnya - 1 bagian digunakan sebagai training set (Bramer, 2007),

Dalam prediksi keberhasilan suatu proses dapat diukur selain dari hasil percobaan yang akurat dan optimal tetapi juga harus diperhitungkan kesalahan (error), berikut adalah error yang sering di gunakan (Bowerman et al, 2005; Santosa, 2007):

1. Mean Squared Error (MSE)

$$MSE = \frac{1}{n} \sum_{i=1}^n (\hat{Y}_i - Y_i)^2 \dots\dots\dots(1)$$

2. Mean Absolute Percentage Error (MAPE)

$$MAPE = \frac{\sum_{i=1}^m APE_i}{m} \dots\dots\dots(2)$$

Dimana

$$APE = \frac{\sum_{i=1}^m |\hat{y}_i - y_i|}{y_i} 100 \dots\dots\dots(3)$$

3. Mean Absolute Deviation (MAD)

$$MAD = \frac{\sum_{i=1}^m |\hat{y}_i - y_i|}{m} \dots\dots\dots(4)$$

4. Sum Square Error (SSE)

$$SE = \sum_{i=1}^n (D_{ij} - f_j(X_i)) \dots\dots\dots(5)$$

5. Root Mean Square Error (RMSE)

$$RMSE = \sqrt{\frac{SSE}{N \cdot K}} \dots\dots\dots(6)$$

6. Normalized Mean Absolute Error (NMAE)

$$NMAE = \frac{MAE}{\bar{P}} = \frac{\sum_{t=1}^N |e(t+k|t)|}{\sum_{t=1}^N P(t)} \dots\dots\dots(7)$$

Pada penelitian ini metode yang akan diusulkan adalah dengan menggunakan Support Vector Machine (SVM) yang akan dioptimasi dengan menggunakan Particle Swarm Optimization (PSO).

Setelah dilakukan proses cleaning, normalisasi, pembagian data, dan juga telah mendapatkan kombinasi-kombinasi optimasi parameter C, γ dan ε, selanjutnya hasil dari proses optimasi tersebut akan di aplikasikan dengan menggunakan algoritma SVMPSO selanjutnya hasil yang didapat akan divalidasi dengan 10 folds x-validation, untuk mendapatkan Root Mean Squared Error (RMSE) dan Normalized Mean Absolute Error (NMRE). Hasil akurasi yang di dapat akan dibandingkan dari masing-masing algoritma, apakah SVMPSO dapat meningkatkan akurasi.

Data yang digunakan dalam penelitian ini adalah data harga saham Yahoo! selama tujuh tahun, yaitu dari 3 Januari 2007 sampai dengan 31 Desember 2013, data yang dikumpulkan adalah harga saham perhari, sehingga terdapat sebanyak 1.762 records. Data tersebut diperoleh dari [finance.yahoo.com](http://finance.yahoo.com). Rata-rata perbulan harga saham Yahoo! dari tahun 2007-2013 digambarkan pada tabel 1.

Tabel 1. Rata-rata perbulan harga saham Yahoo!

Bulan	Tahun													
	2007		2008		2009		2010		2011		2012		2013	
	Open	Close	Open	Close	Open	Close	Open	Close	Open	Close	Open	Close	Open	Close
Jan	28.0	28.3	18.8	19.1	11.8	11.7	15.5	15.0	15.8	16.1	15.5	15.4	19.9	19.6
Feb	30.8	30.8	27.9	27.7	12.6	13.2	15.2	15.3	16.3	16.4	14.8	14.8	21.0	21.3
Mar	31.2	31.2	28.5	28.9	12.7	12.8	16.4	16.5	16.7	16.6	15.3	15.2	23.6	23.5
Apr	28.3	28.0	27.1	27.4	14.1	14.2	17.1	16.5	17.4	17.7	15.5	15.5	24.3	24.7
Mei	28.7	28.7	27.0	26.7	15.1	15.8	15.6	15.3	16.1	16.5	15.2	15.2	26.2	26.3
Jun	27.2	27.1	21.1	20.6	15.8	15.6	13.9	13.8	14.9	15.0	15.6	15.8	25.4	25.1
Jul	23.8	23.2	19.8	19.8	14.7	14.3	13.6	13.8	13.8	13.1	16	15.8	27.9	28.0
Agt	22.8	22.7	19.5	19.3	14.7	14.6	13.1	13.1	13.9	13.6	14.7	14.6	27.3	27.1
Sep	26.4	26.8	17.1	17.3	17.4	17.8	14.2	14.1	13.2	13.1	16.0	15.9	33.0	33.1
Okt	31.5	31.1	13.0	12.8	10.0	15.9	16.3	16.4	16.0	15.6	16.8	16.8	32.4	32.9
Nov	26.9	26.8	10.7	11.5	14.9	14.9	16.2	15.8	15.9	15.7	18.9	18.7	36.9	36.9
Des	23.2	23.2	11.9	12.2	16.9	16.7	16.7	16.6	16.1	16.1	19.4	19.9	40.1	40.4

Sumber : finance.yahoo.com (2014)

Langkah-langkah yang digunakan :

#### 1. Pengumpulan Data

Dalam tahap ini data-data yang dibutuhkan dikumpulkan sesuai dengan kebutuhan agar tujuan penelitian dapat tercapai.

#### 2. Pengolahan Data

Setelah tahap pengumpulan data, data-data tersebut kemudian diolah agar data tersebut dapat di gunakan sebagai data eksperimen, diawali dengan proses *cleaning*, normalisasi, serta menentukan data yang akan di gunakan sebagai *training* dan *testing*.

#### 3. Metode yang digunakan

Dalam tahap ini adalah menentukan metode yang akan digunakan, dimana metode ini merupakan gambaran dari rangkaian kegiatan untuk melakukan pelatihan dan pengujian terhadap data yang sudah dikumpulkan dan diolah.

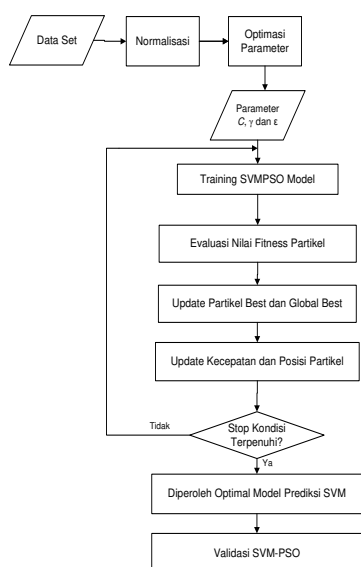
#### 4. Eksperimen dan pengujian model

Dalam tahap ini akan dilakukan eksperimen dan pengujian model terhadap data dengan menggunakan algoritma yang sudah dipilih oleh penulis.

#### 5. Evaluasi

Tahap evaluasi merupakan tahap akhir, dimana dalam tahap ini penulis melakukan evaluasi hasil dari eksperimen dan pengujian model yang digunakan.

Data yang digunakan dalam penelitian ini adalah data harga saham selama tujuh tahun, yaitu dari 1 januari 2007 sampai dengan 31 desember 2013, data yang dikumpulkan adalah harga saham perhari, sehingga terdapat sebanyak 1.762 records. Data tersebut diperoleh dari [www.finance.yahoo.com](http://www.finance.yahoo.com).



Sumber : Kerangka Hasil penelitian (2014)  
 Gambar 2. Metode Yang Diusulkan

Tahap eksperimen yang dilakukan adalah sebagai berikut :

1. Inisialisasi Parameter  $C$ ,  $\gamma$  dan  $\epsilon$

Menentukan nilai parameter  $C$ ,  $\gamma$  dan  $\epsilon$  yang kemudian akan didapat parameter yang optimal untuk mendapatkan hasil terbaik. Semakin besar parameter  $C$  semakin besar finalisasi error untuk memaksimalkan margin (Alamili, 2011) oleh karena itu parameter  $C$  ditetapkan adalah 0 - 100, untuk parameter  $\gamma$  dikarenakan semakin besar  $\gamma$  semakin bagus nilai prediksi pada *training* dan semakin buruk pada validasi (Alamili, 2011) maka ditetapkan *range* untuk  $\gamma$  adalah 0.01 - 1 dan *range* yang sama akan digunakan untuk parameter  $\epsilon$  0.01 - 1. Untuk optimasi parameter pada *Rapidminer* menggunakan operator *Optimize Parameters*, nilai untuk *max generations* adalah 50, *Tournament fraction* adalah 0.25 dan *crossover prob* adalah 0.9, *Population* sebanyak 5.

Dari kombinasi-kombinasi parameter akan di masukkan kedalam algoritma SVM, lalu kombinasi yang memberikan nilai akurasi serta RMSE dan NMAE terbaik maka akan di optimasi dengan menggunakan algoritma PSO.

2. Pemilihan Kernel

Pemilihan kernel berpengaruh terhadap tingkat akurasi serta *root mean squared error*. Kernel yang akan digunakan dalam penelitian ini adalah *Dot*, *Polynomial* dan RBF. Dari parameter-parameter yang didapat dari hasil optimasi parameter akan di uji dengan menggunakan masing-masing algoritma dengan menggunakan kernel tersebut, maka akan terlihat kernel manakah yang paling sesuai dengan data set kita dan mana yang bisa memberikan hasil akurasi yang optimal.

3. Penggunaan Algoritma SVM dan SVMPSO

Setelah optimasi parameter dan juga pemilihan kernel langkah selanjutnya adalah menggunakan algoritma yang sudah di pilih, yaitu SVM dengan nilai parameter yang sama.

Setelah mendapatkan hasil dari penerapan algoritma SVM, selanjutnya adalah penerapan PSO pada SVM dengan menggunakan parameter-parameter yang memberikan hasil terbaik pada penerapan SVM dengan memberikan nilai *Population Size* pada PSO yang berbeda-beda, disini penulis menggunakan *range* 5 - 50 dengan kelipatan sebanyak 5 dan *Max number of generations* sebanyak 50.

HASIL DAN PEMBAHASAN

Penelitian akan dilakukan dengan dua metode, yaitu penelitian dengan menggunakan algoritma *Support Vector Machine* (SVM) dan penggunaan *Particle Swarm Optimization*(PSO) pada SVM. Pengujian model dengan menggunakan kombinasi parameter  $C$ ,  $\gamma$  dan  $\epsilon$  yang diperoleh dari proses optimasi parameter. Masing-masing kombinasi parameter tersebut akan diuji dengan menggunakan kernel *Dot*, *Polynomial* dan *Radial Bases Function*(RBF).

Kombinasi parameter yang di dapatkan adalah sebagai berikut :

Tabel 2.Kombinasi Optimasi Parameter

No.	Gamma	C	Epsilon
1	0.02405156	40.1075386	0.01887688
2	0.09366373	6.16109691	0.06925861
3	0.01824542	96.8492822	0.05620701
4	0.00687633	22.0503833	0.04846603
5	0.06233806	86.3356348	0.05646206
6	0.01785089	86.3468095	0.07355694
7	0.04785582	96.8548177	0.05258414
8	0.08667854	6.15257596	0.06515007
9	0.01149695	96.8476692	0.07165353
10	0.05855341	86.3434104	0.0579597
11	0.01296717	40.0851501	0.01542976
12	0.01485985	96.8428472	0.06416548
13	0.02746757	96.8611303	0.06276041
14	0.1	6.1634302	0.05513132
15	0.01824542	86.3356348	0.05620701
16	0.06233806	96.8492822	0.05646206
17	0.09366373	6.16109691	0.06925861
18	0.01824542	96.8492822	0.05620701
19	0.09618114	6.16952295	0.05955797
20	0.1	6.16880973	0.06072071

21	0.0228764	40.1014465	0.01215962	66	0.1	6.14209442	0.09727257
22	0.01535813	40.0771762	0.01216356	67	0.1	6.14492965	0.09522972
23	0.09193849	6.1664484	0.04778698	68	0.1	6.1381303	0.1
24	0.00999494	40.0986182	0.02017793	69	0.09881227	6.12711906	0.1
25	0.09025234	6.16448801	0.05623495	70	0.0986151	6.14963807	0.09668734
26	0.01359118	40.075911	0.04043826	71	0.09487028	6.13876423	0.1
27	0.1	6.16495729	0.0886845	72	0.09606073	6.142238	0.08777761
28	0.09366373	6.16109691	0.05513132	73	0.1	6.14284083	0.1
29	0.1	6.1634302	0.06925861	74	0.1	6.17220407	0.1
30	0.01296717	40.0851501	0.01542976	75	0.09013238	6.16503751	0.1
31	0.01296717	40.0851501	0.01542976	76	0.1	6.14209442	0.09727257
32	0.09224663	6.15061002	0.08321344	77	0.1	6.14209442	0.09727257
33	0.1	6.17301452	0.05163777	78	0.1	6.14209442	0.09727257
34	0.09243513	40.1016148	0.04540286	79	0.1	6.14209442	0.09727257
35	0.01783987	6.15511181	0.04727452	80	0.1	6.12769323	0.09933874
36	0.01839166	40.0641783	0.04924996	81	0.1	6.13894226	0.08456184
37	0.1	6.1786823	0.05940275	82	0.09057653	6.12647022	0.0920202
38	0.09684942	6.16786295	0.05994631	83	0.08800886	6.15750489	0.1
39	0.08898446	6.16900384	0.05645427	84	0.099098	6.13556706	0.1
40	0.1	6.16205054	0.0873481	85	0.1	6.11866391	0.09508396
41	0.1	6.1634302	0.06925861	86	0.1	6.11900765	0.09226254
42	0.1	6.16880973	0.06072071	87	0.09773283	6.14826915	0.09170352
43	0.09618114	40.075911	0.05955797	88	0.1	6.1227048	0.0920332
44	0.01359118	6.16952295	0.04043826	89	0.1	6.1381303	0.1
45	0.1	6.14209442	0.09727257	90	0.1	6.1381303	0.1
46	0.1	6.18249114	0.08744463	91	0.1	6.12711906	0.09727257
47	0.0938809	6.17454069	0.08837422	92	0.09881227	6.14209442	0.1
48	0.09900437	6.16404779	0.09261282	93	0.09586876	6.10954858	0.1
49	0.08859259	6.17133493	0.08563453	94	0.1	6.13218632	0.1
50	0.08730591	6.16209273	0.09676757	95	0.1	6.1361146	0.09879369
51	0.1	6.17295962	0.07769629	96	0.09956955	6.14871249	0.08898059
52	0.08490413	6.15832061	0.0918869	97	0.08837636	6.13323102	0.1
53	0.0982809	6.14938342	0.07974277	98	0.1	6.13538131	0.1
54	0.1	6.16495729	0.0886845	99	0.08107994	6.13252864	0.1
55	0.1	6.16495729	0.0886845	100	0.1	6.16314022	0.07820982
56	0.09224663	6.16495729	0.08321344	101	0.08405323	6.12668813	0.1
57	0.1	6.15061002	0.0886845	102	0.1	6.1381303	0.1
58	0.08364239	6.1534641	0.09666388	103	0.1	6.1381303	0.1
59	0.09723589	6.13087695	0.09848674	104	0.1	6.1381303	0.1
60	0.1	6.1474898	0.1	105	0.1	6.1381303	0.1
61	0.08843262	6.14269397	0.08388243	106	0.1	6.12125343	0.09482826
62	0.09003008	6.13670093	0.07672718	107	0.0909576	6.16434862	0.08327391
63	0.08887	6.14592493	0.07841052	108	0.1	6.13144394	0.08373482
64	0.08991177	6.1512743	0.1	109	0.08700026	6.14764812	0.1
65	0.1	6.14209442	0.09727257	110	0.1	6.12811242	0.1

111	0.1	6.1251673	0.08914571	124	0.0785737	6.11238682	0.1
112	0.09857413	6.10248144	0.08969158	125	0.1	6.12978878	0.08696708
113	0.1	6.1374344	0.09861191	126	0.1	6.1375058	0.1
114	0.1	6.12815097	0.09710377	127	0.1	6.12058619	0.09529033
115	0.1	6.1381303	0.1	128	0.1	6.13538131	0.1
116	0.1	6.13538131	0.1	129	0.1	6.13538131	0.1
117	0.1	6.13538131	0.1	130	0.1	6.13538131	0.1
118	0.1	6.13538131	0.1	131	0.1	6.13538131	0.1
119	0.1	6.13692476	0.05731598	Sumber : Hasil Olahan (2014)			
120	0.1	6.14297702	0.1	Dari hasil-hasil eksperimen dengan menggunakan kombinasi kombinasi parameter diatas maka didapatkan hasil pengujian terbaik dengan menggunakan kernel-kernel SVM yang terlihat pada tabel3di bawah ini.			
121	0.08981972	6.15061576	0.08875503				
122	0.08269165	6.12921824	0.1				
123	0.1	6.11873681	0.0810105				

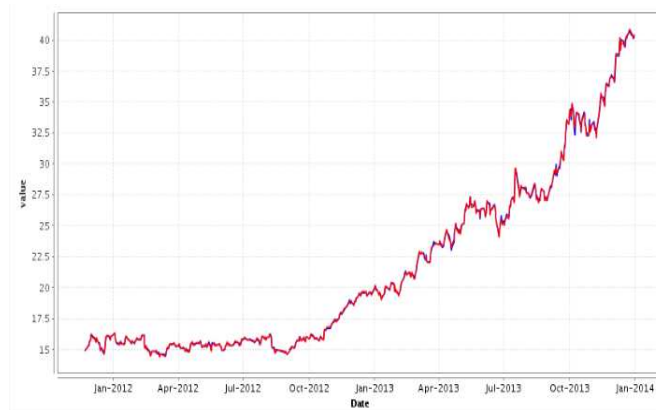
Tabel 3. Hasil Eksperimen Terbaik SVM

	RBF		Dot		Polynomial	
	Trainin ng	Testin g	Trainin g	Testin g	Trainin g	Testin g
Gamma	0.099004366		0.1		0.015358129	
C	6.164047786		6.127119064		40.07717617	
Epsilon	0.092612823		0.097272572		0.012163559	
Bias	23.624	25.368	19.015	21.577	15.319	18.019
RMSE	0.297	0.364	0.187	0.134	4.165	5.725
NMAE	0.033	0.029	0.027	0.017	0.651	0.793
Predicti on Accurac y	94.4	90.5	94.8	94.6	58.6	47.3

Sumber : Hasil Olahan (2014)

Dari hasil eksperimen di atas terlihat bahwa kernel DOT untuk prediksi harga saham memberikan hasil yang lebih baik dari pada kernel RBF dan Polynomial. Dimana pada nilai  $C = 6.127119064$ ,  $\gamma = 0.1$ ,  $\epsilon = 0.097272572$  memberikan nilai akurasi yang terbaik, oleh karena itu untuk penerapan PSO pada SVM penulis akan menggunakan kernel DOT dengan masing-masing nilai parameter tersebut.

Dari hasil eksperimen dengan menggunakan PSO pada SVM dengan kernel DOT, maka didapatkan hasil terbaik yang diperlihatkan oleh tabel-tabel dan grafik-grafik sebagai berikut :



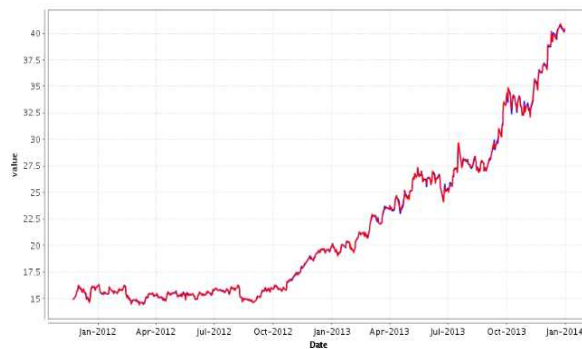
Sumber : Hasil Olahan (2014)

Gambar 3. Hasil Data Testing SVM

Tabel 4. Hasil Eksperimen Terbaik SVMPSO

	SVM(DOT)		SVMPSO(DOT)	
	Training	Testing	Training	Testing
Gamma	0.1			
C	6.127119064			
Epsilon	0.097272572			
Bias	19.015	21.577	19.015	21.577
RMSE	0.187	0.134	0.195	0.133
NMAE	0.027	0.017	0.029	0.016
Prediction Accuracy	94.8	94.6	95.3	95.4
Population Size	15			
Max Generations	50			

Sumber : Hasil Olahan (2014)



Gambar 4. Hasil Data Testing SVMPSO

Dari hasil eksperimen diatas terlihat bahwa PSO dapat meningkatkan nilai akurasi. Terlihat bahwa sebelum menggunakan PSO memberikan nilai akurasi sebesar 94.8% untuk *training* dan 94.6% untuk *testing*, setelah PSO di gunakan maka terjadi peningkatan dengan nilai akurasi sebesar 95.3% untuk *training* dan 95.4% untuk *testing*.

Evaluasi dan Validasi dari harga saham penutupan diperlihatkan pada tabel 5. dibawah ini.

Tabel 5. Perbandingan rata-rata SVM

	RBF		Dot		Polynomial	
	Training	Testing	Training	Testing	Training	Testing
RMSE	0.35195455	0.37051515	0.20474242	0.18581818	27.8832803	5.35325
NMAE	0.04259091	0.03145455	0.0300303	0.02331818	2.88939394	0.81089394

Sumber : Hasil Olahan (2014)

Berdasarkan tabel diatas dapat terlihat bahwa dengan menggunakan kernel DOT baik dilihat pada RMSE dan NMAE memberikan hasil evaluasi yang lebih baik dari pada kernel RBF dan Polynomial.

Hasil Pengujian dengan menggunakan PSO pada SVM dengan menggunakan kernel DOT dapat terlihat pada tabel 4.10. Terlihat bahwa terjadi perbaikan untuk RMSE dari 0.20474242 menjadi 0.1893 pada data *training*, dan dari 0.18581818 menjadi 0.1342 pada data *testing*, untuk NMAE dari 0.0300303 menjadi 0.0274 pada data *training*, dan dari 0.02331818 menjadi 0.0163 pada data *testing*. Terlihat jelas bahwa dengan penerapan PSO pada SVM dapat meningkatkan nilai evaluasi.

Tabel 6. Perbandingan rata-rata SVM dan SVMPSO

	SVM(DOT)		SVMPSO(DOT)	
	Training	Testing	Traini ng	Testi ng
RMS	0.204742	0.185818		0.134
E	42	18	0.1893	2
NMA	0.030030	0.023318		0.016
E	3	18	0.0274	3

Sumber : Hasil Olahan (2014)

Berdasarkan hasil eksperimen yang dilakukan untuk memprediksi harga saham sosial media Yahoo!, dapat disimpulkan bahwa eksperimen dengan menggunakan algoritma SVM dimulai dengan mencari nilai-nilai parameter  $C$ ,  $\gamma$  dan  $\epsilon$



yang dapat memberikan nilai akurasi yang terbaik, dengan cara mencari melalui optimasi parameter dengan memasukkan range 0 – 100 untuk parameter  $C$ , 0 – 0.1 untuk parameter  $\gamma$  dan  $\epsilon$ , nilai untuk *max generations* adalah 50, *Tournament fraction* adalah 0.25 dan *crossover prob* adalah 0.9, *Population* sebanyak 5.

Selanjutnya hasil dari kombinasi parameter-parameter tersebut diuji coba dengan menggunakan algoritma SVM dan menggunakan kernel-kernel yang terdapat pada SVM yaitu, RBF, Dot dan Polynomial. Dari eksperimen tersebut di dapatkan hasil akurasi terbaik dengan menggunakan kernel DOT. Pada data *Training* didapat akurasi 94.8%, RMSE = 0.187 dan NMAE = 0.027 dan pada data *Testing* 94.6%, RMSE = 0.134 dan NMAE = 0.017 dengan nilai  $C=6.1271190636147$ ,  $\gamma = 0.1$  dan  $\epsilon = 0.0972725724082339$ .

Setelah diperoleh nilai parameter yang dapat menghasilkan eksperimen terbaik maka selanjutnya dengan parameter dan kernel yang sama dilakukan uji coba dengan menggunakan algoritma PSO pada SVM. Setelah dilakukan eksperimen dengan menggunakan SVMPSO terjadi peningkatan nilai akurasi dengan menggunakan *max generations* = 50. Maka diperoleh nilai akurasi sebesar 95.3%, RMSE = 0.195 dan NMAE = 0.029 pada data *training*, dan akurasi sebesar 95.4%, RMSE = 0.133 dan NMAE = 0.016 pada data *testing* dengan percobaan menggunakan *population size* = 15.

### KESIMPULAN

Dari hasil penelitian yang dilakukan dari tahap awal hingga pengujian penerapan SVM dan SVMPSO untuk memprediksi harga saham, telah didapatkan jawaban dari pertanyaan-pertanyaan penelitian yang telah diidentifikasi sebelumnya, yaitu :

- Parameter-parameter yang dapat memberikan hasil yang optimal adalah  $C=6.1271190636147$ ,  $\gamma = 0.1$  dan  $\epsilon = 0.0972725724082339$ .
- Dengan menggunakan PSO pada SVM dapat meningkatkan hasil akurasi hal ini diperkuat dengan hasil eksperimen dimana diperoleh nilai akurasi sebesar 95.3% dari sebelumnya 94.8% untuk *training* dan 95.3% dari sebelumnya 94.6% untuk *testing*.
- Dari eksperimen yang dilakukan terlihat bahwa kernel yang memberikan hasil terbaik adalah DOT dengan memberikan nilai akurasi sebesar 94.8% untuk *training*, 94.6% untuk *testing* dibandingkan dengan hasil yang diperoleh dari kernel RBF sebesar 94.4% untuk *training* dan 90.5% untuk *testing* dan kernel Polynomial sebesar

58.6% untuk *training* dan 47.3% untuk *testing*.

Pada penelitian ini secara umum penerapan algoritma SVMPSO dapat meningkatkan akurasi prediksi harga saham, akan tetapi karena keterbatasan penelitian ini perlu disarankan untuk melakukan penelitian selanjutnya untuk mendapatkan akurasi yang lebih baik. Adapun saran-saran yang dapat diberikan, yaitu :

- Penggunaan normalisasi dengan metode yang berbeda terutama untuk kasus seperti prediksi harga saham, agar dapat meminimalisasi nilai RMSE.
- Penggunaan algoritma yang berbeda untuk penelitian sejenis dengan menerapkan nilai parameter-parameter dan iterasi yang berbeda-beda.
- Penelitian ini dapat dikembangkan dengan menggunakan metode optimasi lainnya seperti : *Genetic Algorithm*(GA), *Simulated Annealing*(SA), *Ant Colony Optimization*(ACO) serta *Artificial Bee Colony Algorithm*(ABC).

### REFERENSI

- Achelis S. B. (1995) *Technical Analysis from A to Z*. Probus Publishing, Chicago.
- Alamili, M., (2011). Exchange Rate Prediction using Support Vector Machines. Technische Universiteit Delft.
- Aydin, I., Karakose, M., Akin, E., (2011). A Multi-Objective Artificial Immune Algorithm for Parameter Optimization in Support Vector Machine, *Applied Soft Computing* 11, 120-129.
- Bodt, E., Rynkiewicz, J., Cottrell, M. (2001). *Some known facts about financial data*. European symposium on artificial neural networks 25-27. pp. 223-236.
- Bramer, M., (2007). *Principles of Data Mining*. London : Springer.
- Jung Hsieh, T., Fen Hsiao, H., Chang Yeh, W., (2012). Mining Financial Distress Trend Data using Penalty Guided Support Vector Machines based on Hybrid of Particle Swarm Optimization and Artificial Bee Colony Algorithm, *Neurocomputing* 82, 196-206.
- Kecman, V., (2001). *Learning and Soft Computing*. The MIT Press Cambridge, Massachusetts London, England.
- Kim, K. (2003). Financial time series forecasting using Support Vector Machines. *Neurocomputing* 55, 307-319.
- Larose, D. T., (2005). "Discovering Knowledge in Data," . Canada : Wiley Interscience .

- Lento, C., Gradojevic, N., (2007) "The Profitability Of Technical Trading Rules: A Combined Signal Approach," vol. 23, no. 1, pp. 13-28.
- Liu, L. and Wang, W. (2008). *Exchange Rates Forecasting with Least Squares SVM*. International Conference on Computer Science and Software Engineering.
- Lo, A. W., Mamaysky, H., Wang, J., (2000). "Foundations of Technical Analysis: Computational Algorithms, Statistical Inference, and Empirical Implementation," *The Journal of Finance*, Vol LV, No.4.
- Pilbeam, K. (2010). *Finance and Financial Markets*. Palgrave (ISBN-13 978-0230233218).
- Premanode, B. and Toumazou, C., (2013) "Improving prediction of exchange rates using Differential EMD," *Expert Systems with Applications*, vol. 40, no. 1, pp. 377-384.
- Oanda Corporation (2013). Historical foreign exchange data . Source: <https://fxtrade.oanda.com/trade-forex/fxtrade/historical-data>
- Ratna Wati, D. A., (2011). *Sistem Kendali Cerdas*. Yogyakarta : Graha Ilmu.
- Santosa, B., (2007). *Data Mining Teknik Pemanfaatan Data untuk Keperluan Bisnis*. Yogyakarta : Graha Ilmu.
- Sewell, M. and Shawe-Taylor, J., (2012). "Forecasting foreign exchange rates using kernel methods," *Expert Systems with Applications*, vol. 39, no. 9, pp. 7652-7662.
- Shioda, K., Deng, S. and Sakurai, A. , (2011). "Prediction of Foreign Exchange Market States with Support Vector Machine," *2011 10th International Conference on Machine Learning and Applications and Workshops*, vol. 1, pp. 327-332.
- Supriyanto, H., (2012). "Implementasi Support Vector Machines untuk Memprediksi Arah Pergerakan Harga Harian Valuta Asing (EUR/USD, GBP/USD, dan USD/JPY) dengan Metode Kernel Trick menggunakan Fungsi Kernel Radial Basis Function," pp. 1-4.
- Susanto, S., Suryadi, D., (2010). *Pengantar Data Mining menggali Pengetahuan dari Bongkahan Data*. Yogyakarta : Andi Offset.
- Thissen, U., van Brakel, R., de Weijer, A.P., Melssen, W.J., Buydens, L.M.C. (2003). *Using support vector machines for time series prediction*. *Chemometrics and Intelligent Laboratory Systems* 69, 35 - 49.
- Wei Fei, S., Bin Miao, Y. and Liang Liu, C., (2009). Chinese Grain Production Forecasting Method Based on Particle Swarm Optimization-based Support Vector Machine, *Recent Patents on Engineering*, 3, 8 - 12.
- Yu, L., Wang, S., Huang, W. and Keung Lai, K., (2007). "Are Foreign Exchange Rates Predictable? A Survey From Artificial Neural Networks Perspective," *Scientific Inquiry*, vol. 8, no. 2, pp. 207 - 228
- Yuan Shu, L., Jinn Horng, S., He, M., Fan, P., Wann Kao, T., Khan, M. K., Shine Run, R., Lin Lai, J., Jian Chen, R., (2011). Mutual Funds Trading Strategy based on Particle Swarm Optimization, *Expert System with Applications* 38, 7582 - 7602.
- Yusup, N., Zain, A. M. and Hashim, S. Z. M., (2012). "Overview of PSO for Optimizing Process Parameters of Machining," *Procedia Engineering*, vol. 29, pp. 914-923.
- Zhou, J., Fang, R., Li, Y., Zhang, Y., Peng, B., (2009). Parameters Optimization of Nonlinear Grey Bernoulli Model using Particle Swarm Optimization, *Applied Mathematics and Computation* 207, 292-299.
- Züperl, U., Cüs, F., & Gecevaska, V. (2007). Optimization of the characteristic parameters in milling using the PSO evolution technique. *Journal of Mechanical Engineering* 6, 354-368

#### BIODATA PENULIS



#### **Eka Puspita Sari, M.Kom.**

Lahir di Bukit Tinggi, pada tanggal 10 Mei 1986. Saat ini bekerja sebagai dosen tetap di AMIK BSI Jakarta. Penulis menyelesaikan studi Diploma Tiga (D3) pada tahun 2007 di AMIK BSI Jakarta dengan Program Studi Manajemen

Informatika. Pada tahun 2009, penulis sudah menyelesaikan studi Strata Satu (S1) di STMIK Nusa Mandiri Jurusan Sistem Informatika dengan gelar S.Kom dan studi Strata Dua (S2) di kampus yang sama dengan gelar M.Kom pada tahun 2014.