

Peramalan Harga Saham Menggunakan Metode *Support Vector Machine* (SVM): Studi Kasus Saham di Bursa Efek Indonesia (BEI)

JPMB
61

¹Vincentius Ryan Cokrodiharjo, ^{2*}Dony Abdul Chalid

^{1, 2} Fakultas Ekonomi dan Bisnis, Universitas Indonesia

Paper type
Research paper

Abstract

This study aims to analyze the accuracy of the Support Vector Machine (SVM) model to predict stocks price. A combination of 28 forecasting periods and 30 input period technical indicators was used. Stock transaction data from 31 companies listed on the Indonesian stock exchange and actively traded between March, 2006 and February, 2018 were used. Results show that the highest system performance does not occur when the input period technical indicator is approximately equal to the forecast period. Although, the performance results differ among company stocks. However, the SVM prediction model provides greater profit than buy and hold strategy.

Received: 8 Nov 2020
Accepted: 23 Feb 2021
Online: 26 Feb 2021

Keywords: Stock prices, Prediction, Support Vector Machine, Capital Market, Indonesia

✉ Email korespondensi: donny.abdul@ui.ac.id

Pedoman Sitasi: Cokrodiharjo, V. R., & Chalid, D. A. (2021). Peramalan Harga Saham Menggunakan Metode Support Vector Machine (SVM): Studi Kasus Saham di Bursa Efek Indonesia (BEI). *Jurnal Pasar Modal dan Bisnis*. 3(1), 61 - 74

DOI: <https://doi.org/10.37194/jpmb.v3i1.59>

Publisher:

The Indonesia Capital Market Institute
Indonesia Stock Exchange Building, Tower II, 1st Floor
Jl. Jendral Sudirman Kav. 52-53, Jakarta 12190 - Indonesia



*Jurnal Pasar Modal
dan Bisnis*, Vol 3, No.1,
Februari 2021,
pp. 61 - 74
eISSN 2715-5595

PENDAHULUAN

Peramalan sering diartikan sebagai proses internal untuk membuat prediksi tentang masa depan (Brown, 2014). Secara umum, ada dua metode analisis saham yang digunakan untuk membuat peramalan harga saham, yaitu analisis fundamental dan analisis teknikal. Analisis fundamental fokus pada penentuan nilai wajar suatu bisnis melalui data laporan keuangan perusahaan, sedangkan analisis teknikal berfokus pada pergerakan harga secara statistik. Analisa teknikal berasumsi bahwa harga saham di pasar sudah menunjukkan semua informasi yang tersedia untuk publik. Analisis teknikal mengikuti prinsip bahwa kekuatan permintaan dan penawaran saham di pasar, menentukan harga saham yang terbentuk. Analisis teknikal mengidentifikasi tren pergerakan harga dalam jangka pendek, menengah, dan panjang dan percaya bahwa pergerakan historis akan terulang.

Umumnya peramalan menggunakan metode deret waktu (time-series), dengan melihat data historis. Untuk kondisi pasar yang bergejolak, cara ini kurang bisa memprediksi secara akurat. Akibatnya, metode lain, seperti *machine learning*, telah dikembangkan untuk membantu dalam memprediksi. *Machine learning* adalah bidang ilmu komputer yang menggunakan data dan teknik statistik untuk memberikan kemampuan belajar pada sistem komputer. *Machine learning* menggunakan algoritma yang dapat mempelajari pola, yang mengarah ke prediksi. Pembelajaran yang diawasi adalah jenis machine learning yang digunakan untuk membuat input ke fungsi pemetaan keluaran. Support Vector Machine (SVM) adalah sistem pembelajaran yang menggunakan hipotesis linier ruang dalam ruang fitur berdimensi tinggi, dilatih oleh algoritma pembelajaran teori optimasi, yang mengimplementasikan bias pembelajaran yang diturunkan dari teori pembelajaran statistik (Cristianini dan Shawe-Taylor, 2000). SVM menggunakan fungsi kernel dalam menjalankan proses kerjanya. Beberapa fungsi kernel yang ada adalah linier, Gaussian / kuadrat, dan polinomial kubik. Pemilihan fungsi kernel yang digunakan dalam klasifikasi atau model regresi akan mempengaruhi hasil klasifikasi atau regresi.

Dari studi sebelumnya model prediksi menggunakan SVM menunjukkan tingkat akurasi yang lebih tinggi daripada model *artificial neural network* (Emir, Dinçer, dan Timor, 2012; Nuraini, Jondri, dan Wirayuda, 2015; Shynkevich, McGinnity, Coleman, Belatreche, dan Li, 2017). Meskipun ada juga penelitian lain yang menunjukkan bahwa model *artificial neural network* lebih superior dibandingkan model SVM (Kara, Boyacioglu, dan Baykan, 2011). Keakuratan prediksi juga dipengaruhi oleh faktor dalam pembentukan model prediksi salah satunya adalah waktu input indikator teknikal. Penelitian oleh Shynkevich, *et al.* (2017) menunjukkan bahwa model prediksi paling akurat ketika waktu input indikator teknikal sama dengan periode prediksi. Penelitian sebelumnya oleh Lee (2006) menetapkan bahwa periode input indikator teknis yang optimal adalah tiga hari untuk prediksi jangka pendek dan 20 hari untuk prediksi jangka panjang.

Atas dasar uraian di atas, maka peneliti ingin menganalisis lebih lanjut keakuratan model SVM untuk memprediksi saham di Indonesia. Penelitian sebelumnya dilakukan oleh Kim dan Shamsudin (2008); Yu, Nartea, Gan, dan Yao (2013); dan Shaik dan Maheswaran (2017) mengatakan bahwa pasar Indonesia tidak efisien, sedangkan penelitian yang dilakukan oleh Nikita dan Soekarno (2012); Adrianto dan Miza (2016); Alhashel, Almudhaf, dan Hansz (2018) mengatakan bahwa pasar Indonesia merupakan pasar efisiensi bentuk lemah. Hasil penelitian ini menunjukkan adanya perbedaan kondisi di pasar Indonesia. Penelitian yang dilakukan oleh Alhashel *et al.* (2018) menunjukkan bahwa meskipun Indonesia ada di pasar dengan efisiensi bentuk yang lemah, indikator teknikal memiliki kekuatan untuk membuat prediksi. Selain itu, penelitian yang dilakukan oleh Yu *et al.* (2013) menunjukkan bahwa meskipun Indonesia tidak efisien di pasar, perdagangan teknikal tetap dapat memberikan keuntungan.

Penelitian ini akan menganalisis apakah akurasi tertinggi model prediksi SVM yang digunakan dalam hal ini terjadi pada saat periode waktu input indikator sama dengan waktu prediksi. Penelitian akan menggunakan input indikator teknikal dari tiga hari hingga 30 hari dan periode prediksi dari satu hari hingga 30 hari. Penelitian ini juga akan menganalisis apakah menggabungkan model prediksi dalam penelitian dengan strategi perdagangan sederhana dapat memberikan keuntungan lebih dari pada strategi beli dan tahan (hold and buy strategy).

KAJIAN PUSTAKA

Secara umum, ada dua metode analisis saham yaitu analisis teknikal dan analisis fundamental. Kedua metode tersebut sering kali digunakan untuk mendapatkan informasi tentang saham mana yang menguntungkan. Kedua metode analisis ini memiliki kelebihan dan kekurangan masing-masing. Analisis teknikal berfokus pada penggunaan data historis perdagangan, sedangkan analisis fundamental berfokus pada kinerja perusahaan. Analisis teknikal adalah alat untuk memprediksi pergerakan harga saham di masa depan dengan melihat aktivitas perdagangan seperti pergerakan harga dan volume, melalui statistik atau grafik. Menurut Murphy (1999), analisis teknikal menggunakan filosofi rasionalitas dengan tiga premis: tindakan pelaku pasar bersifat rasional, pergerakan harga mengikuti tren dan ritme tertentu, dan pola tren selalu berulang. Analisis fundamental mengukur nilai intrinsik suatu saham dengan melihat kondisi makro ekonomi seperti kondisi industri atau kondisi mikroekonomi seperti kondisi keuangan perusahaan dan manajemen perusahaan. Tujuan utama dari analisis fundamental adalah untuk melihat apakah harga saham saat ini terlalu tinggi atau terlalu rendah. Analisis fundamental dan analisis teknikal serupa karena keduanya melihat pergerakan saham. Perbedaan antara kedua analisis ini, bagaimanapun, adalah pendekatan yang digunakan; analisis fundamental mempelajari penyebab pergerakan pasar, sedangkan analisis teknikal mempelajari efek perubahan tanpa berusaha memahami penyebabnya.

Salah satu metode yang menggunakan data mining adalah machine learning. Metodologi machine learning menggunakan pasangan input-output sebagai contoh agar komputer memahami hubungan pasangan; ini dikenal sebagai *supervised learning* (Cristianini dan Shawe-Taylor, 2000). Cristianini dan Shawe-Taylor (2000) menyatakan bahwa pasangan masukan-keluaran secara fungsional disebut *training data*, dan pembelajaran dengan *binary output* disebut klasifikasi biner, sedangkan keluaran dengan nilai riil dikenal sebagai regresi. Schölkopf, Smola dan Bach (2002) mendeskripsikan proses klasifikasi sebagai proses untuk menentukan ke dalam kategori data mana bentuk masukan baru cocok.

Support Vector Machine (SVM) merupakan salah satu metode terbaru dari perkiraan harga saham melibatkan *data mining*. *Data mining* adalah proses penting di mana metode cerdas diterapkan untuk mengekstrak pola data (Han, Pei, dan Kamber, 2011). SVM adalah sistem pembelajaran yang menggunakan fungsi linier dari ruang hipotetis dalam ruang fitur berdimensi tinggi yang dilatih dengan algoritma pembelajaran dari teori optimasi dan yang mengimplementasikan bias pembelajaran yang berasal dari teori pembelajaran statistik (*statistical learning theory*) (Cristianini dan Shawe-Taylor, 2000). Implementasi SVM menggunakan prinsip minimalisasi risiko struktural, yang bertujuan untuk meminimalkan kesalahan yang terjadi dan mengklasifikasikannya dengan membuat *hyperplane* untuk memisahkan data menjadi dua kategori. Pengoptimalan *hyperplane* yang benar akan meminimalkan risiko kesalahan klasifikasi data pelatihan dan data pengujian (Cristianini dan Shawe-Taylor, 2000).

Penelitian Huang, Nakamori dan, Wang (2005) menunjukkan bahwa model prediksi yang dibuat dengan menggunakan klasifikasi SVM lebih akurat menentukan pergerakan indeks NIKKEI 225 dibandingkan model lain yang dibentuk dengan menggunakan metode klasifikasi lainnya seperti: *Linear Discriminant Analysis*, *Quadratic Discriminant Analysis* dan *Elman Backpropagation Neural Networks*. Penelitian lain oleh Emir et al. (2012) yang menggunakan data saham di Turki menunjukan model prediksi dengan klasifikasi SVM menunjukkan tingkat akurasi yang lebih tinggi daripada model *artificial neural network*. Penggunaan model prediksi dengan SVM juga dianggap bisa meningkatkan keuntungan strategi investasi yang diambil. Stanković, Marković dan Stojanović (2015) yang menggunakan data saham harian di empat pasar saham negara Eropa timur selama periode 2009 sampai 2013 menunjukkan bahwa hasil strategi perdagangan dengan mengaplikasikan model prediksi SVM bisa memberikan profit lebih besar dibandingkan strategi *Buy & Hold*.

METODE PENELITIAN

Penelitian ini menggunakan data harga saham harian (pembukaan, tertinggi, terendah, dan penutupan) untuk seluruh perusahaan yang aktif melakukan perdagangan di pasar saham Indonesia mulai 1 Maret 2006 hingga 26 Februari 2018. Data harga saham harian dikumpulkan menggunakan *Thomson Reuters Datastream*. Saham yang tidak aktif atau belum diperdagangkan sejak Maret 2006 dikeluarkan. Data tersebut diverifikasi untuk kelengkapan data harga saham harian. Memilih saham dengan batas data yang hilang tidak lebih dari 0,2, terpilihlah 31 saham untuk penelitian ini, seperti yang ditunjukkan pada Tabel 1.

Table 1. Sample Penelitian

No.	Nama Saham	No.	Nama Saham
1	Adhi Karya	17	Indofood Sukses Makmur
2	Aneka Tambang	18	Indosat
3	Astra Argo Lestari	19	Kalbe Farma
4	Astra International	20	Lippo Karawaci
5	Bank Central Asia (BCA)	21	Medco Energi International
6	Bank CIMB Niaga	22	Panin Financial
7	Bank Danamon	23	Perusahaan Gas Negara
8	Bank Mandiri	24	PP London Sumatra Indo
9	Bank Pan Indonesia	25	Semen Gresik
10	Bank Rakyat Indonesia	26	Tambang Batubara Bukit Asam
11	Citra Development	27	Telekomunikasi Indonesia
12	Gajah Tunggal	28	Timah
13	Gudang Garam	29	Unilever Indonesia
14	Holcim Indonesia	30	United Tractors
15	ICT Tunggal Prakarsa	31	Vale Indonesia
16	Indah Kiat Pulp & Paper		

Sumber: diolah

Selanjutnya dilakukan pengolahan data awal guna menyiapkan data yang akan digunakan untuk membuat model prediksi yaitu data masukan. Langkah-langkah berikut dilakukan dalam pemrosesan data awal:

1. Memperbaiki data yang hilang.

Pada tahap seleksi dipilih 31 saham dengan kriteria data hilang 0,2%. Penelitian ini menggunakan metode *last observasi carry forward* (LOCF) atau metode regresi, berdasarkan kondisi kehilangan data untuk masing-masing saham. Jika semua data harga (opening, high, low, dan closing) dalam 1 hari hilang, maka metode regresi digunakan untuk mengganti data yang hilang dengan nilai estimasi berdasarkan informasi lain yang tersedia (Kang, 2013). Jika hanya sebagian dari data harga yang hilang dalam 1 hari, peneliti mengecek terlebih dahulu

untuk melihat apakah sebagian data yang tersedia pada 1 hari tersebut memiliki nilai yang sama dengan hari sebelumnya. Kemudian, metode LOCF digunakan untuk mengganti data yang hilang dengan nilai pengamatan sebelumnya, dengan asumsi nilai keluaran tidak akan berubah dengan data yang hilang (Kang, 2013).

2. Mengatur input berjangka waktu dan periode prediksi.

Dalam penelitian ini, periode input indikator teknikal yang digunakan adalah dari tiga hingga 30 hari perdagangan sebelumnya dan periode prediksi dari satu hingga 30 hari ke depan. Periode ini digunakan untuk semua 31 saham yang dipilih.

3. Transformasi data.

Data harga saham harian diubah menjadi indikator teknikal dengan periode input mulai dari 3 hari hingga 30 hari perdagangan sebelumnya. Ada sepuluh indikator teknikal yang digunakan dalam penelitian ini: SMA, EMA, ATR, ADMI, ROC, RSI, osilator% R William, Stochastic% K, dan Stochastic% D.

4. Data standardisasi.

Karena kesepuluh indikator teknikal digunakan sebagai fitur masukan dalam pembuatan model prediksi menggunakan *machine learning*, maka data harus dilakukan pengecekan ulang menggunakan metode standarisasi agar data memiliki nilai rata-rata nol dan nilai standar deviasi satu. Proses standarisasi data menggunakan Persamaan 1 berikut:

$$x_{new} = \frac{x - \mu}{\sigma}, \quad (1)$$

Di mana,

μ = Rata-rata

σ = Standar Deviasi

5. Pelabelan pergerakan harga Selanjutnya, pergerakan harga saham diberi label “up trend” atau “down trend” untuk dijadikan keluaran model prediksi. Persamaan 3.2 menunjukkan perhitungan proses pelabelan:

$$Labeling = \begin{cases} Up\ Trend, & \text{if } \frac{(C_{t+s} - C_t)}{C_t} > 0; \\ Down\ Trend, & \text{if } \frac{(C_{t+s} - C_t)}{C_t} \leq 0; \end{cases} \quad (2)$$

Di mana:

C_t = Harga penutupan saya pada hari t

C_{t+s} = Harga penutupan saham pada hari $t+s$

s = Periode Prediksi

Pembuatan Model Prediksi SVM

Pembangunan model prediksi SVM menggunakan classification learner function pada software MATLAB. Dalam membuat model prediksi SVM, diperlukan parameter regularisasi yang mengontrol trade-off antara meminimalkan kesalahan klasifikasi dan memaksimalkan margin. Parameter ini dilambangkan dengan huruf c dan dikenal sebagai pengali *Lagrange*, yang menunjukkan vektor pendukung dalam mengoptimalkan pemisahan *hyperplane*. Software MATLAB memiliki fungsi otomatis untuk menentukan nilai parameter c . Fungsi otomatis ini akan membantu mengidentifikasi dan memperkirakan nilai c yang menunjukkan model prediksi yang paling pas. Sebelum dilakukan pemodelan prediktif, data dari tahap pengolahan data awal dibagi menjadi dua bagian yaitu *training data* dan *test data*. Training data digunakan untuk membuat

model prediksi, dan test data akan digunakan sebagai input dalam model prediksi. Model akan menggunakan 2.440 buah *training data* dan 450 buah *test data*. 2.440 entri diproses untuk setiap periode indikator input teknis dan setiap periode prediksi terdiri dari data pelatihan. Di antara data pelatihan, ada dua jenis: indikator teknis dan label pergerakan saham. Indikator teknis digunakan sebagai variabel input, dan label pergerakan saham digunakan sebagai variabel output. Kumpulan *training data* ini kemudian dimasukkan ke dalam perangkat lunak MATLAB untuk melakukan proses pemodelan prediksi SVM. Untuk membuat model prediksi dalam MATLAB, data yang dibutuhkan harus dalam bentuk matriks. Data pelatihan diimpor ke MATLAB membentuk matriks 2440×11 . MATLAB menyediakan fungsi klasifikasi dengan sistem validasi 5 kali lipat, fungsi SVM linier, dan fungsi pemilihan kernel otomatis. Fungsi sistem validasi membantu menghindari terjadinya *overfitting* pada model yang dihasilkan. Prediksi dengan Model SVM Langkah keempat dalam penelitian ini adalah membuat prediksi dengan model SVM. 450 lembar test data tersebut digunakan sebagai data masukan untuk setiap periode waktu masukan, berupa matriks 450×10 dalam MATLAB. Pemrosesan dan Analisis Prediksi Langkah terakhir dalam penelitian ini adalah mengolah dan menganalisis hasil yang dihasilkan oleh model prediksi. Kinerja model prediksi dapat diperiksa dengan dua cara, yaitu dengan melihat besarnya akurasi model prediksi dan keuntungan yang akan dihasilkan dari perdagangan yang diprediksi. Uraian tentang pengolahan dan analisis prediksi untuk setiap bagian adalah sebagai berikut: 1. Akurasi Akurasi arah adalah persentase kecocokan atau kesamaan dalam tren (naik / turun) antara model prediksi dan tren harian aktual dari saham (Mittal dan Goel, 2012). Akurasi persentase yang tinggi berarti model prediksi berhasil memprediksi perubahan dan tren harga saham. Perhitungan akurasi untuk model prediksi ditunjukkan pada Persamaan 3.3:

$$Accuracy = \frac{(Number\ of\ Correct\ Up\ Trend\ Prediction + Number\ of\ Correct\ Down\ Trend\ Prediction)}{Number\ of\ Test\ Data} \times 100\%$$

Setelah menghitung akurasi setiap prediksi saham untuk setiap kombinasi periode waktu input indikator teknis dan setiap periode waktu prediksi, model tersebut diperiksa akurasi lebih dari 50%. Menurut Madge (2015), dalam bentuk teori pasar efisien yang lemah, arah pergerakan harga saham di masa depan tidak dapat diprediksi dengan akurasi lebih dari 50%. Dengan kata lain, prediksi tersebut tidak jauh berbeda dengan tebakan acak sederhana. Untuk menunjukkan hasil penghitungan akurasi digunakan skala tiga warna. Merah menunjukkan akurasi lebih kecil atau sama dengan 50%, hijau menunjukkan akurasi antara 50% dan 75%, dan biru menunjukkan akurasi hasil lebih dari 75%. 1. Keuntungan perdagangan Analisis keuntungan perdagangan dibagi menjadi tiga bagian, tes prediksi profitabilitas, tes aplikasi profit, dan tes aplikasi. Menghitung profitabilitas model prediksi menggunakan dua kondisi. Jika model memprediksi harga akan naik, saham dibeli pada saat ini dan dijual pada saat prediksi. Jika prediksinya adalah harga bergerak turun maka saham dijual pada saat ini dan dibeli pada saat prediksi. Hanya harga penutupan harga saham yang digunakan untuk menghitung profitabilitas model prediksi. Perhitungan keuntungan adalah sebagai berikut pada Persamaan 3.4:

$$R_{t,s} = \begin{cases} \text{if Up Trend Prediction, } (C_{t+s} - C_t)/C_t; \\ \text{if Down Trend Prediction, } (C_t - C_{t+s})/C_t; \end{cases} \quad (3.4)$$

Where:

$R_{t,s}$ = Trading profit

C_t = Closing price on the day t

$C_t + s$ = Closing price on the day $t + s$

s = Prediction period

Aturan perdagangan sederhana digunakan untuk menguji keuntungan dari penerapan model prediksi, dikombinasikan dengan aturan perdagangan sederhana. Aturan pertama adalah jika prediksi harga penutupan naik, periksa harga pembukaan keesokan harinya. Jika harga pembukaan kurang dari harga penutupan hari sebelumnya, maka aplikasi secara otomatis membeli saham di tempat dan menjualnya pada waktu penutupan pada harga penutupan yang diperkirakan. Aturan kedua adalah jika prediksi harga penutupan turun, periksa harga pembukaan di hari berikutnya. Jika harga pembukaan lebih besar dari harga penutupan hari sebelumnya, maka secara otomatis saham tersebut langsung dijual dan dibeli pada saat penutupan pada harga penutupan yang diprediksi. Uji ANOVA dua arah digunakan tanpa interaksi antara jumlah rata-rata pengembalian terakumulasi dalam hasil yang diprediksi dan pengembalian dari strategi beli dan tahan. Tes ini untuk memeriksa apakah penerapan kombinasi model prediktif dan aturan perdagangan sederhana dapat memberikan persentase return yang signifikan dibandingkan dengan menggunakan strategi buy and hold. Hipotesis berikut dibuat agar return dari model prediksi lebih besar daripada strategi *buy and hold*, yaitu:

- *ANOVA Hypothesis column:*
 $H_0: \mu_{\text{model application}} = \mu_{\text{buy and hold}}$, tidak ada perbedaan signifikan antara jumlah keuntungan yang diperoleh dari pengaplikasian model dan strategi *buy and hold*.
 $H_1: \mu_{\text{model application}} \neq \mu_{\text{buy and hold}}$, ada perbedaan signifikan antara jumlah keuntungan yang diperoleh dari pengaplikasian model dan strategi *buy and hold*.
- *ANOVA Hypothesis line:*
 $H_0: \mu_{\text{stock1}} = \mu_{\text{stock2}} = \mu_{\text{stock}} \dots = \mu_{\text{stock21}}$, tidak ada perbedaan signifikan antara jumlah keuntungan yang dihasilkan dari pemilihan saham yang digunakan.
 $H_1: \mu_{\text{stock1}} \neq \mu_{\text{stock2}} \neq \mu_{\text{stock}} \dots \neq \mu_{\text{stock21}}$, ada perbedaan signifikan antara jumlah keuntungan yang dihasilkan dari pemilihan saham yang digunakan.

HASIL DAN PEMBAHASAN

Hasil penelitian menjawab hipotesis yang diajukan berdasarkan teknik analisis yang sudah dijelaskan. Hasil penelitian dilanjutkan dengan pembahasan yang memadai. Keakuratan model prediksi dilihat dengan menghitung seberapa besar hasil prediksi tersebut sama dengan kondisi sebenarnya. Penghitungan akurasi model prediksi dilakukan untuk semua 31 saham periode 4 Maret 2016 hingga 12 Januari 2018 (450 lembar test data yang digunakan).

Dari hasil perhitungan yang ditunjukkan pada Tabel 2, 25 dari 31 saham (80,6%) memiliki rata-rata hasil prediksi akurasi lebih dari 50%. Hasil tersebut menunjukkan bahwa model prediksi SVM dengan menggunakan data input indikator teknikal sebenarnya dapat digunakan untuk memprediksi pergerakan saham di Indonesia. Namun tidak semua saham dapat memanfaatkan model prediksi SVM yang digunakan dalam penelitian karena proses klasifikasi model prediksi SVM yang digunakan membutuhkan informasi dari data historis sebagai training data. Jika data pengujian sangat berbeda dengan data pelatihan, maka akan menyebabkan kesalahan prediksi.

Dilihat dari hasilnya, saham yang memberikan akurasi tertinggi sebagian besar adalah saham yang pergerakan harganya cukup stabil dan bergerak naik dari tahun ke tahun; Data

historis saham tersebut digunakan sebagai training data untuk membuat model prediksi yang baik. Misalnya saham dengan akurasi tertinggi adalah BCA dengan rata-rata akurasi prediksi 67,4%. Pergerakan harga saham dari tahun ke tahun untuk saham ini dapat dilihat pada Gambar 1.

Beberapa saham menunjukkan hasil rata-rata yang rendah atau akurasi kurang dari 50%. Hasil akurasi yang rendah ini berarti model prediksi tidak mampu memprediksi dengan baik saham tertentu selama periode pengujian karena perubahan informasi yang besar. Misalnya, Indah Kiat memiliki hasil akurasi terendah (38,2%). Rendahnya akurasi dalam memprediksi pergerakan harga saham Indah Kiat disebabkan adanya kenaikan harga yang sangat besar mulai tahun 2017.

Tabel 2. Rata-rata Akurasi Model Prediksi

	Nama Saham	Average of Accuracy		Nama Saham	Average of Accuracy
1	Bank Central Asia	67.4%	17	Gudang Garam	53.7%
2	Bank Rakyat Indonesia	65.8%	18	Vale Indonesia	53.4%
3	Bank Mandiri	63.6%	19	Bank Danamon	52.8%
4	Tambang Batubara Bukit Asam	63.0%	20	Semen Gresik	52.5%
5	Unilever Indonesia	60.9%	21	Bank CIMB Niaga	52.0%
6	Bank Pan Indonesia	57.5%	22	Indofood Sukses Makmur	51.2%
7	Kalbe Farma	57.3%	23	Lippo Karawaci	50.7%
8	Indosat	56.7%	24	Holcim Karawaci	50.4%
9	Panin Financial	56.2%	25	PP London Sumatra Indo	50.1%
10	Gajah Tunggal	55.7%	26	Citra Development	50.0%
11	Timah	55.2%	27	Adhi Karya	48.9%
12	Telekomunikasi Indonesia	54.9%	28	Medco Energi International	44.6%
13	Aneka Tambang	54.9%	29	Astra Argo Lestari	42.4%
14	ICT Tunggal Prakarsa	54.2%	30	Perusahaan Gas Negara	42.4%
15	United Tractors	53.9%	31	Indah Kiat Pulp & Paper	38.2%
16	Astra International	53.9%			

Sumber: diolah

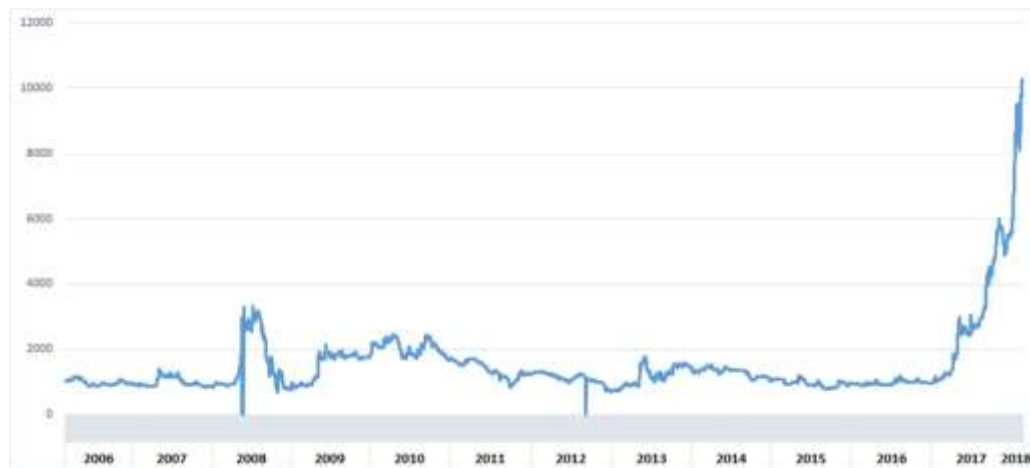
Gambar 1. Pergerakan Harga Saham BCA 1 Maret 2006–26 Februari, 2018



Sumber: diolah

Pergerakan harga saham perusahaan kertas Indah Kiat dari tahun ke tahun dapat dilihat pada Gambar 2. Dibandingkan dengan historis data yang digunakan sebagai training data saham Indah Kiat, pergerakan harga saham cukup stabil. Perbedaan pergerakan yang signifikan ini menyebabkan akurasi prediksi menjadi rendah. Akurasi model prediksi yang diperoleh dari penelitian ini berbeda dengan hasil Shynkevich *et al.* (2017). Penelitian mereka menemukan bahwa model prediksi paling akurat terjadi ketika waktu input indikator teknikal sama dengan periode prediksi. Dalam penelitian ini hasilnya tidak menunjukkan kondisi yang sama. Perbedaan hasil dimungkinkan karena perbedaan kondisi pasar dan saham yang digunakan dalam penelitian. Hasil penelitian menunjukkan bahwa saham yang memiliki rata-rata akurasi tinggi terjadi saat periode prediksi semakin besar karena pergerakan harga saham semakin meningkat sedangkan saham yang memiliki rata-rata akurasi rendah terjadi saat *time frame* prediksi semakin mengecil akibat pergerakan harga saham. Ini bisa sangat berbeda dalam jangka waktu yang lama.

Gambar 2. Pergerakan Harga Saham Indah Kiat 1 Maret 2006–26 Februari, 2018



Sumber: diolah

Tabel 3. Tingkat Keuntungan Rata-rata Model Prediksi

Share Name	Profit	Share Name	Profit
1 Tambang Batubara Bukit Asam	4.7%	17 Panin Financial	0.5%
2 Vale Indonesia	1.9%	18 Perusahaan Gas Negara	0.5%
3 Aneka Tambang	1.9%	19 Astra International	0.4%
4 BCA	1.9%	20 Citra Development	0.3%
5 Bank Rakyat Indonesia	1.9%	21 PP London Sumatra Indo	0.2%
6 Bank Mandiri	1.8%	22 Gajah Tunggal	0.2%
7 Kalbe Farma	1.0%	23 Bank Danamon	0.2%
8 Timah	0.9%	24 Bank CIMB Niaga	0.0%
9 United Tractors	0.9%	25 Holcim Indonesia	-0.3%
10 Unilever Indonesia	0.9%	26 Indofood Sukses Makmur	-0.3%
11 Indosat	0.9%	27 Astra Argo Lestari	-0.4%
12 ICT Tunggal Prakarsa	0.8%	28 Adhi Karya	-0.6%
13 Bank Pan Indonesia	0.8%	29 Lippo Karawaci	-0.7%
14 Semen Gresik	0.7%	30 Indah Kiat Pulp & Paper	-6.9%
15 Telekomunikasi Indonesia	0.6%	31 Medco Energi International	-7.4%

16 Gudang Garam

0.5%

Sumber: diolah

Hasil akurasi tertinggi untuk BCA dan hasil akurasi terendah untuk Indah Kiat. Akurasi hasil saham jangka panjang BCA lebih besar karena pergerakan harga saham BCA terus meningkat dari tahun ke tahun sehingga informasi antara training data dan test data serupa, sedangkan keakuratan hasil saham jangka panjang Indah Kiat lebih rendah karena dari kenaikan harga saham Indah Kiat yang besar pada tahun 2017, sehingga informasi antara *training data* dan *test data* sangat berbeda.

Aplikasi Model Prediksi SVM Bagian ini akan menjelaskan hasil penerapan model prediksi SVM jika dilihat dari sisi besar persentase return. Penjelasan dibagi menjadi dua bagian yaitu menghitung profitabilitas model prediksi, dan pengujian apakah model prediksi memberikan return yang signifikan dibandingkan dengan strategi *buy and hold*. Pengecekan profitabilitas model prediksi dilihat dari rata-rata persentase imbal hasil setiap perdagangan. Perhitungan dilakukan dengan cara menghitung rata-rata besar persentase imbal hasil dari harga penutupan pada saat waktu prediksi dengan harga penutupan saat waktu input. Aturan perhitungannya ada dua yaitu aturan jika prediksi harga naik dan jika prediksi harga turun. Jika prediksi harga naik, maka dilakukan pembelian saham dengan harga penutupan saat waktu input dan dijual pada saat waktu prediksi. Jika prediksi harga turun maka dilakukan penjualan saham dengan harga penutupan saat waktu input dan dibeli pada saat waktu prediksi. Hasil rata-rata profitabilitas model prediksi dapat dilihat pada Tabel 3, yang menunjukkan bahwa 23 dari 31 saham (74,2%) profitabilitas lebih dari 0%. Hasil ini menunjukkan bahwa model prediksi memiliki kemampuan untuk menghasilkan keuntungan jika digunakan dalam perdagangan.

Tahap Selanjutnya adalah menguji apakah model prediksi memberikan *return* yang signifikan dibandingkan dengan strategi *buy and hold* dengan menggunakan uji ANOVA dua arah tanpa interaksi. Sebelum melakukan uji ANOVA, terlebih dahulu dilakukan perhitungan rata-rata persentase akumulasi pengembalian model aplikasi untuk setiap saham dan persentase pengembalian strategi *buy and hold*. Hasil perhitungan dapat dilihat pada Tabel 4 yang menunjukkan bahwa 21 dari 31 saham (67,74%) memiliki rata-rata keuntungan yang lebih besar dengan menggunakan aplikasi model dibandingkan dengan menggunakan strategi *buy and hold*.

Tabel 4. Akumulasi Keuntungan Model Prediksi Dibandingkan Strategi Buy and Hold Selama Periode Penelitian

	Share Name	Average of Profit Accumulation	Buy & Hold			
			Open	Closed	Profit	%Profit
1	Tambang Batubara Bukit Asam	620.0%	1,185	2,890	1,705	143.9%
2	Bank Centra Asia	280.9%	13,600	22,425	8,825	64.9%
3	Bank Rakyat Indonesia	279.4%	2,300	3,540	1,240	53.9%
4	Bank Mandiri	252.1%	4,925	8,100	3,175	64.5%
5	Aneka Tambang	233.3%	406	665	259	63.8%
6	United Tractors	203.5%	14,975	37,025	22,050	147.2%
7	Kalbe Farma	201.0%	1,325	1,695	370	27.9%
8	Indosat	191.2%	5,675	6,000	325	5.7%
9	Unilever Indonesia	175.4%	45,750	54,250	8,500	18.6%
10	Timah	123.0%	620	885	265	42.7%
11	PP London Sumatra Indo	111.0%	1,500	1,390	-110	-7.3%
12	Panin Financial	110.8%	165	234	69	41.8%
13	Astra International	97.0%	6,850	8,175	1,325	19.3%
14	Semen Gresik	92.0%	10,475	10,575	100	1.0%

15	Telekomunikasi Indonesia	85.7%	3,400	4,130	730	21.5%
16	Citra Development	76.0%	1,337	1,175	-162	-12.1%
17	Gudang Garam	72.1%	66,500	83,050	16,550	24.9%
18	Adhi Karya	61.5%	2,640	2,030	-610	-23.1%
19	Indofood Sukses Makmur	39.3%	7,500	7,850	350	4.7%
20	ICT Tunggal Prakarsa	38.0%	20,450	22,700	2,250	11.0%
21	Holcim Indonesia	1.4%	1,015	825	-190	-18.7%

Sumber: diolah

Dari hasil uji ANOVA dua arah tanpa interaksi pada Tabel 5, terdapat dua bagian penjelasan yaitu, ringkasan dan ANOVA. Bagian ringkasan hanya menyajikan deskripsi masing-masing saham (baris) dan strategi yang digunakan (kolom) dengan angka, rata-rata, dan standar deviasi. Bagian ANOVA menyajikan penjelasan apakah hipotesis ditolak atau tidak. Dengan melihat p -value, jika p -value lebih besar dari nilai α (α yang digunakan 0,05) maka keputusan H_0 gagal ditolak dan H_1 ditolak, sedangkan jika p -value lebih kecil dari nilai α (α yang digunakan adalah 0,05), maka keputusan H_0 ditolak dan H_1 gagal ditolak. Melihat nilai F dengan F_{crit} juga dapat menjelaskan apakah hipotesis gagal atau tidak. Jika nilai F lebih kecil dari F_{crit} , maka H_0 gagal ditolak dan H_1 ditolak, sedangkan jika nilai F lebih besar dari F_{crit} , maka H_0 ditolak dan H_1 gagal ditolak. Untuk hasil uji ANOVA kolom menunjukkan bahwa nilai p (0,000017) lebih kecil dari nilai α (0,05) dan nilai F (31,42) lebih besar dari F_{crit} (4,35); oleh karena itu H_0 ditolak dan H_1 diterima. Artinya terdapat perbedaan yang signifikan antara keuntungan yang dihasilkan dari penggunaan model aplikasi dengan strategi *buy and hold*. Hasil uji ANOVA kedua menunjukkan p -value (0,012) lebih kecil dari nilai α (0,05) dan nilai F (2,803) lebih besar dari F_{crit} (2,124), sehingga H_0 ditolak dan H_1 gagal ditolak. Artinya terdapat perbedaan yang signifikan antara besarnya keuntungan yang dihasilkan dari pemilihan saham yang digunakan. Dari kedua hasil tersebut dapat disimpulkan bahwa terdapat perbedaan yang signifikan jumlah keuntungan yang dihasilkan dari penggunaan model aplikasi dan strategi *buy and hold* serta dari pemilihan saham yang digunakan. Hasil penelitian ini sejalan dengan penelitian Stanković et al. (2015) yang menunjukan bahwa strategi investasi dengan menggunakan model prediksi SVM bisa memberikan profit lebih besar dibandingkan strategi *Buy & Hold*.

Tabel 5. Hasil Uji ANOVA Two Factor

SUMMARY	Count	Sum	Average	Variance
TAMBANG BATUBARA BUKIT				
ASAM	2	7.698	3.849	11.620
BCA	2	3.458	1.729	2.333
BANK RAKYAT INDONESIA	2	3.334	1.667	2.543
BANK MANDIRI	2	3.165	1.583	1.759
ANEKA TAMBANG	2	2.971	1.485	1.437
UNITED TRACTORS	2	3.507	1.754	0.158
KALBE FARMA	2	2.289	1.145	1.498
INDOSAT	2	1.969	0.985	1.720
UNILEVER INDONESIA	2	1.940	0.970	1.230
TIMAH	2	1.657	0.829	0.322
PP LONDON SUMATRA INDO	2	1.037	0.519	0.701
PANIN FINANCIAL	2	1.526	0.763	0.238
ASTRA INTERNATIONAL	2	1.164	0.582	0.302
SEMEN GRESIK	2	0.929	0.465	0.414
TELEKOMUNIKASI INDONESIA	2	1.071	0.536	0.206

CITRA DEVELOPMENT	2	0.639	0.320	0.389
GUDANG GARAM	2	0.970	0.485	0.112
ADHI KARYA	2	0.384	0.192	0.358
INDOFOOD SUKSES MAKMUR	2	0.440	0.220	0.060
ICT TUNGGAL PRAKARSA	2	0.490	0.245	0.037
HOLCIM INDONESIA	2	-0.173	-0.087	0.020
<i>Prediction Model Implementation</i>	21	33.507	1.596	1.820
<i>Buy and Hold</i>	21	6.960	0.331	0.210

ANOVA						
Source of Variation	SS	df	MS	F	P-value	F crit
Rows	29.93099598	20	1.4965	2.8033	0.0129	2.1242
Columns	16.77861558	1	16.7786	31.4296	0.0000	4.3512
Error	10.67695061	20	0.5338			
Total	57.38656216	41				

Sumber: diolah

KESIMPULAN

Berdasarkan percobaan yang dilakukan terhadap 31 saham, 25 saham memiliki akurasi rata-rata lebih dari 50%. Akurasi tertinggi dihasilkan oleh saham-saham perbankan besar karena saham-saham tersebut memiliki fundamental dan kinerja yang baik sehingga pergerakan harga sahamnya stabil dan naik dengan mantap. Akurasi terendah dihasilkan oleh saham-saham yang mengalami perubahan informasi atau harga yang besar pada periode data pengujian dibandingkan dengan data pelatihan. Hasil perhitungan akurasi rata-rata model prediksi juga menunjukkan bahwa model prediksi kinerja tertinggi tidak dihasilkan bila periode waktu input indikator teknikal sama dengan periode waktu prediksi seperti yang terdapat pada penelitian sebelumnya.

Penerapan model prediksi SVM dengan kombinasi aturan perdagangan sederhana menunjukkan bahwa 21 saham menghasilkan hasil rata-rata keuntungan dan akumulasi keuntungan yang lebih besar dari pada strategi *buy and hold*. Pengujian ANOVA yang dilakukan pada kelompok saham yang memiliki keuntungan positif atau lebih besar menggunakan kombinasi model prediksi SVM dan aturan perdagangan sederhana dibandingkan dengan strategi *buy and hold* (21 saham), menunjukkan bahwa terdapat perbedaan yang signifikan antara keuntungan yang dihasilkan dari SVM. Hasil penelitian menunjukkan bahwa model prediksi SVM dapat digunakan oleh investor di Indonesia dalam mengambil keputusan untuk membeli atau menjual saham karena dapat memprediksi pergerakan sebagian besar saham di Indonesia. Selain itu, model ini dapat digunakan oleh trader perorangan yang ingin mendapatkan keuntungan lebih besar daripada saat menggunakan strategi beli dan tahan.

Penelitian ini memiliki beberapa keterbatasan. Pertama, kesepuluh indikator teknikal yang digunakan sebagai variabel input hanya terkait dengan harga saham, sedangkan banyak indikator teknikal lainnya yang tidak dipertimbangkan, seperti yang terkait dengan volume perdagangan. Kedua, model prediksi berlaku untuk satu periode observasi; ketika periode pengamatan berubah, proses pembuatan model prediksi harus diulang (proses pelatihan baru harus diselesaikan). Oleh karena itu, prediksi tidak dapat digeneralisasikan. Ketiga, dalam penerapan model prediksi dengan mengungkap aturan perdagangan sederhana, model tidak

melihat jalannya transaksi: apakah saham dimiliki, apakah memberikan keuntungan atau tidak, atau biaya transaksi yang timbul saat melakukan pembelian dan penjualan saham.

Penelitian selanjutnya perlu meningkatkan panjang periode observasi yang meliputi waktu dimana terjadi perubahan yang cukup signifikan dalam pasar seperti saat bullish dan bearish. Hal ini untuk meningkatkan kemampuan prediksi model. Selain itu penambahan variabel input selain harga (seperti volume perdagangan) juga bisa melengkapi model prediksi yang dibangun.

REFERENSI

- Andrianto, Y., & Mirza, A. R. (2016). A testing of efficient markets hypothesis in Indonesia stock market. *Procedia-Social and Behavioral Sciences*, 219, 99-103.
- Alhashel, B.S., Almudhaf, F.W. & Hansz, J.A. (2018). Can technical analysis generate superior returns in securitized property markets? Evidence from East Asia markets. *Pacific-Basin Finance Journal*, 47, 92-108.
- Brown, R.G. (2004). *Smoothing, forecasting and prediction of discrete time series*. New York: Dover Publications inc.
- Cristianini, N. & Shawe-Taylor, J. (2000). *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*. Cambridge: Cambridge University Press.
- Emir, S., Dinçer, H. & Timor, M. (2012). A stock selection model based on fundamental and technical analysis variables by using artificial neural networks and support vector machines. *Review of Economics & Finance*, 2, 106-122.
- Han, J., Pei, J. & Kamber, M. (2011). *Data mining: concepts and techniques*. Amsterdam: Elsevier.
- Huang, W., Nakamori, Y., & Wang, S. Y. (2005). Forecasting stock market movement direction with support vector machine. *Computers & operations research*, 32(10), 2513-2522.
- Kang, H. (2013). The prevention and handling of the missing data. *Korean journal of anesthesiology*. 64(5), 402-406.
- Kara, Y., Boyacioglu, M. A., & 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(5), 5311-5319.
- Kim, J. H., & Shamsuddin, A. (2008). Are Asian stock markets efficient? Evidence from new multiple variance ratio tests. *Journal of Empirical Finance*, 15(3), 518-532.
- Lee, R. S. (2006). iJADE Stock Advisor—An Intelligent Agent-Based Stock Prediction System Using the Hybrid RBF Recurrent Network. *Fuzzy-Neuro Approach to Agent Applications: From the AI Perspective to Modern Ontology*, 231-253.

Madge, S., & Bhatt, S. (2015). Predicting stock price direction using support vector machines. *Independent work report spring*.

Mittal, A., & Goel, A. (2012). Stock prediction using twitter sentiment analysis. *Stanford University, CS229 (2011 <http://cs229.stanford.edu/proj2011/GoelMittal-StockMarketPredictionUsingTwitterSentimentAnalysis.pdf>), 15*.

Murphy, J. J. (1999). *Technical analysis of the financial markets: A comprehensive guide to trading methods and applications*. New York: Prentice Hall Press.

Nikita, M. P., & Soekarno, S. (2012, October). Testing on weak form market efficiency: The evidence from Indonesia stock market year 2008-2011. In *2nd International Conference on Business, Economics, Management and Behavioral Sciences* (pp. 56-60)

Nuraini, F., Jondri, J., & Wirayuda, T. A. B. (2015). Prediksi Pergerakan Indeks Harga Saham Menggunakan Artificial Neural Network Dan Support Vector Machine. *eProceedings of Engineering*, 2(3).

Schölkopf, B., Smola, A. J., & Bach, F. (2002). *Learning with kernels: support vector machines, regularization, optimization, and beyond*. Massachusetts: MIT press.

Shaik, M. & Maheswaran, S. (2017). Market Efficiency of ASEAN Stock Markets. *Asian Economic and Financial Review*, 7(2), 109.

Shynkevich, Y., McGinnity, T. M., Coleman, S. A., Belatreche, A., & Li, Y. (2017). Forecasting price movements using technical indicators: Investigating the impact of varying input window length. *Neurocomputing*, 264, 71-88.

Stanković, J., Marković, I., & Stojanović, M. (2015). Investment strategy optimization using technical analysis and predictive modeling in emerging markets. *Procedia Economics and Finance*, 19, 51-62.

Yu, H., Nartea, G. V., Gan, C., & Yao, L. J. (2013). Predictive ability and profitability of simple technical trading rules: Recent evidence from Southeast Asian stock markets. *International Review of Economics & Finance*, 25, 356-371.