

Analisis Peramalan Harga Penutupan Saham PT. Telekomunikasi Indonesia dengan *Support Vector Machine* (SVM)

Dwi Rahmadini^a, Nur Diana Fahma Salsabila^b, Anang Kunaefi^c, Dwi Rolliawati^d

^{a,b,c,d} Program Studi Sistem Informasi, Universitas Islam Negeri Sunan Ampel, Surabaya, Indonesia

email: ^adwirahmadini1234@gmail.com, ^bnfahmasalsabila@gmail.com, ^cakunaefi@uinsby.ac.id, ^ddwi_roll@uinsby.ac.id

DOI: <https://doi.org/10.21107/edutic.v11i1.22120>

Abstrak

Harga pasar saham menjadi salah satu masalah yang signifikan di pasar finansial karena naik turunnya harga setiap hari. Beberapa faktor seperti lokal dan iklim ekonomi global, kondisi politik, dan aktivitas pasar menjadi dampak yang dapat mempengaruhi harga pasar saham. Menyebabkan pergerakan saham menjadi tidak menentu dan sulit untuk ditebak. Sehingga para investor harus lebih hati-hati dalam membeli saham atau mempertahankan saham yang dimiliki. Oleh karena itu, untuk membantu para investor membuat keputusan yang optimal, dibutuhkan suatu langkah yang tepat seperti memprediksi perilaku harga pasar saham. Penelitian ini memprediksi harga penutupan saham pada PT. Telekomunikasi Indonesia sehingga penelitian ini melakukan prediksi secara univariat. Tujuan pada penelitian ini adalah mengimplementasikan model serta melakukan prediksi harga saham di PT. Telekomunikasi Indonesia. Menggunakan metode SVM yang diuji melalui skenario dalam penginputan *window_size* dan fungsi kernel. Parameter yang digunakan untuk pemodelan adalah parameter C sebesar 100 untuk semua kernel, parameter *degree* sebesar 1 untuk kernel *polynomial*, dan *gamma* sebesar 0.0001 untuk kernel RBF. Sehingga didapatkan pemodelan fungsi kernel yang paling optimal yaitu kernel *polynomial* pada ukuran *window_size* sebesar 3, dengan RMSE sebesar 67.546 dan MAPE sebesar 0.01. Sehingga disimpulkan bahwa performa kernel *polynomial* memiliki kekuatan akurasi yang tinggi.

Kata Kunci: *Prediksi, Harga Penutupan Saham, Support Vector Machine (SVM)*

Abstract

The stock market price is one of the significant problems in the financial market because of the ups and downs of prices every day. Several factors such as the local and global economic climate, political conditions, and market activity are impacts that can affect stock market prices. Causing stock movements to be erratic and difficult to predict. So that investors must be more careful in buying shares or maintaining the shares they own. Therefore, to help investors make optimal decisions, an appropriate step is needed, such as predicting the behavior of stock market prices. This study predicts the closing price of shares at PT. Telekomunikasi Indonesia so that this research makes predictions univariately. The purpose of this study is to implement the model and predict stock prices at PT. Indonesian Telecommunications. Using the SVM method tested through scenarios in the input *window_size* and kernel functions. The parameters used for modeling are C parameter of 100 for all kernels, degree parameter of 1 for polynomial kernels, and gamma of 0.0001 for RBF kernels. So that the most optimal kernel function modeling is obtained, namely the polynomial kernel at a *window_size* of 3, with an RMSE of 67,546 and a MAPE of 0.01. So it is concluded that the performance of the polynomial kernel has a high accuracy power.

Keywords: *Prediction, Close Price Stock, Support Vector Machine (SVM)*



© Author (s)

PENDAHULUAN

Harga pasar saham menjadi salah satu masalah yang signifikan di pasar finansial karena naik turunnya harga setiap hari. Beberapa faktor seperti lokal dan iklim ekonomi global, kondisi politik, dan aktivitas pasar menjadi dampak yang dapat mempengaruhi harga pasar saham (Wei et al., 2019). Menyebabkan pergerakan saham menjadi tidak menentu dan sulit untuk ditebak. Sehingga para investor harus lebih hati-hati dalam membeli saham atau mempertahankan saham yang dimiliki. Oleh karena itu, untuk membantu para investor membuat keputusan yang optimal, dibutuhkan suatu langkah yang tepat seperti memprediksi perilaku harga pasar saham. Bertahun-tahun area ini menjadi perhatian oleh para peneliti dan ahli finansial, sebab memprediksi harga pasar saham secara akurat merupakan tugas yang kompleks karena sifatnya yang non-linear, dinamis, dan sulit ditebak (Petropoulos et al., 2022). Namun, hasil prediksi tersebut dapat mempengaruhi ekonomi. Karena hasil prediksi yang sukses dapat membantu investor untuk membuat keputusan untuk meningkatkan profit dan sebaliknya, investor akan kehilangan profit perusahaan.

Data saham merupakan data *time series* atau deret waktu yang dapat digunakan untuk mempelajari sampel riwayat dari waktu ke waktu. Sehingga prediksi harga saham dapat dilakukan dengan mempelajari data deret waktu atau historis harga saham pada hari-hari sebelumnya (Idrees et al., 2019). Pada penelitian terdahulu salah satu model yang paling umum digunakan adalah ARIMA. Model ARIMA yang digunakan untuk prediksi telah dianggap sangat efektif (Rao et al., 2020). Tetapi keterbatasan pada model ini, ARIMA memiliki beberapa penyimpangan ketika menghadapi masalah non-linear yang kompleks (Ma, 2020). Saat ini dengan memanfaatkan evolusi teknologi komputer, metode untuk memprediksi harga pasar saham sudah banyak berkembang dan salah satunya adalah *machine learning*. Teknik *machine learning* dapat memprediksi harga pasar saham dengan akurasi yang lebih tinggi pada permasalahan linear dan non-linear. Selain itu, performa prediksi lebih efisien dibandingkan dengan ARIMA (Kobiela et al., 2022; Makala & Li, 2021; Pandey & Bajpai, 2019; Wijaya et al., 2010).

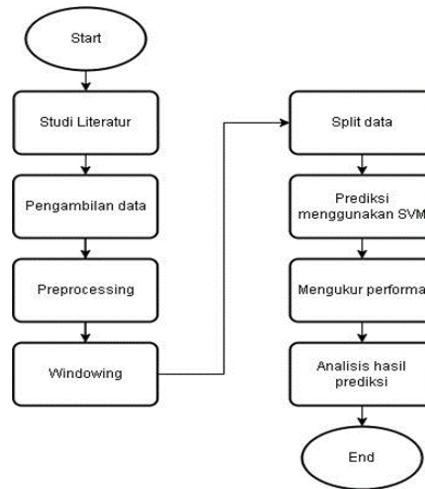
Terdapat berbagai algoritma *machine learning* yang dapat digunakan untuk prediksi harga saham, seperti *Support Vector Machine* (SVM), *K-Nearest Neighbor* (KNN), *Decision Trees*, *Artificial Neuron Network* (ANN), dan lain-lain. Algoritma SVM dianggap metode terbaik dalam memprediksi harga saham dari segi akurasi (Rao et al., 2020; Rouf et al., 2021; Teixeira Zavadzki de Pauli et al., 2020). Hal ini dikarenakan algoritma SVM dapat menghasilkan solusi terbaik (*global solution*) melalui proses pembelajaran yang memisahkan data ke dalam dua kelompok menggunakan garis disebut *hyperplane*. Tidak hanya itu, algoritma ini tidak membutuhkan asumsi-asumsi seperti metode peramalan yang konvensional. Oleh karenanya metode ini sering digunakan dalam melakukan peramalan data deret waktu (generator, 2021; Mustakim et al., 2015). Akan tetapi, metode SVM memiliki kekurangan seperti sulit dipakai dalam jumlah sampel yang besar. Karena waktu pelatihan yang dibutuhkan lebih tinggi (Budianto et al., 2019).

Berdasarkan uraian diatas, penelitian ini mengajukan prediksi harga pasar saham di salah satu perusahaan teknologi di Indonesia, yaitu perusahaan Telekomunikasi menggunakan metode *Support Vector Machine*. Karena di Indonesia sendiri, prediksi harga saham juga masih menjadi tantangan peneliti, terbukti dengan masih banyaknya penelitian dalam prediksi harga saham (Fitriani et al., 2019; Jange, 2021; Karno, 2020; Wijaya et al., 2010). Sehingga diharapkan prediksi harga saham dapat bermanfaat bagi para investor untuk dapat melihat bagaimana prospek investasi saham sebuah perusahaan di masa datang.

METODE PENELITIAN

Penelitian ini dilaksanakan di Universitas UIN Sunan Ampel Surabaya pada Progam Studi Sistem Informasi. Penelitian ini akan menganalisis prediksi harga tutup (*close*) harian pada saham PT Telekomunikasi Indonesia dengan menggunakan bahasa pemrograman *Python*. Adapun penjelasan rinci

tentang langkah-langkah yang diperlukan untuk melakukan penelitian. Ini akan mencakup deskripsi berbagai tahapan penelitian yang akan dilakukan.



Gambar 1. Desain Penelitian

Penelitian ini dimulai dengan melakukan studi literatur, dimana tahapan ini mencari penelitian-penelitian sebelumnya yang relevan. Untuk memahami teori dasar yang berkaitan dengan objek penelitian, maka harus mempelajari beberapa hal yang berkaitan dengan penelitian ini antara lain teori SVM, saham, dan peramalan. Penelitian yang akan dilakukan merupakan penelitian kuantitatif. Penelitian kuantitatif merupakan penelitian dimana hasil dari pengumpulan data akan diukur dengan teknik statistik ataupun komputer. Kemudian data pada penelitian ini adalah data harga saham harian PT.Telekomunikasi Indonesia dengan periode selama 10 tahun yaitu Januari 2012 hingga Desember 2022. Jumlah data yang didapatkan sejumlah 2733 yang diperoleh dari salah satu situs yaitu www.yahoofinance.com. Data tersebut terdiri dari tujuh atribut yang dijelaskan pada Tabel 3.1.

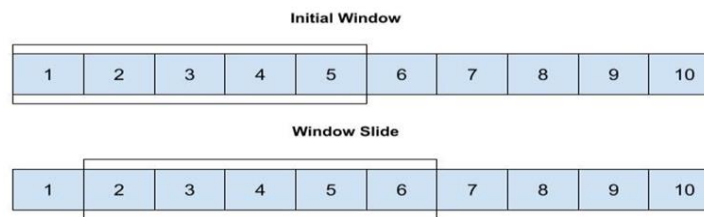
Tabel 1 Variabel Data Saham PT. Telekomunikasi Indonesia

Nama Variabel	Tipe Variabel	Kurs	Keterangan
<i>Date</i>	Numerik	Rupiah (Rp)	Hari, bulan, dan tahun
<i>Open</i>	Numerik	Rupiah (Rp)	Harga pembuka saham
<i>High</i>	Numerik	Rupiah (Rp)	Harga tertinggi saham
<i>Close</i>	Numerik	Rupiah (Rp)	Harga penutup saham
<i>Volume</i>	Numerik	Rupiah (Rp)	Jumlah harga yang diperdagangkan
<i>Adj Close</i>	Numerik	Rupiah (Rp)	Jumlah harga penutup yang telah disesuaikan dari tindakan

Selanjutnya adalah tahap *preprocessing* data, *preprocessing* merupakan tahapan yang dilakukan dengan tujuan menyaring data untuk menghilangkan *noise* dan mentransformasikan data agar lebih sesuai dengan kebutuhan algoritma *machine learning*. Proses ini bertujuan untuk menyiapkan data secara optimal agar dapat diterima dengan baik oleh algoritma yang digunakan dalam analisis. Pada tahap ini, dataset akan diperiksa apakah kumpulan data yang diambil terdapat nilai yang hilang (*missing value*) atau tidak. Jika iya maka nilai diganti dengan menggunakan rata-rata. Selanjutnya adalah menentukan variabel target atau variabel yang ingin diprediksi. Dalam penelitian ini variabel yang akan diprediksi adalah variabel *close* atau variabel harga tutup pada saham. karena harga tutup saham dapat dijadikan sebagai titik referensi dimana perusahaan investor akan menggunakan variabel ini untuk mengevaluasi harga saham mendatang.

Setelah variabel target atau variabel yang ingin diprediksi ditentukan, tahap selanjutnya yaitu melakukan proses *windowing*. Proses ini merupakan proses perkiraan sementara atas nilai sebenarnya

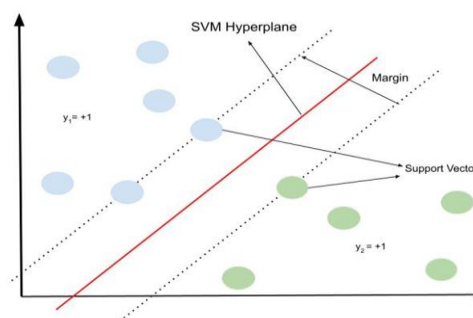
dari data *time series*. Teknik ini digunakan untuk mempermudah model dalam mempelajari data, data harus ditransformasi menjadi *multiple sample*. *Windowing* memiliki ukuran dan segmen, dimana mereka terus meningkat hingga mencapai error yang lebih sedikit (BenYahmed et al., 2015). Setelah memilih segmen pertama, segmen berikutnya dipilih. Proses ini diulang sampai semua data deret waktu tersegmentasi. Proses ini ditunjukkan pada Gambar 2 dengan ukuran *window* adalah 5. *Sliding window* mengakumulasi data deret waktu historis untuk memprediksi di hari berikutnya (Mozaffari et al., 2014). Setiap angka (1,2,3...10) merepresentasikan observasi harian dari data deret waktu. Awalnya *window* telah mencakup dari 1 hingga 5 menunjukkan bahwa data historis 5 hari digunakan untuk memprediksi, kemudian *window* bergeser ke kanan satu hari untuk mencakup 5 hari berikutnya (dari 2 hingga 6). Proses ini akan dilanjutkan sampai waktu data dari periode waktu tertentu yang telah dipertimbangkan untuk tujuan percobaan telah habis (Vafaeipour et al., 2014).



Gambar 2. Proses *Windowing* (Hota et al., n.d.)

Tahap selanjutnya yaitu *split data*, Pada tahap ini akan dilakukan pengelompokan data menjadi data *training* dan data *testing*. Pembagian akan dilakukan dengan proporsi 80% untuk data *training* dan 20% untuk data. Tujuan dari pembagian tersebut adalah agar model yang dibentuk dapat belajar dengan baik menggunakan data *training* dan memberikan hasil peramalan yang optimal saat diuji dengan data *testing*.

Selanjutnya adalah memprediksi dengan *machine learning SVM*. SVM Merupakan salah satu mesin pembelajaran yang menggunakan prinsip *Structural Risk Minimization (SRM)* untuk mengurangi batas atas kesalahan generalisasi, bukan mengurangi kesalahan pelatihan. Prinsip ini mengacu pada fakta bahwa kesalahan generalisasi terjadi karena beberapa *training error* dan interval yang bergantung pada dimensi (vapnik-Chervonenkis) (Amin & Hoque, 2019). SVM didasarkan pada konsep pengklasifikasi linear, yang memisahkan kasus-kasus yang dapat dipisahkan secara linear. Akan tetapi, SVM telah didesain untuk dapat menangani masalah non-linear dengan menggunakan konsep kernel pada ruang berdimensi tinggi (Sutrisno, 2018). Di ruang berdimensi tinggi, *hyperplane* sering digunakan untuk memaksimalkan jarak antar kelas data. Linear Support Vector Machine (SVMs) bekerja dengan menempatkan *hyperplane* optimal sebagai pemisah untuk dua kelas di ruang input. Konsep ini terdiri dari *pattern* atau pola, margin, dan *support vector*. Pola adalah anggota dari dua kelas yang dipisahkan oleh *hyperplane* dengan nilai 1 dan kelas lain dengan nilai -1. Margin adalah jarak antara *hyperplane* dengan *pattern* terdekat dari masing masing class. Pola terdekat tersebut disebut support vector (Sutrisno, 2018). Konsep ini diilustrasikan pada Gambar 3.



Gambar 3. Ilustrasi SVM (Yang & Prayogo, 2022)

Berdasarkan pada uraian sebelumnya, konsep SVM yaitu untuk mencari hyperplane yang terbaik. Gambar 3 menggambarkan sejumlah pola yang merupakan bagian anggota dari dua buah kelas. Pola yang termasuk kelas -1 diwakili oleh lingkaran hijau, dan yang termasuk kelas +1 diwakili oleh lingkaran biru. Garis tebal berwarna merah pada Gambar 3 menggambarkan *hyperplane* terbaik yang memisahkan kedua kelas, dan dalam ilustrasi tersebut terdapat tiga pola yang berada di posisi terdekat dengan garis *hyperplane*. Karena SVM memisahkan variabel menjadi dua kelas, maka persamaan yang didapatkan yaitu:

$$b = -\frac{1}{2}(w \cdot x^+ + w \cdot x^-) \tag{1}$$

$$w = \sum_{i=1}^n \alpha_i y_i x_i \tag{2}$$

Dimana nilai bias ditandakan sebagai b, $w \cdot x^+$ adalah nilai bobot untuk kelas data positif dan sebaliknya, w adalah bobot vektor, dan α_i untuk nilai bobot data ke-i (Sutrisno, 2018).

Selain permasalahan yang linear, SVM juga dapat bekerja pada permasalahan yang nonlinear dengan memanfaatkan fungsi kernel. Prinsip dasar dari penggunaan kernel ini adalah untuk mengubah vektor input ke dalam ruang fitur yang memiliki dimensi yang lebih tinggi. Data yang tidak dapat dipisahkan secara linier diasumsikan bahwa kelas-kelas dalam ruang input tidak dapat dipisahkan secara sempurna. Keunggulan dari mesin vektor pendukung berasal dari kemampuannya untuk menerapkan pemisahan linier ke data masukan non-linier dimensi tinggi. Hal ini dicapai melalui penggunaan fungsi kernel. Efektivitas SVM sangat dipengaruhi oleh jenis fungsi kernel yang dipilih dan diimplementasikan sesuai dengan karakteristik data (Dou et al., 2019). Beberapa kernel yang umum digunakan dalam SVM antara lain:

1. Linear kernel

$$K(x, y) = x^T \cdot y \tag{3}$$

Dimana x dan y adalah vektor ruang input (Schölkopf et al., 2004).

2. Radial Basic Function (RBF) kernel

$$K(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right) \tag{4}$$

Dimana $\|x_i - x_j\|$ adalah jarak Euclidean antara dua nilai data dalam ruang fitur yang berbeda dan σ (sigma) digunakan untuk menentukan bobot kernel RBF di SVM. Parameter ini, perlu disesuaikan untuk memberikan hasil yang lebih akurat. Nilai default sigma adalah 1 (Schölkopf et al., 2004).

3. Polynomial kernel

$$K(x, y) = (x \cdot y + c)^d \tag{5}$$

Dimana x dan y adalah vektor dalam ruang input, yaitu vektor fitur yang dihitung dari sampel pelatihan atau uji, c adalah parameter bebas, dan d adalah derajat polynomial (Shashua, 2009).

Prediksi dengan menggunakan SVM pada penelitian ini memanfaatkan semua fungsi kernel untuk memprediksi harga saham hari kedepan. Dalam fungsi kernel terdapat tiga parameter yang akan digunakan dan masing-masing nilai parameter yang diambil mengacu pada jurnal (Yudhawan & Purwaningsih, 2020), yaitu parameter C adalah 100, gamma adalah 0.0001, dan degree adalah 1. Adapun penjelasan terkait masing-masing parameter (Dou et al., 2019), yaitu:

1. Regulasi (C)

Parameter C merupakan parameter yang bekerja untuk mengoptimalkan SVM supaya menghindari misklasifikasi dalam dataset training. Parameter ini digunakan untuk semua kernel pada SVM.

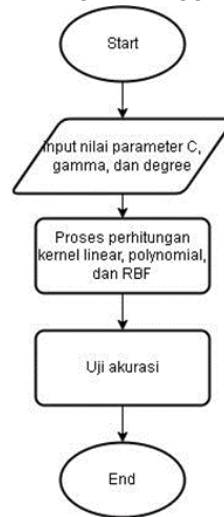
2. Gamma

Parameter gamma digunakan untuk menentukan seberapa jauh pengaruh dari satu sampel pada data training. Nilai gamma yang rendah menandakan titik yang jauh untuk menentukan garis hyperplane

berada dan demikian sebaliknya, nilai parameter gamma yang tinggi menandakan titik dekat untuk menentukan garis *hyperplane* berada. Parameter gamma digunakan pada kernel RBF, polynomial, dan sigmoid.

3. *degree*

Parameter *degree* merupakan parameter yang hanya digunakan pada kernel polynomial. Pada dasarnya parameter ini digunakan untuk menemukan hyperplane untuk membagi data. Adapun langkah-langkah dalam melakukan prediksi harga saham dengan menggunakan SVM dijelaskan pada Gambar 4.



Gambar 4. Langkah-Langkah Memprediksi dengan SVM

Tahap terakhir yaitu, tahap evaluasi. Evaluasi merupakan tahap dimana tahap pengujian selesai dilakukan. Terdapat macam-macam metode yang dapat digunakan untuk mengevaluasi presisi dan akurasi dari metode yang digunakan, namun pada penelitian ini metode yang digunakan adalah *Root Mean Square Error (RMSE)* dan *Mean Square Percentage Error (MAPE)*.

1. RMSE

RMSE adalah jumlah kesalahan kuadrat atau perbedaan pada nilai aktual dan nilai perkiraan. Semakin rendah nilai RMSE, maka semakin baik model prediksi yang dibuat (Hodson, 2022).

2. MAPE

MAPE merupakan ukuran keakuratan prediksi yang dibuat dengan menggunakan metode statistik. Ini dihitung sebagai rata-rata nilai absolut dari persentase kesalahan prediksi, dan memberikan indikasi berapa banyak kesalahan perkiraan dibandingkan dengan nilai sebenarnya. Interpretasi MAPE dapat dilihat pada Tabel 2.

Tabel 2 *Range Nilai MAPE (Lewis, 1982)*

MAPE	Forecasting Power
<10%	Highly accurate power
10%-20%	Forecasting is good
20%-50%	Reasonable Forecasting
>50%	Weak and inaccurate forecast

HASIL DAN PEMBAHASAN

Pada bagian ini akan menguraikan hasil berdasarkan alur penelitian yang telah dipaparkan pada metodologi penelitian. Alur penelitian dimulai dari tahap pengumpulan data hingga tahap analisis hasil prediksi. Penelitian ini memiliki sifat univariat, karena variabel yang akan digunakan yaitu ‘close’ sebagai variabel target atau yang akan diprediksi. Pemilihan ini didasarkan karena harga penutupan

sebagai titik referensi untuk menentukan jual beli saham dan evaluasi kinerja investasi suatu perusahaan. Penelitian ini juga dilakukan iterasi *trial and error* dalam proses penginputan *window_size*.

Sebelum memprediksi harga penutupan saham, diperlukan gambaran umum tentang data harga penutupan PT. Telekomunikasi Indonesia. Hal ini bertujuan untuk memahami karakteristik data yang telah diambil. Gambaran umum pada harga penutupan saham perusahaan telkom dapat diamati pada Tabel 3.

Tabel 3 Deskripsi data

Variabel	Close
Count	2733
Mean	3289.047
StD	878.661
Min	1350
25%	2700
50%	3400
75%	4020
Max	4800

Tabel 3 ini memberikan informasi statistik tentang data harga penutupan PT Telekomunikasi Indonesia. *Count* mencerminkan jumlah data yang ada. *Mean* menunjukkan nilai tengah dari distribusi harga penutupan, sedangkan standar deviasi mengukur sejauh mana data tersebar dari rata-ratanya. Nilai minimum dan maksimum mengindikasikan rentang harga penutupan yang diamati, sementara kuartil pertama (25%) menunjukkan nilai di bawah 25% data, median (50%) merupakan nilai tengah, dan kuartil ketiga (75%) menunjukkan nilai di bawah 75% data. Dengan tabel ini, dapat memiliki gambaran tentang karakteristik data harga penutupan PT Telekomunikasi Indonesia.

Selanjutnya, *preprocessing* yang dilakukan untuk menyeleksi kembali data yang akan digunakan dan mengeliminasi data yang tidak sesuai. Sehingga data yang akan digunakan lebih mudah untuk diproses oleh machine learning SVM. Pada penelitian ini *preprocessing* data yang dilaksanakan adalah melakukan cek *missing value*. Proses cek *missing value* digunakan agar data dapat digunakan pada proses selanjutnya. Baris yang memiliki nilai yang hilang diganti dengan hasil *mean* dari sepuluh baris sebelum *missing value*.

Kemudian untuk melakukan prediksi dengan *machine learning* SVM, data yang sudah di *preprocessing* akan dibagi menjadi dua bagian, yaitu data *training* dan data *testing*. Pembentukan data *training* dilakukan untuk meningkatkan kinerja dari model SVM yang akan dibuat. Hasil Pembagian data ditampilkan pada Tabel 4.

Tabel 4 Pembagian data

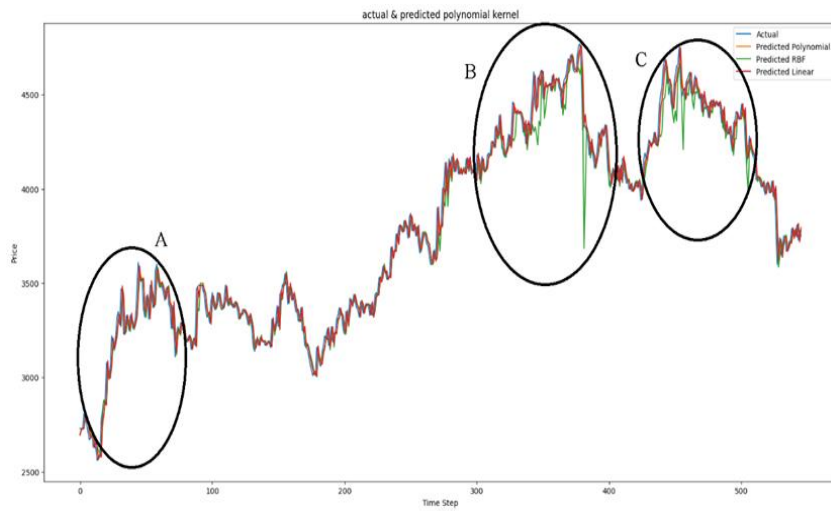
Data	Jumlah Data	Persentase
Training	2179	80%
Testing	545	20%
Total	2733	100%

Data pada Tabel 4 akan dilakukan proses pembentukan beberapa sampel melalui *windowing*. Proses *windowing* ini dapat mempermudah model dalam mempelajari data dan membantu memprediksi nilai berikutnya. Pada penelitian ini diterapkan beberapa proses *windowing* dengan ukuran *window_size* yang dipilih menggunakan bilangan yang dimulai dari tiga hingga batas maksimum yaitu sepuluh. Nilai ukuran ini akan melalui proses *trial dan error* untuk mencari nilai akurasi tertinggi dan tingkat error yang rendah. Berikut hasil dari salah satu percobaan *windowing* yaitu dengan ukuran lima yang ditampilkan pada Tabel 5.

Tabel 5 Hasil *Windowing* Data

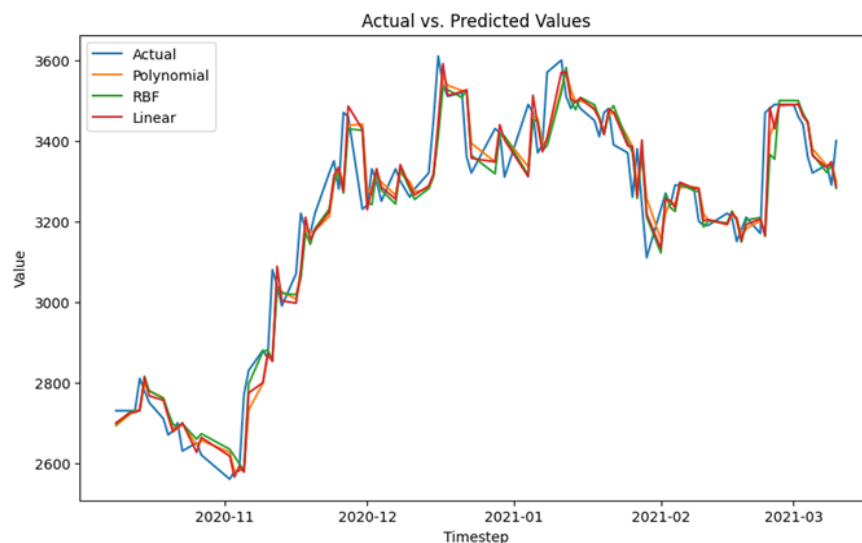
Data-1	Data-2	Data-3	Data-4	Data-5	Label
1410	1410	1420	1410	1390	1410
1410	1420	1410	1390	1410	1420
1420	1410	1390	1410	1420	1400
1410	1390	1410	1420	1400	1400
1390	1410	1420	1400	1400	1390

Hasil dari proses *windowing* akan digunakan untuk dipelajari oleh SVM dengan fungsi kernel linear, *polynomial*, dan RBF. Performa model yang dikembangkan akan diukur menggunakan teknik akurasi RMSE dan MAPE. Hasil nilai MAPE akan diinterpretasikan berdasarkan Tabel 3. Hasil prediksi untuk semua kernel dengan ukuran *window_size* sebesar tiga ditunjukkan pada Gambar 5.



Gambar 5. Perbandingan Aktual dan Prediksi Semua Pemodelan Kernel

Berdasarkan pada Gambar 5 terdapat beberapa garis fluktuatif yang masing-masing telah disimbolkan menggunakan huruf abjad. Pada simbol A harga fluktuatif terjadi dalam rentang *timestep* antara 0 sampai dengan 100. Dalam simbol tersebut jarak ketiga pemodelan saling berhimpitan sehingga sulit untuk menentukan pemodelan yang mana yang mendekati dengan nilai aktual. Sehingga grafik pada interval tersebut akan diperbesar dan ditampilkan pada Gambar 6.



Gambar 6. Perbandingan Aktual dan Prediksi Interval 0-100

Berdasarkan Gambar 4.3 garis yang merepresentasikan aktual adalah garis yang berwarna biru. Sehingga sesuai pada Gambar 4.3 pemodelan yang mendekati atau yang memiliki jarak terkecil dengan garis aktual merupakan pemodelan *polynomial* yang direpresentasikan dengan garis berwarna kuning. Dalam beberapa titik garis *polynomial* memiliki jarak terkecil dibandingkan dengan pemodelan lainnya, sehingga jarak tersebut diartikan sebagai pemodelan *polynomial* memiliki nilai yang mendekati dengan nilai aktual.

Berdasarkan lingkaran yang bersimbolkan dengan huruf B pada Gambar 6, pemodelan RBF merupakan pemodelan yang memiliki jarak yang terbesar dari pemodelan lainnya. Sehingga pada *timestep* antara 300 hingga 400 pemodelan RBF memiliki perbedaan yang signifikan dengan harga aktual. Mengartikan dalam interval *timestep* tersebut pemodelan RBF memiliki tingkat kesalahan yang lebih besar dibandingkan pemodelan lainnya. Kemudian garis linear dan garis *polynomial* memiliki jarak yang saling berhimpitan. Jarak tersebut memberikan petunjuk bahwa prediksi antara linear dan *polynomial* memiliki jarak nilai yang tidak jauh. Dan kedua garis pemodelan tersebut memiliki jarak yang berdekatan dengan garis aktual. Membuat kedua pemodelan tersebut memiliki nilai yang mendekati pada nilai aktual.

Lalu, berdasarkan lingkaran yang bersimbolkan dengan huruf C pada Gambar 6. Garis pemodelan RBF memiliki perbedaan yang tipis dengan nilai aktual. Sehingga dari ketiga garis fluktuatif dapat disimpulkan bahwa pemodelan RBF memiliki kekuatan akurasi yang lebih rendah dari pemodelan linear dan *polynomial*. Kemudian, untuk mengetahui lebih detail untuk mencari pemodelan yang paling optimal antara linear dengan *polynomial* maka akan dilakukan analisis pada RMSE dan MAPE di setiap masing-masing hasil dari pemodelan kernel.

Tabel 6 Hasil RMSE Setiap Model Kernel

<i>window_size/</i> <i>model kernel</i>	3	4	5	6	7	8	9	10
Linear	68.100	68.765	69.965	70.824	71.055	72.538	73.253	73.709
<i>Polynomial</i>	67.546	68.232	69.220	69.605	70.357	70.977	71.467	71.942
RBF	86.867	106.746	129.993	151.830	172.833	191.768	206.629	234.506

Berdasarkan pada Tabel 6, didapatkan hasil RMSE yang paling optimal yaitu pada ukuran *window_size* tiga dimana memiliki hasil error yang terendah di antara ukuran *window_size* lainnya. Juga model kernel yang memiliki tingkat error yang paling rendah adalah model *polynomial* pada ukuran *window_size* 3 dengan nilai 67.546 yang diikuti oleh model linear dan model RBF sebagai performa model terendah. Untuk hasil MAPE akan dipaparkan pada Tabel 7.

Tabel 7 Hasil MAPE Setiap Model Kernel

<i>window_size/</i> <i>model kernel</i>	3	4	5	6	7	8	9	10
Linear	0.0134	0.0136	0.0138	0.0139	0.0140	0.0142	0.0143	0.0144
<i>Polynomial</i>	0.0131	0.0134	0.0136	0.0137	0.0139	0.0140	0.0141	0.0142
RBF	0.0155	0.0179	0.0206	0.0229	0.0255	0.0281	0.0302	0.0344

Berdasarkan dari Tabel 6 dan Tabel 7 telah didapatkan hasil akurasi yang paling optimal yaitu hasil akurasi pada ukuran *window_size* 3 pada pemodelan *polynomial* dengan hasil RMSE dan MAPE yang paling kecil, yaitu sebesar 67.546 untuk RMSE dan 0.0134 untuk MAPE, dan apabila diinterpretasikan untuk hasil MAPE diartikan sebagai *forecast is good* karena hasilnya berada di antara

10%-20%. Dan untuk ukuran *window_size* terakhir yaitu 10 merupakan ukuran yang memperoleh hasil performa yang kurang optimal. Dalam ukuran tersebut kernel yang memperoleh performa yang terendah, yaitu kernel RBF. Dimana kernel RBF memperoleh nilai RMSE sebesar 206.629 dan MAPE sebesar 0.0302. Namun meskipun model RBF memiliki performa yang paling rendah diantara kedua model lainnya, kekuatan model RBF masih dianggap bagus berdasarkan hasil interpretasi, dimana model RBF ini memasuki kategori sebagai *reasonable forecast* sehingga model ini masih dapat diterapkan sebagai pertimbangan. Dan berdasarkan pada setiap ukuran, kernel polynomial merupakan model prediksi yang paling optimal, diikuti oleh kernel linear, dan kernel RBF yang merupakan model prediksi yang paling rendah.

Dilihat dari hasil ukuran *window_size* 3 hingga 10. Dapat disimpulkan bahwa semakin besar jumlah *window_size* pada proses *windowing* maka akurasi untuk prediksi harga saham harian PT.Telkom Indonesia semakin rendah. Dan begitupun sebaliknya, semakin kecil jumlah *window_size* yang digunakan pada *windowing* maka akurasi untuk prediksi harga saham harian PT.Telkom Indonesia semakin tinggi. Hal ini dapat dilihat dari nilai performa yang semakin menurun seiring bertambahnya jumlah *window_size* yang digunakan. Hasil analisis tersebut menyerupai dengan hasil yang didapatkan pada penelitian (Liu et al., 2021), dimana penelitian ini mencari jumlah ukuran window yang optimal dan mendapatkan hasil bahwa semakin kecil ukuran window yang digunakan maka hasil performa pada penelitian tersebut semakin bagus.

KESIMPULAN

Setelah dilakukan 4 skenario pengujian didapatkan hasil performa SVM model tiga kernel terbaik dengan parameter C sebesar 100, gamma 0.0001, dan *degree* sebesar 1. Tingkat error RMSE dan MAPE yang dihasilkan adalah 67.546 dan 0.01. Hasil ini didapatkan oleh model kernel polynomial pada skenario dimana ukuran *window_size* yang digunakan sebesar tiga. Disamping itu, penggunaan *windowing* juga mempengaruhi tingkat akurasi dalam suatu model untuk memprediksi saham. Semakin tinggi ukuran *window_size* yang digunakan untuk proses *windowing* maka performa model yang diciptakan semakin rendah. Berdasarkan hasil analisa prediksi dengan menggunakan support vector machine, terdapat berbagai saran untuk penelitian selanjutnya, yaitu:

1. Dapat menggunakan variabel lainnya seperti variabel *open*, *high*, dan *low* atau mengembangkan model prediksi dengan variabel multivariat.
2. Dapat mencoba untuk menggunakan *machine learning hybrid* untuk memprediksi harga saham.
3. Dapat mencoba untuk menggunakan metode *deep learning* untuk memprediksi saham.

UCAPAN TERIMA KASIH

Penulis menyampaikan ucapan terimakasih yang sebesar-besarnya kepada dosen dan teman-teman penulis yang telah mendukung dan membantu hingga penulisan artikel selesai.

DAFTAR PUSTAKA

- Amin, M. A. A., & Hoque, Md. A. (2019). Comparison of ARIMA and SVM for Short-term Load Forecasting. *2019 9th Annual Information Technology, Electromechanical Engineering and Microelectronics Conference (IEMECON)*, 1–6. <https://doi.org/10.1109/IEMECONX.2019.8877077>
- BenYahmed, Y., Abu Bakar, A., RazakHamdan, A., Alshareef, A., & Abdullah, S. M. S. (2015). *Adaptive sliding window algorithm for weather data segmentation*. 80, 322–333.
- Budianto, A., Ariyuana, R., & Maryono, D. (2019). PERBANDINGAN K-NEAREST NEIGHBOR (KNN) DAN SUPPORT VECTOR MACHINE (SVM) DALAM PENGENALAN KARAKTER PLAT KENDARAAN BERMOTOR. *JIPTEK : Jurnal Ilmiah Pendidikan Teknik dan Kejuruan*, 11(1), Article 1. <https://doi.org/10.20961/jiptek.v11i1.18018>

- Dou, J., Yunus, A. P., Bui, D., Sahana, M., Chen, C., Zhu, Z., Wang, W.-D., & Pham, B. (2019). Evaluating GIS-Based Multiple Statistical Models and Data Mining for Earthquake and Rainfall-Induced Landslide Susceptibility Using the LiDAR DEM. *Remote Sensing*, 11, 638. <https://doi.org/10.3390/rs11060638>
- Fitriani, R. R., Ernastuti, E., & Swedia, E. R. (2019). ALGORITMA LEARNING VECTOR QUANTIZATION DAN FUZZY K-NN UNTUK PREDIKSI SAHAM BERDASARKAN PESAING. *Jurnal Ilmiah Teknologi dan Rekayasa*, 24(1), Article 1. <https://doi.org/10.35760/tr.2019.v24i1.1929>
- generator, metatags. (2021). Penerapan Algoritma Support Vector Machine Untuk Model Prediksi Kelulusan Mahasiswa Tepat Waktu | Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi). <http://jurnal.iaii.or.id/index.php/RESTI/article/view/3007>
- Hodson, T. O. (2022). Root-mean-square error (RMSE) or mean absolute error (MAE): When to use them or not. *Geoscientific Model Development*, 15(14), 5481–5487. <https://doi.org/10.5194/gmd-15-5481-2022>
- Hota, H. S., Handa, R., & Shrivasa, A. K. (n.d.). Time Series Data Prediction Using Sliding Window Based RBF Neural Network.
- Idrees, S. M., Alam, M. A., & Agarwal, P. (2019). A Prediction Approach for Stock Market Volatility Based on Time Series Data. *IEEE Access*, 7, 17287–17298. <https://doi.org/10.1109/ACCESS.2019.2895252>
- Jange, B. (2021). Prediksi Harga Saham Bank BCA Menggunakan Prophet. *Journal of Trends Economics and Accounting Research*, 2(1), Article 1.
- Karno, A. S. B. (2020). Prediksi Data Time Series Saham Bank BRI Dengan Mesin Belajar LSTM (Long ShortTerm Memory). *Journal of Informatic and Information Security*, 1(1), Article 1. <https://doi.org/10.31599/jiforty.v1i1.133>
- Kobiela, D., Krefta, D., Król, W., & Weichbroth, P. (2022). ARIMA vs LSTM on NASDAQ stock exchange data. *Procedia Computer Science*, 207, 3836–3845. <https://doi.org/10.1016/j.procs.2022.09.445>
- Lewis, C. D. (1982). *Industrial and business forecasting methods: A practical guide to exponential smoothing and curve fitting*. Butterworth-Heinemann.
- Liu, X., Yuan, X., Liu, C., Ma, H., & Lian, C. (2021). Analysis of Stock Price Data: Determination of The Optimal Sliding-Window Length. *Frontiers in Physics*, 9. <https://doi.org/10.3389/fphy.2021.741106>
- Ma, Q. (2020). Comparison of ARIMA, ANN and LSTM for stock price prediction. *E3S Web of Conferences*, 218, 01026.
- Makala, D., & Li, Z. (2021). Prediction of gold price with ARIMA and SVM. *Journal of Physics: Conference Series*, 1767(1), 012022. <https://doi.org/10.1088/1742-6596/1767/1/012022>
- Mozaffari, L., Mozaffari, A., & Azad, N. (2014). Vehicle speed prediction via a sliding-window time series analysis and an evolutionary least learning machine: A case study on San Francisco urban roads. *Engineering Science and Technology, an International Journal*, 6. <https://doi.org/10.1016/j.jestch.2014.11.002>
- Mustakim, M., Buono, A., & Hermadi, I. (2015). Support Vector Regression untuk Prediksi Produktivitas Kelapa Sawit di Provinsi Riau. *SITEKIN: Jurnal Sains, Teknologi dan Industri*, 12(2), Article 2. <https://doi.org/10.24014/sitekin.v12i2.990>
- Pandey, V., & Bajpai, A. (2019). Predictive Efficiency of ARIMA and ANN Models: A Case Analysis of Nifty Fifty in Indian Stock Market. <https://www.semanticscholar.org/paper/Predictive-Efficiency-of-ARIMA-and-ANN-Models%3A-A-of-Pandey-Bajpai/0e8395852a4aaab67a50643b45137630de634adb>

- Petropoulos, F., Apiletti, D., Assimakopoulos, V., Babai, M. Z., Barrow, D. K., Taieb, S. B., Bergmeir, C., Bessa, R. J., Bijak, J., Boylan, J. E., Browell, J., Carnevale, C., Castle, J. L., Cirillo, P., Clements, M. P., Cordeiro, C., Oliveira, F. L. C., Baets, S. D., Dokumentov, A., ... Ziel, F. (2022). Forecasting: Theory and practice. *International Journal of Forecasting*, 38(3), 705–871. <https://doi.org/10.1016/j.ijforecast.2021.11.001>
- Rao, P. S., Srinivas, K., & Mohan, A. K. (2020). A Survey on Stock Market Prediction Using Machine Learning Techniques. In A. Kumar, M. Paprzycki, & V. K. Gunjan (Eds.), *ICDSMLA 2019* (Vol. 601, pp. 923–931). Springer Singapore. https://doi.org/10.1007/978-981-15-1420-3_101
- Rouf, N., Malik, M. B., Arif, T., Sharma, S., Singh, S., Aich, S., & Kim, H.-C. (2021). Stock Market Prediction Using Machine Learning Techniques: A Decade Survey on Methodologies, Recent Developments, and Future Directions. *Electronics*, 10(21), Article 21. <https://doi.org/10.3390/electronics10212717>
- Schölkopf, B., Tsuda, K., Vert, J.-P., & others. (2004). *Kernel methods in computational biology*. MIT press.
- Shashua, A. (2009). Introduction to Machine Learning: Class Notes 67577. *CoRR*, abs/0904.3664. <http://arxiv.org/abs/0904.3664>
- Sutrisno, I. P. dan M. F. dan S. (2018). Penerapan Metode Support Vector Machine (SVM) Pada Klasifikasi Penyimpangan Tumbuh Kembang Anak. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 2(10), 3163–3169.
- Teixeira Zavadzki de Pauli, S., Kleina, M., & Bonat, W. H. (2020). Comparing Artificial Neural Network Architectures for Brazilian Stock Market Prediction. *Annals of Data Science*, 7(4), 613–628. <https://doi.org/10.1007/s40745-020-00305-w>
- Vafaeipour, M., Rahbari, O., Rosen, M., Fazelpour, F., & Ansarirad, P. (2014). Application of sliding window technique for prediction of wind velocity time series. *International Journal of Energy and Environmental Engineering*, 5. <https://doi.org/10.1007/s40095-014-0105-5>
- Wei, Y., Qin, S., Li, X., Zhu, S., & Wei, G. (2019). Oil price fluctuation, stock market and macroeconomic fundamentals: Evidence from China before and after the financial crisis. *Finance Research Letters*, 30, 23–29. <https://doi.org/10.1016/j.frl.2019.03.028>
- Wijaya, Y. B., Kom, S., & Napitupulu, T. A. (2010). Stock Price Prediction: Comparison of Arima and Artificial Neural Network Methods - An Indonesia Stock's Case. *2010 Second International Conference on Advances in Computing, Control, and Telecommunication Technologies*, 176–179. <https://doi.org/10.1109/ACT.2010.45>
- Yudhawan, D. H., & Purwaningsih, T. (2020). Developing support vector regression model to forecast stock prices of mining companies in Indonesia. *Jurnal Informatika*, 14(2), 43–48.