

ANALISIS TIME SERIES MENGGUNAKAN MODEL HYBRID ARIMA-SVR PADA HARGA BITCOIN

Wickly Gusthvi¹, Afrionaldi A. Roza², Afse Weldi Mallisa³, Aryanto⁴

^{1,2,3,4}Program Studi Statistika Fakultas MIPA Universitas Cenderawasih, Indonesia
e-mail: wickly23@gmail.com¹, afrionaldiroza19@gmail.com², Weldi1825@gmail.com³

ABSTRAK

Isu resesi global yang kemungkinan akan terjadi pada tahun 2023, membuat ketidakpastian pasar. Salah satu contoh pasar yang relative akan terdampak resesi adalah pasar Bitcoin. Mata uang virtual yang satu ini memiliki harga yang relative fluktuatif dan sulit ditebak sehingga diperlukan pendekatan secara ilmiah melalui analisis harga dari Bitcoin itu sendiri. Pendekatan analisis secara ilmiah ini diharapkan dapat diaplikasikan untuk memberikan gambaran harga Bitcoin di tahun 2023, sekaligus memberikan argumen ilmiah bagi investor dalam pengambilan keputusan investasi di pasar Bitcoin sehingga tidak terjadinya pelemahan pasar sebagai dampak dari resesi global. Oleh karena itu, menjawab permasalahan tersebut, dengan melihat pola data dengan rentang waktu dari 13 November 2017 hingga 17 Oktober 2022, serta asumsi adanya komponen linier dan nonlinier pada data harga Bitcoin, dan tujuan untuk peramalan harga Bitcoin itu sendiri, maka peneliti memutuskan untuk melakukan pendekatan melalui analisis deret waktu dan juga peramalan dengan membandingkan model ARIMA dengan model *Hybrid* ARIMA-SVR, kemudian memodelkan data Bitcoin menggunakan model *Hybrid* ARIMA-SVR. Proses penelitian dimulai dengan membentuk model ARIMA dari data training Bitcoin dan menentukan residual ARIMA. Apabila residual ARIMA nonlinier, maka dapat dibentuk model SVR dari residual ARIMA. Selanjutnya kemudian membentuk model *Hybrid* ARIMA-SVR. Hasil dari penelitian ini yaitu model ARIMA (2,1,1) dan parameter SVR yaitu parameter C , γ , dan ϵ berturut-turut sebesar 1, 0, dan 0.26. Serta dengan parameter yang ada membentuk model *Hybrid* ARIMA-SVR dan diperoleh nilai MAPE model ARIMA untuk data training dan testing sebesar 0.37348% dan 0.22356%, sedangkan untuk *Hybrid* ARIMA-SVR sebesar 0.36511% dan 0.21229%. Pada bagian terakhir, sebagai simulasi dilakukan prediksi harga Bitcoin beberapa hari kedepan dari tanggal 18 Oktober 2022 hingga 17 November 2022 dengan model tersebut.

Kata Kunci: Bitcoin, ARIMA, SVR, Hybrid ARIMA-SVR

1. PENDAHULUAN

Bitcoin adalah mata uang virtual yang mulai diperkenalkan pada 3 Januari 2009 oleh Satoshi Nakamoto. Bitcoin adalah mata uang virtual yang menggunakan teknologi *peer-to-peer* yaitu teknologi yang membuat bitcoin tidak memerlukan otoritas pusat atau bank sentral seperti mata uang pada umumnya. Pengelolaan transaksi dan penerbitan bitcoin dilakukan secara kolektif di dalam jaringan. Bitcoin tidak dikendalikan oleh siapapun sehingga memungkinkan semua orang dapat mengambil bagian di dalamnya bahkan ikut mengembangkan bitcoin. Transaksi menggunakan bitcoin bersifat *irreversible* (tidak dapat dikembalikan) dan juga *anonymous* (tidak diketahui).

Di Indonesia sendiri mata uang ini sedang diminati banyak orang sebagai salah satu alternatif investasi. Harga bitcoin yang fluktuatif, membuatnya banyak digunakan sebagai sarana mencari profit. Para investor mendapatkan keuntungan dari bitcoin dengan membelinya menggunakan mata uang rupiah pada saat harga mata uang bitcoin sedang turun, dan menjualnya kembali ke mata uang rupiah ketika harganya sedang naik. Meskipun demikian, adanya isu resesi yang akan terjadi pada tahun 2023 mengakitbatkan ketidakpastian pada harga Bitcoin dan relative menyebabkan melemahnya pasar Bitcoin juga, sehingga diperlukan solusi yang dapat mengurangi risiko kerugian bagi pemilik aset Bitcoin. Salah satu solusi yang dapat dilakukan adalah dengan melakukan prediksi harga Bitocin di masa mendatang menggunakan analisis deret waktu dan peramalan. Analisis deret waktu adalah metode kuantitatif untuk menentukan pola data masa lalu yang telah dikumpulkan secara teratur.

Pada analisis deret waktu, data dapat memiliki komponen linier maupun non linier, namun tidak menutup kemungkinan pada beberapa kasus data dapat memiliki kedua komponen tersebut secara bersamaan. Data Bitcoin diasumsikan memiliki kedua komponen tersebut, sehingga diperlukan model yang dapat mengantisipasi data yang memiliki komponen linier dan nonlinier sekaligus, sehingga diperlukan model kombinasi dari model linier dan nonlinier yaitu, *Hybrid ARIMA–SVR*. Model *Autoregressive Integrated Moving Average* (ARIMA), yang terdiri dari metode *Autoregressive* (AR) dan *Moving Average* (MA) dengan proses diferensiasi sebagai model linier, sedangkan metode *Support Vector Regression* (SVR) sebagai model nonlinier. Oleh karena itu, penelitian ini bertujuan untuk membandingkan metode ARIMA dengan metode *Hybrid ARIMA–SVR* pada data Bitcoin. Selain itu, untuk memodelkan data Bitcoin menggunakan model *Hybrid ARIMA–SVR*. Studi kasus pada penelitian ini menggunakan data harga penutupan Bitcoin harian dari 13 November 2017 sampai 17 Oktober 2022. Penentuan parameter terbaik model ARIMA menggunakan estimasi parameter yang signifikan dan asumsi residual *white noise*, sedangkan parameter model SVR menggunakan metode *Grid Search* dengan Kernel *Radial Basis Function* (RBF). Proses dari penelitian ini dimulai dengan pembagian data *training–testing* dan uji stasioneritas. Kemudian dibentuk model ARIMA terbaik dan menentukan residual ARIMA. Apabila residual ARIMA nonlinier, maka dapat dibentuk model SVR dari residual ARIMA. Model *Hybrid ARIMA–SVR* diperoleh dari penggabungan model ARIMA dan SVR. Apabila residual *Hybrid ARIMA–SVR* linier, maka model *Hybrid ARIMA–SVR* merupakan model terbaik

2. METODE PENELITIAN

Secara sistematis penelitian ini dibagi ke dalam beberapa tahapan penelitian yaitu, yakni mengumpulkan data, dilanjutkan dengan mengolah data, lalu usulan model, kemudian melakukan eksperimen dan pengujian, dan pada tahap akhir dilakukan evaluasi dan model dilakukan tahap validasinya.

2.1 Tahap Pengumpulan Data dan Pengolahan Data

Data yang digunakan pada penelitian ini adalah data sekunder. Data berupa harga bitcoin selama 1815 hari mulai dari tanggal 13 November 2017 sampai dengan 17 Oktober 2022 atau selama periode 5 tahun, yang diperoleh dari <https://id.investing.com/crypto/bitcoin/btc-idr-historical-data>. Analisis data dilakukan dengan menggunakan model *Hybrid ARIMA-SVR*, yang diimplementasikan menggunakan bahasa pemrograman *Python* dan *R*.

2.2 Teori Pendukung

STASIONERITAS DATA

Model stasioner pada data runtun waktu merupakan suatu model yang seluruh sifat-sifat statistiknya tidak berubah dalam pergeseran waktu. Stasioneritas data dibagi menjadi dua, yaitu stasioner dalam varians dan rata-rata. Pengujian data stasioner dalam varians menggunakan transformasi Box–Cox. Persamaan transformasi Box–Cox sebagai berikut:

$$Z_t = \begin{cases} \ln(Y_t), & \lambda = 0 \\ \frac{Y_t^\lambda - 1}{\lambda}, & \lambda \neq 0 \end{cases}$$

dengan λ adalah parameter transformasi, Y_t adalah data runtun waktu ke- t , dan Z_t adalah hasil transformasi data runtun waktu ke- t . Transformasi Box-Cox dapat dijelaskan dengan Tabel 1.

Tabel 1. Transformasi Box-Cox

Nilai λ	-1.0	-0.5	0	0.5	1.0
Transformasi	$\frac{1}{Y_t}$	$\frac{1}{\sqrt{Y_t}}$	$\ln(Y_t)$	$\sqrt{Y_t}$	Y_t

Apabila nilai λ bernilai satu, maka data tidak perlu melakukan transformasi. Data runtun waktu dengan λ sebesar satu disebut stasioneritas dalam varians. Cara untuk menguji stasioneritas data dalam ratarata adalah dengan menguji *Augmented Dickey-Fuller (ADF)* yang merupakan suatu uji akar unit. Apabila data tidak stasioner dalam rata-rata, maka dilakukan proses diferensiasi. Diferensiasi dapat dilakukan hingga beberapa periode sampai data runtun waktu bersifat stasioner. Notasi penting dalam proses diferensiasi adalah operator backward shift (B). Notasi B ditunjukkan pada persamaan berikut:

$$\begin{aligned} BY_t &= Y_{t-1} \\ BBY_t &= B^2Y_t = Y_{t-2} \\ B^dY_t &= Y_{t-d} \end{aligned}$$

Data runtun waktu yang telah didiferensiasi pada orde ke-2 ditunjukkan dengan persamaan berikut:

$$Y_t'' = Y_t' - Y_{t-1}' = (Y_t - Y_{t-1}) - (Y_{t-1} - Y_{t-2}) = Y_t - 2BY_t + B^2Y_t = (1 - B)^2Y_t$$

Jika data telah stasioner pada diferensiasi orde ke-d, maka persamaan diferensiasi yang diperoleh adalah sebagai berikut:

$$Y_t^{(d)} = (1 - B)^dY_t$$

dengan $Y_t^{(d)}$ merupakan data runtun waktu yang melalui proses diferensiasi orde ke-d.

UJI LINIERITAS

Data runtun waktu terdiri dari komponen linier dan nonlinier, apabila setelah dilakukan pemodelan linier pada data, residual dari model tersebut bersifat nonlinier. Pada penelitian ini, setelah data dimodelkan dengan ARIMA, dilakukan pengujian linieritas dari residual ARIMA menggunakan uji Terasvirta. Berikut ini adalah hipotesis dari uji Terasvirta:

H0: residual tidak mengandung pola nonlinier

H1: residual mengandung pola nonlinier

Statistik uji:

$$F_{hitung} = \frac{(SSR_0 - SSR_1)/m}{SSR_1/(n - p - m - 1)}$$

dengan SSR_0 adalah jumlah kuadrat residual, SSR_1 adalah jumlah kuadrat residual dengan pendekatan ekspansi taylor. n merupakan banyaknya data, p merupakan banyaknya variabel prediktor, dan m merupakan prediktor tambahan dari hasil pendekatan ekspansi taylor. Kriteria pengambilan keputusan adalah H0 ditolak apabila nilai F_{hitung} lebih besar disbanding $F_{(a,m,n-p-m-1)}$, atau nilai p -value kurang dari nilai taraf signifikansi (5%).

MODEL AUTOREGRESSIVE INTEGRATED MOVING AVERAGE (ARIMA)

Model ARIMA merupakan kombinasi dari model AR dengan orde p , MA dengan orde q , dan proses diferensiasi (d). Metode AR menyatakan suatu ramalan sebagai fungsi dari nilai-nilai sebelumnya dan nilai residual saat ini dari runtun waktu tertentu. Persamaan AR(p) adalah sebagai berikut:

$$Y_t = \phi_1Y_{t-1} + \phi_2Y_{t-2} + \dots + \phi_pY_{t-p} + \varepsilon_t$$

Metode MA menyatakan suatu ramalan sebagai fungsi dari nilai-nilai residual dari runtun waktu sebelumnya dan saat ini. Persamaan MA(q) adalah sebagai berikut:

$$Y_t = \varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2} - \dots - \theta_q\varepsilon_{t-q}$$

Secara umum, model ARIMA (p,d,q) untuk data runtun waktu dinyatakan pada Persamaan (1)

$$\phi_p(B)(1 - B)^dY_t = \theta_q(B)\varepsilon_t \quad (1)$$

dengan $\phi_p(B) = 1 - \phi_1B - \phi_2B^2 - \dots - \phi_pB^p$ adalah koefisien AR dan $\theta_q(B) = 1 - \theta_1B - \theta_2B^2 - \dots - \theta_qB^q$ adalah koefisien MA. $(1 - B)^d$ merupakan diferensiasi orde ke-d, dengan adalah operator *backward shift*, serta ε_t adalah residual data runtun waktu ke-t.

MODEL SUPPORT VECTOR REGRESSION (SVR)

Support Vector Regression (SVR) adalah algoritma yang dapat mengenali pola data runtun waktu yang dapat memprediksi komponen nonlinier dari data runtun waktu dengan fungsi kernel. Ide dasar dari SVR yaitu dengan membagi data menjadi data training dan testing. Data training bertujuan untuk menentukan suatu fungsi dengan batasan deviasi tertentu, sehingga dapat menghasilkan prediksi yang mendekati nilai aktual. Misalkan terdapat m data training, (x_i, y_i) dengan $i = 1, 2, \dots, m$. Data input merupakan data output adalah berdasarkan yang bersesuaian. Fungsi regresi dari metode SVR untuk komponen nonlinier dituliskan pada Persamaan

$$f(x_i) = \sum_{i=1}^m (a_i - a_i^*) \varphi(x_i) \varphi(x) + b, \quad (2)$$

dengan $f(x_i)$ merupakan fungsi SVR, a_i dan a_i^* merupakan koefisien lagrange, $\varphi(x_i)$ merupakan fungsi yang memetakan x_i , $\varphi(x)$ merupakan fungsi yang memetakan x , serta b adalah bias. x_i merupakan data input ke- i , sedangkan x merupakan data input untuk memetakan fungsi ke ruang fitur.

Fungsi regresi $f(x_i)$ diasumsikan dapat menempatkan seluruh titik (x_i, y_i) dengan akurasi ε . Pada umumnya $\varphi(x_i)$ pada Persamaan (2) tidak diketahui. Oleh karena itu, $\varphi(x_i)$ dapat diatasi dengan kernel yang ditunjukkan pada Persamaan (3).

$$K(x_i, x) = \varphi(x_i) \cdot \varphi(x) \quad (3)$$

Metode kernel adalah perkalian skalar dari $\varphi(x_i)$ dan $\varphi(x)$ dalam ruang fitur. Kernel RBF merupakan fungsi kernel yang digunakan ketika terdapat komponen nonlinier pada data. Fungsi kernel RBF dijelaskan pada Persamaan (4).

$$K(x_i, x) = \exp(-\gamma \|x_i - x\|^2) + b, \text{ dengan } \gamma > 0 \quad (4)$$

dengan $K(x_i, x)$ merupakan kernel RBF dan merupakan parameter gamma. Apabila fungsi kernel RBF pada Persamaan (3) dan (4) disubstitusikan ke Persamaan (2), maka menghasilkan fungsi regresi SVR pada Persamaan (5).

$$f(x_i) = \sum_{i=1}^m (a_i - a_i^*) \exp(-\gamma \|x_i - x\|^2) + b \quad (5)$$

dengan $0 < a_i \leq C, 0 < a_i^* \leq C$

SVR dengan fungsi kernel RBF memiliki tiga parameter yaitu parameter cost (C), gamma (γ), dan epsilon (ε). Metode yang digunakan untuk mencari parameter SVR yang optimal adalah metode *grid search*. *Grid search* merupakan metode untuk menentukan parameter optimal untuk model SVR. Metode *grid search* melalui dua tahap pengerjaan, yaitu tahap *loose grid* dan *finer grid* [10]. *Loose grid* adalah proses pemilihan nilai C dan γ dengan pangkat bilangan bulat. Pemilihan parameter C dan γ menggunakan urutan angka eksponensial. Misalnya rentang parameter C adalah $2^{-5}, 2^{-3}, \dots, 2^{15}$, sedangkan rentang parameter γ adalah $2^{-15}, 2^{-13}, \dots, 2^5$. Penentuan parameter ε menggunakan rentang dari 0,01 sampai 0,1 yang merupakan bilangan real. *Finer grid* merupakan proses tahapan lanjutan dari *loose grid* dengan pemilihan nilai C dan γ yaitu nilai pangkat berada pada rentang parameter optimum dari tahap *loose grid*. Parameter dari tahap *loose grid* tetap digunakan pada tahap *finer grid*. Pemilihan parameter optimum berdasarkan pada nilai *error* terendah

MODEL HYBRID ARIMA-SVR

Model *Hybrid ARIMA-SVR* merupakan kombinasi model ARIMA dan model SVR. Model *Hybrid ARIMA-SVR* dapat digunakan apabila diasumsikan model data runtun waktu terdiri dari komponen linier dan nonlinier. Model *Hybrid ARIMA-SVR* dinyatakan dalam persamaan berikut:

$$Y_t = L_t + N_t$$

Dengan Y_t adalah data runtun waktu ke- t , L_t merupakan komponen linier ke- t (ARIMA) dan N_t merupakan komponen nonlinier ke- t (SVR). Setelah menentukan model ARIMA, kemudian diperoleh residual ARIMA yang diasumsikan mengandung komponen nonlinier. Persamaan residual ARIMA sebagai berikut:

$$\varepsilon_t = Y_t - \hat{L}_t$$

Dengan ε_t merupakan residual ARIMA dan \hat{L}_t merupakan nilai estimasi dari model ARIMA dalam waktu ke- t . Nilai residual yang memuat komponen nonlinier dimodelkan dengan SVR. Ketika model ARIMA dan SVR telah diperoleh, maka hasil prediksi dengan model *Hybrid ARIMA-SVR* ditunjukkan dalam Persamaan (6)

$$\hat{Y}_t = \hat{L}_t + \hat{N}_t \quad (6)$$

<https://ejournal.fmipa.uncen.ac.id/index.php/CJSDS>

Dengan \hat{N}_t nilai estimasi dari model SVR dalam waktu ke-t dan \hat{Y}_t adalah nilai estimasi dari model *Hybrid ARIMA-SVR*.

AKURASI METODE PERAMALAN

Akurasi metode peramalan bertujuan untuk mengevaluasi hasil dari prediksi yang telah dilakukan. Penelitian ini menggunakan *Mean Absolute Percentage Error* (MAPE) sebagai akurasi dari metode peramalan. Berikut persamaan untuk menentukan nilai MAPE:

$$MAPE = \frac{\sum_{t=1}^N \left| \frac{Y_t - \hat{Y}_t}{Y_t} \right|}{n} \times 100\%$$

dengan \hat{Y}_t adalah hasil prediksi data runtun waktu ke-t dan n adalah banyaknya data. Berdasarkan penilaian MAPE, kesimpulan akurasi metode peramalan dijelaskan pada Tabel 2.

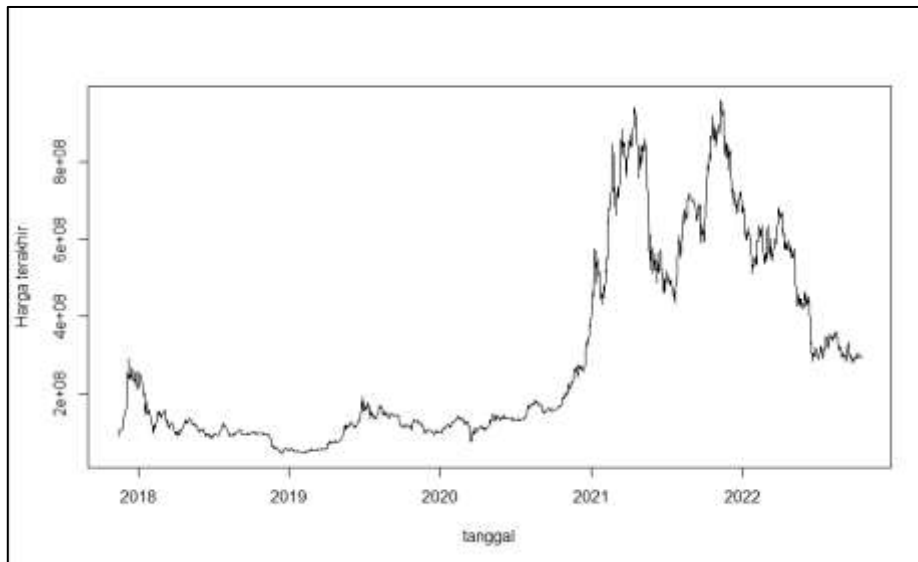
Presentae MAPE	≤ 10%	11% – 20%	21%-50%	>50%
Tingkat Akurasi	sangat akurat	akurat	kurang akurat	Tidak akurat

Tabel 2 Kriteria MAPE

3. HASIL DAN PEMBAHASAN

3.1 Studi Kasus

Data yang digunakan pada penelitian ini merupakan data sekunder, yaitu data harga penutupan Bitcoin (BTC) dari website *Investing.com*. Rentang waktu BTC dari 13 November 2017 sampai 17 Oktober 2022. Data BTC dinyatakan dalam variabel Y_t .



Gambar 1 Plot Data BTC

3.2 Pembentukan Model ARIMA

Langkah awal untuk membentuk model ARIMA yaitu menguji kestasioneritas data. Uji stasioner dalam varians dapat dilakukan dengan cara mencari nilai lambda (λ) data *training* BTC. Nilai λ yang diperoleh sebesar 0.163, sehingga perlu dilakukan transformasi. Data yang telah ditransformasi disebut Z_t . Setelah data telah stasioner dalam varians, maka dapat dilanjutkan pengujian kestasioneran data dalam rata-rata.

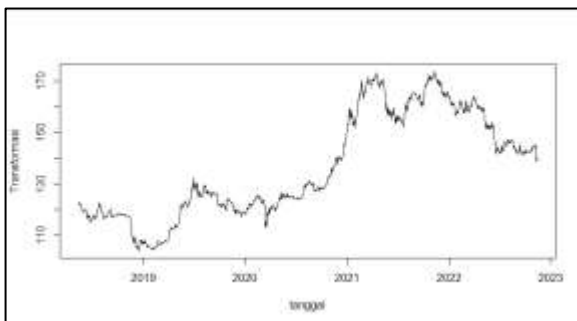
Uji stasioner dalam rata-rata dapat dilakukan dengan uji *Augmented Dickey-Fuller* (ADF) yang dijelaskan pada Tabel 3.

Statistik Uji	Nilai	
	Sebelum Diferensiasi	Sesudah Diferensiasi
Dickey-Fuller	-1.876	-10.8
<i>p-value</i>	0.6308	0.001

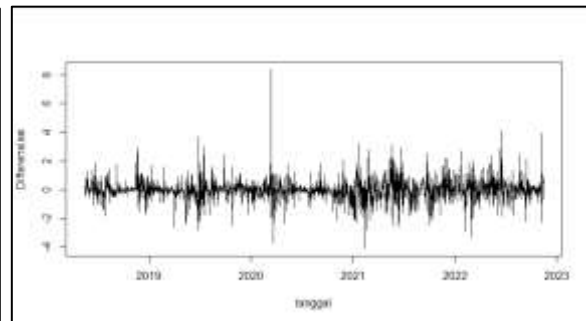
Tabel 3 Uji ADF Data Training Bitcoin

Tabel menunjukkan nilai *p-value* pada data *training* BTC sebelum diferensiasi lebih besar dibandingkan nilai α yaitu 0,05, sehingga H_0 tidak ditolak. Kesimpulannya adalah data tidak stasioner dalam rata-rata. Apabila data tidak stasioner dalam rata-rata, maka dilakukan diferensiasi untuk menstasionerkan data. Setelah melakukan diferensiasi, nilai *p-value* pada uji ADF lebih kecil dibandingkan nilai α , sehingga H_0 ditolak. Kesimpulannya adalah data *training* BTC telah stasioner dalam rata-rata. Plot data setelah ditransformasi dan diferensiasi ditunjukkan pada Gambar

Gambar 2 (a) Plot tranformasi BTC

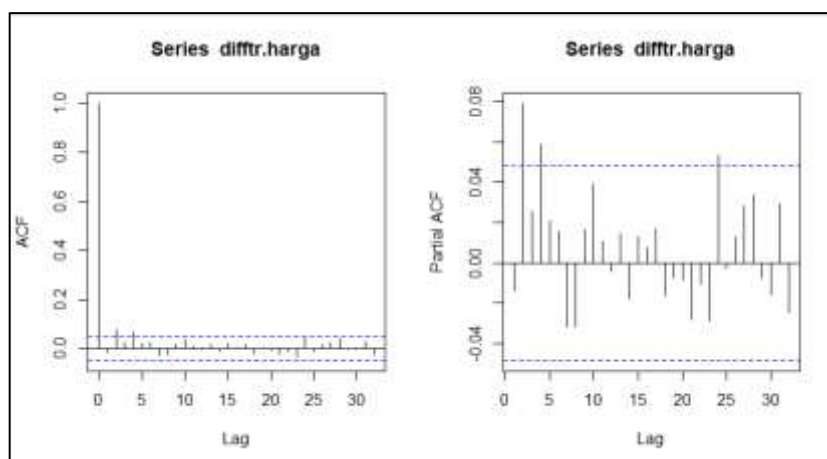


Gambar 2 (b) Diferensiasi BTC



Gambar 2(a) merupakan fluktuasi dari transformasi data *training* BTC, sedangkan Gambar 2(b) merupakan grafik diferensiasi dari hasil transformasi data *training* BTC. Berdasarkan Gambar 2(b) terlihat bahwa data telah berfluktuasi di sekitar suatu nilai rata-rata, memiliki varians yang konstan, serta tidak bergantung pada waktu.

Identifikasi model ARIMA didasarkan pada plot korelogram ACF dan PACF dari data BTC yang telah diferensiasi yang ditunjukkan dengan Gambar 3.



Gambar 3 Korelogram ACF dan PACF

Gambar 3 menunjukkan korelogram ACF dari data BTC yang dapat dilihat bahwa ACF terputus setelah lag ke-2. Korelogram PACF pada Gambar 3 menunjukkan bahwa PACF terputus setelah lag ke-2. Oleh karena itu, beberapa model ARIMA yang diperoleh dari korelogram ACF dan PACF adalah ARIMA (0,1,2), ARIMA (2,1,0), ARIMA (2,1,2), dan ARIMA (2,1,1). Estimasi parameter model ARIMA dinyatakan signifikan apabila seluruh parameter memiliki nilai $Pr(>|z|)$ kurang dari nilai taraf signifikan (0.05). Parameter model ARIMA dijelaskan pada Tabel 4.

Tabel 4. Estimasi Parameter Model Arima

No	Model	Parameter	Estimasi	Pr(> z)	Kesimpulan
1	ARIMA(0,1,2)	θ_1	-0.01601	0.5186	Tidak signifikan
		θ_2	0.07168	0.0021	Signifikan
2	ARIMA(2,1,0)	Φ_1	-0.01275	0.6051	Tidak Signifikan
		Φ_2	0.07926	0.0013	Signifikan
3	ARIMA(2,1,2)	Φ_1	0.20091	0.4421	Tidak Signifikan
		Φ_2	0.47399	0.0650	Tidak Signifikan
		θ_1	-0.21280	0.4352	Tidak Signifikan
		θ_2	-0.38897	0.1439	Tidak Signifikan
4	ARIMA(2,1,1)	Φ_1	0.58972	0.0001	Signifikan
		Φ_2	0.08466	0.0012	Signifikan
		θ_1	-0.60807	0.0001	Signifikan

Berdasarkan estimasi parameter dan uji diagnostik residual, maka ARIMA (2,1,1) merupakan model ARIMA terbaik. Persamaan model ARIMA (2,1,1) adalah:

$$\hat{L}_t = 0.58972(Y_{t-1}) - 0.58972(Y_{t-2}) + 0.08466(Y_{t-2}) - 0.08466(Y_{t-3}) + Y_{t-1} - 0.60807(\varepsilon_{t-1}) + 0.60807(\varepsilon_{t-2})$$

Kemudian *invers* nilai transformasi dengan λ sebesar 0.163. apabila nilai $Y_t = \sqrt[0.163]{0.163(Z_t) + 1}$, Persamaan dari model ARIMA(2,1,1), ditunjukkan pada persamaan (7):

$$\hat{L}_t = 0.58972(\sqrt[0.163]{0.163(Z_{t-1}) + 1}) - 0.58972(\sqrt[0.163]{0.163(Z_{t-2}) + 1}) + 0.08466(\sqrt[0.163]{0.163(Z_{t-2}) + 1}) - 0.08466(\sqrt[0.163]{0.163(Z_{t-3}) + 1}) + \sqrt[0.163]{0.163(Z_{t-3}) + 1} - 0.60807(\sqrt[0.163]{0.163(Z_{t-1}) + 1}) + 0.60807(\sqrt[0.163]{0.163(Z_{t-1}) + 1}). \quad (7)$$

dengan persamaan residual model ARIMA adalah $\varepsilon_t = Y_t - \hat{L}_t$. Z_t merupakan data transformasi, Y_t merupakan data asli, dan \hat{L}_t merupakan model ARIMA.

Pembentukan Model SVR

Residual ARIMA diubah ke dalam bentuk *time lag*. Pembagian *time lag* dibagi menjadi satu sampai 2 *lag*. Pemilihan *time lag* terbaik menggunakan *trial and error* dengan akurasi RMSE. Berdasarkan nilai RMSE, penduga model ke dua memiliki nilai RMSE terkecil, sehingga nilai *lag* kedua digunakan sebagai data *input* pada model SVR. Setelah menentukan *time lag* residual, maka dilanjutkan dengan uji linieritas residual dengan uji Terasvirta. Hasil dari uji Terasvirta dijelaskan pada Tabel 5.

Tabel 5. Uji Terasvirta Residual ARIMA

Statistik Uji	Nilai
<i>p-value</i>	0.04342

Nilai *p-value* pada uji Terasvirta sebesar 0.04342 yang lebih kecil dari nilai taraf signifikansi sebesar 0,05. Kesimpulannya adalah residual ARIMA mengandung pola nonlinier. Oleh karena itu, dapat dilanjutkan dengan metode SVR.

Parameter SVR yaitu *cost* (C), *gamma* (γ), dan *epsilon* (ε) ditentukan dengan metode *grid search*. Pada tahap *hyper grid*, rentang parameter *cost* dan *gamma* yaitu 0, 1, 2, 3, ..., 50, sedangkan parameter *epsilon*

yaitu 0.01, 0.02, ..., 5.00. Penentuan parameter optimal menggunakan nilai residual terkecil. Nilai residual kombinasi parameter SVR dapat dilihat pada Tabel 5.

Tabel 6. Rentang Tahapan *Hyper Grid* Metode *Grid search*

Parameter	Rentang Nilai
C	0, 1, 2, 3, ..., 50
γ	0, 1, 2, 3, ..., 50
ε	0.01, 0.02, ..., 5.00

Rentang nilai parameter yang digunakan pada tahapan *finer grid* dapat dilihat pada Tabel 6. Melalui proses *grid search* pada tahapan *hyper grid* diperoleh nilai optimal dari parameter model SVR adalah $C = 1$, $\gamma = 0$, dan $\varepsilon = 0.26$. Nilai parameter yang diperoleh menggunakan metode *grid search* pada tahapan *finer grid* merupakan parameter optimal yang digunakan pada model SVR. Sehingga model SVR terbaik yang diperoleh adalah model SVR menggunakan kernel *Radial Basic Function* (RBF) dengan parameternya adalah $C = 1$, $\gamma = 0$, dan $\varepsilon = 0.26$. Model SVR ditunjukkan pada Persamaan (8).

$$\hat{N}_t = \sum_{t=3}^{1634} (\alpha_t - \alpha_t^*) \exp(0\|x_t - x\|^2) + b, \text{ dengan } 0 < \alpha_t \leq 0.26, 0 < \alpha_t^* \leq 0.26. \quad (8)$$

HYBRID ARIMA-SVR

Setelah menentukan parameter terbaik dari model ARIMA dan SVR, selanjutnya adalah menggabungkan kedua model tersebut. Berdasarkan Persamaan (7) dan (8), maka Hybrid ARIMA–SVR ditunjukkan pada Persamaan (9).

$$\hat{Y}_t = \hat{L}_t + \hat{N}_t$$

$$\hat{Y}_t = 0.58972 \left(\sqrt[0.163]{0.163(Z_{t-1}) + 1} \right) - 0.58972 \left(\sqrt[0.163]{0.163(Z_{t-2}) + 1} \right) + 0.08466 \left(\sqrt[0.163]{0.163(Z_{t-2}) + 1} \right) - 0.08466 \left(\sqrt[0.163]{0.163(Z_{t-3}) + 1} \right) + \sqrt[0.163]{0.163(Z_{t-3}) + 1} - 0.60807 \left(\sqrt[0.163]{0.163(Z_{t-1}) + 1} \right) + 0.60807 \left(\sqrt[0.163]{0.163(Z_{t-1}) + 1} \right) + \sum_{t=3}^{1634} (\alpha_t - \alpha_t^*) \exp(0\|x_t - x\|^2) \quad (9)$$

dengan $0 < \alpha_t \leq 0.26, 0 < \alpha_t^* \leq 0.26$

Interpretasi Persamaan (9) adalah komponen linier pada data *training* BTC yang dimodelkan dengan ARIMA. Nilai BTC waktu ke- t dipengaruhi oleh nilai BTC satu sampai tiga hari sebelumnya dan residual nilai BTC satu sampai dua hari sebelumnya. Komponen nonlinier dimodelkan oleh SVR. Nilai BTC waktu ke- t dipengaruhi nilai residual ARIMA ke- t , dengan kernel Radial sebesar (γ) 0, dan koefisien *lagrange* diantara 0 sampai 0.26.

Setelah memperoleh model *Hybrid* ARIMA–SVR, dilakukan kembali pembagian *time lag* dan pengujian linieritas residual dari *Hybrid* ARIMA–SVR. Uji terasvirta yang dijelaskan pada Tabel 7.

Tabel 7. Uji Terasvirta Residual *Hybrid* ARIMA–SVR

Statistik Uji	Nilai
<i>p-value</i>	0.1342

Berdasarkan Tabel 7, kesimpulan yang diperoleh yaitu residual dari *Hybrid* ARIMA–SVR telah linier, sehingga model *Hybrid* ARIMA–SVR tersebut merupakan model terbaik.

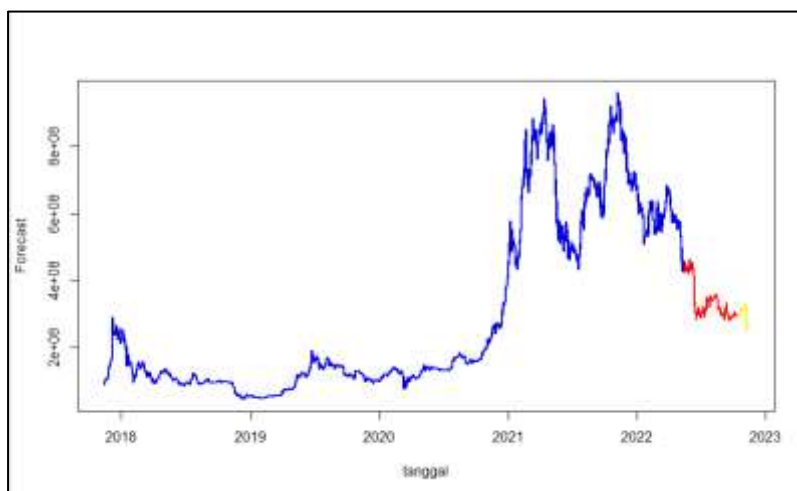
PERBANDINGAN MODEL ARIMA DAN HYBRID ARIMA-SVR

Perbandingan model ARIMA dan *Hybrid* ARIMA–SVR menggunakan nilai MAPE yang dijelaskan pada Tabel 5.

Tabel 8. Nilai MAPE Model ARIMA dan *Hybrid* ARIMA–SVR

Model	Data training	Data testing
ARIMA	0.37348%	0.22356%
<i>Hybrid</i> ARIMA-SVR	0.36511%	0.21229%

Tabel 8 menunjukkan nilai MAPE pada model ARIMA dan *Hybrid* ARIMA–SVR. Berdasarkan Tabel 8, nilai MAPE model *Hybrid* ARIMA–SVR lebih kecil dibandingkan model ARIMA, sehingga model *Hybrid* ARIMA–SVR merupakan model terbaik. Grafik prediksi dengan *Hybrid* ARIMA–SVR ditampilkan pada Gambar 4.



Gambar 4 Grafik Hasil Prediksi *Hybrid* ARIMA–SVR

Gambar 4 menunjukkan fluktuasi BTC apabila diprediksi dengan model *Hybrid* ARIMA–SVR. Data *training* dan *testing* berturut-turut ditunjukkan dengan warna biru dan merah, sedangkan data prediksi 30 hari kedepan dari tanggal 18 Oktober 2022 hingga 17 November 2022 ditunjukkan dengan warna kuning. Pola dari data *training* dan *testing* BTC mengikuti pola data asli (berwarna hitam) dengan baik.

4. SIMPULAN DAN SARAN

Metode *Hybrid* ARIMA–SVR merupakan metode gabungan dari ARIMA dan SVR, dengan metode ARIMA bertujuan untuk mengatasi komponen linier dan metode SVR yang mengatasi komponen nonlinier. Studi kasus yang digunakan pada penelitian adalah Bitcoin dengan rentang 13 November 2017 sampai 13 November 2022. Kesimpulan yang diperoleh yaitu model *Hybrid* ARIMA–SVR, dengan ARIMA (2,1,1) dan parameter optimal SVR yaitu parameter C sebesar 1, parameter γ sebesar 0, dan parameter ε sebesar 0.26. Model *Hybrid* ARIMA–SVR merupakan model terbaik, karena memiliki nilai MAPE yang lebih baik dibandingkan model ARIMA. Nilai MAPE yang diperoleh dari model *Hybrid* ARIMA–SVR untuk data training dan data testing berturut-turut sebesar 0.36511% dan 0.21229%.

DAFTAR PUSTAKA

- [1]. Wei WWS. Time Series Analysis: Univariate and Multivariate Methods. Second Edition. New York: Addison Wesley; 2006.
- [2]. Purnama DI. Prediksi Harga Emas Saat Pandemi Covid-19 menggunakan Model Hybrid Autoregressive Integrated Moving Average–Support Vector Regression. Jambura Journal of Mathematics. 2021; 3(1):52–65.
- [3]. Fitria IM, Darwis S. Pemodelan menggunakan Autoregressive Integrated Moving Average (ARIMA) dan Support Vector Regression (SVR) pada Vibrasi Bearing. Prosiding Statistika. 2020; 6(2):49–56.
- [4]. Zhang Y, Luo L, Yang J, Dunhu L, Kong R, Feng Y. A Hybrid ARIMA–SVR Approach for Forecasting Emergency Patient Flow. Journal of Ambient Intelligence and Humanized Computing. 2018; 10(1):3315–3323.
- [5]. Rusdi. Uji Akar-Akar Unit dalam Model Runtun Waktu Autoregresif. Statistika. 2011; 11(2):67–78.

- [6]. Prabowo H, Suhartono, Prastyo, DD. The Performance of Ramsey Test, White Test, and Terasvirta Test in Detecting Nonlinearity. *Inferensi*. 2020; 3(1):2721–3862.
- [7]. Yasin H, Prahutama A, Utami TW. Prediksi Harga Saham menggunakan Support Vector Regression dengan Algoritma Grid Search. *Media Statistika*. 2014; 7(1):29–35.
- [8]. Sriyana, Martha S, Sulistyaningsih E. Prediksi Nilai Tukar Dolar Amerika Serikat Terhadap Rupiah Dengan Metode Support Vector Regression (SVR). *Bimaster*. 2019; 8(1): 1–10.
- [9]. Vapnik VN. *The Nature of Statistical Learning Theory*. New York: Springer; 1995.
- [10]. Purnama DI, Hendarsin OP. Prediksi Jumlah Penumpang Berangkat melalui Transportasi Udara di Sulawesi Tengah menggunakan Support Vector Regression (SVR). *Jambura Journal of Mathematics*. 2020; 2(2):49–59.
- [11]. Hsu CH, Chang CC, Lin CJ. *A Practical Guide to Support Vector Classification*. Taipei: Department of Computer Science National Taiwan University; 2016.
- [12]. Yahoo Finance. Histori Harga Penutupan Indeks Harga Saham Gabungan [Internet]. 2021 [cited 2021 Nov 2]. Available from: [finace.yahoo.com/quote/%5EJKSE/history/](https://finance.yahoo.com/quote/%5EJKSE/history/).
- [13]. Clara Vista Magdalena Sihombing, Shantika Martha, Nur'ainul Miftahul Huda. Analisis Metode Hybrid Arima–Svr Pada Indeks Harga Saham Gabungan. *Jurusan Matematika FMIPA Untan, Pontianak. Bimaster*. 2022; 11(3):413-442
- [14]. Investing.com. Histori Harga Bitcoin Rupiah Indonesia [Internet]. 2022 [cited 2022 Nov 1]. Available from: <https://id.investing.com/crypto/bitcoin/btc-idr-historical-data>